HUMAN RETROVIRUSES and AIDS 1994

*A COMPILATION AND ANALYSIS OF
NUCLEIC ACID AND AMINO ACID SEQUENCES*

Editors

Gerald Myers
Theoretical Division
Los Alamos National Laboratory

Bette Korber
Theoretical Division
Los Alamos National Laboratory

Simon Wain-Hobson
Laboratory of Molecular Retrovirology
Pasteur Institute

Kuan-Teh Jeang
Molecular Virology Section
National Institutes of Health

Louis E. Henderson
Frederick Cancer Research Facility
National Cancer Institute

George N. Pavlakis
Frederick Cancer Research Facility
National Cancer Institute

Database and Analysis Staff

Ethan Allen, Jonathan Bingham, Barton Bryant, Charles Calef, Kersti MacInnes
Los Alamos National Laboratory

This publication is being funded by the Division of AIDS, National Institute of Allergy and Infectious Diseases, through an interagency agreement with the U.S. Department of Energy.

*Published by Theoretical Biology and Biophysics
Group T-10, Mail Stop K710
Los Alamos National Laboratory, Los Alamos, New Mexico 87545 U.S.A.*

LA-UR 94-4171

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ff

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

CONTENTS

More detailed tables of contents are to be found within the various parts of the compendium. The following provides merely an overview.

ACKNOWLEDGMENTS	ii
INTRODUCTION	iii
GLOSSARY and LANDMARKS	v
PART I. NUCLEIC ACID ALIGNMENTS AND SEQUENCES	
Introduction	I-1
Contents	I-3
A. HIV-1 Sequences	
Nucleotide Alignments and Consensus Sequences	I-A-2
Entries	I-A-147
B. HIV-2/SIV Sequences	
Nucleotide Alignments and Consensus Sequences	I-B-1
Entries	I-B-56
C. SIVAGM, SIVMND, and SIVSYK Sequences	
Nucleotide Alignments and Consensus Sequences	I-C-1
Entries	I-C-20
PART II. AMINO ACID ALIGNMENTS	
Introduction	II-1
Contents	II-2
A. HIV-1 Sequences	
Amino Acid Alignments and Consensus Sequences	II-A-2
B. HIV-2/SIV Sequences	
Amino Acid Alignments and Consensus Sequences	II-B-1
C. SIVAGM, SIVMND, and SIVSYK Sequences	
Amino Acid Alignments and Consensus Sequences	II-C-1
PART III Analysis	
Contents	III-1
A. Gag Proteins	III-2
B. HIV-1 Tat: Structure and Function	11
C. HIV-1 and HIV-2 Sequence Subtypes	23
D. HIV Synonymous and Nonsynonymous Substitution Frequencies	34
E. Global Variation in the HIV-1 V3 Region	37
F. Mutations in HIV-1 RT and Protease	93
PART IV. RELATED SEQUENCES	
Introduction	IV-1
Contents	IV-2
HIV and Cellular Factors	3
Sequence Entries	13
PART V. DATABASE COMMUNICATIONS	
Introduction	V-1
Roadmap to FTP Server	V-2

ACKNOWLEDGMENTS

The HIV Sequence Database and Analysis Project is funded by the Vaccine and Prevention Research Program of the AIDS Division of the National Institute of Allergy and Infectious Diseases (Dr. James Bradac, Project Officer) through an interagency agreement with the U.S. Department of Energy.

We thank the many researchers who have made their sequences available prior to publication. Of particular note, the Institute of Tropical Medicine in Antwerp allowed us to use their unpublished O group sequences for the figure shown on the cover of this compendium.

The summarization of RT and protease mutations by John Mellors, Raymond Schinazi, and Brendan Larder is updated from their publication in *International Antiviral News*. We thank them and the publishers for providing us with this section. Furthermore, we are pleased to report that John Mellors has agreed to come aboard as a database editor in 1995, for the purpose of continuing this effort.

We are grateful to Kim Millman, Andrew Farmer, Miranda McEvilly, and Anne Yesley, research assistants who have other roles in our group, for assisting the HIV Database staff with this publication.

INTRODUCTION

This compendium and the accompanying floppy diskettes are the result of an effort to compile and rapidly publish all relevant molecular data concerning the human immunodeficiency viruses (HIV) and related retroviruses. The scope of the compendium and database is best summarized by the five parts that it comprises: (I) Nucleic Acid Alignments and Sequences; (II) Amino Acid Alignments; (III) Analysis; (IV) Related Sequences; and (V) Database Communications. Information within all the parts is updated at least twice in each year, which accounts for the modes of binding and pagination in the compendium.

While this publication could take the form of a review or sequence monograph, it is not so conceived. Instead, the literature from which the database is derived has simply been summarized and some elementary computational analyses have been performed upon the data. Interpretation and commentary have been avoided insofar as possible so that the reader can form his or her own judgments concerning the complex information. In addition to the general descriptions below of the parts of the compendium, the user should read the individual introductions for each part.

Part I. Nucleic Acid Alignments and Sequences. Annotated nucleic acid sequences of HIV and SIV are presented in a form close to that of the GenBank Sequence Library. Our few modifications of standard GenBank format were instituted to better serve the particular community for which this database is intended.

The LOCUS name or identifier of an entry may differ slightly from that found in the GenBank or EMBL libraries, but the ACCESSION numbers are identical for entries in all three databases. Thus each entry is universally and uniquely traceable. Large sets of PCR-derived sequences are represented using a single sequence entry followed by an alignment of all the sequences in the set. When available, the entire range of ACCESSION numbers for the PCR set is usually indicated within the representative entry. The SOURCE line provides information, when available, about the infectivity or biological activity of the molecular clone from which a sequence has been derived. REFERENCES are limited to literature or personal communications having authority for the original sequence data; references that review sequence information, or that shed light upon the function or variation of coding and regulatory sequences, are listed in Part V.

Entries in Part I are annotated within the sequence, while their GenBank or EMBL-formatted versions on the floppy diskettes make use of FEATURES tables. The hard-copy annotation includes coding regions, regulatory structures, splice sites, and other features of functional significance. The authority for this annotation is largely invariance, the recurrence of patterns such as TATAA and AATAAA. Although our practice has been to conservatively annotate, we caution the user against docility: sequence information regarding transcripts, for example, is far from certain or complete at this time. Part I is divided into three subsections, A, B, and C, concerned with HIV-1s, HIV-2/SIVs, and SIVAGMs, SIVMND and SIVSYK.

Part II. Amino Acid Alignments. This section contains in alignment the amino acid sequences of all known coding regions and open reading frames of HIV-1, HIV-2, and SIVs. Beginning in 1993, Part II is broken down into subsections A, B and C, for HIV-1s, HIV-2/SIVs and SIVAGMs respectively. Consensus sequences and AACC consensus-like patterns are given with their respective alignments. Protein processing sites are annotated when known. The reader should consult the introduction to Part II for further explanation of the presentation and annotation of the amino acid sequences.

Part III. Analysis. This section is open-ended with the constraint that the sequence analyses and compilations be basic and of interest to the diversity of users. Currently the analyses include phylogenetic trees, HIV V3 Loop amino acid sequences, summaries of protein variability and molecular evolutionary patterns of the various regions of the HIV and SIV genomes, and envelope glycosylation patterns.

Introduction

Part IV. Related Sequences. Heretofore, this section of the compendium has featured HIV related viral sequences—of nonprimate lentiviruses and the human T-cell lymphotropic viruses. Beginning in 1993, Part IV entries include, with greater emphasis, coding sequences for cellular proteins involved with HIV pathogenesis. Entries are presented in the same format as Part I entries.

Part V. Database Communications. In previous years this part contained a printed supplemental reference list which has been omitted this year to save space. As usual, diskettes of sequences are provided. The floppy diskettes contain the nucleic acid sequences from Part I and Part IV and the translated amino acid sequences of coding regions. A number of sequences not printed in this compendium may be found on the diskettes. Those sequences are listed in the table of contents of Part I. Beginning in 1993, sequence alignments are also included to meet user requests. For the most current information regarding database files, see the README file on each diskette. Nucleotide entries are presented in GenBank format for North American users and in EMBL format for European users (unless otherwise requested). Similarly, amino acid sequences are in either PIR or Swiss-Prot format. The diskettes themselves are either 5.25" IBM-DOS format or 3.5" Macintosh format, depending upon what has been requested. If there is any trouble using these files with software designed to work with the format we have sent, please let us know the name of the program you are using and the file that it could not handle.

A comprehensive compilation of the nucleic acid and protein sequences published in the Human Retroviruses and AIDS Database since 1987 is available in the NCBI Data Repository through Internet FTP (File Transfer Protocol). To transfer data over the Internet via anonymous FTP type: `ftp.ncbi.nlm.nih.gov` or `ftp.130.14.20.1`. Enter anonymous for the login name and your electronic mail address for the password. Change directory to 'repository/aids-db' to download the entries from the AIDS database.

The NCBI also offers several CD-ROMs containing HIV and SIV sequence information. The ENTREZ:Sequences CD contains the GenBank, NCBI backbone, PIR, SWISS-PROT, and MEDLINE databases with retrieval software for users with Macintosh and PC-compatible computers running Microsoft Windows. The second CD, NCBI-GenBank, will contain the complete GenBank database in the currently distributed flatfile format. Each release of this CD will contain a full GenBank release with new data incorporated from direct submissions, the NCBI Backbone (journal scanning), and the EMBL and DDBJ databases. The third CD, NCBI-Sequences, will include the integrated dataset provided on the Entrez:Sequences CD in the ASN.1 standard data description format. Questions about services or software provided by the NCBI may be directed to the NCBI via phone: (301) 496-2475, fax: (301) 480-9241, or email: `info@ncbi.nlm.nih.gov`.

We are prepared to quickly enter both protein and nucleotide sequences into the Human Retroviruses and AIDS database, and in the case of nucleotide sequences, oversee their entry into the GenBank and EMBL libraries. Submission of unpublished sequences is invited and encouraged. Sequence data or inquiries regarding the database should be addressed to

Gerald Myers
Theoretical Division
T-10, MS K710
LANL
Los Alamos, NM 87545

(505)-665-0480; fax (505)-665-3493
e-mail: `glm%termin@lanl.gov`

A short glossary follows.

GLOSSARY
HIV/SIV PROTEINS

Name	Typical Size	Function	Localization
GAG	p55/p17/p24/p7 p6/p2/p1	structural capsid proteins	virion
POL	p66/p51/p32	viral enzymes	virion
ENV	gp120/gp41	viral surface (SU) and transmembrane (TM) glycoproteins	plasma membrane, virion envelope
TAT	p16/p14	viral transactivator	primarily in nucleolus/nucleus
REV	p19	RNA transport, stability and utilization factor (phosphoprotein)	primarily in nucleolus/nucleus
VIF	p23	promotes virion infectivity	Golgi?
VPR	p10-15	unknown, also called 'rap' (rapid)	virion
VPU	p16	promotes extra- cellular release of viral particles (phosphoprotein)	integral membrane protein
NEF	p27-p25	unknown (myristoylated protein)	cytoplasm, plasma membrane
VPX	p12-16	unknown, appears to be a duplication of VPR in HIV-2 and SIVs only	virion
TEV	p28	tripartite tat-env-rev protein (found only in HXB2; also named trv)	primarily in nucleolus/nucleus

LANDMARKS ON THE HIV-1 GENOMIC RNA

PROTEINS

RNA SITES

SPICE SITES AND EXONS

Landmarks

LANDMARKS:

HIV GENOMIC STRUCTURAL ELEMENTS

- LTR** Long terminal repeat, the DNA sequence flanking the genome of integrated proviruses. It contains important regulatory regions, especially those for transcription initiation (upstream and downstream promoters) and polyadenylation.
- TAR** Target sequence for viral transactivation, the binding site for Tat protein and for cellular proteins; consists of approximately the first 45 nucleotides of the viral mRNAs in HIV-1 (or the first 100 nucleotides in HIV-2 and SIV.) TAR forms a hairpin stem-loop structure with a side bulge; the bulge is necessary for Tat binding and function. May also be a DNA regulatory site.
- RRE** Rev responsive element (also named CAR), an RNA element encoded within the env region of HIV-1. It consists of approximately 200 nucleotides (positions 7327 to 7530 from the start of transcription in HIV-1HXB2R.) RRE is necessary for Rev function; it contains a high affinity site for Rev; in all, approximately seven binding sites for Rev exist within the RRE RNA. Other lentiviruses (HIV-2, SIV, visna, CAEV) have similar RRE elements in similar locations within env, while HTLVs have an analogous RNA element (RXRE) serving the same purpose within their LTR; RRE is the binding element for Rev protein, while RXRE is the binding element for Rex protein. RRE (and RXRE) form complex secondary structures important for specific protein binding.
- CRS** cis-acting repressive sequences postulated to inhibit structural protein expression in the absence of Rev. One such site was mapped within the pol region of HIV-1. The exact function has not been defined; splice sites have been postulated to act as CRS sequences.
- INS** Instability RNA sequences found within the structural genes of HIV-1 and of other complex retroviruses. Multiple INS elements exist within the genome and can act independently; one of the best characterized elements spans nucleotides 414 to 631 in the gag region of HIV-1HXB2R.

GENES AND GENE PRODUCTS

- GAG** group specific antigens or capsid proteins. The precursor is the p55 myristoylated protein, which is processed to p17 (Matrix) p24 (Capsid) and p7 (NucleoCapsid) proteins by the viral protease. Other small proteins are generated from the gag polyprotein by protease cleavage.
- POL** generates the viral enzymes protease, reverse transcriptase and integrase after processing of a gag-pol precursor polyprotein by the viral protease; the Gag-Pol precursor is produced by ribosome frameshifting.
- ENV** viral glycoproteins produced as a precursor (gp160) and processed to the external glycoprotein gp120 and the transmembrane glycoprotein gp41. The mature proteins are held together by noncovalent interactions; as a result, a substantial amount of gp120 is released (shed) into the medium. gp120 contains the binding site for the CD4 receptor.
- TAT** Transactivator of HIV gene expression. One of the two necessary viral regulatory factors (Tat and Rev) for HIV gene expression. Two forms are known, Tat-1exon (minor form) of 72 amino acids and Tat-2exon (major form) of 86 amino acids. The electrophoretic mobility of these two forms in SDS gels is anomalous, with apparent sizes of approximately 16 kD and 14 kD for Tat-2exon and Tat-1exon, respectively. Low levels of both proteins are found in persistently infected cells. Tat has been localized primarily in the nucleolus/nucleus by immunofluorescence. It acts by binding to the TAR RNA element and activating transcription initiation and/or elongation from the

LTR promoter. It is the first eukaryotic transcription factor known to interact with RNA rather than DNA and may have similarities with prokaryotic anti termination factors.

REV The second necessary regulatory factor for HIV expression. A 19 kD phosphoprotein, localized primarily in the nucleolus/nucleus, Rev acts by binding to RRE and promoting the nuclear export, stabilization and utilization of the viral mRNAs containing RRE. Rev is considered the most functionally conserved regulatory protein of lentiviruses (although it is mutationally quite unstable).

VIF Viral infectivity factor, typically 23 kD. Required for the efficient transmission of cell-free virus in tissue culture. In the absence of Vif the produced viral particles are defective, while the cell-to-cell transmission of virus is not affected significantly. Accumulates in the Golgi system and is not found in the virion.

VPR Virion associated protein of unknown function found in HIV-1, HIV-2, SIVmac and SIVmnd; typically 15 kD. Probably homologous to VPX of HIV-2/SIVs.

VPU "oUr" protein, promotes extracellular release of viral particles. Found only in HIV-1. Integral membrane phosphoprotein of 16 kD; similar to M2 protein of influenza virus; may be involved in env maturation; not found in the virion.

NEF Approximately 27 kD non-virion protein found in the cytoplasm of infected cells. It is myristoylated and associated with the inner plasma membrane. One of the first HIV proteins to be produced in infected cells, it is the most immunogenic of the accessory proteins. During tissue culture propagation, Nef is dispensable and probably counter-selected. On the contrary, SIV Nef is required for viral propagation in animals. Nef affects the levels of CD4 and may modulate intracellular signaling pathways.

VPX Virion protein of 12 kD found only in HIV-2/SIVs. Function unknown. Vpu/Vpx may be used to distinguish between HIV-1 and HIV-2 infection.

TEV (also named *tnv*) Tripartite 28 kD viral phosphoprotein produced by some HIV-1 strains. Found primarily in the nucleolus/nucleus. Tev contains the first exon of Tat, a small part of Env and the second exon of Rev. It exhibits both Tat and Rev functions and can functionally replace both essential regulatory proteins of HIV-1. It is produced very early in infection.

STRUCTURAL PROTEINS/VIRAL ENZYMES The products of gag, pol and env genes, which are essential components of the retroviral particle.

REGULATORY PROTEINS Tat and Rev proteins of HIV/SIV and Tax and Rex proteins of HTLVs. Modulators of transcriptional and posttranscriptional steps of virus gene expression; essential for virus propagation.

ACCESSORY OR AUXILIARY PROTEINS additional virion and non-virion-associated proteins produced by HIV/SIV retroviruses: Vif, Vpr, Vpu, Vpx, Nef. Their exact function is mostly unknown. Although the accessory proteins are in general not necessary for viral propagation in tissue culture, they have been conserved in the different isolates; this conservation and experimental observations suggest that their role in vivo is very important.

COMPLEX RETROVIRUSES Retroviruses regulating their expression via viral factors and expressing additional proteins (regulatory and accessory) essential for their life cycle. By one interpretation, all retroviruses are "complex" and HIV is said to be merely one of the "more complex."

Nucleic Acid Alignments and Sequences

Nucleotide sequence alignments were generated using the PIMA program developed by Randy Smith and Temple Smith (*Protein Engineering* 5:35, 1992). By this approach, a protein pattern is created which then serves as a template for the nucleotide sequence alignment. Hand-editing, when required, was done with the MASE program (D. V. Faulkner and J. Jurka, *TIBS* 13:321, 1988). For information concerning either of these programs, contact Dr. Randall Smith, Institute of Molecular Genetics, Baylor College of Medicine, Houston, Texas, 713-798-4735.

With few exceptions, only full-length coding sequences were included in these alignments. Mixed case consensus sequences are used as the reference sequences for each alignment. Upper case letters indicate 100% conservation of nucleotide bases in a given position of the alignment, and lower case letters represent bases conserved in at least 50% of the sequences. To reduce the alignment uncertainty with the large HIV-1 set, sequence subtypes (see below and Part III) were constituted and consensus sequences were generated for each. Divergent sequences that could not be easily assigned to any given subtype are shown at the bottom of each block. Large deletions, if present in only one sequence (e.g., the HIV2nhz *nef* coding region sequence), were also excluded from the consensus sequences.

In the 1994 compendium, HIV-1 alignments are presented for the *gag*, *tat*, and *env* coding sequences only, as little new information was obtained over the other coding regions. HIV-2 and related SIV sequence alignments are over *gag*, *pol*, and *env* coding sequences only. SIV-AGM sequence alignments are shown for the LTR and *env* region. All amino acid sequence alignments are given in Part II and all nucleotide sequence alignments are available on the FTP server. The 1993 compendium presented alignments for all coding sequences.

Prior to the 1992 database compendium, HIV-1 sequences had been roughly classified as 'U.S.' and 'African.' In light of many new HIV-1 *gag* and *env* sequences, it became more useful, starting with the 1992 compendium, to categorize HIV-1 sequences into five sequence subtypes, depending upon the coding sequence. HIV-1 subtypes now number eight (A through H). Collectively, these are called group M sequences, as they are significantly distinct from group O sequences. The bases for this classification, discussed at greater length in Part III, are:

- i) subtypes are approximately equidistant from one another in *env* (a "star phylogeny");
- ii) the *env* phylogenetic tree is for the most part congruent with *gag* phylogenetic trees;
- iii) two or more samples are required to define a sequence subtype.

Subtype naming problems have arisen for several reasons. For example, viral samples such as MAL and TN243 appear to be hybrid, clustering with one sequence subtype in *gag* and another sequence subtype in envelope. We ask database users to be patient with these problems, most of which are unavoidable (recombination) or transitory (same name for different variants). In the case of hybrid viruses, gene sequences are aligned in this compendium according to the best fit with one of the existing subtypes.

The majority of HIV-1 variants being characterized in the centers of the pandemic fall into one of the five sequence subtypes, A through E, based upon *env* coding sequences. The discovery of the MVP5180 virus in Cameroon, its close relationship to ANT70 and the likelihood that 5% to 10% of infected Cameroonians carry this virus, is an interesting exception. These viral sequences are classified as group O.

There are at least five sequence subtypes for the HIV-2s, A through E.

The authority for some of the annotation is limited largely to invariance—the recurrence of patterns such as AATAAA, for example. The reader should be cautious in drawing upon this information.

Due to space limitations, only certain sequences are reported as entries in this section. The Table of Contents lists other sequences, in addition to these, that are available on diskettes and the FTP server. Short sequences, typically generated through PCR, of necessity must usually be presented as aligned sets rather than as GenBank-formatted entries, however accession numbers can usually be provided for the sequences that appear in alignment. Sequence entries from previous years are presented in the 1990–1993 compendiums. The Tables of Contents for those follow directly below.

PART I HIV and SIV Nucleic Acid Sequences

Introduction	I-1
Contents(1994)	I-3
Contents(1993)	I-7
Contents(1992)	I-10
Contents(1991)	I-12
Contents(1990)	I-14
A. HIV-1 Sequences	
Nucleotide alignments of HIV-1 <i>gag</i>	I-A-2
Nucleotide alignments of HIV-1 <i>tat</i>	I-A-50
Nucleotide alignments of HIV-1 <i>env</i>	I-A-56
HIVAMK101 (U08192)	147
HIVAMK82 (U08193)	148
HIVCA1 (X80438)	149
HIVCAR286A (pending)	155
HIVCAR4017 (pending)	156
HIVCAR4020 (pending)	157
HIVCAR4023 (pending)	158
HIVCAR4039 (pending)	159
HIVCAR4054 (pending)	160
HIVCAR4067 (pending)	161
HIVCAR4071 (pending)	162
HIVCAR4081 (pending)	163
HIVCAR423A (pending)	164
HIVCARELO (pending)	165
HIVCARGAN (pending)	166
HIVCARMBA (pending)	167
HIVCARSAS (pending)	168
HIVCI47 (X72049)	169
HIVD808 (U07012)	184
HIVDJ2581 (L11763)	189
HIVDJ2582 (L22939)	190
HIVDJ2591 (L11764)	192
HIVDJ2592 (L22940)	193
HIVDJ263A (L22941)	195
HIVDJ264A (L23064)	197
HIVDJ373A (L23065)	199
HIVGER (U11549)	201
HIVHYDA (L29091)	202
HIVK1241 (L11769)	203
HIVK124A2 (L22942)	204
HIVK986 (U12984)	206
HIVKENYA (L22943)	210
HIVN759 (U15576)	212
HIVRJ484 (U00409)	215
HIVRJSP3 (U08975)	219
HIVSE364A (L22944)	224
HIVSE3651 (L11797)	226
HIVSE365A2 (L22945)	227

HIVSM1451 (L11803)	229
HIVSM145A2 (L22946)	230
HIVUG2661 (L11798)	232
HIVUG266A2 (L22947)	233
HIVUG2681 (L11799)	235
HIVUG268A2 (L22948)	236
HIVUG269A (L22949)	238
HIVUG273A(L22957)	240
HIVUG2741 (L11801)	242
HIVUG274A2 (L22950)	243
HIVUG275A (L22951)	245
HIVYAU (X80020)	247
HIVYH1911 (L11783)	249
HIVYH191A2 (L22952)	250
HIVYI5251 (L11792)	252
HIVYI525A2 (L22953)	253
HIVZAM181 (L03705)	255
HIVZAM18A (L22954)	256
HIVZAM184 (L22955)	258
HIVZAM201 (L03707)	260
HIVZAM20A (L22956)	261

HIV-1 sequences stored on diskettes and FTP server only

HIVBR0141 (U08796)
HIVBR0204 (U08797)
HIVBR0259 (U15121)
HIVRW0914 (U08793)
HIVTH0065 (U08810)
HIVTH0224 (U15120)
HIVTH0266 (U08802)
HIVTH1412 (U08801)
HIVUG0205 (U08794)
HIVUG0219 (U08804)
HIVUG0314 (U08795)
HIVUG0317 (L34667)
HIVUG0378 (U15119)

B. HIV-2 and Related SIV Sequences

Nucleotide alignments of HIV-2 and SIV <i>gag</i>	I-B-1
Nucleotide alignments of HIV-2 and SIV <i>pol</i>	I-B-11
Nucleotide alignments of HIV-2 and SIV <i>env</i>	I-B-32
HIV2EHOA (L14545)	56
HIV222381 (L33077)	60
HIV222382 (M87138)	61
HIV260415K1 (L33076)	62
HIV260667K1 (L33078)	63
HIV27312A1 (L33079)	64
HIV27312A2 (L33092)	65
HIV27312A3 (L33088)	66
HIV27312A4 (L33094)	67

HIV27810A1 (L33080)	68
HIV27924A1 (L33081)	69
HIV2ALI (L25445)	70
HIV2CI171 (X78505)	72
HIV2D1024 (U07104)	74
HIV2D1071 (U07105)	75
HIV2D766 (U07106)	76
HIV2D808 (U07107)	77
HIV2D868 (U07108)	78
HIV2FA1 (L33082)	79
HIV2FA2 (L33089)	80
HIV2FA3 (L33095)	81
HIV2FO784 (L33083)	82
HIV2JA1 (L33085)	83
HIV2JA2 (L33090)	84
HIV2JA3 (L33096)	85
HIV2ONFT1 (L33086)	86
HIV2ONFT2 (L33091)	87
HIV2ONFT3 (L33097)	88
HIV2ONFT4 (L33084)	89
HIV2PA1 (L33087)	90
HIV2PA2 (L33093)	91
SIVCPZGAB2 (U11495)	92
SIVHUB670 (S67406)	93
SIVSMCI2 (X78508)	94

HIV-2/SIV sequences stored on diskettes and FTP server only

HIV2CAM1E (U05359)
HIV2CAM2E (U05354)
HIV2CAM3E (U05355)
HIV2CAM4E (U05356)
HIV2CAM5E (U05357)
HIV2CAM6E (U05358)
HIV2CBL21E (U05350)
HIV2CBL22E (U05351)
HIV2CBL23E (U05352)
HIV2CBL24E (U05353)

C. SIVAGM and SIVMND Sequences

Nucleotide alignments of AGM <i>tr</i>	I-C-1
Nucleotide alignments of AGM <i>env</i>	I-C-8
SIVSAB1C (U04005)	20
SIV2010E (U10897)	24
SIV2010G (U10899)	25
SIVGRI2E (U03995)	26
SIVGRI2L (U04007)	27
SIVGRI3E (U03994)	28
SIVGRI3L (U04006)	29
SIVSAB2E (U03996)	30
SIVSAB3E (U03997)	32

Contents (1994)

SIVSAB3L18 (U04012)	34
SIVSAB4E (U03998)	36
SIVSAB4L10 (U04014)	38
SIVSABD37P (U04018)	40
SIVTAN17E (U04000)	41
SIVTAN1E (U03999)	42
SIVTAN40E (U04001)	43
SIVTAN49E (U04002)	44
SIVVER1E (U04003)	45
SIVVER266E (U10896)	46
SIVVER2E (U04004)	47
SIVVER2L (U04017)	48
SIVVER385E (U10898)	49

SIVAGM sequences stored on diskettes and FTP server only

SIVSAB1L15 (U04008)
SIVSAB2L16 (U04010)

PART I HIV and SIV Nucleic Acid Sequences

A. HIV-1 Sequences

Nucleotide alignments of HIV-1 sequences	I-A-9
HIVANT70C (L20587, M31171)	173
HIVMVP5180 (L20571)	177
HIVBCSG3C (L02317)	181
HIVBZ162 (L11751)	185
HIVBZ167 (L11752)	186
HIVBZ190 (L11753)	187
HIVBZ200 (L11754)	188
HIVCI20 (L11755)	189
HIVCI32 (L11756)	190
HIVCI4 (L11757)	191
HIVCI51 (L11758)	192
HIVCI59 (L11759)	193
HIVDJ258 (L11763)	194
HIVDJ259 (L11764)	195
HIVG109 (L11765)	196
HIVG141 (L11766)	197
HIVIC144 (L11767)	198
HIVK112 (L1168)	199
HIVK124 (L11769)	200
HIVK29 (L11770)	201
HIVK31 (L11771)	202
HIVK7 (L11772)	203
HIVK88 (L11773)	204
HIVK89 (L11774)	205
HIVK98 (L11775)	206
HIVLBV105 (L11776)	207
HIVLBV23 (L11777)	208
HIVLBV217 (L11778)	209
HIVLBV2310 (L11779)	210
HIVPH136 (L11780)	211
HIVPH153 (L11781)	212
HIVSE365 (L11797)	213
HIVSM145 (L11803)	214
HIVTN238 (L11760)	215
HIVTN240 (L11761)	216
HIVTN245 (L11762)	217
HIVUG266 (L11798)	218
HIVUG268 (L11799)	219
HIVUG270 (L11800)	220
HIVUG274 (L11801)	221
HIVUG280 (L11802)	222
HIVVI174 (L11782)	223
HIVVI191 (L11783)	224
HIVVI203 (L11784)	225
HIVVI205 (L11785)	226
HIVVI310 (L11786)	227
HIVVI313 (L11787)	228

HIVVI32 (L11788)	229
HIVVI325 (L11789)	230
HIVVI354 (L11790)	231
HIVVI415 (L11791)	232
HIVVI525 (L11792)	233
HIVVI557 (L11793)	234
HIVVI57 (L11794)	235
HIVVI59 (L11795)	236
HIVVI69 (L11796)	237
HIV17M1234 (L04602)	238
HIVA101MA (M96458, M96493-M96496)	239
HIVA151RW (M96459)	240
HIVA161RW (M96460)	241
HIVA171RAW (M96461)	242
HIVA81SE (M96456)	243
HIVB54CH (M96497, M96457, M96499, M96498)	244
HIVLAIB2M2 (M96466)	245
HIVLAIB8M2 (M96464, M96463, M96465, M96467-M96470)	246
HIVLAIC319 (M96481, M96471-M96480, M96482-M96492)	247
HIVDLC23 (M96462)	250
HIV109A11X (L16594, L16591-L16593, L16595-L16620)	251
HIVNET12 (L01282, L01283-L01297)	253
HIVAMTAN (L01298, L01299-L01339)	254
HIVZ6555 (L19623, L19624-L19636)	259
HIVFLPI01 (L22590, L22591-L22605)	261
HIVFLPID (L22606)	264
HIV1789V (M93383, M93364-M93382)	265
HIVP896 (M96155)	269
HIVB10561 (L19226, L19225-L19237, L19240-L19246, L20963)	271
HIVC11893X (L11457, L11458-L11525)	277
HIVM12199 (L15721, L15722-L15735)	281
HIVRA14018 (L19571, L19572-L19579)	283
HIVMA236X (L15143, L15138-L15142, L15144-L15187)	285
HIVMB293X (L15098, L15085-L15097, L15099-L15137)	292
B. HIV-2 and Related SIV Sequences	
Nucleotide alignments of HIV-2 and SIV sequences	1-B-1
HIV2UC1 (L07625)	82
HIV21971P6 (L16971)	86
HIV25132P1 (L16970)	87
HIV25132E2 (L16963, L16970-L16971)	88
HIV2GB1221 (L10638)	90
HIV2GB1222 (L10637)	91
HIV2GH2P22 (D10458)	92
HIV2S52907 (S52907)	93
HIV21888A (L08462)	94
SIVSMP209 (L20008, L20009, L20098-L20099)	96
C. SIVAGM and SIVMND Sequences	
Nucleotide alignments of AGM sequence	1-C-1
SIVAGMT17 (L19250)	44
SIVAGMT27 (L19251)	45

SIVAGMT40 (L19252)	46
SIVAGMT49 (L19253)	47
SIVAGMT9 (L19254)	48
SIVAGMTB05 (M81070)	49
SIVAGMTB14 (M80208)	50
SIVAGMTB30 (M81071)	51
SIVAGMTB53 (M81072)	52
SIVAGMSD29 (M81073)	53
SIVAGMSD30 (M81074)	54
SIVAGMSD37 (M81075)	55
SIVAGMSD45 (M81076)	56
SIVSYK (L06042)	57

PART I HIV and SIV Nucleic Acid Sequences

A. HIV-1 Sequences

Nucleotide alignments of HIV-1 sequences	I-A-9
HIVCAM1 (D10112)	118
HIVYU10 (M93259 M89974)	122
HIVYU2 (M93258 M89974)	126
HIVGUN (M59192 M59193)	130
HIVD31 (X61240 X16109)	132
HIVSF2B13 (L07422)	136
HIVNH5 (L07423)	138
HIVOLNEF1 and related sequences (X63040-X63045)	141
HIVGD110 and related sequences (M92995)	144
HIVH5986 and related sequences (L07145)	147
HIVZAM18 (L03705)	170
HIVZAM19 (L03706)	171
HIVZAM20 (L03707)	172
HIVTB1321 (L03696)	173
HIVTB1322 (L03697)	174
HIVTN2431 (L03702)	175
HIVTN2432 and related sequences (L03698-L03704)	176
HIVT7787 and related sequences (L07442)	183
HIVNOF (L07426)	187
HIVD687 (L07652)	188
HIVD747 (L07653)	189
HIVD757 (L07654)	190
HIVD760 (L07655)	191
HIVD1044 (L07651)	192
HIVM6506 and related sequences (L15731)	193
HIVZ6555 and related sequences (L07465)	196
HIVU3079 and related sequences (M98912)	199
HIVU44342 (M98408-M98416)	212
HIVSIMI (L07421)	218
HIVKIGA93 and related sequences (L07082)	220

B. HIV-2 and Related SIV Sequences

Nucleotide alignments of HIV-2 and SIV sequences	1-B-1
HIV2FO784 (M87115-M87117; M87111-M87114; M87106-M87110; M87069-M87105)	61
HIV22238 (M87135-M87139; M87118-M87134)	78
HIV27312A (M87144-M87145; M87140-M87143)	85
HIV2STV (M86924 M31113)	89
HIV2ISYR (J04498)	91
HIV2GB12 and related sequences (L10638)	95
HIV2GHLTR (M69108)	98
SIVSMMLIB1 (M62651)	99
SIVSMM9 (M80194)	100

C. SIVAGM and SIVMND Sequences

Nucleotide alignments of AGM sequences	1-C-1
--	-------

PART I. HIV and SIV Nucleic Acid Sequences

A. HIV-1 Sequences

Nucleotide alignments of HIV-1 sequences	I-A-3
HIVLAI (K02013; M64186-190; M64194-198; M64203-223)	70
HIVNL43 (M19921)	76
HIVJB01 (X57447)	80
HIVJB02 (X57450)	83
HIVJB03 (X57465)	89
HIVBAL (M68893)	92
HIVADA (M60472)	95
HIVSWB40 (M27951; M27946)	96
HIVSWB88 (M27952; M31462; M27947; M58206)	97
HIVM77, HIVV77, HIVT77 (see M61277-M61407, M36997)	100
HIVM82, HIVV82, HIVT82 (M61365, M36997, M61402, M77602, see M61277-M61407)	103
HIVNJA (M61028-M61045)	117
HIVU455 (M62320)	122
HIVCBL4 (M63474; M63475)	126
HIVSF170 (four segments) (M66533-M66535; M81729)	127
HIVABA2 (M75686)	132
HIVACP1 (M80660)	135
HIVACH9 (M80661)	136
HIVJM (M80662)	137
HIVWM (M80663)	138
HIVHP200 (M81692)	139
HIVHP1089 (M81693)	140
HIVHP1098 (M81694)	141
HIVJRFL (pending, and M74978-M75007)	142
HIVSPBL14 (M74960; M74955-M74977)	148
HIVSPBR1 (M74964)	149
HIVTIBL7 (M75016; M75008-M75028)	152
HIVTIBR1 (M75022)	153
HIVRJS (M37491; M37573-M37577)	156
HIVSB (three segments) (M77228-M77230)	158
HIVFLD (Florida Dental Case V3V4V5 Sequences)	
(M90847-M90966; M91084-M91115; M91121-M91149)	
(M91153-M91156; M92100-M92150)	161
HIVHEI (X61351-X61359; X64746-X64745)	306
HIVYU10 (M89974)	324
HIVAM113 (M74591-M74640)	327
HIVAM49501 (M74641-M74684)	328
HIVMA208 (M79342-M79354; M90025-M90046)	333
HIVMOM1 (two segments) (M76789-M76799; M76900-M76912)	346
HIVTOT1 (two segments) (M76842-M76862; M76942-M76947)	346
HIVMOM2 (two segments) (M76800-M76818; M76913-M76925)	351
HIVTOT2 (two segments) (M76863-M76875; M76948-M76963)	351
HIVMOM3 (two segments) (M76819-M76841; M76926-M76941)	356
HIVTOT3 (two segments) (M76876-M76899; M76964-M76973)	356

B. HIV-2 and Related SIV Sequences

Nucleotide alignments of HIV-2 and SIV sequences	I-B-1
HIV2D194 (J04542; X52223)	58
HIV2D205 (X16109; X61240)	62
SIVMM251 (M19499; X06393; M15897; M16125; Y00283; X06879; Y00294; Y00295; X06391)	66
SIVMM239 (M33262)	71
SIVMM1A1 (M76764)	76
SIVMM32H (D01065)	80
Nucleotide alignments of SIVMM	84
SIVSTAK (three segments) (X57323; X60667; X60668)	109
HIVCAM2 (D00835)	112

C. SIVAGM and SIVMND Sequences

Nucleotide alignments of AGM sequences	I-C-1
SIVAGM677 (M66437)	22
SIVMND121 (M62883; M62884)	26
SIVAGM381 (M73537; M73538)	27
SIVAGM384 (M73539)	27
SIVAGM385 (M81730-M81733)	28
SIVAGM386 (M81734)	29

PART I. HIV AND SIV NUCLEIC ACID SEQUENCES

A. HIV-1 Sequences

Nucleotide alignments of HIV-1 sequences	I-A-3
HIVBRU (K02013)	60
HIVNL43 (M19921)	64
HIVHXB2R (K03455 M38432)	68
HIVPCV12 (M11840)	72
HIVMN (M17449)	74
HIVSC (M17450)	78
HIVJH3 (two segments)(M21137 M21138)	80
HIVBRVA (M21098)	83
HIVSF2 (K02007)	85
HIVHAN (EMBL assignment)	89
HIV0YI (M26727)	93
HIVCDC45I(two segments)(M13136 M13137)	97
HIVNY5(K03346)	101
HIVJFL (M31451)	102
HIVTRA (two segments)(M27956 M27948)	104
HIVBAN (two segments)(M27953 M27944)	105
HIVBOS1 (M27910)	106
HIVPHIL1 (M27911)	107
HIVMAB (M27955)	108
HIVSWB88 (three segments)(M27952 M31462 M27947)	109
HIVSWB40 (two segments)(M27951)	111
HIVWMI2 (two segments)(K03457 M12507)	112
HIVRF (M17451 M12508)	115
HIVMAL (K03456)	119
HIVELI (K03454)	123
HIVZ2Z6 (M22639)	127
HIVZ6 (K03458)	131
HIVZ321 (M15896 J00001)	134
HIVJY1(J03653)	136
HIVZ3 (K03347)	138
HIVNDK (M27323)	140
HIVANT70 (M31171)	144
HIVJRCSE (M38429)	145
HIVJREFL (pending and M74978-M75007)	149
HIVMFA (M33943)	153
HIVNITE (M36998)	155
HIVSF33 (M38427)	156
HIVSF162 (M38428)	158
HIVALA1 (M38430)	160
HIVNY5CG (M38431)	162
SIVCPZCG (X52154)	166

B. HIV-2 and related SIV sequences

Nucleotide alignments of HIV-2 and SIV sequences	I-B-1
HIV2ROD (M15390)	46
HIV2NIHZ (J03654)	50
HIV2ISY (J04498)	54
HIV2ST (M31113)	58
HIV2GH1 (M30895)	62
HIV2BEN (M30502)	66
HIV2D205 (X16109)	70
HIV2D194 (J04542)	73
SIVMM142 (M16403 Y00277)	77
SIVMM251 (M19499X06393)	81
SIVMM239 (M33262)	86
SIVMMP11 (M23164)	90
SIVMNE (M32741)	91
SIVSMMH4 (X14307)	95
SIVSMMM7 (M27256)	99
SIVSMMPBJ (M31325)	100

C. SIVAGM and SIVMND Sequences

Nucleotide alignments of AGM sequences	I-C-1
SIVAGMTYO (X07805)	19
SIVAGM385 (M21311)	23
SIVAGM266 (M22974)	24
SIVAGM155 (M29975)	25
SIVAGM677 (M29973)	29
SIVAGM90 (M33718)	30
SIVAGM692 (two segments)(M33719 M29974)	31
SIVAGM3 (M30931)	33
SIVMNDGB1 (M27470 X15781)	37

This page intentionally left blank.

HIV1 GAG

CONSENSUS.A	ATGGGTGCGAGAGcGTCagTatTaAgtGggggaaaATTaGATycaTgGgAgAaaATTcgGtTaaGGCCaGg	71
HIV0455	-----C-AA-----T-----	71
HIVMAL	-----C-----	71
HIVVI59	--A-----C-A-----A-----	71
HIVVI310	-----C-----AAG-----	71
HIVVI57	-----C-A-----	71
HIVK112	-----	71
HIVK89	-----G-----	71
HIVK29	-----	71
HIVK124	-----C-----	71
HIVK7	--A-----C-----C-----AG-A-----	71
HIVK98	-----A-----	71
HIVK89	-----C-----A-G-----	71
HIVVI32	-----A-----	71
HIVVI415	-----	71
HIVCI4	-----G-----C-----	71
HIVG141	-----C-----G--AT--A--A--C-----	71
HIVLBV23	-----C-----A-G--A-----	71
HIVTN243	-----A-----C-----	71
HIVTN245	-----A-----C-----	71
HIVTN240	-----A-----C-----	71
HIVCI20	-----G-----	71
HIVCI59	-----G-----	71
HIVLBV2310	-----T-----	71
HIVCI51	-----A-----T-----	71
HIVLBV105	-----G--C-----A-----C-----	71
HIVCI32	-----G-----CG-----	71
HIVIC144	-----C-A-----G-----	71
HIVDJ258	-----T-----G-----	71
HIVQM238	-----C-----A-----C-----	71
HIVUG266	-----T-----C-A--G-----A-----A-----	71
HIVVI354	-----A-C--C-----A-A-G-----G-----	71
CONSENSUS.B	ATGGGTGCGAGAGCGTCaGTaTaaGcGggGGagaATTaGATagaTGGgAaAaaATTcGGtTaaGGCCaGg	71
HIVSF2	-----G-----A-----	71
HIVBZ167	-----G-----G-----	71
HIVPH153	-----GA-----C-----	71
HIVPH136	-----C-----G-----	71
HIVBZ200	-----C-----A-----	71
HIVTB132	-----C-----A-G-----C-----	71
HIVBZ190	-----A-----A-----	71
HIVLA1	-----C-----	71
HIVKXB2R	-----C-----	71
HIVMN	-----G-----C-----C-----	71
HIVJH3	-----C-----G-----	71
HIVJRCSE	-----G-----G-----	71
HIVQY1	-----A-----	71
HIVNY5CG	-----G-----A-----	71
HIVNL43	-----G-----A-----	71
HIVCDC4	-----	71
HIVHAN	-----C-A-----	71
HIVCAM1	-----A-----	71
HIVRF	-----C-A-----C-A-----A-----	71
HIVD31	-----	71
HIVBH102	-----C-----	71
HIVPV22	-----C-----	71
HIVJRFL	-----A--G--A-----	71
HIVUG280	-----C-----	71
HIVYU2	-----T-C--G-----AG-----	71
HIVBCSG3C	-----G--A--G--AG--G-----	71

CONSENSUS .C	ATGGGTGCGAGAGCGTCAaTATTAAGAGGcGgAAAAATTAGATacaTGGGAaAaAATTAGGtTAAGGCCAGG	71
HIVUG268	-----G-----C-----A-----	71
HIVSM145	-----G-----G-----G-----	71
HIVZAM18	-----A--G-----	71
HIVZAM19	-----N--G-----G-----	71
HIVZAM20	-----G-----A-----G-----	71
HIVDJ259	-----A-----G-C-----A-C-----	71
HIVVI313	-----C-----G-----A-----A-----	71
CONSENSUS .D	ATGGGTGCGAGAGCGTCAaGTATTAAGCGGGGAaAATTaGATgaaTGGGAaAaAATTCgGTACCGGCCAGG	71
HIVELI	-----A-----	71
HIVZ226	-----G-----C-----	71
HIVNDK	-----AC-----G-----	71
HIVVI205	-----C-----C-----	71
HIVG109	-----C-----C-----G-----	71
HIVK31	-----A-----	71
HIVUG274	-----	71
HIVUG270	-----G-----C-----	71
HIVSE365	-----C-----CG-----	71
HIVVI203	-----C-----TCG-----	71
CONSENSUS .F	ATGGGTGCGAGAGCGTCAGTATTAAAGCGGgGGAAAAATaGATgCATGGGAAAAAATTCggTTAAGgCC?GG	70
HIVVI174	-----A-----G-----	71
HIVVI69	-----A-----A--G-----	71
HIVBZ162	-----A-----C-----A-----	71
HIVVI325	-----G--A-----A-----	71
CONSENSUS .G	ATGGGTGCGAGAGCGTCAGTATTAAAGCGGGGAAAA?TAGATG?ATGGGAAAAAATTCggTTAAGgCC?GG	67
HIVLBV217	-----T-----A-----A--C-----	71
HIVVI191	-----C-----C-----C--T-----	71
CONSENSUS .H	ATGGGTGCGAGAGCGTCAGTATT?AGCGGGGAA?ATTAGATGC?TGGGAGAAAAATTCGGTT?AGGCCAGG	67
HIVVI525	-----G-----G-----T-----A-----	71
HIVVI557	-----A-----A-----A-----G-----	71
CONSENSUS .O	ATGGGTGCGAG?GCGTCTGTGTTGACAGG?AG?AAATTGGATGCATGGGAAC?AATTAGGTAA?GCCAGG	66
HIVAMT70	-----T-----A--C-----A-----A-----	71
HIVMVP51	-----A-----G--T-----G-----G-----	71
SIVCPZGAB	ATGGGTGCGAGAGCGTCAGTTCTAACAGGGGAAAAATTAGATCGCTGGGAAAAAGTTCGGCTTAGGCCCGG	71

HIV1 GAG

```

CONSENSUS .A gGGAAagAAAAaTATagacTgAAaCAT?TAGTaTgGGCaaGCagggAGcTgGAAagATTcgCACTPaAcC 141
HIVU455 -----C-----T-----A--A----- 142
HIVMAL -----T----- 142
HIVV159 A-----C-----T--G-A-----A----- 142
HIVV1310 -----C-----A-G--C--G----- 142
HIVV157 -----C----- 142
HIVK112 -----A-C-----C-----A--T----- 142
HIVK88 -----A-----T-----C--T----- 142
HIVK29 -----A-----T-----T----- 142
HIVK124 -----G-----T-----T----- 142
HIVK7 -----C----- 142
HIVK98 -----T--G-----A--A-----A----- 142
HIVK89 -----A-----T----- 142
HIVV132 -----A-----T-----C----- 142
HIVV1415 -----G-----A-----C-----C----- 142
HIVCI4 -----A-----C-A-----G--A----- 142
HIVG141 -----CC--GG--AG--A--A-----A-----C--T--G--T----- 142
HIVLBV23 -----A-----C-----T----- 142
HIVTN243 -----GA--G--GA-----T-----A--T-A----- 142
HIVTN245 -----GA--G--AGA-----T-----A--T-A----- 142
HIVTN240 -----GA-----G-----T-----A--T-A----- 142
HIVCI20 -----A-----C-----T----- 142
HIVCI59 -----A-----C----- 142
HIVLBV2310 -----A-----C----- 142
HIVCI51 -----G--C-----A-----C----- 142
HIVLBV105 -----A-----C-----T----- 142
HIVCI12 -----A--A-----T----- 142
HIVCI144 -----A-----C-----T----- 142
HIVDJ258 -----A-----C-----G----- 142
HIVCM238 -----G-----AGA-A-----T-----C-A--T-A----- 142
HIVUG266 -----A-----C--G-A-----A--A----- 142
HIVV1354 -----G--CACA-A-----A--G-----A--T----- 142

```

/<- INS-1

```

CONSENSUS .B gGGAAAGAAAAaTATaattTAAaACATATAGTATGGGCAAGCAGggAgCTAGAACGaTTcGCagTtaAtC 142
HIVSP2 -----G-----C----- 142
HIVBZ167 -----GG-----G-----A----- 142
HIVPH153 -----C----- 142
HIVPH136 ----- 142
HIVBZ200 -----GC-----C----- 142
HIVTB132 -----G----- 142
HIVBZ190 -----GG--C--C-----A----- 142
HIVLAT ----- 142
HIVHXB2R ----- 142
HIVMN -----G-----C----- 142
HIVJH3 -----T-----T----- 142
HIVJRCSE A-----G-----T-----C----- 142
HIVQVI -----C--C-----A----- 142
HIVNY5CG -----C--G--C-----A----- 142
HIVNL43 -----C-----C----- 142
HIVCDC4 -----C--CG--G-----A-----C----- 142
HIVHAN -----C----- 142
HIVCAH1 ----- 142
HIVRF -----G--G-----T--T--C----- 142
HIVD31 -----G-----A-----T----- 142
HIVBH102 ----- 142
HIVPV22 ----- 142
HIVJREFL -----G-----T----- 142
HIVUG280 ----- 142
HIVYU2 -----C--G-----G----- 142
HIVBC6G3C -----C----- 142

```

CONSENSUS.C	gGGAAaGAAAcacTATATGaTaAAACAccTAGTaTGGGCAAGCAGGGAGCTgGaAAGATTTCACCTTaACC	142
HIVUG268	-----TG-----G-----C-----G-----T	142
HIVSM145	-----	142
HIVZAM18	-----	142
HIVZAM19	-----T-----A-----	142
HIVZAM20	-----T-----TT-----	142
HIVDJ259	A---G---G---C---C---A---	142
HIVVI313	-----G-----G-----	142
CONSENSUS.D	aggAaagAAAAAATATaaCTaAAACATaTagTATGGGCAAGCAGgGagCTagaACGATTtTgCACTTAATC	142
HIVELI	-----G-----A-----	142
HIVZ236	-----G-----C-----	142
HIVNDK	-----GC-----T-GA-----A-----	142
HIVVI205	-AA--GC-----G-----A-----	142
HIVG109	G-----C-----A-----	142
HIVK31	G--C-C-----A-----	142
HIVUG274	G-----C-----G-----A-----	142
HIVK3270	G-----C-----C-----	142
HIVSE365	-----G-----A-----	142
HIVVI203	-----G--G-----A-----	142
CONSENSUS.F	gGGaA?gAAAAAATaCagaaTgAAgCATCTAgTATGGGCAAGCAGGGAGcTAGAAcGATTTGCA?TT?AtC	138
HIVVI174	-G-A-----A-----A-G-----	142
HIVVI69	-G-A-----C-A-----A-----AT-----A-----C-G-----	142
HIVBZ162	-A-----T-----A-A-----	142
HIVVI325	A---GC-----GC-A-A-----C-A-C-----	142
CONSENSUS.G	GGGAAGAAA?AATATAGAAT?AA?CAT?TAGTATGGGCAAGCAGGGAGCTGGA?A?ATTTGCACCTTAACC	132
HIVLBV217	-----A-----G-A-T-----G-G-----	142
HIVVI191	-----C-----A-G-C-----A-A-----	142
CONSENSUS.H	GGAAGAGAAA?AATATAG?CT?AAACAT?TAGTATGGGCAAGCAG?GAGCTGGAAAGATTTCACCT?AACC	132
HIVVI525	-----C-----G-A-C-----A-----T-----	142
HIVVI557	-----A-----A-G-T-----G-----C-----	142
CONSENSUS.O	ATCTAAAAAG??ATATAG?CTAAAACA?TTAGTATGGGCAAGCAG?GAGCTGGAAAGAT?CGCATGTAATC	131
HIVANT70	-----AA-----A-----C-----A-----T-----	142
HIVMVP51	-----GC-----G-----T-----G-----A-----	142
SIVCPZGAB	GGGAAGAAAAAGATATATGATGAAACATTTAGTATGGGCAAGCAGGGAGCTGGAAAGATTTCGCATGTGACC	142

HIV1 GAG

CONSENSUS.A CtcGccTTtTagAaACAgcaGaagGctGTcAacaaATaTggaACAgTAcAaACAgcTcTcaagACAGga 212
 HIVU455 -G-----G-----G-----C-----G-----A-----C-----C----- 213
 HIVM4L -G-----G-----C-----A-----A-----C-----A-----C----- 213
 HIVVI59 -A-----A-----A-----A-----A-----C-----A-----C----- 213
 HIVVI310 -A-G-----A-----A-----C-----C-----A-----A----- 213
 HIVVI57 -A-----A-----C-----C-----A-----A-----C-----A----- 213
 HIVK112 -G-A-----A-----A-----A-----A-----C-----A----- 213
 HIVK88 -C-----A-----A-----A-----A-----C-----A-----A----- 213
 HIVK29 -A-----A-----A-----A-----A-----C-----A-----A----- 213
 HIVK124 -A-----A-----A-----A-----A-----C-----A-----A----- 213
 HIVK7 -A-----A-G-----A-----A-----A-----C-----A----- 213
 HIVK98 -G-----A-----G-----G-----A-----A-----C-----A----- 213
 HIVK89 -A-----G-----A-----A-----A-----C-----A-----G----- 213
 HIVVI32 -C-G-----AA-----G-----G-----G-----G----- 213
 HIVVI415 -G-----A-----G-----C-----C-----GC-----G----- 213
 HIVCI4 -A-----A-----G-----C-----C-----C-----C----- 213
 HIVG141 -C-----C-----G-----T-----G-----C-----A-----C-----C-----T----- 213
 HIVLBU23 -G-----A-----G-----G-----A-----G-----A-----C-----G----- 213
 HIVTN243 -G-----A-----A-----A-----G-----A-----A-----C----- 213
 HIVTN245 -G-----C-----AAG-----G-----A-----A-----A----- 213
 HIVTN240 -T-----A-----A-----G-----A-----A-----A----- 213
 HIVCI20 -C-----G-----G-----G-----C-----A-----G-----CA-----G----- 213
 HIVCI59 -G-----T-----C-----C-----C-----G-----G----- 213
 HIVLBU2310 -G-----T-----T-----A-----A-----G-----G----- 213
 HIVCI51 -A-----A-----A-----A-----A-----A-----A----- 213
 HIVLBU105 -G-----C-----C-----A-----C-----T-----G-----T-----A----- 213
 HIVCI32 -T-----C-----A-----A-----C-----A-----GG----- 213
 HIVCI44 -G-T-----C-----A-----A-----A----- 213
 HIVDJ258 -G-----GT-----A-----G-----A-----GG-----A----- 213
 HIVCM238 -T-----A-----A-----G-----A-----A----- 213
 HIVUG266 -A-----AG-----C-----A-----A-----C-----C----- 213
 HIVVI354 -CG-T-----A-----G-----A-----A-----A----- 213

CONSENSUS.B CtgGCCtTgtTaGAaaCatCagaaGGcTGtaGACaaATAcTgggACAgcTACaaCCatcCCTTcAGACaGGA 213
 HIVSF2 -C-----T-----G-----G-----T----- 213
 HIVB2167 -AC-----GG-----C-----G-----T----- 213
 HIVPH153 -G-----T-----A-----G----- 213
 HIVPH136 -G-----T-----A-----G----- 213
 HIVB2200 -C-----TC-----C-----A-----A-----A----- 213
 HIVTB132 -A-G-----G-----A-----AG----- 213
 HIVB2190 -G-----G----- 213
 HIVLAT - 213
 HIVHXB2R - 213
 HIVMN - 213
 HIVJH3 -A-----G----- 213
 HIVJRC5F -T-----A-----A----- 213
 HIVOYI -T-----T-----A----- 213
 HIVNY5CG -T-----G-----A----- 213
 HIVNL43 -T-----G----- 213
 HIVCDC4 -T-----A-----T----- 213
 HIVHAN -G-----A-----G----- 213
 HIVCAM1 -T-----G-----G----- 213
 HIVRF -A-----T-----G-----G-----G-----G----- 213
 HIVD31 -T----- 213
 HIVBH102 - 213
 HIVPV22 - 213
 HIVJREFL -T-----A-----T-----G-----A----- 213
 HIVUG280 -A-----G-----G----- 213
 HIVYU2 -G----- 213
 HIVBCSG3C -C-----G-----G----- 213

CONSENSUS.C	CTGGCCCTTTAGAGACatCAGaaGGCTGTaaACAAATAATaAaaACAGCTaCAaCCAgCTcTTCAGACAGGa	213
HIVUG268	-----A-----G-----G	213
HIV9M145	-----G-----G-----	213
HIVZAM18	-----T-A-----	213
HIVZAM19	---T-----	213
HIVZAM20	-----G---C-----G-----T-----	213
HIVDJ259	--A-----G-----	213
HIVVI313	-----C-----G-G-----A-----	213
CONSENSUS.D	CTGGCCCTTTAGAGACatCAGaaGGCTGTaaACAAATAATaGgaCAGCTACaaCCAgcTaTtcagaCAGGA	213
HIVELI	-----A-----G-----	213
HIVZ226	-----T-----C-G-----	213
HIVNDK	-----T-----A-----	213
HIVVI205	-----G-----A-----	213
HIVG109	-----G-----A-----C-AA-C---T-----	213
HIVK31	-----A-C-----G-----C-----	213
HIVUG274	-----A-----G-----C-A-----	213
HIVUG270	-----A-C-----G-----T-----	213
HIVSE365	-----T---G-----	213
HIVVI203	-----A-----T-----	213
CONSENSUS.F	CtgGcCTTcTAGAaACatCAGAAGGCTGT?g?aaaATAATagGACAgcTACAAcCATCCCTTCAGACAGGA	207
HIVVI174	---T-----C-----A-G-----C-----A-----	213
HIVVI69	-----A-A-----	213
HIVB2162	-----CAA-----	212
HIVVI325	-CA-----T---G---A-----C-GC-----A---A-----	213
CONSENSUS.G	CTGGCCCT?TTAGAAACAACAGAAGG?TGTCAACAAATA?TG??ACAGTTGCAACCA?CTCTCCAGACAGGA	197
HIVLEV217	-----C-----T-----C-CA-----T-----	213
HIVVI191	-----T-----C-----A-AQ-----G-----	213
CONSENSUS.H	C?G?CCTTTTAGA?ACA?CAGAAGGCTGTCTACA??TAATAGAACAGCTACA?CCAGCTCTTAAGACAGGA	196
HIVVI525	-T-G-----A---T-----GA-----G-----	213
HIVVI557	-C-A-----C---G-----AT-----A-----	213
CONSENSUS.O	CTG??CTA?TAGAAACTGCAGA?GGTA?TGAG?A?CTGCTACAGCAGTTAGAGCCAGCTCTCAAGACAGGG	195
HIVANT70	---AG---C-----G---A---A-G-----	213
HIVMVP51	---GT---T-----A---C---C-A-----	213
SIVCPZGAB	CCGGGCTAATGGAAAGTAAGGAAGGATGTACTAAATGTTTACAACAATTAGAGCCAGCTCTCAAAACAGGC	213

HIV1 GAG

CONSENSUS . A ?CAGaAGAAcTtAaaTCaTTAtttaaCagTAGCaacCCtTatTGcgTaCatcaaaggATAgA?gTAAa 281
 HIVU455 A-----G-----A-----GT-----T-----T----- 284
 HIVMAL T-----A-----A-----T-----T-----T----- 284
 HIVV159 A-----G-----G-----GT-----C-----TA----- 284
 HIVV1310 A-----A-G-----A-----GT-----T-----C-----G----- 284
 HIVV157 A-----GT-----C-----G----- 284
 HIVK112 A-----G-----T----- 284
 HIVK88 A-----G-----C-----T-----A-----T----- 284
 HIVK29 A-----G-----GT-----C-----G----- 284
 HIVK124 A-----A-G-----G-----GT-----T-----C-----G----- 284
 HIVK7 A-----GG-----A-----T-----G-CA-----G----- 284
 HIVK98 A-----G-G-A-----C-----T----- 284
 HIVK89 A-----G-----GG-----CAA-A-T----- 284
 HIVV132 A-----A-----A-----GT-----T-----C-----T----- 284
 HIVV1415 A-----A-----A-----T-----G-----T----- 284
 HIVC14 A-----G-----A-C-----C-----T----- 284
 HIVG141 T-----C-C-A-----T-----GC-A-AA----- 284
 HIVLBV23 A-----A-----CA-----GT-----G-----T----- 284
 HIVTN243 T-----GT-----C-----G----- 284
 HIVTN245 T-----T-----GG-G-C-----G----- 284
 HIVTN240 T-----GG-----C-----G----- 284
 HIVC120 T-----A-----T-GG-----CA----- 284
 HIVC159 T-----G-A-----GT-----T-GG-----CA----- 284
 HIVLBV2310 T-----A-----T-GG-----C----- 284
 HIVC151 T-----G-----T-GG-----CA----- 284
 HIVLBV105 T-----G-----T-----TG-T-----C-----G-----G----- 284
 HIVC132 T-----G-----A-----A-----T-GG-T-----G-----GA----- 284
 HIVC144 T-----G-----A-----T-----A-----CA-----G----- 284
 HIVD3258 T-A-----A-----CA-----GT-----T-GG-----CA----- 284
 HIVCN238 T-----G-----GT-----C-----G----- 284
 HIVUG266 G-----G-C-A-----GT-----GA----- 284
 HIVV1354 T-----T-----A-----A-T----- 284

- - - INS-1 - - -

CONSENSUS . B tCAGaAGAAcTtagATCATTataTAaCaCagTAGCaacCCtTATTGTGtgCATCAAaggATAGaggTAAa 284
 HIVSF2 -----A-----T----- 284
 HIVB2167 -----A-----A-----T-----G----- 284
 HIVPH153 A-----AC----- 284
 HIVPH136 -----A-----C-----A-----T----- 284
 HIVB2200 -----GT-----C-----A-----T----- 284
 HIVTB132 -----A-----GT-----A-----A----- 284
 HIVB2190 -----A-----T-----G-----GT-----CA----- 284
 HIVLA1 -----T-----A----- 284
 HIVHX82R -----A----- 284
 HIVMN -----A-----A-----A----- 284
 HIVJH3 -----A-----T-----A-----A----- 284
 HIVJRCSP -----C-----A-----A----- 284
 HIVOY1 -----A-----T-----T-----A----- 284
 HIVNY5CG -----G-----T-----GT-----T----- 284
 HIVNL43 -----A-----GT-----T----- 284
 HIVCDC4 -----G----- 284
 HIVHAN -----A----- 284
 HIVCAM1 -----G-----A-----T----- 284
 HIVRF -----A-----G-----A-----AT-----G----- 284
 HIVD31 -----T-----C----- 284
 HIVBH102 -----A----- 284
 HIVPV22 -----A----- 284
 HIVJRF4 -----A----- 284
 HIVUG280 -----G-T-----T----- 284
 HIVYU2 -----G-----C-----A-----A----- 284
 HIVBCSG3C -----A-----A-----GT-----T-----T----- 284

CONSENSUS.C	ACAGAGGAACCTTAGATCATTatstAAacCACTAGCAactCTcTatTGTGTACATgaa?agATAgaggTACg	283
HIVUG268	-----T-----A--GG-----A-----	284
HIVSM145	-----TC-----A-----	284
HIVZAM18	-----A-----TC-----T-----A--GG-----G--T--ACA-----	284
HIVZAM19	-----A-----C--C--G-----GTC-----A--N--T--ACA-----	284
HIVZAM20	-----C--GC-----	284
HIVDJ259	-----A-----T-----C--AG-----A--A	284
HIVVI313	-----A-----T-----C--T-----GA-----A-----	284
CONSENSUS.D	CCagAaGAactTAAATCATTAT?TAATACAgTAGCAACCTCTATTGcGTACATgaeaggATAgAGTaaa	283
HIVELI	A-----G-----A-----A--G--A-----T-----	284
HIVZ226	-----G-----T-----	284
HIVNDK	-----A--G-----A-----	284
HIVVI205	-----A-----T-----A-----G-----	284
HIVG109	-----G--C-----A-----A--C-----	284
HIVK31	-----G--A-----A-----C-----C--G-----A-----C	284
HIVUG274	-----G--A--C-----T-----A-----C	284
HIVUG270	-----G--G-----T-----T-----A-----GC	284
HIVSE365	-----G-----A-----A-----A-----	284
HIVVI203	A-----T-----C-----	284
CONSENSUS.F	TCAGAAAGAGCTTAgATCATTaT?tAATACAgTAG?AgtcCTCTATTtGTACATCAaAgGgTAGAG?TAAa	275
HIVVI174	-----G--T-----A--T-----A-----A-----	284
HIVVI69	-----A-----C--T-----A-----G-----	284
HIVBZ162	-----A-----C-----G-----	283
HIVVI325	-----A-----TC-----T--AC-----GC-----G--A-----A--G	284
CONSENSUS.G	ACAGAGGAGG?TTAAATCA?TAT?TAAT?CA?TAGCA??TCTCT?TGT?TACATCAAAGGATAGaggTAAA	258
HIVLBV217	-----C-----T--T--G--G-----GT--GG--A-----G-----	284
HIVVI191	-----A-----C--A--A--A--AC--AT--G-----A-----	284
HIVTAIG	-----T-----	7
CONSENSUS.H	ACAGAAGAACTTCAATCATTATTTAATCTATTAGCA?TCTCTATTG?GTACATCAAAGAATAGAT?TAAA	263
HIVVI525	-----GT-----T-----A-----	284
HIVVI557	-----AC-----C-----G-----	284
CONSENSUS.O	TCAGA??TCTG?A?TC?CTCTGGAACGCAATAG?AGT?CTCTGGTG?GTTCAACACAGAT?T?A?AT?TG	252
HIVANT70	-----CAG--C--G--A-----T--G-----T-----A--A--TG--	284
HIVMVP51	-----GGA--A--A--T-----C--A-----C-----T--G--C--CC--	284
SIVCP2GAB	TCAGAAGGACTGCGGTCTTGTTTAACACTCTGGCAGTACTGTGCTGCATACATAGTGACATCACTGTAGA	284

HIV1 GAG

```

CONSENSUS .A aGACaCCAagGAaGCTtTagAtAaaaTAGAGGaAaTaaAAAatAagagcaagcaaaag???????????? 339
HIVU455 -----A-----G-----A-----G----- 342
HIVMAL -----GC-----G-----ACACAGCAGGCAG 355
HIVVI59 -----A-----C-----A-----G-----T----- 342
HIVVI310 -----T-----G----- 342
HIVVI57 -----T-----G----- 342
HIVK112 G-----C----- 342
HIVK88 -----G-----C-----C----- 342
HIVK29 -----G-----C-----T-----G----- 342
HIVK124 -----C-----C-----G-----A----- 342
HIVK7 -----A-----C----- 342
HIVK98 -----C----- 342
HIVK89 -----C-----G-----A-----A-----A----- 342
HIVVI32 -----C-----A-----T----- 342
HIVVI415 -----A-----A-----A----- 342
HIVCI4 -----C----- 342
HIVG141 -----C----- 336
HIVLBV23 -----G-----G-----CC----- 342
HIVTN243 -----G-----C----- 342
HIVTN245 -----G-----C----- 342
HIVTN240 -----G-----C----- 342
HIVCI20 -----C-----T----- 336
HIVCI59 -----A-----C-----T-----G-----A----- 342
HIVLBV2310 -----C-----T-----A----- 342
HIVCI51 -----C-----GC-----CAA-----A----- 339
HIVLBV105 -----A-----C----- 342
HIVCI32 -----C-----ACACAGCAGGCAG 355
HIVIC144 -----C----- 342
HIVBJ258 -----C-----C-----G----- 342
HIVCM238 -----G-----C----- 342
HIVUG266 -----G----- 342
HIVVI354 -----A-----A----- 342

```

INS-1 ->/

```

CONSENSUS .B AGACACCAagGAAGCTTTAGA?AagaTaGAGGAGAGCAaaaCAaaAGTAagaAAAAa 341
HIVSF2 -----G-----G----- 342
HIVBZ167 -----G-----G-----G-----G----- 342
HIVPH153 -----A-----C-----G----- 342
HIVPH136 -----G----- 342
HIVBZ200 -----C-----A-----G-----G----- 342
HIVTB132 -----G-----A-----G----- 342
HIVBZ190 -----T-----C-----T-----G----- 342
HIVLAI -----C----- 342
HIVHXB2R -----C----- 342
HIVMN -----G-----A----- 342
HIVJH3 -----G----- 342
HIV3RCSF -----A-----C-----T-----G----- 342
HIVVOY1 -----T----- 342
HIVNY5CG -----C-----T-----G----- 342
HIVNL43 -----C-----T-----G----- 342
HIVCDC4 -----A-----C-----A----- 342
HIVHAN -----C-----G-----A-----C-----G----- 342
HIVCAM1 -----C-----G-----A-----G----- 342
HIVRF -----C----- 342
HIVB31 -----T----- 342
HIVBH102 -----C----- 342
HIVPV22 -----C----- 342
HIVJRFL -----A-----GC----- 342
HIVUG280 -----C-----G----- 342
HIVYU2 -----G----- 342
HIVBCSG3C -----G-----C----- 342

```

CONSENSUS.C	AGACACCAAGGAAGCctTAGACaagaTgAGGAAGACAAAACAAaagTCAGCAAAA? ? ? ? ?	341
HIVUG268	-----G-T-----	342
HIVSM145	-----	342
HIVZAM18	-----G-T-----ACAAAA	348
HIVZAM19	-----N-----G-----GT-----	342
HIVZAM20	-----G-----T-----	342
HIVDJ259	-----A-AC-----T-----	342
HIVVI313	-----A-A-----	342
CONSENSUS.D	agAcACCAAGGAAGCtTAGAAaAgaTgGAGGAAGAcCAAAACAAAAGTAAGAAaAG? ? ? ? ?	341
HIVELI	-----G-----	342
HIV2226	-----C-AAAG-----	345
HIVNDK	-----A-----G-----	342
HIVVI205	-----C-----AC-----AAGA-----	345
HIVGI09	-----GCACAGCAA-----	351
HIVK31	-A-----C-----A-----C-----A-----	342
HIVUG274	-----C-----A-----C-----	342
HIVUG270	-----A-----C-----A-----A-----G-----C-----A-----	342
HIVSE165	-----G-----G-----	342
HIVVI203	G--T-----C-AT-----	342
CONSENSUS.F	GGACACCAAGGAAGCtTTAGAcAagCTAGAGGAaGAACAAAACAAAAGTCAGCAAAAAG	333
HIVVI174	-----A-----	342
HIVVI69	-----G-----G-----	342
HIVB2162	-----	341
HIVVI325	-----C-----A-----A-----A-----	342
CONSENSUS.G	AGACACCAAGGAAGCtCTAGAGGAAGTGGAAaAagTcCAAAAaGAACAGTCAGCAAAA	316
HIVLBV217	-----AGG-----	342
HIVVI191	-----C-----T-C-----G-----A-----	342
HIVTAIG	-----A-----G-----	65
CONSENSUS.H	AGACACCAAGGAAGCtTTA? ? ? AAG? TA? AGGA? G? ACAAAACAA? AG? CAG? A? A? ?	308
HIVVI525	-----AGT-----A--G-----A-T-----C--A--C-A-AA-----	342
HIVVI557	-----GAG-----T-A-----G-C-----A-C--A-C-GG-----	342
CONSENSUS.O	AGATAC?CAGCAGGCAATACAAAAGTTAAAG... GAAGTAATGG? ? AGCAGGAAGTC? GC?G	307
HIVANT70	-----G-----GG-----G-----	343
HIVMVP51	-----A-----CA-----A-----	343
SIVCP2GAB	AGACACACAGAAAGCtCTAGAACAGCTAAAGCGCATCATGGAGAACACACAGAGCAAAACTGAAAGTAACT	355

HIV1 GAG

CONSENSUS.A	?????acacagcaggCaGcaGCT	g?cACA...ggaAaC???	agcagcaa?gtc?????	379
HIVU455		AA.....G.....		372
HIVKAL	CAGCTG.....	-C.....AA.....	GT.....	402
HIVVI59		AA.....	A-G.....	384
HIVVI310	-A.....A.....	-A.....	G.....	381
HIVVI57	-A.....	-A.....	A.....	384
HIVK112	-A.....	-AT.....-A-G.....	G.....	384
HIVK88	-A.....	-A.....	G.....	384
HIVK29	-A.....	-CT.....	G.....	384
HIVK124	-A.....	-A.....	G.....	384
HIVK7	-A.....	-A.....	A.....	384
HIVK98	GA.....A.....	-A.....	G.....	384
HIVK89	-A.....	-AT.....	-A-G-T.....	384
HIVVI32	G.....A.....	-A.....	-G-A.....	384
HIVVI415	-A.....	-G.....	C.....	384
HIVCI4		-A.....	AC-G.....	378
HIVG141		-C.....-G.....	-G-GTAC.....	372
HIVLBV23		-A.....	G.....A.....	384
HIVTN243		-G.....-G.....	-A.....	384
HIVTN245		-G.....-G.....	-A.....	384
HIVTN240		-G.....-G.....	-A.....	384
HIVCI20	A.....	-C.....-G.....		360
HIVCI59		-C.....-G.....	G.....	375
HIVLBV2310		-C.....-G.....		375
HIVCI51		-CA.....-G.....		369
HIVLBV105		-A.....-G.....	-A-G-G-AGCCAA	390
HIVCI32	CAGCTG-C.....A.....	-C.....-G-AAT.....		390
HIVCI44	-T.....	-T.....-G.....		375
HIVD3258	-G.....	-C.....-G.....		375
HIVCM238	-T.....	-G.....-G.....	-A.....	384
HIVUG266		AA.....	-A-G.....	384
HIVVI354		-C.....-A.....	-G.....	384
CONSENSUS.B	gcacagcaAgcaGCAgct?????gaCaCA	...GgAaAC?????????agcAgccAgGTC		383
HIVSF2	GCAGCT-G.....			390
HIVB2167	-A...GCT.....G.....	C.....		387
HIVPH153		G.....	T.....T.....	384
HIVPH136			AG.....	384
HIVB2200	-C.....		A.....	384
HIVTB132		-A.....-A.....	-A-T.....	384
HIVB2190	A.....		A.....	384
HIVLAI		C.....		384
HIVHB28		-C.....	-AT.....	384
HIVMN		AGAGGAAAC.....	A.....	393
HIVJH3			A.....	384
HIVJRCSE				384
HIVOI	-A.....			384
HIVNV5CG				384
HIVNL43			-A.....	384
HIVCDC4				384
HIVHAN	-A.....	-G.....	-A-A.....	384
HIVCAN1	-G.....	-G.....		384
HIVRF			G.....	384
HIVD31	-C.....-CG.....	-G.....	-A.....	384
HIVBH102		-C.....	-A-T.....	384
HIVPV22		-C.....	-T.....	384
HIVJRF1	-G.....			378
HIVUG280	-A.....-G.....	-G.....		384
HIVYU2				384
HIVBCSG3C				384

CONSENSUS.C	ACACagcagGcAaaAgcg...gctGAC???.GGaaAg.....GTC.....	374
HIVUG268	-A-C---T.....-AAA.....	369
HIVSM145	-.....-.....-.....	375
HIVZAM18	-.....-C.....	381
HIVZAM19	-A.....-A.....	375
HIVZAM20	-A.....-A.....	375
HIVDJ259	-A---G-A-A.....-AAG.....	378
HIVVI313	-GC.....-AAT.....	378
CONSENSUS.D	gcACAGCAAGcAaCAGCT....GAcaca...agaaac.....AgCAGCCAGGTC.....	383
HIVELI	-G.....-G.....-A.....	384
HIVZ226	-G.....-G-G.....-A.....	387
HIVNDK	A.....-G.....-T.....	375
HIVVI205	-G.....-AG.....	387
HIVG109	-.....-A.....	393
HIVK31	-A.....	384
HIVUG274		384
HIVUG270		384
HIVSE365	-T.....-T.....G.....-A.....	384
HIVVI203		384
CONSENSUS.F	acAcAGCAag?ggCAGCT....gACAAA...GGg.....GTC.....	362
HIVVI174	-G.....-A.....	369
HIVVI69	-A.....-T.....	372
HIVB2162	-A.....-C.....	371
HIVVI325	G.....-AAA.....-A.....	372
CONSENSUS.G	AcACAgcaGgca?????....GAgAAA...GGAAAC.....AgCAGcCAAGTC.....	352
HIVLBV217	-T.....-T.....-G.....-A.....	379
HIVVI191	-AAC.....-G.....-T.....	375
HIVTAIG	-GCAAGT.....	107
CONSENSUS.H	ACACAGCA?GCAACAG?T....GATAA?...G?AAA?...G?CAA?AAG?TC.....	342
HIVVI525	-A.....-C.....-G.....-A---A---G.....	384
HIVVI557	-G.....-G.....-A.....-G---T.....-G---C---A.....	384
CONSENSUS.O	A?GCGGCTAAGGAA.....GA?ACA....AGC.....?C?AGGCAAG?C?.....	336
HIVANT70	-C-----C-----G-A-----G-G.....	378
HIVMVP51	-G-----A-----C-T-----A-A.....	378
SIVCPZGAB	CAGGAAGCCGTGAAGGGGGAAGCCAGTCAAAGGCGCTAGTGCCTCTGCTGGCATT.....	408

HIV1 GAG

		p17 \/ p24	
CONSENSUS.A	?????AGcCAaaAaTACCCcTATAGTgCAaaAaTgcACAgGgCAAAatgatAcacCAG?ccTtATCACcTAg		443
HIVU455	T.....C.....CCAG.....C.....G.....		437
HIVMAL	T.....C.....G.....A.....		467
HIVVI59	C.....AG.....C.....A.....		449
HIVVI1310	T-T-C.....C.....G.....C.....T.....		446
HIVVI57	C.....G.....G.....		449
HIVK112	C.....AA.....G.....		449
HIVK88	C.....A.....G.....		449
HIVK29	C.....C.....T.....G.....		449
HIVK124	C.....AGT.....G.....		449
HIVK7	G.....A.....C.....AA.....G.....		449
HIVK98	C.....AA.....G.....		449
HIVK89	T.....A.....G.....		449
HIVVI32	T.....		449
HIVVI415	G.....C.....G.....A.....		449
HIVCI4	T.....A.....G.....C.....T.....A.....		443
HIVG141	T.....C.....G.....G.....A.....		437
HIVLBV23	C.....G.....T.....C.....		449
HIVTN243	G.....C.....T.....		449
HIVTN245	GC.....C.....T.....		449
HIVTN240	GC.....C.....T.....		449
HIVCI20	C.....T.....A.....G.....		425
HIVCI59	A.....C.....C.....A.....G.....A.....		440
HIVLBV2310	C.....C.....A.....G.....		440
HIVCI51	T.....C.....T.....A.....		434
HIVLBV105	AATGTC.....T.....G.....C.....C.....TG.....		461
HIVCI32	CCT.....G.....G.....C.....G.....C.....		455
HIVIC144	C.....A.....		440
HIVDJ258	G.....C.....A.....G.....		440
HIVCM238	G.....C.....T.....		449
HIVUG266	T.....C.....A.....A.....A.....C.....T.....C.....G.....G.....		449
HIVVI354	T.....C.....G.....A.....G.....		449
CONSENSUS.B	AGCCaAAAATACCCcTATAGTgCAGaAaccTccAGGGgCAAAATGGTACATCaggCCaTATCACcTAg		448
HIVSF2	A.....		455
HIVBZ167	A.....G.....A.....		452
HIVPH153	A.....		449
HIVPH136	A.....C.....		449
HIVBZ200	T.....T.....A.....		449
HIVTB132	A.....TA.....G.....		449
HIVB2190	A.....A.....A.....		449
HIVLAI	A.....		449
HIVHXB2R	A.....		449
HIVMN	C.....A.....G.....		458
HIVJH3	A.....		449
HIVJRCSE	G.....		449
HIVOI1	A.....T.....C.....		449
HIVNY5CG			449
HIVNL43			449
HIVCDC4	T.....		449
HIVHAN	A.....A.....A.....		449
HIVCAM1	A.....		449
HIVRF	T.....A.....		449
HIVD31	C.....		449
HIVBH102	A.....		449
HIVPV22	A.....		449
HIVJRF1	A.....G.....		443
HIVUG280	A.....A.....		449
HIVYU2	A.....		449
HIVBCSG3C	G.....A.....C.....		449

		p17 \ / p24	
CONSENSUS.C		AGTCAAAATTATCCTATAGTgCagAAATCTcCAAGGGCAAATGgTACacCAGgCCaTaTCACCTAG	439
HIVUG268		-----C-----	434
HIVSM145		-----A-----	440
HIVZAM18		-----C-----	446
HIVZAM19		-----C-----	440
HIVZAM20		-----T-----A-----T-----T-----	440
HIVDU259		-----T-----A-----C-----C-----	443
HIVVI313		-----A-----C-----G-----	443
CONSENSUS.D		AGCCAAAATTATCCTATAGTgCAaAAACcTaCagGGGCAaATGGTACacCAgGcCaTaTCACCTAG	448
HIVELI		-----G-----T-----	449
HIVZ226		-----G-----T-----	452
HIVNDK		-----C-----G-----T-----	440
HIVVI205		-----G-----A-----T-----	452
HIVQ109		-----G-----C-----G-----	458
HIVK31		-----T-----G-----	449
HIVUG274		-----A-----T-----	449
HIVUG270		-----A-----T-----	449
HIVSE365		-----C-----A-----C-----	449
HIVVI203		-----C-----G-----C-----	449
CONSENSUS.F		AGTCaAAATTAACCCPATAGTACA?AATcTtCAGGGaCAAATGGTACA?CAGgC?aTATCACCTAG	424
HIVVI174		-----G-----A-----T-----T-----	434
HIVVI169		-----A-----C-----C-----	437
HIVB2162		-----G-----T-----T-----T-----	436
HIVVI325		-----G-----A-----G-----G-----C-----C-----CC-----	437
CONSENSUS.G		AGtCAAAATTATCCTATAGTGCAGaAtgcaCAAGGGCAAAtGgTACACCAGcCCaTaTcCACCTAG	417
HIVLBV217		-----C-----	443
HIVVI191		-----T-----A-----	440
HIVTAIG		-----C-----NNATC-----CC-A-----G-----	172
CONSENSUS.H		AGTCAAAATTATCC?ATAGTACA?AATGCCCAAGG?CAG??GGTACACCAGGCCATATCACCTAG	402
HIVVI525		-----C-----G-----G-----AT-----	449
HIVVI557		-----T-----A-----A-----CC-----	449
CONSENSUS.O		?GTCAAAATTACCCATAGTA?CAAAAGC?CAGGGACAAATGGTACATCA?GCCATCTCCCCAG	397
HIVANT70		-----G-----T-----G-----G-----	443
HIVMVP51		-----A-----A-----A-----A-----	443
SIVCP2GAB		AGTGGAAATTACCCCTCGTACAAATGCTCAAGGGCAGATGOTGCATCAGGCCATCTCACCCAG	473

HIV GAG

CONSENSUS. A	gACtTtTgAAtGCATgGGTgAAAGtAaTAgAagAaAagGcTTTcAgcCcAGAAgTaATACCCAAtGTTtCtCaG	514
HIVU455	---C-----G-----C-----T---	508
HIVMAL	-----G-----	538
HIVVI59	---C-A-C---A-----G-----	520
HIVVI310	-----G-----G-----G-----T---	517
HIVVI57	-----G-----	520
HIVK112	-----A-----G-----	520
HIVK88	-----A-----	520
HIVK29	-----A-----	520
HIVK124	---C-----A-----	520
HIVK7	---A---A-A-----T-----T---	520
HIVK98	---A---A-----A-----	520
HIVK89	---A-A-----A-----	520
HIVVI32	-----T-----C-----	520
HIVVI415	-----A-----	520
HIVCI4	-----A-----T---	514
HIVG141	---A-----	508
HIVLBV23	---CC-----C-----T-----A---	520
HIVTN243	A-----A-----G-----G-T-A-----	520
HIVTN245	A-----G-----G-----T-A-----	520
HIVTN240	A-----G-----G-----T-A-----	520
HIVCI20	-----G-----G-----TA---	496
HIVCI59	-----G-----TA---	511
HIVLBV2310	-----G-----T---	511
HIVCI51	---A---G-C-----G-----T-----T---	505
HIVLBV105	A---AG---A-G-G-----G-G---A-----	532
HIVCI32	-----A-----G-----G-----TA---	526
HIVIC144	---A-----G-----T---	511
HIVDJ258	-----G-----G-----TA---	513
HIVCM238	A-----A-----G---A-----G-T-A-----	520
HIVUG266	---C-A---A-A-----G---G-----T---	520
HIVVI354	---A---A-----A-G-----T-----	520
CONSENSUS. B	AACTTTAAATGCATGGGTAAAGTAgTaGAAGAGaAAGGcTTTcAGCCCAGAAgTaATACCCATGTTTcCAG	519
HIVSF2	-----A-----	526
HIVB2167	-----GA-----	523
HIVPH153	-----	520
HIVPH136	-----	520
HIVB2200	-----G-----	520
HIVTB132	-----A-----G-----	520
HIVB2190	-----A-----	520
HIVLAJ	-----G-----	520
HIVHXB2R	-----G-----	520
HIVMN	-----	529
HIVJH3	-----T-----	520
HIVJRC5F	---A-----C-----	520
HIVQYI	-----	520
HIVNY5CG	-----	520
HIVNL43	-----	520
HIVCBC4	---A-----T-----G---	520
HIVHAN	-----G-----	520
HIVCAM1	-----	520
HIVRF	-----T-----	520
HIVD31	-----	520
HIVBH102	-----	520
HIVPV22	-----G-----	520
HIVJRFL	-----C-----	514
HIVUG280	-----T-----	520
HIVYU2	---G-----G-----	520
HIVBC3G3C	-----	520

HIV1 GAG

CONSENSUS.C	AACCTTGAATGCATGGGTaaaAGTAATAGAGGagAAGGCTTTCAGCCcagag?taatacccatgtttacag	509
HIVUG268	-----A-----G-----	505
HIVSM145	---C-----A---C---T---G---A-----	511
HIVZAM18	-----GA---A---C---T---A-----	517
HIVZAM19	---A-----G-----NA-----	489
HIVZAM20	-----A---AGAG-A-----	511
HIVDJ259	-----AG-----	514
HIVVI313	-----G-----G---G-----	514
CONSENSUS.D	aACCTTgAAcGCATGGGTaaaAGTAATAGaaGA?AAGGctTTcAgcCCaGAaGTaAtACCCATGTTTcCAG	518
HIVELI	-----A-----	520
HIVZ226	-----A-----	523
HIVNDK	-----A---C---G-----	511
HIVVI205	---A---T---G---A---C---G---C-----	523
HIVG109	-----G-----A---G-----G---	529
HIVK31	G---C-----G---G---T-----	520
HIVUG274	---T-----G---G---T-----A---	520
HIVUG270	-----G---G---T-----	520
HIVSE365	G-----C-----G---A-----G-----A---	520
HIVVI203	-----G---A-----AT-----	520
CONSENSUS.F	AACCTTTAAATGCATGGGTaaAgGTgATAGAAAGagAAGGCTTTLAGCCcAGAAGTAATACCCATGTTcTCAG	495
HIVVI174	-----A-----C-----	505
HIVVI69	-----T-----	508
HIVBZ162	-----	507
HIVVI325	---G---A---A---A---C-----	508
CONSENSUS.G	aACCTTgAAATGCATGGGTAAAAGTAgTAGAAGAAAAGgcCTTCAGTCCAGAAGTAATACCCATGTTTcCAG	488
HIVLBV317	G---A-----	514
HIVVI191	-----A-----AA-----	511
HIVTAIG	-----	243
CONSENSUS.H	GACCTTAAATGCATGGGTAAAAGTAGTAGAAGA?AAGGCTTTTAGCCcAGAAGTAATACC?ATGTTTTCAG	471
HIVVI525	-----A-----T-----	520
HIVVI557	-----G-----C-----	520
CONSENSUS.O	GACTTPTAAATGCATGGGTAAAGGCAGTAGAAGA?AA?GCCTTTAAACCTTGAAT?ATTCCATGTT?ATGG	464
HIVANT70	-----A---A-----C-----C-----	514
HIVNVP51	-----G---G-----T-----	514
SIVCPZGAB	AACCTTGAATGCATGGGTAAAAGTGGTAGAGGAAAAGGCATTAGCCCTGAGGTAAATACCAATGTTTTCAG	544

HIV1 GAG

CONSENSUS .A	caTTaTCAGAgGGaGCGcACccCaCAagATtTaAAatagATGcTgAAcAtAOTggggGGaCacCAgGCAGCt	585
HIVU455	-----TG-----	579
HIVMAL	-----G-----T-A-----	609
HIVVI59	A-----A-----A-TG-----A-A-----	591
HIVVI310	-----A-----T-----A-----C-----	588
HIVVI57	-----A-----G-----G-----	591
HIVKI12	-----A-----A-----	591
HIVK88	-----A-----	591
HIVK29	-----A-----T-----	591
HIVKI24	-----A-----	591
HIVK7	-----A-----G-----A-----	591
HIVK98	-T-----A-----G-----	591
HIVK89	-----A-----T-----	591
HIVVI32	---G-----A-----A-T-----	591
HIVVI415	-----A-----T-----A-----	591
HIVCI4	-----A-----T-----A-----	585
HIVGI41	-----A-----T-----A-----	579
HIVLBV23	-----A-----T-----A-----	591
HIVTN243	-----A-----T-----A-----	591
HIVTN245	-----A-----T-----A-----	591
HIVTN240	-----T-A-----T-----A-----	591
HIVCI20	-----A-----T-----A-----	567
HIVCI59	-----G-----A-----A-----	582
HIVLBV2310	-----G-----A-----A-----	582
HIVCI51	-----G-----A-----A-----	576
HIVLBV105	-----T-----A-----A-----	603
HIVCI32	-----A-----G-----C-CC-----T-A-----C-----A-----A-----C-----	597
HIVIC144	-----G-----A-----A-----A-----	582
HIVDJ258	-----C-----A-----A-----A-----	582
HIVCM238	-----A-----T-----A-----A-----	591
HIVUG266	-----A-----T-----TG-----A-AA-----	591
HIVVI354	-----A-C-----A-----A-----TAAA-----T-----	591
CONSENSUS .B	CAITTAACAGAGGGaGCGCACCCCaCAAGATtTAaAcACcATGcTAAACACAGTGGGGGACaLCAAGCAGCc	590
HIVSF2	-----T-----C-----	597
HIVB2167	-----T-----C-----	594
HIVPH153	-----T-----	591
HIVPH136	-----T-----	591
HIVB2200	-----	591
HIVTB132	-----	591
HIVB2190	-----G-----	591
HIVLAI	-----	591
HIVHXB2R	-----	591
HIVMN	-----G-----	600
HIVJH3	-----G-----	591
HIVJRC5F	-----T-----	591
HIVQYI	-----G-----C-----	591
HIVNY5CG	-----T-----	591
HIVNL43	-----T-----	591
HIVCD4	-----C-----	591
HIVHAN	-----	591
HIVCAM1	-----	591
HIVRF	-----	591
HIVD31	-----	591
HIVBH102	-----	591
HIVPV22	-----	591
HIVJRFL	-----	585
HIVUG280	-----	591
HIVYU2	-----C-----	591
HIVBCSG3C	-----T-----	591

CONSENSUS .C	cattatcagaaggagccacccacagaatttaaacacccatgttaaatacagtggggggacatcaagcagcc	580
HIVUG268	-----C-----	576
HIVSM145	-----G-----	582
HIVZAM18	-----	588
HIVZAM19	-----	489
HIVZAM20	-----C-----C-----	582
HIVDJ259	-----	585
HIVVI313	-----G-----A-----	585
CONSENSUS .D	CATtATCAGAAGGAGCCACCCACAAgATTtAAAcACCATGcTAAAcACAGtGGGGGGaCATCAaGCAGCc	589
HIVELI	-----	591
HIVZ226	-----T-----	594
HIVNDK	-----T-----	582
HIVVI205	-----T-----G-----T-----T-----	594
HIVGI09	---G---C---T---T---C---T---	600
HIVK31	-----G-----G-----	591
HIVUG274	-----G-----	591
HIVUG270	-----G-----	591
HIVSE365	---C---G---T---	591
HIVVI203	-----T-----	591
CONSENSUS .F	CATTATCAGAAgg?GCCACtCCACAAGATtTAAACACCATGcTAAA?ACAGTGGGGGGACATCAaGCAGC?	563
HIVVI174	-----G---C-----T-----T---	576
HIVVI69	-----A-----C-----G---C---	579
HIVB2162	-----G-----T---T---C---	578
HIVVI325	---G---T---A---C-----T---	579
CONSENSUS .G	CATTATCAGAAggAGCCACcCCACAAGATtTAAATAcCatGcTAAAcACAGTGGGGGGgCATCAAGCAGCT	559
HIVLBV217	-----T---C-----T-----A-----	585
HIVVI191	-----	582
HIVTAIG	-----G-----C-T-TA-TC-----	314
CONSENSUS .H	CATTATCAGAAggAGCCACCCACAAAGATtTAAATGCCATG?TAAATA?AGTGGGGGG?CATCAAGCAGCC	538
HIVVI525	-----C-----C-----G-----	591
HIVVI557	-----T-----T-----T-----A-----	591
CONSENSUS .O	CATT?TCAGA?GG?GC??T??CCTATGATAT?AATAC?ATGCT?AATGCCATAGG?GGACA?CAAGG?GCT	522
HIVANT70	---G---G---A---AA-TT-----T-----T---A-----A---T---A---	585
HIVNVP51	---A---A---G---TG-CC-----C---C---G-----G---C---G---	585
SIVCP226AB	CTCTATCAGAGGGAGCCCTTACCTCAGGATGTTAACACCATGTtTAAATGCAGTAGGGGGACATCAAGGAGCC	615

HIV GAG

```

CONSENSUS . A ATGCAaATGtTaaaagAcaCcaTcAATGagGAaGCTGCAGaaTgGGAcAag?tACatCCAATcCatGCAGG 655
HIVU455 -----G-----T-----G----- 650
HIVMAL -----C-----G----- 680
HIVV159 -----A-----A-----T-----G-----A----- 662
HIVV1310 -----G-----G----- 659
HIVV157 -----T----- 662
HIVK112 -----T----- 662
HIVK88 -----T----- 662
HIVK29 -----T----- 662
HIVK124 -----A-----T----- 662
HIVK7 -----C-----A-----A-----T----- 662
HIVK98 -----C-----A-----T-----T----- 662
HIVK89 -----T----- 662
HIVV132 -----G-----G-----T----- 662
HIVV1415 -----A-----T----- 662
HIVC14 -----G-----G----- 656
HIVG141 -----C-----AC----- 650
HIVLBV23 -----T-----G-----G-----G----- 662
HIVTN243 -----A-----T-----G-----C----- 662
HIVTN245 -----A-----T-----G-----C----- 662
HIVTN240 -----A-----T-----G-----C----- 662
HIVC120 -----G-----A-----A-----G-----C----- 638
HIVC159 -----G-----G-----G-----G----- 653
HIVLBV2310 -----AG----- 653
HIVC151 -----G-----G-----AAC----- 647
HIVLBV105 -----AC-----G----- 674
HIVC132 -----T-----AG----- 668
HIVC144 -----G-----G-----G----- 653
HIVDJ258 -----G-----G----- 653
HIVCM238 -----A-----T-----G-----C----- 662
HIVUG266 -----C-----G-----A-----A-----T-----G-----A-----A----- 662
HIVV1354 -----A-----A-----G----- 662

CONSENSUS . B ATGCAAAATGtTAAAAGagACcATCAATGAGGAAGCTGCAGAAATgGGATAGAtTgCatCCAGtgCatGCAGG 661
HIVSF2 -----T-----G----- 668
HIVB2167 -----T----- 665
HIVPH153 ----- 662
HIVPH136 ----- 662
HIVB2200 -----A-----A----- 662
HIVTB132 ----- 662
HIVB2190 ----- 662
HIVLAI -----G----- 662
HIVHXB2R -----G----- 662
HIVMN ----- 671
HIVJH3 -----A-----C-----G----- 662
HIVJRCSE -----C-----A----- 662
HIVQYI -----A----- 662
HIVNY5C3 ----- 662
HIVNL43 ----- 662
HIVCDC4 ----- 662
HIVHAN -----A-----C----- 662
HIVCAM1 ----- 662
HIVRF -----T----- 662
HIVD31 ----- 662
HIVBH102 -----G-----A----- 662
HIVPV22 -----G----- 662
HIVJRFL -----C----- 656
HIVUG280 ----- 662
HIVYU2 ----- 662
HIVBCSG3C -----A----- 662

```

CONSENSUS.C	atgcaaattgttaaaagataccatcaatgaggaggctgcagaatgggatag?ttacatccagt?catgcagg	649
HIVUG268	-----C-----G-----G-----	647
HIVSM145	-----A-----A-----A-----	653
HIV2AM18	-----G-----A-----A-----	659
HIVZAM19	-----C-----AA-----A-----	489
HIVZAM20	-----C-----AA-----A-----	653
HIVDJ259	-----G-----G-----G-----	656
HIVVI313	-----A-----G-----G-----	656
CONSENSUS.D	ATGCAAATGCTAAAGAGACCATCAATGAGGAAGCTGCAGAAATGGGATAGgcTACATCCAGtgCATGCAGG	660
HIVELI	-----C-----A-----T-----	662
HIV2226	-----C-----G-----T-----	665
HIVNDK	-----C-----C-----C-AT-----	653
HIVVI205	-----C-----T-----C-----	665
HIVG109	-----C-----C-----C-----	671
HIVK31	-----C-----C-----C-----	662
HIVUG274	-----A-----G-----G-----	662
HIVUG270	-----A-----A-----A-----	662
HIVSE365	-----A-----AT-----AT-----	662
HIVVI203	-----A-----AT-----AT-----	662
CONSENSUS.F	ATGCAAATGCTTAAAGACACCATCAATGAGGAAGCTGCAGAAATGGGACAGaTTACatCCAGtGCA?GCAGG	633
HIVVI174	-----C-----G-----G-----	647
HIVVI69	-----T-----T-----T-----	650
HIVB2162	-----C-----C-----C-----	649
HIVVI325	-----T-T-----G-----G-----C-----T-----	650
CONSENSUS.G	ATGCAAATGCTAAAGGATACCATCAATGAGGAAGCTGCAGAGTGGGACAGGATACATCCACCACAgGCAGG	630
HIVLBV217	-----T-----A-----A-----	656
HIVVI191	-----G-----G-----G-----	653
HIVFAIG	-----TC-----A-----A-----A-----	385
CONSENSUS.H	ATGCAAAT7?TAAAGATACATCAATGAGGAAGCTGCAGAAATGGGATAGG7TACATCCAGTGCATGCAGG	606
HIVVI525	-----AC-----C-----C-----	662
HIVVI557	-----GT-----G-----G-----	662
CONSENSUS.O	TTACAAGTG7T?AAGGAAGTAATCAATGAGGAAGCAG?AGA?TGGGATAGAACTCA?CCACCA?C7T?GG	584
HIVANT70	-----C-A-----T-----G-----C-----C-GG-A-----	656
HIVMVP51	-----T-G-----C-----A-----T-----G-AA-G-----	656
STVCPZGAB	ATGCAAGTGCTAAAGGAAGTCATTAATGAGGAAGCCGCCGAGTGGGATCGCTTACATCCCACTCATGCTGG	686

HIV1 GAG

CONSENSUS .A	GCCTattcCaCcagGcCagaTgAGaGAACCAaggGGaAGtGacATAGCAGGaacTACTAGtAcCctTcaaG	726
HIV0455	-----C-----A-----C-----G-----	721
HIVMAL	-----C-----A-----	751
HIVVI59	-----AA-----	733
HIVVI310	---G-----A-----G-----C-----C-----	730
HIVVI57	-----A-----G-----C-----C-----	733
HIVK112	-----C-----	733
HIVK88	-----C-----	733
HIVK29	-----C-----	733
HIVK124	-----C-----	733
HIVK7	---C-----A-----C-----	733
HIVK98	-----C-----	733
HIVK89	-----A-----C-----	733
HIVVI32	-----A-----C-----	733
HIVVI415	-----A-----AC-----	733
HIVCI4	-----T-----	727
HIVG141	---GC-----A-----	721
HIVLBV23	---TA-----C-----G-----GC-----	733
HIVTN243	-----G-----	733
HIVTN245	-----T-----G-----	733
HIVTN240	-----G-----	733
HIVC120	-----G-----G-----A-----	709
HIVCI59	-----G-----G-----A-----	724
HIVLBV2310	-----G-----A-----	724
HIVCI51	-----G-----GT-----	718
HIVLBV105	-----T-----	745
HIVCI32	---G-----T-----G-----	739
HIVIC144	-----A-----G-----A-----	724
HIVDJ258	-----A-----G-----	724
HIVCM238	-----G-----	731
HIVUG266	-----T-----A-----G-----	733
HIVVI354	-----A-T-----A-----T-----AC-----	733
CONSENSUS .B	GCCTaTTgCaCCAGGcCagATGAGAGAACCAAGGGaAGTgACATAGCAGGaaCTACTAGTACCCCTCAGG	732
HIVSF2	-----A-----	739
HIVB2167	-----A-----	736
HIVPH153	-----	733
HIVPH136	---G---G-----	733
HIVB2200	-----	733
HIVTB132	-----	733
HIVB2190	-----	733
HIVLAI	-----	733
HIVHXB2R	-----	733
HIVMN	-----A-----	742
HIVJH3	-----	733
HIVJRC5F	-----G-----	733
HIVQYI	-----	733
HIVNY5CG	-----	733
HIVNL43	-----	733
HIVCDC4	-----	733
HIVHAN	-----	733
HIVCAN1	-----	733
HIVRF	-----T-----C-----	733
HIVD31	-----	733
HIVBH102	-----	733
HIVPV22	-----	733
HIVJRF1	-----A-----	727
HIVUG280	-----G-----	733
HIVYU2	-----	733
HIVBCSG3C	---G-----	733

CONSENSUS .C	gcctgttgccaccaggccaaaTgAGAGaCCaAGGGGAAGTGACATAGCAGGAACtACTAGTAcCCCTTCAGG	720
HIVUG268	-----A-----	718
HIVSM145	-----C-----	724
HIVZAM18	---AA-----T-----	730
HIVZAM19	-----	541
HIVZAM20	-----	724
HIVDJ259	---A-----A-----	727
HIVVI313	---C-----A-----G-----	727
CONSENSUS .D	gCCTtATTGCACCAGGcCagATGAGAGaCCaAGGGGAAGTGATATAGCAGGAACtACTAGTAcCCCTTCAGG	731
HIVELI	-----	733
HIVZ226	-----	736
HIVNDK	---G-----A-----	724
HIVVI205	-----	736
HIVG109	-----	742
HIVK31	-----A-----A-----	733
HIVUG274	A-----A-A-----C-----	733
HIVUG270	-----A-----G-----C-----T-----	733
HIVSB365	-----	733
HIVVI203	-----	733
CONSENSUS .F	gCCTAtCCCAcCAGGCCAAGT?AGgGAACCTAGGGGgAGTGATATAGCtGGAACTACTAGtACCCCTTCAGG	703
HIVVI174	-----A-----C-----	718
HIVVI169	---A-----G-----	721
HIVB3162	A-----A-----	720
HIVVI325	-----G-A-----G-----A-----C-C-----	721
CONSENSUS .G	gCCTATTcCACCAGGcCAGATAAGsGA?CCaAGGGGAAGTGATATAGCAGGgACTACTAGTAcCCCTaCAGG	700
HIVLSV217	-----C-----G-----G-----	727
HIVVI191	A-----G-A-G-----	724
HIVTAIG	---G-----T-----T-----	456
CONSENSUS .H	GCCTATTCCACCAGGCCA?ATGAGAGAACCAAGGGGAAG?GATATAGCAGGAACtACTAGTAcCCCTTCAGG	675
HIVVI525	-----A-----C-----	733
HIVVI557	---G-----T-----	733
CONSENSUS .O	GCCGTT?CCACCAGGcCAGATAAGGGAACCAACAGGAAGTGACMTTgCTGG?ACAACtAGCAC?CAGCAAG	652
HIVANT70	---G-----G-----C-----	727
HIVMVP51	---A-----A-----A-----	727
SIVCP2GAB	TCCGATAGCCcCAGGACAGCTAAGGGAGCCaAGGGGTAGTGATATAGCAGGGACCCACCAGCACACTGCAGG	757

HIV1 GAG

CONSENSUS. A	AACAAaTaggeTgGATgACaagc.....	AAtCCACCTatCCCaGTgGGagAcATcTATaaaAgaTgg	788
HIVU455	-----G-----	-----G-----	783
HIVMAL	-----C-----	-----C-----	813
HIVVI59	-----A--A-----G-----	-----C-----A--T-----G-----	795
HIVVI310	-----C-----	-----C-----G-----A--A-----G-----	792
HIVVI57	-----C-----	-----C-----G-----G-----	795
HIVK112	-----C-----	-----C-----	795
HIVK88	-----G-----	-----C-----	795
HIVK29	-----G-----	-----C-----	795
HIVK124	-----T-----	-----C-----	795
HIVK7	-----T-CA-----	-----A-----	795
HIVK98	-----T-CA-----	-----C-----	795
HIVK89	-----C-----G-----	-----C-----G-----	795
HIVVI32	-----C-----	-----C-----G-----T-----	795
HIVVI415	-----C-----G-----	-----C-----A-----	795
HIVCI4	-----G-----	-----G-----G-----G-----G-----	789
HIVG141	-----C-----	-----C-----A-----	783
HIVLBV23	-----CA-----	-----G--A-----	795
HIVTN243	-----A-----	-----G-----	795
HIVTN245	-----A-----	-----G-----	795
HIVTN240	-----A-----	-----G-----	795
HIVCI20	-----G-----	-----A-----	771
HIVCI59	-----G--A-----	-----G--A-----	786
HIVLBV2310	-----A-----	-----A-----	786
HIVCI51	-----C-----	-----C-----A-----A-----	780
HIVLBV105	-----A--P-----	-----C--G-----	807
HIVCI32	-----C-----A-----	-----C-----G--A-----A-----	801
HIVCI144	-----T-----	-----A-----	786
HIVDJ258	-----A-----	-----G-----	786
HIVCM238	-----G-CA--A-----A-----	-----C-----A-----A-----G-----	795
HIVUG266	-----G-CA--A-----G-----	-----C-----A-----A-----G-----A-----	795
HIVVI354	-----A--A-----	-----A--A-----	795
CONSENSUS. B	AaCAAATAGGaTGGATgACAAaT.....	AATCCACCTatCCCaGTaGGAgAAATcTATAAAAGaTgG	794
HIVSP2	-----A-----	-----A-----	801
HIVB2167	-----A-----	-----A-----	798
HIVPH153	-----A-----	-----A-----	795
HIVPH136	-----A-----	-----A-----	795
HIVB2200	-----A-----	-----A-----	795
HIVTB132	-----A--C-----	-----T--C-----	795
HIVB2190	-----T-----	-----T-----	795
HIVLAI	-----T-----	-----T-----	795
HIVMXB2R	-----T-----	-----T-----	795
HIVM1	-----G-----	-----G-----	804
HIVJH3	-----G-----	-----G-----	795
HIVJRCSE	-----G-----	-----G-----	795
HIVQY1	-----C-----	-----C-----	795
HIVNY5CG	-----C-----	-----C-----	795
HIVNL43	-----C-----	-----C-----	795
HIVCDC4	-----C-----	-----C-----	795
HIVHAN	-----A-----	-----A-----	795
HIVCAM1	-----G-----	-----G-----	795
HIVRF	-----G-----	-----G-----	795
HIVD31	-----G-----	-----G-----	795
HIVBH102	-----T-----	-----T-----	795
HIVPV22	-----T-----	-----T-----	795
HIVJRP1	-----T-----	-----T-----	789
HIVUG280	-----T-----	-----T-----	795
HIVYU2	-----T-----	-----T-----	795
HIVBCSG3C	-----T-----	-----T-----	795

CONSENSUS . C	AACCAATAgcaTGGATGACaa?T.....AACCCaCCTaTCCAGTgGGAGAcATCTATAAAAGATGG	781
HIVUG268	-----AAT-----G-----A-----	780
HIVSM145	-----G-----A-----G-----	786
HIVZAM18	---G---A---G---	792
HIVZAM19	---G---A---G---	803
HIVZAM20	-----GC-----G-----A-----	786
HIVDJ259	-----G-----C-----A-----	789
HIVVI313	-----A-----A-----	789
CONSENSUS . D	AaCAAAaTAG?atCGATGACAgc.....AAcCCaCCcATCCcAGTAggAgAAATcTAtAAAaGATGG	792
HIVBLJ	-----C-----T-----C-----	795
HIVZ226	-----C-----C-----	798
HIVNDK	-----C-----C-----	786
HIVVI205	-----C-----C-----AA-A-----	798
HIVGL09	-----CC-----GC-----A-----A-----	804
HIVK31	-G-----G-----A-----C-----	795
HIVUG274	-----G-----G-----	795
HIVUG270	-----G-----G-----	795
HIVSE365	-----G-----AT-----C-----	795
HIVVI203	-----G-----T-----	795
CONSENSUS . F	AACAAATAcasTGOATGACAAGC.....AAcCCACCTgTCCCAOTGGGAGaaATCTATAAAAGATGG	765
HIVVI174	-----A-----A-----	780
HIVVI69	-----G-----T-----A-----C-----	783
HIVBZ162	-----AC-----	782
HIVVI325	-----AC-----	783
CONSENSUS . G	AACAAATaAGATGGATGACCAGc.....AATCCACCTATCCcAGTgGGAGaAATTTATAAAAGATGG	762
HIVLBV217	-----C-----T-----C-----	789
HIVVI191	-----C-----T-----C-----	786
HIVTAIG	-----A-----	518
CONSENSUS . H	AACAAATAGCATGGATGACA?GC.....AATCC??CTATCCcAGTGGGAGACATCTATAAAAGATGG	734
HIVVI525	-----G-----AG-----	795
HIVVI557	-----A-----CC-----	795
CONSENSUS . O	AGCAAAAT?T?TGGACTAC?AG?...??C?AAC???CTATCCcAGTAGGAGACATCTATAGAAAATGG	706
HIVAMT70	-----CAC-----C-G-----C-C---CAAC-----	792
HIVMB51	-----ATA-----T-A...GGGG-T-----T-----	792
SIYCP2GAB	AACAAATTGGGTGGACAAACAGCA.....AATCCTCCcATCCcAGTAGGGGATGTTTATAGAAAGATGG	819

HIV1 GAG

CONSENSUS .A	ATAATccTgGGaTTAAATAAAATAGTaAGaATGTATAGcCCTgttAGCATttTgGAcATAAgaCAAgggCC	859
HIVU455	-----C-----	854
HIVMAL	-----C-----	884
HIVVI59	-----A-----A-----G-----A-----	866
HIVVI310	-----C-----A-----A-----	863
HIVVI57	-----A-----	866
HIVK112	-----T-----A-----	866
HIVK88	-----TG-----A-----A-----	866
HIVK29	-----T-----A-----	866
HIVK124	-----A-----	866
HIVK7	-----A-----T-----A-----	866
HIVK98	-----T-----A-----	866
HIVK89	-----T-----A-----	866
HIVVI32	-----A-----	866
HIVVI415	-----T-----G-----AC-----C-----TG-----A-----	866
HIVCI4	-----G-----	860
HIVG141	-----C-----	854
HIVLBV23	-----A-----G-----A-----	866
HIVTN243	-----A-----	866
HIVTN245	-----	866
HIVTN240	-----	866
HIVCI20	-----G-----T-----C-----A-----	842
HIVCI59	-----G-T-----C-----	857
HIVLBV2310	-----G-T-----C-----	857
HIVCI51	-----G-----C-----	851
HIVLBV105	-----G-----C-----AG-----	878
HIVCI32	-----C-----C-----T-----	872
HIVCI144	-----G-----C-----	857
HIVDJ258	-----C-----	857
HIVCM238	-----A-----	866
HIVUG266	-----A-C-----A-----A-----	866
HIVVI354	-----C-----	866
CONSENSUS .B	ATAATccTgGGaTTAAATAAAATAGTaAGaATGTATAGcCCTaccAGCATtTgGAcATAAgaCAAGGaCC	865
HIVSF2	-----G-----T-----T-----	872
HIVB2167	-----G-----T-----T-----	869
HIVPH153	-----A-----G-----T-----T-----	866
HIVPH136	-----TA-----G-----T-----T-----A-----	866
HIVB2200	-----G-----G-----	866
HIVTB132	-----A-----G-----T-----	866
HIVB2190	-----A-----A-----	866
HIVLAI	-----	866
HIVHXB2R	-----	866
HIVMN	-----G-----T-----	875
HIVJH3	-----G-----T-----	866
HIVJRCSP	-----G-----G-----GT-----	866
HIVQYI	-----	866
HIVMYSCG	-----	866
HIVNL43	-----	866
HIVCDC4	-----C-T-----A-----	866
HIVHAN	-----T-----A-----	866
HIVCAMI	-----T-----C-T-----G-----	866
HIVRF	-----T-----C-T-----	866
HIVD31	-----	866
HIVBH102	-----	866
HIVPV22	-----G-----G-----	866
HIVJRFL	-----G-----	860
HIVUG280	-----T-----G-----T-----A-----	866
HIVYU2	-----T-----T-----	866
HIVBCSG3C	-----G-----P-----T-----T-----	866

CONSENSUS.C	ATAATTCCTGGGGTTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTCGGACATAAAACAAGGGCC	852
HIVUG268	-----	851
HIVBM145	-----A-----	857
HIVZAM18	-----C-----	863
HIVZAM19	-----G-----C-----G-----	674
HIVZAM20	-----C-----	857
HIVDJ259	-----G-----	860
HIVVI313	-----	860
CONSENSUS.D	ATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTCGGACATAAGACAAGGACC	863
HIVELI	-----TG-----	866
HIVZ226	-----G-----	869
HIVNDK	-----G-----	857
HIVVI205	-----A-----A-----T-----G-----	869
HIVGI09	-----A-----	875
HIVK31	-----A-----	866
HIVUG274	-----A-----T-----A-----	866
HIVUG270	-----A-----	866
HIVSS365	-----G-----T-----	866
HIVVI203	-----T-----	866
CONSENSUS.F	ATCATCCTAGGATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTCGGACATAAGACAAGGGCC	836
HIVVI174	-----	851
HIVVI69	-----	854
HIVBZ162	-----	853
HIVVI325	---A---G---T---C---A---	854
CONSENSUS.G	ATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTCGGATTAAAGACAAGGGCC	832
HIVLBV217	-----G-----C-----	860
HIVVI191	-----C-----T-----	857
HIVTAIG	-----	560
CONSENSUS.H	ATAATCCTGGGATTAAATAAATAGTAAGAATGTATAG?CCTGTCAGCATTTCGGACATAA?ACAAGG?CC	801
HIVVI525	-----G-----T-----A-----A-----	866
HIVVI557	-----A-----C-----G-----G-----	866
CONSENSUS.O	ATAGTG?TAGGACTAAACAAAATGGTAAAAATGTACAG?CCAGTGAGCATCTTAGATATTA?GCAGGGACC	774
HIVAM70	-----T-----C-----A-----	863
HIVMVP51	-----C-----T-----G-----	863
SIVCPZGAB	GTAATTTTAGGGTTAAACAAAGTGGTTAGAAATGTATGCCCCAGTAAGTATCCTTGACATCAGACAAGGGCC	890

HIV1 GAG

CONSENSUS .A	aaAaGAaCCcTTcagaGAcTATGTaGAtAagTTCtTAAaactcTcAGAGCtgAgcAaGCTaCaCagGagG	930
HIVU455	-----G--T-----A-----T-----	925
HIVMAL	---G---T--T-----AG-----C--A-----T-----	955
HIVVI59	-----T-----AG-----C--A-----T-----	937
HIVVI310	-----T-----A-----G-----CA-----	934
HIVVI57	G-----T-----OGT-----T-----C-----	937
HIVK112	-----T-----T-----C-----	937
HIVK88	-----T-----T-----C-----	937
HIVK29	-----T-----T-----C-----	937
HIVK124	-----T-----T-----C-----	937
HIVK7	-----T-----T-----C-----	937
HIVK98	-----G-----T-----G-----A-----T-----T-----	937
HIVK89	-----G-----T-----G-----A-----T-----T-----	937
HIVV132	-----T-----A-----A-----A-----	937
HIVVI415	-----T-----T-----T-----T-----	937
HIVCI4	-----T-----T-----TG-----A-----	931
HIVGI41	-----T--T--G-----	925
HIVLBV23	-----T-----T-----T-----	937
HIVTN243	-----A-----G--A-----	937
HIVTN245	-----A-----G--A-----	937
HIVTN240	-----A-----G--A-----	937
HIVCI20	-----T-----T-----T--G-----A-----G-----	913
HIVCI59	-----T-----T-----T--G-----AT-----G-----	928
HIVLBV2310	-----T-----T-----GT--T--G-----A-----G-----	928
HIVCI51	---G-----T-----T-----T--G-----A-----G-----	922
HIVLBV105	-----T-----C-----T--A-----C-----	949
HIVCI32	-----T-----T-----A-----T--G-----A-----T-----	943
HIVCI144	-----T-----T-----G-----A-----T--G-----	928
HIVDI258	-----T-----T-----T--G-----A-----G-----	928
HIVCN238	-----A-----G--A-----	937
HIVUG266	-----T--A--A-----A-----T-----	937
HIVVI354	-----G--A-----A-----G-----A-----	937
CONSENSUS .B	aaAgGAaCCcTTTAGAGAcTATgTAGAccGgTTCtATAAaACtCTAAGAGCcGAgCAaGCTcCACAGGagG	936
HIVSF2	-----T-----A-----	943
HIVB2167	-----G-----T-----	940
HIVPH153	-----T-----	937
HIVPH136	-----A-----	937
HIVB2200	-----T-----G-----	937
HIVT2132	-----G-----C-----	937
HIVB2190	-----T-----A-----	937
HIVLAI	-----A-----	937
HIVHXB2R	-----	937
HIVH8N	-----	946
HIVJH3	-----	937
HIVORCSP	-----C-----A-----	937
HIVQYI	-----T-----	937
HIVNY5CG	-----	937
HIVNL43	-----A-----A-----	937
HIVCDC4	-----A-----G-----T-----	937
HIVHAN	-----A-----T-----C-----A-----A-----	937
HIVCAM1	-----A-----T--A-----	937
HIVRF	T-----T-----	937
HIVD31	-----A-----	937
HIVBH102	-----A-----T-----	937
HIVPV22	-----A-----C-----	937
HIVJ8FL	-----	931
HIVUG280	-----	937
HIVYU2	-----T-----	937
HIVBCSG3C	-----T-----T-----A-----T-----	937

CONSENSUS.C	AAAGGAACCCCTTAGAGACTATGTAGATCGGTTCTTTAAACCTTAAGAGCTGAACAAGCTaCaCAAGATG	923
HIVUG268	-----C-----	922
HIVSM145	-----C-----	928
HIV2AM18	-----T-----	934
HIV2AM19	-----C-----	745
HIV2AM20	-----C-----	928
HIVDJ259	-----T-----C-----AT-T-----	931
HIVVI313	-----C-----	931
CONSENSUS.D	AAAggAACCCCTTAGAGACTATGTAGA?CGGTTCTATAAAACCTTAAGAGCTGAGCAAGCTcCACAgGATG	933
HIVELI	-----T-----C-----	937
HIVZ2Z6	-----T-----C-----A-----	940
HIVNDK	-----T-----C-----	928
HIVVI205	---AA---T-----C-----	940
HIVB109	---A-----C-----	946
HIVK31	-----T-----T-----T-----A-----	937
HIVUG274	-----T-----T-----G-----	937
HIVUG270	---A-----T-----A-----G-----A-----	937
HIVSE365	---A-----T-----T-----A-----A-----G-----	937
HIVVI203	-----T-----T-----A-----A-----G-----	937
CONSENSUS.F	AAAAGAACCCCTTAGAGACTATGTAGACAGGTTCTTTAAACCTTAAGAGCTGAGCAAGCTACACAGGAAG	906
HIVVI174	-----T-----	922
HIVVI69	-----A-----C-----G-----	925
HIVB2162	---G-----C-----	924
HIVVI325	---T-----T---C-----T-----A-----C-----	925
CONSENSUS.G	AAAAGAACCCCTTAGAGATTATGTAGATAGGTTCTTTAAAC??TGAGAGCTGAGCAAGCTACACAGGA?G	900
HIVLBV217	-----TT-----	931
HIVVI191	-----AC-----A-----	928
CONSENSUS.H	AAAAGAACCCCTTAGAGACTATGTAGA?AGGTTCTTTAAA?CTTTAAGAGCTGA?CAAGC?ACACAGGA?G	867
HIVVI525	-----C-----A-----G-----C-----T-----	937
HIVVI557	-----T-----G-----A-----T-----G-----	937
CONSENSUS.O	AAA?GAACCTTTAGAGA?TATGTAGATCGGTT?TACAAAACATTAAGAGCTGAGCAAGCTACTCAAGAAG	841
HIVANT70	---G-----T-----C-----C-----	934
HIVMVP51	---A-----C-----T-----T-----	934
SIVCPZGAB	AAAGGAGCCCTTTAGAGATTATGTGGACAGATTCTACAAAACCTTAAGGGCTGAGCAAGCTTCACAGGAAG	961

HIV1 GAG

CONSENSUS .A	TAAAAaacTggATGACaGAaaACctTgcTggTCCAAAAatCCgAAatCCAGAcTGTAAgtCCCApTttTAAGAgca	1001
HIVU455	-----T-----A-----A-----	996
HIVMAL	-----T-----A-----A-----	1026
HIVVI59	-----T-----A-----A-----	1008
HIVVI310	-----A-----T-----A-----	1005
HIVVI57	-----T-----T-----T-----	1008
HIVK112	---GGT---A-A---A---A---T---G---	1008
HIVK88	---GGT---A-A---A---A---T---G---	1008
HIVK29	---GGT---A-A---A---A---T---G---	1008
HIVK124	---GGT---T---A-A---A---A---T---G---	1008
HIVK7	---GGT-A---A-A---A---A---T---A---	1008
HIVK98	---GGT---A-A---A---A---T---A---	1008
HIVK89	---T---A---A---A---T---A---	1008
HIVVI32	---GGT---T---A---A---T---A---	1008
HIVVI415	---GGT---A---A---A---T---A---	1008
HIVCI4	---T---A---A---A---T---A---	1002
HIVG141	---T---A---A---A---T---A---	996
HIVLBV23	-----A---T---A---A---A---	1008
HIVTN243	-----A---A---A---A---A---	1008
HIVTN245	-----A---A---A---A---A---	1008
HIVTN240	-----A---C---A---A---A---	1008
HIVCI20	-----A---A---A---A---A---	984
HIVCI59	-----A---A---A---A---A---	999
HIVLBV2310	-----A---A---A---A---A---	999
HIVCI51	-----A---A---A---A---A---	993
HIVLBV105	---G-T---C---C---T---C---T---AA---	1020
HIVCI32	-----A---A---A---A---A---	1014
HIVIC144	-----A---A---A---A---A---	999
HIVDJ258	-----A---A---A---A---A---	999
HIVCM238	-----A---A---A---A---A---	1008
HIVUG246	---A---A---A---A---T---C---C---	1008
HIVVI354	---T-A---C---A---T---A---C---G---	1008
CONSENSUS .B	TAAAAAAatcGGatGACAGAAACCTTgtTGGTCCaAAATgCgAAcCCAGATTGTAAgACTATTTTAAAGCA	1007
HIVSF2	-----A-----A-----A-----	1014
HIVBZ167	-----A-----A-----A-----	1011
HIVPH153	-----C-----A-----A-----	1008
HIVPH136	-----A-----A-----A-----	1008
HIVBZ200	-----A-----A-----A-----	1008
HIVTB132	-----C-----A-----A-----	1008
HIVBZ190	-----T-----A-----A-----	1008
HIVLAI	-----A-----A-----A-----	1008
HIVHXB2R	-----A-----A-----A-----	1008
HIVRN	---CC---C-----A-----A-----	1017
HIVJH3	-----A-----A-----A-----	1008
HIVJRC6F	-----A-----A-----A-----	1008
HIVQYI	-----A-----A-----A-----	1008
HIVNY5CG	-----A-----A-----A-----	1008
HIVNL43	-----A-----A-----A-----	1008
HIVCDC4	-----T-----A-----A-----	1008
HIVHAN	-----T-----A-----A-----	1008
HIVCAM1	-----A-----A-----A-----	1008
HIVRF	---CC---A-----A-----A-----	1008
HIVD31	-----A-----A-----A-----	1008
HIVBH102	-----A-----A-----A-----	1008
HIVPV22	-----A-----A-----A-----	1008
HIVJRF1	-----A-----A-----A-----	1002
HIVUG280	-----T-----A-----A-----	1008
HIVYU2	-----A-----A-----A-----	1008
HIVBCSG3C	-----A-----A-----A-----	1008

CONSENSUS.C	TAAAAAATGGGATGACAGACACCTTGTGGTCCAAAATGCGAACCAGATTGTAAAGACCATTCTAAAGAGCA	994
HIV00268	-----	993
HIVSM145	-----	999
HIVZAM18	-----A-----	1005
HIVZAM19	-----A-----	816
HIVZAM20	-----C-----C-----	999
HIVDJ259	-----	1002
HIVVI313	-----	1002
CONSENSUS.D	TaAAAAaATGGGATGAC?GAAMCCTTGTGGTcCAAAAATGCGaAACCAGATTGtAaaaActATCTTAAAAGCA	1003
HIVELI	-----A-----A-----C-G-----	1008
HIVZ236	-----GG-----A-----A-----G-C-----	1011
HIVMDK	-----C-----A-----A-----G-C-----	999
HIVVI205	-G-----C-----A-----A-----	1011
HIVG109	-----C-----C-----	1017
HIVK31	-----C-----C-----T-----	1008
HIVUG274	-----T-----T-----G-----	1008
HIVUG270	-----T-----T-----	1008
HIVSE365	-----T-----A-----	1008
HIVVI203	-----C-----	1008
CONSENSUS.F	TAAAgggTTGGATGACAGACACctTGTtGgTCCAAAAtGCGaAACCAGATTGTAAgACCATT?TTAAAAGCa	976
HIVVI174	-----T-----A-----G-----G-----A-----C-----	993
HIVVI69	-----A-----G-----G-----A-----C-----	996
HIVB2162	-----T-----T-----	995
HIVVI325	---AAA-----C-----C-A-----T-----G	996
CONSENSUS.G	TAAAA?CTGGATGACAGACAC?TTGTGGTCCAAAATGCGAACCAGATTGTAAAGACCATTCTTAA?AGCA	967
HIVLBV217	---AA-----T-----A-----	1003
HIVVI191	---GG-----C-----G-----	999
CONSENSUS.H	T?AA???TTGGATGACAGACACcCTTGTGGTCCAAAATGCGAATCCAGATTG?AAGA?TATTTTAA?AGCA	931
HIVVI525	-A--GAA-----C--A-----A-----	1008
HIVVI557	-G--AGG-----T--C-----G-----	1008
CONSENSUS.O	TAAA?AATGGATGACAGAAAC??TGCTTGITCA?AAT?C?AA?CCAGATTG?AAACA?ATT?TGAAA?CA	901
HIVANT70	---A-----TC-----A--G-C-T-----C-----G--T-----T--	1005
HIVMVP51	---G-----CT-----G--T-A-C-----T--A--C-----G--	1005
SIVCPZGAB	TGAAAAATTGGATGACTGACACTCTCTTAGTGCGAGAATGCAAACCCAGATTGTAAAGCAAATTCTAAAAGCA	1032

HIV1 GAG

CONSENSUS.A	TTAGGa?cagggGCTaCaTTAgAaGAAATGATGACaCaTgcCAGggAGTggGaGgACCcgGcCAtAAaGC	1071
HIVU455	-----GC-----	1067
HIVMAL	-----C-----	1097
HIVVI59	---A-GT-A-----A-----G---	1079
HIVVI310	---A---A-A-----A-----	1076
HIVVI57	---A-----A-----	1079
HIVK112	---A---A-----A-----	1079
HIVK88	---A-----A-----	1079
HIVK29	---G-----C-----	1079
HIVK124	---G---G-----A-T-----	1079
HIVK7	---A---C-----A---A---C---	1079
HIVK98	---A-----C-----	1079
HIVK89	---A-----G-----C-----	1079
HIVVI32	---C-----T-----	1079
HIVVI415	---A-----A-----	1079
HIVCI4	---A-----A-----	1073
HIVG141	---C-----T---T-----	1067
HIVLBV23	---G-----G-----	1079
HIVTN243	---A---A-----TA-----	1079
HIVTN245	---T---A-----TA-----	1079
HIVTN240	---A---A-----TA-----	1079
HIVCI20	---C-----T-----T-----	1055
HIVCI59	---GC-----T---T-----T-----	1070
HIVLBV2310	---C---C-----T-----T-----	1070
HIVCI51	---C-----G---T---T-----	1064
HIVLBV105	---G---C---P-----T-----	1091
HIVCI32	---G---C-----T-----T-----	1085
HIVIC144	---C-----T-----T-----	1070
HIVDJ258	---C-----G---T---T-----T-----	1070
HIVCM238	---A---A-----TA-----	1079
HIVUG266	---A---C---AA-----AA---AA---A-----	1079
HIVVI354	C-----G-C-----A-----A-----A-----	1079
CONSENSUS.B	TTGGaCCaGcaGCTACAcTAGAaGAAATGATGACaCaTgcCAGGGaGTgGGaGGaCCcgGcCAtAAaGC	1078
HIVSF2	-----G-----G-----T-----	1085
HIVBZ167	---T-----G-----G-----T-----	1082
HIVPH153	---A-----T-----	1079
HIVPH136	-----AA-----	1079
HIVBZ200	-----T-----	1079
HIVTB132	-----T-----G-----	1079
HIVBZ190	-----G-----	1079
HIVLAI	-----G-----G-----	1079
HIVHX92R	---G-----A-----G-----	1079
HIVMN	-----T---T-----	1068
HIVJH3	---C-----	1079
HIVJRC5F	-----	1079
HIVQYI	-----G-----	1079
HIVNY5CG	-----G-----	1079
HIVNL43	---G---G-----G-----	1079
HIVCDC4	-----G-----	1079
HIVHAN	---A-----G-----	1079
HIVCAM1	-----G-----	1079
HIVRF	-----A---G---A-----	1079
HIVD31	-----G-----	1079
HIVBH102	---G-----A-----G-----	1079
HIVPV22	-----C-----G-----	1079
HIVJREFL	-----C-----	1073
HIVUG280	-----T-----	1079
HIVYU2	-----G-----	1079
HIVBCG3C	---G-----G-----C-----	1079

CONSENSUS.C	TTAGGACCAAGGGCTTCACTAGAGAAATGATGACAGCATGTCAGGGAGTGGGAGGACCTAGCCACAAAGC	1065
HIVUG268	---G---A-----G-----	1064
HIVSM145	-----GC-----	1070
HIVZAM18	-----C---A-----	1076
HIVZAM19	-----C---A-----	887
HIVZAM20	-----A-----	1070
HIVDJ259	---G---C-----G-----	1073
HIVVI313	---G---C-----G-----	1073
CONSENSUS.D	TTGGGACCAgTgGCTACATAGAGAAATGATGACAGCATGTCAGGGAGTGGGgGgGCCCCaStCATAAAGC	1073
HIVSLI	-----CA---C-----G---C-----	1079
HIVZ236	-----CA---C-----G---C-----	1082
HIVNDK	-----CA---C-----C-----G---G-C-----	1070
HIVVI205	-----CA---C---A-----A-----G---C-----	1082
HIVG109	-----CA-----	1088
HIVK31	-----G-----A---G-----	1079
HIVUG274	-----G-----A-----	1079
HIVUG270	-----C-----G-----G-----	1079
HIVSE365	---A---G-----C---G---A-----	1079
HIVVI203	---G-----	1079
CONSENSUS.F	TTGGGACCAAGGGGCTaCAcTAGAGAAATGATGACAGCATGTCAGGGAGTGGGAGGACCTGGCCATAAGGC	1047
HIVVI174	-----	1064
HIVVI169	A-----	1067
HIVB2162	-----	1066
HIVVI325	-----T---T---G-----A-----A---	1067
CONSENSUS.G	TTAGGACCAAGGAGCTACACTAGAGAAATGATGACAGCATGTCAGGGAGTGGGAGGACCC?GCCATAAAGC	1037
HIVLBV217	-----A-----	1073
HIVVI191	-----G-----	1070
CONSENSUS.H	TTAGGA??AGGGGCTTCAATAGAGAAATGATGACAGC?TG?CAGGGAGTGGGAGG?CCT?GCCA?AAAGC	995
HIVVI525	-----AC-----A---T-----A---A---C-----	1079
HIVVI557	-----CA-----C---C-----G---G---T-----	1079
CONSENSUS.O	TTAGG?CCAG?AGCTAC?TTAGAGGA?ATGATGGTAGCCTGTCA?GGAGTAGGAGG?CCAACPCA?AAGGC	965
HIVANT70	-----G---G-----C-----G-----G-----A-----T-----	1076
HIVMVP51	-----A---A---T-----A-----A-----G-----C-----	1076
SIVCPZGAB	TTGGGACCTGGGGCCACCC?AGAGGAAATGATGACAGCTTGTCAAGGAGTAGGAGGACCATCTCATAAAGC	1103

HIV1 GAG

p24 \ /		
CONSENSUS .A	AagggTtTTGCTGagGCAATgAGtCAAGtacaacAT?caAac.....?????????aTAATGATgC	1123
HIVU455	-----GA--G-	1120
HIVMAL	--A-----C--C-AC-A-T-CT.....GCTGCC-----	1156
HIVVI59	--A-----A--A-----GA--G-	1132
HIVVI310	--A-C-----A-----G-C-----A-----T-	1129
HIVVI57	-----GC-----	1132
HIVK112	---A-----A-----	1132
HIVK88	-----G--A--A-----	1132
HIVK29	-----G--A--A-----	1132
HIVK124	--A-C-----C--C-AC-A-AAT-TA.....AATGCTGCTG--	1141
HIVK7	--A--A-----C-----GA--T-	1132
HIVK98	-----A-----C-----GG-----	1132
HIVK89	-----A-----G--A--A-----	1132
HIVVI32	-----C-----GG-----	1132
HIVVI415	-----A-----	1132
HIVCI4	-----A-C-----	1126
HIVG141	--A--A--A-----A--G-C-----	1120
HIVLBV23	-----A-----A-----G-----	1132
HIVTN243	-----C-C-C-----G--T-----	1132
HIVTN245	-----C-C-C-----G--T-----	1132
HIVTN240	-----A-----C-C-C-----G-CT-----	1132
HIVCI20	-----GT-C-----	1108
HIVCI59	-----GT-C-----	1123
HIVLBV2316	-----GA-C-----	1123
HIVCI51	-----GT-C-----	1117
HIVLBV105	---C---A-----C-----	1144
HIVCI32	-----GT-C-----	1138
HIVIC144	-----GT-C-----	1123
HIVDJ258	-----G-----AT-C-----	1123
HIVCM238	-----C-C-C-----GG--GT-----	1132
HIVUG266	-CA-----A-----GA--G-----	1132
HIVVI354	-CA-----A-----CA--A--G-----	1132
CONSENSUS .B	AAGAgTTTTGGCTGAAGCaATGAGCcAAGtaacaaat.....???tCAgctaccataATGaTgC	1134
HIVSF2	-----C-----C-----A-----	1141
HIVB2167	-----A-----C-----AGT-----	1141
HIVPH153	-----C-----C-----G--C-----	1135
HIVPH136	-----G--G-----GCC-----	1135
HIVB2200	-----A-----T-----AG-----	1135
HIVTB132	-----G-----G-----	1135
HIVB2190	-----	1135
HIVLAI	-----	1135
HIVHXB2R	-----	1135
HIVMN	-----G--G-----	1144
HIVJH3	-----A-----	1135
HIVJRCSP	-----C-----TC--G-----	1135
HIVQYI	-----TC--G-----	1132
HIVNY5CG	-----C-----A-----	1135
HIVNL43	-----C-----A-----	1135
HIVCDC4	-----	1135
HIVMAN	-----GG-----G-----	1135
HIVCAM1	-----	1135
HIVRF	-----A-----C-----	1135
HIVD11	-----TG-----	1135
HIVBH102	-----A-----	1135
HIVPV22	-----A-----	1135
HIVJRF1	-----C-----	1129
HIVUG280	-----A-----C--T--A-----	1135
HIVYU2	-----	1135
HIVBC5G3C	-----G-----C-----	1135

p24 \/			
CONSENSUS.C	AAGAGTGTtTGCTGAGGCAATGAGCCAAgCAAaCAatacAAac.....	..ATAtTGATGC	1118
HIVUG268	-----T-----	-----	1117
HIVSM145	-----T-----	-----C-----	1123
HIVZAM18	-----A-----	-----C-----	1129
HIVZAM19	-----A-----G-----CN.	N-----	940
HIVZAM20	-----A-----G-G-----T.	G-----	1123
RIVDJ259	-----C-----C-----C-----	-----	1126
HIVVI313	-----C-----T-----	-----	1126
CONSENSUS.D	AAGAGTtTtTGCTGAGGCAATGAGCCAAgCa.....	..acaasT.....tca?at?ctgC?atAATgATGC	1129
HIVELI	-----C-----	-----GT-A--A-AGC-----	1138
HIV2226	-----C-----	-----GC-G-C--AG-----	1141
HIVNDK	-----T-----	-----GC-A--AG-----	1129
HIVVI205	-----G-----	-----GC-A--AG--A-----	1141
HIVG109	-----G-----	-----A--G--C-----	1147
HIVG31	-----	-----GG-A--A--T-----	1135
HIVUG274	-----C-----C-----	-----G--A--A--T-----	1138
HIVUG270	-----C-----C-----T-----	-----G--A--G--T-----	1138
HIVSE365	-----C-----G-----	-----GC-G--C-----	1138
HIVVI203	-----C-----A-----	-----A--A--GT--C-----	1138
CONSENSUS.F	aAGAGTtTtTGCTGAGGCAATGAGCCAAgCa.....	..ACAAAT.....?????gCA?CCATAaTGATGC	1099
HIVVI174	-----	-----ACAGCA--G-----	1123
HIVVI69	-----A-----	-----G-----	1120
HIVB2162	G-----	-----A--T--C-----	1119
HIVVI325	-----A-----G--T-----	-----A--A-----	1120
CONSENSUS.G	AAGAGTtTtTAGCTGAAGC?ATGAGCCA?GCA.....	..TCAGGT.....??AGCA?CAGCCAT?ATGATGC	1090
HIVLBY217	-----G-----G-----	-----AC--A--A--A-----	1132
HIVVI191	-----A-----A-----	-----GT--G--G-----	1129
CONSENSUS.H	AA?AGTtTtTGCTGAGGCAATGAGCCAAgTA??...	..ACAAAT.....?????GCAG?PAT?ATGATGC	1044
HIVVI525	-----G-----	-----ACACAT--CC--A-----	1138
HIVVI557	-----A-----AAT	-----TT--G-----	1135
CONSENSUS.O	?A?A?T?CTAGCAGAAgCAATGGCT?C?GCCAGCAAGATTT?AAAGGAGGATACACAGCAGTA?TCATGC		1029
HIVANT70	C-G-G-G-----A-A-----G-----		1147
HIVMVP51	A-A-A-A-----T-T-----A-----		1147
SIVCPZGAB	AAGGCTTCTAGCTGAAGCCATGTCAATGGTTGAG...AATCAA.....	..GGGAGAGCAGATGTTTCTTCC	1165

HIV1 GAG

CONSENSUS.A	AgAgaGGcaAttTt...aggGGccagaaaaa??atTAAgTGTtTCAAcTGTGGCAAAAGaGGaCacCTA	1188
HIVU455	-----C-G-----	1185
HIVMAL	-----T-----A-----	1221
HIVVI59	-----A-----	1197
HIVVI110	---G---A---A---A---T---	1194
HIVVI157	-----G-G-----	1197
HIVKI112	-----C-----G-----	1197
HIVKB8	-----G-----	1197
HIVK29	-----A-----G-----	1197
HIVKI124	---A---CA-G-A-ATC---A---G---A---	1209
HIVK7	---AA---A---CG---T-T---	1197
HIVK98	-A-----A---G---T-T---	1197
HIVK89	---G---A---G---	1197
HIVVI132	-A-----A-C---GGCT-G-A-AGA---	1200
HIVVI1415	---AA---G---C---T---	1197
HIVCI4	---A-----T---	1191
HIVG141	---A-----A---	1185
HIVLBV23	-----C-----T---	1197
HIVTN243	-----A---A---C---	1197
HIVTN245	---A---A---C---G---	1197
HIVTN240	---C---A---C---T-G---	1197
HIVCI20	-----A-G-C---A-A---	1173
HIVCI59	---G---G-C---A---	1188
HIVLBV2310	-----G-C---A---	1188
HIVCI51	-----G-C---A---	1182
HIVLBV105	---A-----ATC-----T---	1212
HIVCI32	---C---A---A-G-AGTTCG-----G---A---	1206
HIVCI144	-----G-C---A---	1188
HIVDJ258	---G---G-C---A---	1188
HIVCM238	---A---G---C---C---G---	1197
HIVUG266	---AA-----	1197
HIVVI354	---A---A---A---	1197
CONSENSUS.B	AgAgaGGCAATTT...AggaacccaagAAAgactgTtAAgTGTtTCAAcTGTGGCAAAAGaGGgCacatA	1202
HIVSF2	-----T-----	1209
HIVB2167	---A-----A---T---	1209
HIVPH153	-----AC-----A---	1203
HIVPH136	-----A-T---	1203
HIVB2200	-----T---	1203
HIVTB132	---A---A---T---	1203
HIVB2190	-----A-----A---	1203
HIVLAI	---A-----T---	1203
HIVHB2R	-----T-----C---	1203
HIVMN	---T---T-A-C---C---	1212
HIVJH3	-----T-A-----T---	1203
HIVJRCSP	---C---A-----	1203
HIVQYI	---A---A-----	1200
HIVNY5CG	-----A---G-A---	1203
HIVNL43	---A-----	1203
HIVCDC4	---C---GA---G---	1203
HIVHAN	-----	1203
HIVCAM1	-----C---TC---	1203
HIVRF	---A---T---G---A-T---C---T---	1203
HIVD31	-----	1203
HIVBH102	-----TG-----C---	1203
HIVPV22	-----TG-----C---	1203
HIVJRP6	---C---T---	1197
HIVUG280	---AC---T---C---T---	1203
HIVYU2	-----A-----	1203
HIVBCSG3C	-----T---	1203

```

CONSENSUS.C AgAgAaGCAATTTT...AAaGGccCtAaaagAATgTTAAATGTTTCAAcTGTGGCAaGGAAGGGCACaTA 1186
HIVUG268 -----G-----G-----A-----T-----C----- 1185
HIVSM145 -A--G-----T-----A-----A----- 1191
HIVZAN18 -----A-----G----- 1197
HIVZAN19 -----A--G-N----- 1208
HIVZAN20 -----G-----C----- 1191
HIVDJ259 --A-----G-----G-A-C-A-----C----- 1194
HIVVI313 --A-----G-----A-----T----- 1194

CONSENSUS.D AgAgAGgcAATTTt...AAgGGcCcAagaAaaattaTTAAGTGTTCACAcTGTGGCAaaGAAGGgCACata 1197
HIVELI -----T----- 1206
HIVZ226 -----C----- 1209
HIVNDK -----G-----G-----C----- 1197
HIVVI205 -----A-----A-----A----- 1209
HIVG109 --A--T-----A-----G-C-----G----- 1215
HIVK31 -----A--G--C-----A--C----- 1203
HIVUG274 -----T-----G--C----- 1206
HIVUG270 -A--C--C-----T--AG-G--AC-----C----- 1206
HIVBE365 -----T-A--G--CG----- 1206
HIVVI203 -----C-----A--G-----G--A--C----- 1206

CONSENSUS.F AGAaAAGtAAcTTT...AAGGGCCAAAGAaAATgTTAAATGTTTAAPTGTGGCAaAgaaGGACAcTA 1167
HIVVI174 ----- 1191
HIVVI69 -----T-----C-----G----- 1188
HIVB2162 -----AC----- 1187
HIVVI325 ---G-G-A-----A--A-----C-----G--G--T--- 1188

CONSENSUS.G A?AAA?GC???TFT...AAGGGCCCGA?AA?AA?TATTAAAGTGTTC?ACTGTGGCAAGGAAGGACATCTA 1149
HIVLBV217 -G--G--AAT-----G--A--A-----G----- 1200
HIVVI191 -A--A--GGG-----A--G--C-----A----- 1197

CONSENSUS.H AGAAAGG?AACTT?...AAGGG?CAAAGAAGAATT?TTAAATGTTTCAA?TGTGGCAAGGAAGGACAC?TA 1106
HIVVI525 -----C-----T-----T-----G-----C-----A----- 1206
HIVVI557 -----T-----C-----C-----A-----T-----C----- 1203

CONSENSUS.O AAAGAGGGCA?AAT...CCAA?TAG?AAAGG?C?ATAAAATG?TTCAA?TGTGGAAAAGAGGGACATATA 1089
HIVANF70 -----A-----T--G-----AA-T-----T-----C----- 1215
HIVMVP51 -----G-----A--A-----GC-C-----C-----T----- 1215

SIVCP2GAB AGAAGGGACAAGGT...GCGGGCCCAAAAGAAAAATAAAATGCTTTAATTGTGGAAAGAGGTCACTTA 1233

```

HIV1 GAG

CONSENSUS .A	GCcAgAAATTCAggGCCCTAggAAAAagGGCTGTTCGAAATGTCgggAagGAaGGaCACCAAATgAAaGA	1259
HIVU455	---A---T-----C---A-----	1256
HIVMAL	-----	1292
HIVV159	-----A-----A-----A-----A-----	1268
HIVV1310	-----A-----T-G-----T-----	1265
HIVV157	-----G-----G-----T-----	1268
HIVK112	-----G-----G-----T-----	1268
HIVK88	-----G-----G-----T-----	1268
HIVK29	---T-A-----A-----G-----	1268
HIVK124	--A-A-----T-----G-----A-G-----	1280
HIVK7	-----A-----A-----A-----A-----T-----	1268
HIVK98	-----G-----G-----T-----	1268
HIVK89	---T-----G-----G-----	1268
HIVV132	-----G-----G-----	1271
HIVV1415	-----A-----G-----	1268
HIVC14	-----G-----	1262
HIVG141	-----	1256
HIVLBV23	-----	1268
HIVTN243	-----A-----T-----G-----T-----	1268
HIVTN245	-----A-----T-----T-----T-----	1268
HIVTN240	-----A-----T-----C-----T-----	1268
HIVC120	-----A-----G-----	1244
HIVC159	-----A-----G-----	1259
HIVLBV2310	--A-----A-----G-----G-----	1259
HIVC151	-----A-----G-----G-----	1253
HIVLBV105	-----A-----A-----T-----	1283
HIVC132	--A-A-----A-----	1277
HIVC144	-----A-----	1259
HIVDJ258	-----A-----T-----A-----A-----G-----	1259
HIVCM238	-----A-----A-----T-----T-----	1268
HIVUG266	-----A-----A-----A-----AA-----A-----A-----	1268
HIVV1354	-----A-----A-----A-----C-----T-----A-----A-----	1268
CONSENSUS .B	GCcAaAAATTCAGGGCCCTAGgAaaAaGGGCTGTTCGAAaATGTGGAAagGAAGGACACCAAATGAAaGA	1273
HIVSF2	-----G-----G-----	1280
HIVB2167	-----G-----	1280
HIVPH153	---G-----A-----A-----G-----	1274
HIVPH136	-----G-----	1274
HIVB2200	--A-----GA-----	1274
HIVTB132	-----G-----	1274
HIVB2190	---C-----	1274
HIVLAI	---G-----	1274
HIVHXB2R	---G-----	1274
HIVMN	-----G-----	1283
HIVJH3	---G-----	1274
HIVJRC5F	---G-----	1274
HIVQY1	-----T-----G-----	1271
HIVNY5CG	-----	1274
HIVNL43	-----	1274
HIVCDC4	--T-G-----A-----G-----	1274
HIVHAN	---G-----T-----G-----	1274
HIVCAM1	---T-----T-----	1274
HIVRF	-----	1274
HIVD31	-----G-----	1274
HIVBH102	---G-----	1274
HIVPV22	---G-----	1274
HIVJRF1	---G-----	1268
HIVUG280	-----	1274
HIVYU2	-----T-----	1274
HIVBCS03C	---G-----A-----G-----C-----	1274

CONSENSUS.C	GCcAgAAaTTGCAGGGCCcCCTaGGAAAAgGGCTGTTGGAAATGTGGAA?GGAAGGACACCAAAATGAAAGA	1256
HIVUG268	-----A-----G-----	1256
HIVSH145	--T-----T-----A-----	1262
HIVZAM18	--A-----A-----A-----	1268
HIVZAM19	--A--T-----A-----N-----	1079
HIVZAM20	-----T--C-----G-----	1262
HIVDJ259	-----G-----	1265
HIVVI313	-----A-----	1265
CONSENSUS.D	GCaAaAAATTCcAggCCcCCTAg?AaaAaGGGCTGTTggAaATGtGGAAagGAAGGaCACCAAAATgAAaGA	1267
HIVELI	-----G-----G-----C-A-----	1277
HIV2226	-----G-G-----C-A-----G-----	1280
HIVMOK	-----A-----C--G-----	1268
HIVVI205	-----AA-----G-----A-----	1280
HIVG109	-----A-----G--G-----A--G-----	1286
HIVK31	-----T--G--G-----G-----	1274
HIVUG274	-----T--T--G-----	1277
HIVUG270	-----T--A-----AA-----G-----A-----	1277
HIVSE365	-----G-----A--T--A-----C-----G-----	1277
HIVVI203	-----C-----A-----	1277
CONSENSUS.F	GCCAAaAaATTGcAGGGCCCCCTAGAAAAAAGGGCTGTTGGAAATGTGGaAgaGAaGGACACCAAAATGAAAGA	1238
HIVVI174	-----G-----	1262
HIVVI59	-----G-----	1259
HIVB2162	-----T-----	1258
HIVVI325	-----C-----G-AG-----	1259
CONSENSUS.G	GCCAGAAaATTGCAGGGCCcCCTAGGAAAA?GGGCTGTTGGAAATGTGGAA?GGA?GGACATCAAAATGAAAGA	1217
HIVLBV217	-----G-----A--G-----	1271
HIVVI191	-----A-----G--A-----	1268
CONSENSUS.H	GC?AGAAaATTGCAGGGCCCCCTAGAAAAAAGGG?TGTGGAAATGTGGAA?GGAAGGACATCAG?TGAAAGA	1172
HIVVI525	--T-----C-----T--G-----A-----	1277
HIVVI557	--C-----A-----C--A-----C-----	1274
CONSENSUS.O	GCAA?AAA?TGTCAGGACACCTAG?AAAA?GGGTTGCTGGAAATGTGGACAGGAAGGTCACCAAAATGAAAGA	1156
HIVANT70	-----G--T-----G--A-----	1286
HIVMVP51	-----A--C-----A--G-----	1286
STVCPZGAB	GCTAGAAaATTGTAAAGCACCAAGAAGAAAAAGGCTGTGTGGCGGTGTGGCAAGAGGGACACCAAAATGAAAGA	1304

HIV1 GAG

		pol cds ->	
CONSENSUS.A	cTgcAcT...??gAg...aGACAGGCTAATTTTtTAGggGaaaAtTcTgCcCTTCCagCAAgGGg...AGGC		1318
HIV0455	-----A-----		1315
HIV0461	-----CA-----A-----		1351
HIVV159	-----C-----		1327
HIVV1310	-----C-----		1324
HIVV157	-----C-----		1327
HIVK112	-----A-----		1327
HIVK98	-----A-----G-----G-----		1327
HIVK29	-----C-----C-----		1327
HIVK124	T-----A-----CA-----A-----		1339
HIVK7	-----A-----G-----A-----		1327
HIVK98	T-----G-----		1327
HIVK89	A-----A-----G-----		1327
HIVV132	-----C-----AAT-----C-----A-----		1333
HIVV1415	-----C-----T-----		1327
HIVC14	-----C-----		1321
HIVG141	-----A-----A-----		1315
HIVLBV23	-----A-----A-----		1327
HIVTN243	-----A-----A-----		1327
HIVTN245	-----A-----A-----		1327
HIVTN240	-----A-----A-----		1327
HIVC120	-----A-----		1303
HIVC159	-----A-----		1318
HIVLBV2310	-----A-----		1318
HIVC151	T-----T-----		1312
HIVLBV105	-----C-----		1342
HIVC132	-----A-----C-----		1336
HIVC144	-----A-----		1318
HIVD3258	-----C-----		1318
HIVC230	-----A-----A-----		1327
HIVG266	-----A-----AA-----		1327
HIVV1354	-----G-----A-----A-----		1327
CONSENSUS.B	tTgTact...??gAg??gAGACAGGCTAATTTTtAGGgAAgATcTGGCCCTTCCcaCAAGGGa...AGGC		1332
HIVSF2	-----C-----T-----		1339
HIVB2167	-----T-----A-----T-----		1339
HIVPH153	-----A-----A-----G-----		1333
HIVPH136	-----A-----A-----G-----		1333
HIVB2200	C-----C-----A-----A-----T-----A-----G-----		1333
HIVTB132	-----A-----A-----A-----		1333
HIVB2190	-----T-----A-----A-----G-----		1333
HIVLAI	-----T-----		1333
HIVHX82R	-----T-----		1333
HIVMN	-----TG-----		1341
HIVJH3	-----A-----A-----TC-----		1333
HIVJRC5F	G-----T-----		1333
HIVQY1	-----		1330
HIVNY5CG	-----		1333
HIVNL43	-----		1333
HIVCDC4	-----		1333
HIVHAN	-----C-----A-----A-----G-----		1333
HIVCAM1	-----A-----A-----G-----		1333
HIVRF	-----C-----AAT-----GGAC-----A-----		1339
HIVD31	-----T-----		1333
HIVBH102	-----T-----		1333
HIVPV22	-----T-----		1333
HIVJRFL	-----T-----		1327
HIVUG280	-----A-----AAT-----T-----G-----		1336
HIVYU2	-----T-----		1333
HIVBCSG3C	-----A-----		1333

pol cds ->			
CONSENSUS.C	CTGTACT.....GAG...AGGCAGGCTAATTTTATAGGGAaaATCTGGCCTTCCaCAAGGGG...AGGC		1315
HIVUG268	-----A-----C-----G-----C-A-----		1315
HIVSM145	-----C-----C-----C-----C-----		1321
HIVZAM18	-----C-----C-----C-----C-----		1327
HIVZAM19	-----C-----C-----C-----C-----		1338
HIVZAM20	T-----A-----A-----A-----		1321
HIVDJ259	-----A-----A-----A-----A-----		1324
HIVVI313	-----T-----T-----T-----T-----		1324
CONSENSUS.D	CTGCACt.....Gaa...AGACAGGCTAATTTTATAGGGAaaATCTGGCCTTCCaCAAGGga...aGGC		1326
HIVELT	-----G-----G-----G-----G-----		1336
HIVZ226	-----G-----G-----G-----G-----		1339
HIVNDK	-----G-----G-----G-----G-----		1327
HIVVI205	-----T-----T-----T-----T-----		1339
HIVGI09	-----A-----A-----A-----A-----		1345
HIVK31	-----G-C-----G-C-----G-C-----		1333
HIVUG274	G-----G-----G-----G-----		1336
HIVUG270	-----A-----A-----A-----A-----		1336
HIVSE365	-----G-----G-----G-----G-----		1336
HIVVI203	-----G-----G-----G-----G-----		1336
CONSENSUS.F	CTGCACt.....Gaa...aGACAGGCTAATTTTATAGGGAaaATCTGGCCTTCCaCAAGGGG...AgGC		1297
HIVVI174	-----G-----G-----G-----G-----		1321
HIVVI69	-----G-----G-----G-----G-----		1318
HIVB2162	-----G-----G-----G-----G-----		1317
HIVVI325	-----C-----G-----G-----A-----A-----		1318
CONSENSUS.G	CTGCACA.....GAG...AGACAGGCTAATTTTATAGGGA?AAT?TGGCCTTCCaCAAGGGG...AGGC		1374
HIVLSV217	-----A-----A-----A-----A-----		1330
HIVVI191	-----G-----T-----T-----T-----		1327
CONSENSUS.H	?TGCACA.....GA?...AGACAGGCTAATTTTATAGGGAaaAT?TGG?CT?CCA?CAAGGGG...AGGC		1225
HIVVI525	T-----G-----C-----C-----T-----G-----		1336
HIVVI557	C-----A-----T-----T-----C-----A-----		1333
CONSENSUS.O	TTGCA?A.....AATGGAA?ACAGGCAAAATTTTATAGG?AA?TACTGGCCTCCGGGGGGCAGG...AGGC		1214
HIVANI70	-----G-----A-----C-----A-----		1348
HIVMVP51	-----A-----G-----G-----G-----		1348
STVCPZGAB	CTGCACA.....GGA...AGACAGGTGAATTTTATAGGGAaaAGGTGGCCTTCCCGCAGCGGG...AGGC		1363

I-A-42
DEC 94

CONSENSUS .A		CagGaaAtTTtCctCAGagCAGAcca.	Gag	1347
HIVU455	---	-O-----	---	1344
HIVNAL	---	-G-----C-T-----	---	1380
HIVV159	---	-----G-	---	1356
HIVV1310	---	-----A---TTG	---	1353
HIVV157	---	-----C-----G-	---	1356
HIVK112	---	-----A---TG	--A	1356
HIVK88	---	-----C---A---TG	---	1356
HIVK29	---	-----A---T-	---	1356
HIVK124	---	---G-C-C-T-----	--A	1368
HIVK7	---	-----	---	1355
HIVK98	---	-----G-	---	1356
HIVK89	---	-----A---TG	---	1356
HIVV132	---	-----G-	---	1362
HIVV1415	---	-----C-----G-	---	1356
HIVCI4	---	---G-----	---	1350
HIVG141	---	---G-C-----G-	---	1344
HIVLBV23	---	-----	---	1356
HIVTN243	---	-G-G-----	---	1356
HIVTN245	---	-G-G-----	---	1356
HIVTN240	---	-G-G-----	---	1356
HIVCI20	---	-----	--A	1332
HIVCI59	---	-----G---G-	--A	1347
HIVLBV2310	---	-----C-----G-	--A	1347
HIVCI51	---	-----G-	--A	1341
HIVLBV105	---	---G---T-----A-	---	1371
HIVCI32	---	-----G-	--A	1365
HIVIC144	---	-A-----A-G-	--A	1347
HIVDJ25B	---	-----A-----	---	1347
HIVCM238	---	-G-G-----	---	1356
HIVUG266	---	-A-G-C-----	---	1356
HIVV1354	---	---G-----TC--A----	---	1356
CONSENSUS .B		cagGgAAatTTtCcttCAGagCAGaccA?????????????????????????????.....?????????GAg		1361
HIVSF2	---	-----	---	1368
HIVB2167	---	-----C-C-----	---	1368
HIVPH153	---	-----	--A	1362
HIVPH136	---	-----C---G-	---	1362
HIVB2200	---	-----C---A---G-	---	1362
HIVTB132	---	---A-----	---	1362
HIVB2190	---	---C-----	---	1362
HIVLA1	---	-----GAGCCACAGCCCCACCATTCTTCAG.....AGCAGACCA	---	1398
HIVHXB2R	---	-----	---	1362
HIVMN	---	. C-----C-----A-	---	1368
HIVJH3	---	-----	---	1362
HIVJRC5F	---	-----C-----	---	1362
HIVOYI	---	-----A-----	---	1359
HIVNY5CG	---	-----	---	1362
HIVNL43	---	-----	---	1362
HIVCD4	---	-----	---	1362
HIVHAN	---	---A-----	---	1362
HIVCAM1	---	-----	---	1362
HIVRF	---	-----C-----	---	1368
HIVD31	---	-----C-----	---	1362
HIVBH102	---	-----GAGCCAACAGCCCCACCATTCTTCAG.....AGCAGACCA	---	1398
HIVPV22	---	-----GAGCCAACAGCCCCACCATTCTTCAG.....AGCAGACCA	---	1398
HIVJRPL	---	-----	---	1356
HIVUG280	---	-----	---	1365
HIVYU2	---	---A-----	---	1362
HIVBCS03C	---	-----C-----T-	---	1362

CONSENSUS.C	CAGGGGAATTTCCTtCAGAgCAGgCCa????????????????????????????????gag	1344
HIVG3268	-----A-----	1344
HIVSH145	-----A-----GAGCCAACAGCCCCACCAGCAGAGAGCTTCAGGGTCAGGCCA--	1392
HIVZAM18	-----GAGCCAACAGCCCCCT-----AGGCCA--	1377
HIVZAM19	-----N-----	1167
HIVZAM20	-----C--A-----GAGCCAACAGCCCCACCAGCAGAGAACTTCAGG.	1380
HIVDJ259	-----A--T-----	1353
HIVVI313	-----C--A--A-----	1353
CONSENSUS.D	CaGGgAACTT?CtTCagAGCAgACca.....GAG	1354
HIVBLI	-G-----T--C--A-----	1365
HIVZ2Z6	-G-----T-----	1368
HIVNDK	-G-----T--T-----	1356
HIVVI205	-G-----T--T-----	1368
HIVG109	-----A--C-----	1374
HIVK31	-T--A--C-----G-----	1362
HIVUG274	-----C-----TG-----	1365
HIVUG270	-----C-----	1365
HIVSE365	-----T--C-----	1365
HIVVI203	-----C-----	1365
CONSENSUS.F	C?GG?AATTTtCT?CAGAgCAGaCCA.....GAG	1323
HIVVI174	-C--A-----T-----	1350
HIVVI69	-A--G-----T-----	1347
HIVB2162	-C--A--C--C--A--G-----	1346
HIVVI325	-A--G-----C-----	1347
CONSENSUS.G	C?GGGAATTTCTTCAGa?CAG??CA.....GAG	1299
HIVLBV217	-A-----A--AC-----	1359
HIVVI191	-G-----T--GA-----	1356
CONSENSUS.H	CAGGGGAATTTtCT?CAGAGCAGGCCA.....GAG	1253
HIVVI525	-----T-----	1365
HIVVI557	-----C-----	1362
CONSENSUS.O	CAGGCAATTATGTGCAGa?AC?AG??.....??C	1237
HIVANT70	-----G--C--CA-----CA-	1377
HIVMVP51	-----A--A--TG-----TC-	1377
SIVCPZGAB	CAGGCAACTTTGTGCAGAACGAACA.....GAA	1392

HIV1 GAG

CONSENSUS.A	CCaaCAGCCcCCaGCAGAg	a?cttt???gGgatgggggaagAgataac	1393
HIVU455	-----A-----	-----T-----A-----G-----	1391
HIVMAL	-----G-----C-----T-----T-----G-----A-----	-----G-----C-----T-----T-----G-----A-----	1427
HIVVI59	-----T-----G-----A-----AA-----A-----A-----G-----	-----T-----G-----A-----AA-----A-----A-----G-----	1403
HIVVI310	-----T-----G-----C-----	-----G-----C-----	1400
HIVVI57	-----CT-----A-----G-----	-----CT-----A-----G-----	1403
HIVK112	-----T-----GG-----	-----T-----GG-----	1403
HIVK88	-----T-----G-----G-----	-----T-----G-----G-----	1403
HIVK29	-----T-----G-----G-----	-----T-----G-----G-----	1403
HIVK124	-----G-----C-----T-----T-----G-----	-----G-----C-----T-----T-----G-----	1415
HIVK7	-----T-----A-----A-----A-----G-----	-----T-----A-----A-----A-----G-----	1403
HIVK98	-----T-----G-----G-----	-----T-----G-----G-----	1403
HIVK89	-----T-----G-----G-----	-----T-----G-----G-----	1403
HIVVI32	-----T-----G-----G-----C-----	-----T-----G-----G-----C-----	1409
HIVVI415	-----CT-----C-----G-----	-----CT-----C-----G-----	1401
HIVCI4	-----G-----C-----T-----T-----G-----	-----G-----C-----T-----T-----G-----	1397
HIVG141	-----T-----G-----C-----T-----T-----G-----	-----T-----G-----C-----T-----T-----G-----	1391
HIVLBV23	-----TGA-----CTTT-----	-----TGA-----CTTT-----	1406
HIVTN243	-----A-----GA-----GG-----G-----C-----	-----A-----GA-----GG-----G-----C-----	1403
HIVTN245	-----A-----A-----GG-----	-----A-----A-----GG-----	1403
HIVTN240	-----A-----A-----GG-----	-----A-----A-----GG-----	1403
HIVCI30	-----G-----A-----	-----G-----A-----	1379
HIVCI59	-----A-----A-----C-----	-----A-----A-----C-----	1394
HIVLBV2310	-----G-----	-----G-----	1394
HIVCI51	-----G-----T-----	-----G-----T-----	1388
HIVLBV105	-----CT-----C-----G-----T-----C-----A-----G-----AG-----	-----CT-----C-----G-----T-----C-----A-----G-----AG-----	1418
HIVCI32	-----G-----C-----A-----A-----C-----	-----G-----C-----A-----A-----C-----	1412
HIVIC144	-----G-----A-----A-----C-----	-----G-----A-----A-----C-----	1394
HIVDJ258	-----G-----	-----G-----	1394
HIVCM238	-----A-----A-----GG-----	-----A-----A-----GG-----	1403
HIVUG266	-----T-----G-----A-----AAA-----G-----	-----T-----G-----A-----AAA-----G-----	1403
HIVVI354	-----G-----C-----T-----C-----A-----G-----	-----G-----C-----T-----C-----A-----G-----	1403
CONSENSUS.B	CCaaCAGCCcCCaCAGaagAG	AgCtTC...aGGTttgggGaaGagAcaac	1408
HIVSF2	-----G-----A-----	-----G-----A-----	1415
HIVB2167	-----T-----	-----T-----	1415
HIVPH153	-----CA-----	-----CA-----	1409
HIVPH136	-----G-----C-----A-----G-----TG-----	-----G-----C-----A-----G-----TG-----	1409
HIVB2200	-----G-----C-----	-----G-----C-----	1409
HIVTB132	-----G-----	-----G-----	1409
HIVB2190	-----C-----T-----	-----C-----T-----	1445
HIVLAI	-----C-----T-----	-----C-----T-----	1409
HIVHKB2R	-----C-----T-----	-----C-----T-----	1415
HIVMN	-----T-----G-----	-----T-----G-----	1409
HIVJH3	-----C-----G-----	-----C-----G-----	1409
HIVJRC8P	-----G-----	-----G-----	1409
HIVDYI	-----C-----G-----	-----C-----G-----	1406
HIVNY5CG	-----	-----	1409
HIVNL43	-----G-----C-----	-----G-----C-----	1409
HIVCDC4	-----CA-----G-----	-----CA-----G-----	1409
HIVHAN	-----A-----	-----A-----	1415
HIVCAM1	-----G-----	-----G-----	1409
HIVRF	-----C-----T-----	-----C-----T-----	1445
HIVD31	-----C-----T-----	-----C-----T-----	1445
HIVBHI02	-----A-----G-----	-----A-----G-----	1403
HIVFV22	-----T-----C-----	-----T-----C-----	1412
HIVJREFL	-----G-----A-----	-----G-----A-----	1409
HIVUG280	-----T-----G-----	-----T-----G-----	1409
HIVYU2	-----T-----G-----	-----T-----G-----	1409
HIVBCSG3C	-----T-----G-----	-----T-----G-----	1409

CONSENSUS.C	CCAAcAGCCcCACCAGCAGAG????????????????AGCTTC...aGGTTC...GAGGaGaCaAC	1388
HIVUG268	-----CCAAcAGCCcCACCAGCAGAG-----	1409
HIVSM145	-----	1436
HIVZAM18	-----G-----	1421
HIVZAM19	---N---N---N-----	1211
HIVZAM20	-----	1424
HIVDJ259	-----G-----C-G-----	1397
HIVVI313	-----	1397
CONSENSUS.D	CCaACAGCCCCCACCAGCAGAG.....AGcTTC...GGGTTcGGgGAggAgAaaC	1401
HIVELI	-----A-----	1412
HIVZ226	-----A-----	1415
HIVNDK	-----	1403
HIVVI305	-----A-----	1415
HIVG109	-----A-G-----	1421
HIVK31	---G-----T-----	1409
HIVUG274	-----A-----	1412
HIVUG3270	-----	1412
HIV8E365	-----CG-----	1412
HIVVI203	-----C-A-----A-G-----	1412
CONSENSUS.F	CCaAcAGCCCCCaCCAGCAGAG.....AgCTTc...GGGTTC?GaGAGCAGATAAC	1369
HIVVI174	-----G-----	1397
BIVVI69	---T-----G-----T-----	1394
HIVBZ162	-----A-----A-----	1393
HIVVI325	---G-----A-----G-G-----	1394
CONSENSUS.G	CCA?CAGCCCCACC?GCAGA?.....agcctc...gggttcgggGAGGAGATA?C	1343
HIVLBV217	---T-----C-----G-----	1406
HIVVI191	---A-----T-----A-----	1388
CONSENSUS.H	CCAACAGCCCCCACCAGCAGAG.....AGCTT?...GGGTT?GGAGAGGA?ATGAC	1297
HIVVI525	-----T-----C-----G-----	1412
HIVVI557	-----C-----T-----A-----	1409
CONSENSUS.O	CCATCAGCCCCACC?ATGGAG.....GAGG?AGTGAA	1267
HIVANT70	-----G-----A-----	1409
HIVMVP51	-----A-----C-----	1409
STVCPZGAB	CCAACGGCCCCCACCATAGAG.....AGTTAT...GGGTACCAGGAGGAAGAGAA	1439

HIV1 GAG

CONSENSUS.A	c...tCctct.....ccgaagcagGagCagaagac.....?????aagggaaccag??tcc?c	1434
HIVG455	-----C-----G-----A-----T-----G-----AC-----	1432
HIVMAL	A...C-----A-----A-----A-----TT-TA-----	1468
HIVVI59	-----C-----A-----A-----CT-----	1429
HIVVI310	-----A-----A-----G-----TCC-----T-----	1444
HIVVI57	-----C-----A-----G-----GBC-----A-----	1447
HIVK112	-----C-----G-----ACC-----A-----	1447
HIVK88	-----C-----G-----ACC-----A-----	1447
HIVK29	TC-----G-----GCC-----A-----	1447
HIVK124	A...C-----A-----A-----C-----T-TA-----	1456
HIVK7	-----C-----A-----CCA-----	1444
HIVK98	-----C-----G-----ACCA-----	1444
HIVK89	-----C-----T-----CG-----G-----	1441
HIVVI32	-----C-----G-----AGC-AA-----	1453
HIVVI415	-----C-----G-----GCC-----A-----	1438
HIVCI4	-----C-----G-----CC-----	1432
HIVG141	-----C-----A-----A-----G-----AAAGAC-----T-TA-----	1438
HIVLBV23	-----C-----A-----A-----G-----CCC-TT-----	1450
HIVTN243	-----TA-----TA-----G-TCC-----T-----	1447
HIVTN245	A-----TA-----T-----TCC-----T-----	1447
HIVTN240	-----TA-----TCC-----T-----	1447
HIVCI20	-----A-C-GG-----G-TATA-----T-----	1423
HIVCI59	-----A-C-GG-----G-T-TA-----T-----	1438
HIVLBV2310	-----A-C-GG-----G-TATA-----T-----	1438
HIVCI51	-----C-----T-C-----C-GG-----GG-TATAC-----T-----	1432
HIVLBV105	-----C-----G-----T-TA-----T-----	1462
HIVCI32	-----CAGGG-----G-TATA-----T-----	1456
HIVIC144	-----C-----A-C-GG-----G-TATA-----T-----	1438
HIVD3258	-----C-----A-C-GG-A-----C-G-TATA-----T-----	1438
HIVCN238	-----TA-----T-----CCC-----T-----	1447
HIVUG266	-----C-----A-----A-----AC-----	1444
HIVVI354	-----C-----G-C-----CA-GG-A-----G-G-TA-----	1444
CONSENSUS.B	aactCCCTcT????????????caGAAGCAGGAgccgatAGAC.....aAgGaactgTAT???c	1452
HIVSF2	-----A-----A-----T-----	1459
HIVB2167	-----A-----A-----T-----	1459
HIVPH153	-----C-----A-----	1453
HIVPH136	-----A-----A-----	1453
HIVB2200	C-----C-----A-----A-----G-A-G-----CCCT-----	1453
HIVTB132	-----A-----A-----A-----	1453
HIVB2190	Q-----A-----G-----C-----AA-----	1453
HIVLAI	-----	1489
HIVXB2R	-----C-----	1453
HIVMN	-----A-CAGAAGCAGGAGA-----A-----C-----	1471
HIVJH3	-----C-----G-----	1453
HIVJRC5F	-----CAGAAGCAGGAG-----T-----	1465
HIVYI	-----C-----G-----	1450
HIVNY5CG	-----	1453
HIVNL43	-----G-----	1453
HIVCDC4	-----G-----	1453
HIVHAN	T-----	1450
HIVCAM1	-----AA-----	1453
HIVRF	T-----	1456
HIVD31	-----T-----	1453
HIVBH102	-----C-----	1489
HIVPV22	-----C-----	1489
HIVJRFL	-----A-----	1447
HIVUG280	-----TA-----G-----T-----	1456
HIVYU2	-----	1453
HIVBC5G3C	-----C-----G-GA-----A-----	1453

CONSENSUS.C	c...CCcgCT.....CcGAAGCAGGAGCcgAaAGAc??????????agggaa?????????C	1423
HIVUG268	T...--C-----	1444
HIVSM145	-----T-----A-----	1465
HIVZAM18	-----T-----	1456
HIVZAM19	-----N-----T-----	1246
HIVZAM20	-----G-----TCCTACAGGGGAA	1468
HIVDJ259	-----C-----A-----TA-----AAGGAACCTAC-A--G--	1444
HIVVI313	-----T-----A-----	1432
CONSENSUS.D	c...CCctCt.....cagAAaCagGAGCagAAaGAc??????????aagGAactgtat...C	1443
HIVEL1	-----A-----	1453
HIVZ226	-----	1456
HIVNDK	-----	1444
HIVVI205	-----A-----AA--	1456
HIVG109	-----C-----A-----AAGGAAAAGAC	1474
HIVK31	A.....GA	1450
HIVUG274	A.....G-CA-G--	1453
HIVUG270	A.....G--	1453
HIVSE366	-----	1453
HIVVI203	-----C-----GC--G-----CA--G--G-----G--A--	1450
CONSENSUS.P	C...CCCTCT.....CCGAaGCAGGAGCaGAA?GAc.....gAGGgACTGctatcCTC	1412
HIVVI174	-----G-G-----	1441
HIVVI69	-----G-----C-----	1438
HIVB2162	-----A--A-----C-----	1437
HIVVI325	-----G-----AC--A-----A--A--A--GG--T--	1438
CONSENSUS.G	C...?CCTC?.....CCGAAGCAGGA?CC?agggac.....AAGGA?CTATATcctc	1381
HIVLBV217	---C---C-----G--A-----G-----	1441
HIVVI191	---T---T-----A--G-----A-----	1432
CONSENSUS.H	C...CCC7CT.....CC?AA?CAGGAG?7?A?GGA?.....???GAA.....?CTC	1323
HIVVI525	---T---G--G-----CCC-G--C-----AAG-----C---	1450
HIVVI557	---C---A--A-----AGG-A--T-----GCA-----T---	1447
CONSENSUS.O	G...G?ACA?.....GAGAA?CAG??7?CAGAA?GGG.....G??C?G?A?GAGCT?TA?...C	1301
HIVANT70	---G---G-----GAA-----A-----GC-C-A-C-----C--T---	1456
HIVMVP51	---A---A-----AGT-----G-----AT-A-G-A-----G--C---	1456
SIVCP2GAB	G...AGCCAG.....GAGAAAGAGGGGAGAGAGC.....AGTC?ATATCCAC	1480

HIV1 GAG

CONSENSUS.A	CtTtAgttTCCCTCAAAATCACTCTTTGGCAACGACCCcTTGTCACAG	1481
HIVU455	-----C-----AG-----	1479
HIVMAL	-----C-----AG-----	1515
HIVVI59	---C-A-----	1449
HIVVI310	-C-----	1464
HIVVI57	-----	1467
HIVK112	-C-C-A-C-----	1467
HIVK88	-----	1467
HIVK29	-C---AC-----	1467
HIVK124	-----C---G-----	1476
HIVK7	---C-----	1464
HIVK98	-----	1464
HIVK89	---C-A-----	1461
HIVVI32	-----	1473
HIVVI415	-----	1458
HIVCI4	---C-A-----	1452
HIVQ141	-----C-----	1458
HIVLBV23	---C---C-----	1470
HIVTN243	-----	1485
HIVTN245	-----	1467
HIVTN240	-----	1467
HIVCI20	-----C-----	1443
HIVCI59	---AC-----	1458
HIVLBV2310	---AC-----	1458
HIVCI51	-----C-----	1452
HIVLBV105	-C---AC-----	1482
HIVCI32	---C-----	1476
HIVIC144	---AC-----	1458
HIVDJ258	-C---AC-----	1458
HIVCM238	-----	1467
HIVUG266	---A-----	1464
HIVVI354	---ACC-----	1464
CONSENSUS.B	CtTtAgCtLCCCTCAGaTcACTCTTTGGCAaCGACCCcTcGTCACaa	1499
HIVSF2	---A-----	1506
HIVBZ167	-----A-----	1479
HIVPH153	---C-----A-T-	1473
HIVPH136	-----A-----	1473
HIVBZ200	-C-----	1473
HIVTB132	---A-G---A-----	1491
HIVBZ190	-C---G-----	1473
HIVLAI	---A-----	1536
HIVKX2R	---A-----G-----	1500
HIVMN	-----A-----A-T-	1518
HIVJH3	-----CC	1500
HIVJRC8F	---A-----	1512
HIVQYI	---A-C-----A-----	1497
HIVNY5CG	-----G-----	1500
HIVML43	-----G-----	1500
HIVCDC4	-----A-----	1500
HIVHAN	-C---C---A-----G-----	1497
HIVCAM1	-----	1500
HIVRF	-----A-----A-----G	1503
HIVD31	-----	1500
HIVBH102	---A-----	1536
HIVPV22	---A-----	1536
HIVJRF1	---A-----	1494
HIVUG280	---C-----	1476
HIVYU2	-----G-----	1500
HIVBCSG3C	-C-----	1500

CONSENSUS.C	CCTTAACTCCCTCAAAATCACTCTTTGGCAACGACCCCT	1462
HIVUG268	-----T-----	1464
HIVSM145	-----	1485
HIVZAM18	-----	1476
HIVZAM19	-----	1285
HIVZAM20	-----G-----	1488
HIVDJ259	-----G-----G-----	1464
HIVVI313	-----	1452
CONSENSUS.D	CtTtAgCttCCCTCAAAATCACTCTTTGGCAACGACCCCTtGTcGAA	1489
HIVELI	-----A-----G-----	1500
HIVZ226	---C-A-G-----T-----	1503
HIVNDK	-----C-----	1491
HIVVI205	-----A-C-----	1476
HIVGI09	-C--A--G--	1494
HIVK31	-C-----	1470
HIVUG274	-----	1473
HIVUG270	---C---G---	1473
HIVSE365	-----	1473
HIVVI203	--G-----	1470
CONSENSUS.F	CcTTAgCtTCCCTCAAATCA	1432
HIVVI174	-----C-----	1461
HIVVI69	-----	1458
HIVBZ162	-----	1457
HIVVI325	-T---A-----	1458
CONSENSUS.G	CTTTA?PTTCCCTCAAATCA	1399
HIVLBV217	-----AC-----	1461
HIVVI191	-----GT-----	1452
CONSENSUS.H	C?TFA?CTTCCCTCAAATCA	1341
HIVVIS25	-C--G-----	1470
HIVVI557	-T---A-----	1467
CONSENSUS.O	C?TTGGCTCCCTCAAATCCCTCTTTGGGACAGACCAATAG	1341
HIVANT70	-G-----	1494
HIVMVP51	-A-----	1497
SIVCPZGAB	CA...ACCTCCCTCAAATCACTCTTTGGGACGACCCCTCATCCCAG	1524

HIV1 TAT

	<= vpr	
CONSENSUS.A	ATGGA?CC?GTAGATCCTAaCCTAGAGCCCTGGAAaCA?CCGGGAAGTCAGCCTACAACtgcCTTGTAGCAAg	69
HIV0455	---G--A-----A--C-----C-----C-----A	72
HIVSF170	---T--A-----T-----C-----A	72
HIV1UG273A	---T--G-----T--T-----A-----C-----A	72
HIV1UG275A	---T--G-----T-----C-----A	72
HIV1DJ263A	---G--G-----G-----C-----C-----A	72
HIV1DJ264A	---G--A-----G-----C-----T	72
HIV1DJ258A	-----	16
CONSENSUS.B	ATGGA?CCAGTAGATCCTAgacTAGAGCCCTGGaAGCATCCAGGAAGTCAGCCTAageCtgcttGtAccaat	72
HIVLAI	-----A-----C-----C-----	72
HIVTH475A	-----A-----C-----	72
HIVHXB2R	-----A-----	72
HIVNL43	-----A-----	72
HIVN9	-----C-----	72
HIVSC	---T-----AG-----TC-----	72
HIVBAL1	-----A-----C-----	72
HIVJRC5F	-----C-----	72
HIVJRF1	-----C-----C-----	72
HIVNH5	-----C-----C-----	72
HIV0Y1	-----A-----A-----	72
HIVSP2	---AT-----G-----A-----	72
HIVNY5CG	-----A-----A-----	72
HIVSP162	---T-----A-----	72
HIVCAM1	-----C-----G-----	72
HIVCDC4	-----A-----	72
HIVSP33	---AC-----G-----C-----C-----	72
HIVRJS	-----A-----AG-----	72
HIVHAN	-----G-----CC-----	72
HIVD31	-----G-----	72
HIVRF	-----A-----	72
HIVYU2	---AC-----G-----A-----	72
HIVBCSG3C	-----C-----A-----A-----	72
CONSENSUS.C	atgGa?ccagTAGATCCTAaaCTAGAGCCCTGGAAaCCATCCaGGAaGTCAGCCTAAAACtgcCTTGTActAag	71
HIV1SM145A	-----T-----A-----A-----T	37
HIV1ZAM184	---T-----C-----G-----C-----	72
HIV1DJ259A	TCC-TCTAGA---G-----A-----	72
HIV1DJ373A	---G-----C-----A-----	72
HIV1UG268A	-----C-----	32
CONSENSUS.D	atgGatccagTAGATCCTAaccTAGAGcCCTGGaACCATCCAGGaAGTCAGCCTAggActgcCTTGTAAaAag	72
HIVELI	-----A-----A-----G-----G-----	72
HIV2226	-----T-----T-----G-----T-----	72
HIVNDK	---T-----G-----T-----	72
HIVMAL	---T-----G-----G-----T-----	72
HIV1SE365A	-----A-----C-----	72
HIV1UG274A	TCC-TCTAGA---GA-----C-----A-----	72
CONSENSUS.O	ATGGATCCAGTAGATCCTGAG?TCCCCCTTGGCA?CACCCCTGG?AG??AGCCCCA?A?CCCTTGTATAAT	65
HIVAMT70	---G-----C-----A--TC-----G--T-----	72
HIVMVP5180	---A-----T-----G--CA-----A--C-----	72
CONSENSUS.CPZ	ATGGA?CC??TAGA??CTGA?????CC?TGG?ACACCC?????T??CCA??AAC????TG?AATAA?	40
SIVCPZGAB	---T--AA---TC---CCTGGAG--C--AA---AGGCAG-CAG--AG--AGTT--T---T	72

		\3'sj\3'sj -> rev cds	
CONSENSUS.A	TGTTACTGTAAAatGTGTGCTggCATTGccaa?tttTGCTTTCTgaaacAAAAGGCTTAGGCATcTCCTATGGC	140	
HIV0455	-----T-----G-----C-----A-----T-----	144	
HIVSP170	-----G-----AT-----C-GC-----ACG-----	144	
HIV1UG273A	-----G-----AT-----C-G-----A-----	144	
HIV1UG275A	-----A-----AT-----TAC-G-----	144	
HIV1DJ263A	-----A-C-----T-T-G-----G-----	144	
HIV1DJ264A	-----A-C-----T-G-----G-----	144	
HIV1DJ258A	-----AA-C-----T-G-----G-----	8B	
CONSENSUS.B	TGCTATTGTAAAaagTGTGCTtTCATTGccAagGttTGTtTCacaAcAAAAGgCTTAGGCATcTCCTATGGC	144	
HIYLA1	-----C-----	144	
HIVTH475A	-----TG-----C-----	144	
HIVHKB2R	-----T-----C-----	144	
HIVNL43	-----TG-----C-----	144	
HIVNM	-----A-----C-----	144	
HIVSC	-----	144	
HIVBAL1	-----C-----T-----	144	
HIVJRCSP	-----C-----	144	
HIVJRF1	-----C-----C-----T-----	144	
HIVNH5	-----A-----	144	
HIVQY1	-----G-----C-----C-----T-----A-----T-----	144	
HIVSF2	-----T-C-CG-----G-----	144	
HIVNY5CG	-----G-----C-----	144	
HIVSF162	-----T-----	144	
HIVCAM1	-----C-----T-----	144	
HIVCD4	-----G-A-----C-----	144	
HIVSF33	-----C-----T-----A-----	144	
HIVRJS	-----	144	
HIVHAM	-----C-----G-----A-----	144	
HIVD31	-----A-C-----A-----	144	
HIVRF	-----A-----C-----TT-----	144	
HIVYU2	-----T-----A-----	144	
HIVBCSG3C	-----A-----TG-----	144	
CONSENSUS.C	TGTTaTtGTAAAaaaTGTaGcTatCATtGtCtAGTtTGCTTTcagAcAAAAGGcTTAGGCATtTCCTATGGC	143	
HIV1SM145A	-----CTC-----C-----C-----	109	
HIV1ZAM184	-----G-----T-----C-C-----G-----TA-AC-----C-----	144	
HIV1DJ259A	-----T-----	144	
HIV1DJ373A	-----T-G-----T-----A-----T-AC-----C-----	144	
HIV1UG268A	-----T-G-----T-----A-----T-AC-----C-----	104	
CONSENSUS.D	TGc?ATTGTAAAAGTGTGTGCTATCATTTGCCAagGttTGCTTCaTAaAgAAAAGGCTTAGGCATcTCCTATGGC	143	
HIV1UG269A	-----T-----	20	
HIVELY	-----C-----C-----T-AC-----	144	
HIVZ226	-----C-----	144	
HIVNDR	-----C-----	144	
HIVHSL	-----T-----C-----A-G-----	144	
HIV1SE365A	-----T-----T-----AA-----A-G-----	144	
HIV1UG274A	-----CT-----C-----T-G-----	144	
CONSENSUS.O	TGCTATTCGCAAAAGATGCTGCTATCATTCCTATGTTGTTTC?AA?FAAGGCTTTGGGAATCTCC?ATGGC	132	
HIVANT70	-----GT-----GA-----T-----	144	
HIVNVP5180	-----AC-----AG-----C-----	144	
CONSENSUS.CPZ	TG?TA?TG?AAA??TGTGTETA?CA?TG?????TGCTTCACAAAAAAGGCTT?GG?AT?TCCTATGGC	94	
SIUCPZGAB	--C--T--C--GCT-----T--C--TATATAT-----A--A--C-----	144	

HIV1 TAT

	1/3' s1	
CONSENSUS.A	AGGAAGAAGCGG.....AGaC7cCGACGAGaGaaCTCCTCAaaGcagTaAGGATCaTCAAAaTCCtTaTAcCA	205
HIVU455	-A-----A-C-----C-----G-----C-T-----	210
HIVSF170	-----AG-----A-----	210
HIV1UG273A	-----T-AG-----A-----	210
HIV1UG275A	-----AG-AG-----A-----T-A-----	210
HIV1DJ263A	-----G-----G-TC-C-----G-----	210
HIV1DJ264A	-----G-----G-TC-C-----G-----T-----	210
HIV1DJ258A	-----G-----G-TC-C-----C-----G-----	154
CONSENSUS.B	AGGAAGAAGcCG.....AGaCAGCGAGaaGAGCTCcTcaagaCAGTcAGacTCATCAAgtttCTcTATcCA	210
HIVLAI	-----C-----G-----	210
HIVTH475A	-----A-GA-----C-----	210
HIVHXB2R	-----A-GA-----C-----	210
HIVNL43	-----A-GA-----C-----	210
HIV2M	-----G-----C-----	210
HIV8C	-----A-----C-----	210
HIVBAL1	-----G-----	210
HIVJRCSP	-----C-----C-----	210
HIVJREFL	-----C-----	210
HIVNH5	-----	210
HIVVOYI	-----A-----	210
HIVSF2	-----G-----C-----	210
HIVNY5CG	-----G-----	210
HIVSF162	-----C-----G-GT-----C-----	210
HIVCAM1	-----A-----AG-----C-----	210
HIVCDC4	-----A-----A-----CA-----	210
HIVSF33	-----G-----AG-----G-CAG-----A-----A-----	210
HIVRJS	-----A-----A-----A-----C-----	210
HIVHAN	-----C-----	210
HIVD31	-----G-----	210
HIVRF	-----A-----G-C-----G-----C-----T-----	210
HIVYU2	-----C-----G-----AG-----	210
HIVBCSG3C	-----C-----	210
CONSENSUS.C	AGGAAGAAGCGG.....AGACagCGACGAagcgCTCCTccaagCAGtGAGGATCATCAaaATCCtTATatCA	209
HIV1SM145A	-----AG-----	175
HIV12AM184	-----CC-----AT-----A-A-----C-----C-----	210
HIV1DJ259A	-----G-C-----	210
HIV1DJ373A	-----T-----T-----	210
HIV1UG268A	-----A-----T-----	170
CONSENSUS.D	AGGAAGAAGCGG.....AGACAGcGACGaagAcCTcCTcAagGCggTCAppCTCATCAAGtTcCTaTACCA	209
HIV1UG269A	-----T-----A-----	86
HIVEL1	-----G-----	210
HIV22Z6	-----A-----T-----A-----A-----	210
HIVNDK	-----A-----A-----	210
HIVMAL	-----G-----AA-----A-----C-----	210
HIVISE365A	-----A-----AA-----A-----T-G-----	210
HIV1UG274A	-----C-----A-A-----A-----A-----	210
CONSENSUS.O	AGGAAGAAGCGA.....??GAC?A?CAGCTGCTGCAAGC?ATCCAGAT?ATAAAGATCCTGTACCA	187
HIVANT70	-----G-----G-C-----C-----C-----	204
HIVMVP5180	-----AGAA-----C-G-----T-----A-----	207
CONSENSUS.CPZ	AGGAAGA??GA?C??A??A??GACG?AC??CT?CTG?A?GCAGT?AAAATAATCAAGAT?CT?TA?CA	143
SIVCPZGAB	-----AGC-----A-C-----CA-GAC-----C-OG-----C-G-----A-----T-A-C-----	213

		/- env -> (for more information, see the ENV CDS that follow)		
		exon\exon		
CONSENSUS.A	AAGCAACCCCTaCCCCaaaaCcccaagGGgtC???	CCGacAGgCCCCGaaAGGAATCGaAGAGaAGGTGGagAGC	274	
HIVU455	-----G-----T-----GA-----T-----G-----A-----G-----		279	
HIVSF170	-----T-----A-----		279	
HIV1UG273A	-----A-----A-----A-----T-----C-----		279	
HIV1UG275A	-----A-----A-----T-----G-----A-----		279	
HIV1DJ263A	-----ACC--AG--A--A-----A-----G-----C-----		279	
HIV1DJ264A	-----ACC--AG--AA-----G-----G-----		279	
HIV1DJ258A	-----ACC--TGG--AA--CAC-----A-----G-----C-----		226	
CONSENSUS.B	AAGCAACCC?CctCCCCagccccgAGGGGAC...	CCgacagCCCCGaaGgAatcGAAGaAGaAGGTGGAGaga	278	
HIVLAI	-----A-----A-----A-----A-----		279	
HIVTH475A	-----A-----AT-----A-----		279	
HIVHXB2R	-----A-----A-----A-----		279	
HIVNL43	-----A-----AT-----A-----		279	
HIVNM	-----G--C--TT-----		279	
HIVSC	-----G--AG--A-----		279	
HIVBAL1	-----A-----T--T-----G-----		279	
HIVJRC3F	-----T-----AA-----		279	
HIVJRF1	-----T-----		279	
HIVNH5	-----T-----CGACA-----		279	
HIVQY1	-----G--A--A-----		279	
HIVSF2	-----G-----T-----C-----		279	
HIVNY5CG	-----G-----A-----		279	
HIVSF162	-----G--T-----A-----		279	
HIVCAM1	-----G--T-----TT--A-----G-----		279	
HIVCDC4	-----T-----AA-----C-----G-----		279	
HIVSF33	-----T--A-----A-----G-----		279	
HIVR78	-----G-----A-----		279	
HIVHAN	-----T-----AA-----CGACA-----A-----		279	
HIVD31	-----G-----C-----		279	
HIVRF	-----A-----G-----		279	
HIVYU2	-----A-----T--A-----C-----		279	
HIVBCSG3C	-----G--A-----A-----		279	
CONSENSUS.C	AAGCAACCCCTaCCCCaaaaCCCCgagGGGgC...	CCGACAGGCCcGgAGgAATCGAAGAGaAGGTGGAGAGC	278	
HIV1SM145A	-----T-----C-----C-----AGA--T--T-----A-----G-----		244	
HIV1ZAM1B4	-----T-----C-----C-----AGA--T--T-----A-----G-----		279	
HIV1DJ259A	-----T-----G-----A-----A-----		279	
HIV1DJ373A	-----T-----G-----A-----A-----		279	
HIV1UG268A	-----C-----A-----		239	
CONSENSUS.D	aAgCAACCCCTCTCCCAgcccCGAGGGGgAC???	CCGACAGGCCCGAAGgAatAGAAGAAGaAGGTGGAGAGC	278	
HIV1UG269A	-----		155	
HIVELI	-----C-----		279	
HIV8226	-----		279	
HIVNDK	G-----T-----A-----		279	
HIVMAL	G-----A--A-----CAC-----		282	
HIV1SE365A	--A-----A-----		279	
HIV1UG274A	-----A-----A-----		279	
CONSENSUS.O	?AGCAATCCC???	CC?CACC???	AGGAAG...CAG?AACGCCAGGAAGAACAGGAG?AG?AGGTGGAGAG	244
HIVANT70	A-----CAA--AT--AAG-----G-----G--G-----		273	
HIVMP5180	G-----TGT--CA--GGC-----A-----A--A-----		276	
CONSENSUS.CP2	AA?????CC?TA?CC?A?CCAG??GGA?C...	?AG????GC??G?AGAAAT???	GAAG?AGGTGB???	181
SIVCP2GAB	--GCAAC--T--T--C--GT--AG--A...	A--GAAG--TC--G-----CGAC--G-----CGAGC		282

HIV1 TAT

CONSENSUS.A	AAGaCAGAGACAGATCgattCgcTTAG	301
HIVU455	---G--A-----	303
HIVSF170	---G-----A-	303
HIV1UG273A	-----	306
HIV1UG275A	-----A----	306
HIV1DJ263A	-----AGGG--A--	306
HIV1DJ264A	---A-----CG-G--A--	306
HIV1DJ258A	-----A--A-A----	253
CONSENSUS.B	GAGaCAGAGaCagATCCggtcgattAG	305
HIVLAI	-----AT-----	306
HIVTH475A	-----AT-----	306
HIVHXB2R	-----AT-----	306
HIVNL43	-----AT-----	306
HIVMN	-----C-----	303
HIVSC	-----	303
HIVBAL1	-----C--	303
HIVJRC8F	-----A-A--	303
HIVJRF1	-----A--	303
HIVNH5	-----G-----C--	303
HIVQYI	-----AA--	303
HIVSF2	-----T--	303
HIVNY5CG	-----A-C--	303
HIVSF162	-----A--C--	303
HIVCAM1	-----C--G--	303
HIVCDC4	---G-----G--AC--	303
HIVSF33	-----AC--	303
HIVRJS	-----C--	303
HIVHAN	-----G-----T--	303
HIVC31	-----AA--	303
HIVRF	-----CG-TGC--	306
HIVYU2	-----C--	303
HIVBCSG3C	-----	303
CONSENSUS.C	AAGaCAGAGaCAGATCcaatlCGATTAG	305
HIV1SH145A	-----A-----	271
HIV12AM184	---G-----G-----	306
HIV1DJ259A	-----G-----G-----	306
HIV1DJ373A	---A--G-----C-----	306
HIV1UG268A	-----G-GA-----A	266
CONSENSUS.D	aAGGCAGAGaCAGATCcaTlcGATTggTGA	308
HIV1UG269A	-----G-----A--	182
HIVELI	G-----G-GA-----C--	309
HIVZ2Z6	G-----	309
HIVNDK	G-----	309
HIVMAL	-----G-----A-----	312
HIV1SE365A	-----A-----	306
HIV1UG274A	-----G-----	306
CONSENSUS.O	?AG?CAGGCCCA?GTGGA?A?OCT?GCC?C?AGG?TTC?TGC?AC?GTTGTACACGGA?CTCAG?ACAATAA	303
HIVANT70	A--G-----G-----T-C--C--G-A--G-----C--T-----C--A--	342
HIVMVP5180	G--A-----A-----C-G--T--A-C--A-----A--A-----T-----G-----	348
CONSENSUS.CPZ	AAGACAG?A?CAG?T?T?T?G?T?T?T?	196
SIVCPZGAB	-----A-G--A-CAGT-AGA--AG	309

This page intentionally left blank.

HIV1 ENV

CONSENSUS .A	ATGAgAGtGAtGGgGata...caga?gAatT?ccaacaCTtgTgGAgA????????????	43
HIVSF1703	-----T--C-G--A--A-T-----	45
HIVU455	-----G--A--CTTG-----	45
HIVZ321	---A---A---GG--GG--A--G---A---	45
HIVRW020	-----G-----C-G-GG--A-----	41
HIVUG031	---G-----G--GG--G--A-T---	41
HIVNI	-----A--A--C--T--A-----	45
HIVRW009	-----C---TT--T--A--A-----G	41
HIVUG275A	---A--C---C---T--G--A--G---	45
HIVUG273A	-----A-----G--GG--G--T---	45
HIVV1191A	-----A-----T--GG-----AG	45
HIVDJ264A	-----G--A--C--T--C--A---	45
HIVDJ263A	-----G--G--A--C--T--A--A---	45
HIVDJ258A	-----A--A--C--T--A-----	45
HIVCARGAN	-----A--A-----T-----	45
HIVCARSA5	-----A--C---A--ACTCT-C-----	45
HIVCAR4054	-----A--A--C---G--GG--CA-----A-	45
HIVCAR206A	-----A--C---G--GG--A-----	45
HIVCAR4023	-----A--A-----A--A--G-----TGGGGGTGGAAA	48
HIVCAR423A	-----A--A-----A--A--G-----TGGGGGTGGAAA	48
HIVKENYA	-----C---G--G--G--A-T-----	45
CONSENSUS .B	ATGGCAATGAgAGtGAagGggatc...aggaagaatTatCagcacTtgTggaga????????????????	51
HIVJRC5F	-----A-----	45
HIVJHFL	-----G--T-----A-----	45
HIVALA1	-----G-----G-----	45
HIVBRVA	-----A-----TGG-----	48
HIVJH3	-----A-----	45
HIVBAL1	---C--A---G-----G-----	45
HIVYU2	---C--C--A---G-----A-----	45
HIVMN	-----G--G-----G--G-----	45
HIVHB2R	-----A-----A-----TGGGGGTGGAGA	48
HIVLAI	-----A-----A-----TGGGGGTGGAAA	48
HIVNL43	-----A-----G-----TGGGGGTGGAAA	48
HIVMFA	-----A-----A-----TGGGGGTGGAAA	48
HIVCAM1	---C---A---G-----G--A-----	45
HIVNY5CG	---C---C---C-----	45
HIVADA	---A-----A-----	45
HIVJFL	---A-----G-----A-----	45
HIVSIMB4	-----G-----	45
HIVD31	---A-----A-----	45
HIVSF162	-----A-----	45
HIVBCSG3C	-----A-----	45
HIVOXI	---C--C--G---C-----G-----	45
HIVSF33	---C--G--A--C-----TG-----	45
HIVCDC4	---C-----G-----	51
HIVSF2	---A-----C-----G-----	45
HIVSF2B13	---A-----A--A--C-----	45
HIVHAN	---A-----A-----G-----	45
HIVRF	---T--A--G-----G-----A-----	45
HIVWMJ2	---T--G-----G-----T-----	45
HIVTB132	---A-----G-----	45
HIVBR020	---A--G---C--GG---G--A---	41
HIVTH014	---C-----G--A--A-----	41
HIVTH026	---C-----T--A--G-----	41
HIVUS711	A-----C-----A-----	41
HIVUS712	---C--A-----G--A--G-----	41
HIVUS715	---C-----G--A--G-----A---	41
HIVUS716	---C-----G-----G--A---	41
HIVHA593	---C--A--G---C--G-----A---	41
HIVHA594	---C-----A-----	41
HIVHA596	---C-----A-----	41
HIVHA599	---C-----G-----	41
HIVHA651	---C---A--G---C--GG---C---	41
HIVRJS	---C--G--A---C--GG-----	45
HIVSC	---ATCA...G--A--G-----	45
HIVBR014	---C-----G-----G--TT-G	41
HIVBR021	-----G-----G-----	45
HIVNA208	---G-----G-----TGGGGCACCATGCTCCT	62
HIVP896	-----A-----GG-----	42

CONSENSUS.C	ATGAGAGTGAAGGGGATA...ctGagGAATg?CaACaaTGGtGGATA.....	44
HIVMA959	-----C-C-----G-----	41
HIVMA960	-----TC-----G-----	41
HIVZAM20A	-----T-----	45
HIVD1024	-----AT-----G-----	45
HIVSM145A	-----G-----	45
HIVZAM18A	-----A-----G-----	45
HIVDJ259A	---A---T---C---A---C---G---C---GG---G---	45
HIVDJ373A	---T---T---C---A---T---C---T---	45
HIVSE364A	---C---G---T---T---C---	45
HIVUG268A	---T---A---T---	45
CONSENSUS.D	ATGAGAGTGAAGGGGATA...gagAggAATTaTCA?cacttgTgGaga????????????????	44
HIVJY1	---T---G---AG---T---A---A---	45
HIVNDK	---C---A---G---AA---A---	45
HIVMAL	---C---C---AA---G---	45
HIVELI	---C---G---A---G---AA---G---A---	45
HIVZ226	---G---G---AA---A---	45
HIVUG024	---G---C---G---T---A---A---	45
HIVUG021	---G---AG---G---G---T---G---	41
HIVUG038	---T---AT---GAG---A---	45
HIVUG046	---T---G---G---TC---GAGAGGAATCAGCACTCGTT	62
HIVUG266A	---A---G---CC---AA---A---T---	45
HIVUG269A	---G---	45
HIVUG274A	---G---G---T---	45
HIVSE365A	---C---G---A---A---G---	45
HIVCAR4020	---G---A---A---	45
CONSENSUS.E	ATGAGAGTGAAGGGGATA...cagatgAatTggCaaACTtgTGGaaA????????????	45
HIVTH966	-----	41
HIVTH975	-----	41
HIVTH022	-----	41
HIVTH243	-----	45
HIVTH006	-----	41
HIVTH011	-----	45
HIVCARELO	---G---C---	45
HIVCAR4017	---A---TAT---CACTTG---AG---TGGGG---	48
HIVCAR4039	---G---G---A---A---	45
HIVCAR4071	---G---	45
HIVCARMBA	---G---C---G---	45
CONSENSUS.F	ATGAGAGTGAAGGGGGATG...CAGAGGAA?TGGCA?CACTTGGGGAAA.....	43
HIVBZ163A	---C---G---	45
HIVBZ126A	---T---A---	45
CONSENSUS.G	ATGAGAGTGAAGGGGATA...CaGAGGAATTGGCAacACTTgTGGAAa.....	45
HIVV1525A	---G---C---A---T---	45
HIVLBV217	---C---T---	45
HIVCAR4067	---G---A---	45
CONSENSUS.O	ATGACAGTGCATATGAAA...GcAATGgaGAAGAgAACAgGAAGTtAtGG.....	48
HIVANT70	---T---A---A---	48
HIVMVP5180	---T---A---AC---C---	48
HIVVAU	---C---T---G---	48
CONSENSUS.CPZ	ATGA????A?G??GA??...????????????????????T??...	9
SIVCP20AB	---AAGTA-T-QA---AG...AAGAAGAGAGACTGGAAACAGCT-ATCC.....	45
CONSENSUS.U	ATGAGAGTGAATGGTGAATA...CaGAGGAATTgtCaaCacTt?TgGagA.....	43
HIVZAM184	---G---G---C---A---A---G---T---	45
HIVCAR4081	---A---A---C---C---G---CA---	45
HIVK124A	---G---G---G---T---A---T---G---	45
HIV23	---A---A---A---A---G---A---	45

HIV1 ENV

	signal peptide \ / gp120	
	<- vpu	
CONSENSUS.A	TGGGG?ectatgaTctTtgGgatgataataATttgTa?Tgct???cagaa???aA?ttGTGGG	97
HIVSF1703	A-----G-A-T-----C--G-----A-AGC...-A-----	103
HIVU455	A-----G--T-----A--A...-C--C-A-----	100
HIVZ321	G--T-----C--T--G-G-----G-----C-----	103
HIVRW020	A-----G--C--G-G--A-G...-G-----C-----	99
HIVUG031	A-T-----G-----C--G--AGAA--C--A-----	102
HIVNI	T-G-AT--A--T--A--G-----A--C-A-----	100
HIVRW009	A-----G--T-----G--G-A-C--C-----	99
HIVUG275A	A-----G--C-----G-A--G--G-C-----	103
HIVUG273A	A--A--G--TT--C--C--G-A--G-A--C-----	103
HIVVI191A	A-T-----C--T--GGG--A-----T-----	100
HIVDJ264A	T-TG--A--A--T--A--G-----A-----G-----	100
HIVDJ263A	T-G--C--A--T--CA--G-----A-----G-----	100
HIVDJ258A	T-G--A--A--T-A--A--G-----A-----G-GA-----	100
HIVCARGAN	G-----T-A-----A--TG--T--G--C--C-AC-----	103
HIVCAR5AS	G-T-----C-----C-----G-----A-----	103
HIVCAR4054	C-CCTC-----A--T--G--C--A--AA-----C-----	103
HIVCAR286A	CCTC-----A--T-G--G--A--A--A--C-----C-----	103
HIVCAR4023	C-GC--C--C--AT-G--G--C--A--A--GACA--GAC--A-----	109
HIVCAR423A	C-GC--C--C--AT-G--G--C--A--T--C--A-----	106
HIVKENYA	A-----G-----C-----G-----C-----	103
CONSENSUS.B	???????tGGGGCaccaTGCTCctTGGgatgttgATGAtctgTAgTgct???gcagaa...aaattgtGGG	109
HIVJRC5F	G-----T--T-----CA--A-----T-----G-----	103
HIVJRF1	G-----T-----A--A-----T-----G-----	103
HIVLA1	T-----A-----A-----A-----	103
HIVBRVA	G--TG-----A-----A--C-----	106
HIVJH3	G-----A-----A-----C-----	103
HIVBAL1	T-----A--A-----A-----A-----	103
HIVYU2	G-----T-----A-----C-----	103
HIVMN	G-----T-A--A-----A-----	103
HIVHKB2R	A-----A-----A-----	106
HIVLAI	A-----A-----A-----	106
HIVNL43	A-----A-----A-----	106
HIVMFA	T-----A-----A-----T-----	106
HIVCAM1	A-----A-----A-----C-----	103
HIVNY5CG	G-----A-----A-----C-----	103
HIVADA	T-----A-----A-----T-----	103
HIVJFL	A-----A-----G-----	103
HIVSIMI84	A-----A-----T-----C-----	103
HIVD31	T-----A-----A-----	103
HIVSF162	G-----T-----T-----T-----	103
HIVBCSG3C	T-----A-----A-----GG-----T-----	103
HIVYOI	A-----A-----A-----T-----	103
HIVSF33	G-----A-----A-----T-----	103
HIVCDC4	A-----A-----C-----C-----	109
HIVSF2	T-----A-----A-----	103
HIVSF2B13	T-T-----A-----AA-----	103
HIVHAN	G-----T-----A-----A-----T-----G-----	103
HIVRF	G-----T-----A-----G-----G--G-C-----	103
HIVMMJ2	G-----T-----G-----T-----C-----	103
HIVTB132	T-----A--A-----TA-----T-----	103
HIVBR020	G--T-T-----C-----AA--C-----C-----	99
HIVTH014	A-----A-----A-----G-----	99
HIVTH026	A-----A-----A-----C-----	99
HIVUS711	A-----A-----A-----T-----	99
HIVUS712	T-----A--A-----C-----A-AA-----G--C-----	99
HIVUS715	T-----A--A-----A-----A-----G-----	99
HIVUS716	TG-----A--A-----A-AA-----C-----	96
HIVHA593	C-A--T-----A--C-----C-----	99
HIVHA594	T-----T-----T-----C-----	99
HIVHA596	T-----T-----T-----C-----	99
HIVHA599	TG-----T-----A-----G--T-----	99
HIVHA651	A--A-----A--AG-----A-----C-----	96
HIVRJS	G-----A-----A-----T-----	103
HIVGUN	G-----G-----A-----C-----	18
HIVSC	AGA-----T-----A-----A-----	103
HIVBR014	T-----C-----GAA--A--C-----	102
HIVBR021	T-----C-----GAA--A--C-----	106
HIVMA208	TGGGATA-----A-----T-----T-----	127
HIVP096	G-----T-----GCAAA-----AC-----	103

	signal peptide \ / gp120	
	<- vpu	
CONSENSUS.C	tGGGGcATCTTAGGCTTTTGGATgttaatTGAttTgTAaTGtG??gtgggg...aACTTGTGGG	102
HIVMA959	-C-----G-----TG-----	99
HIVMA960	-A-----A-G-----T-----	99
HIVZAM20A	-----G-----G-----	103
HIVD1024	-----A-A-----	103
HIVBR025	-----AA-----A-----	13
HIVSM145A	-----G-G-----GA-----	103
HIVZAM18A	-G-----AC-A-----GTA-GA-AC-----	106
HIVDU259A	-A-----GT-----G-----	97
HIVDU373A	G-A-----C-GT-----G-----A-----	103
HIVSE364A	-----CG-T-----A-----G-----	103
HIVUG268A	-A-----A-----A-----	103
CONSENSUS.D	???????TGGGGcA?caTgcPCCtTcgggaTgtTgATgAcaTgTAgtg?c???gcAGaa...?a?ttGTGGG	98
HIVJY1	-T-----A-----G-----G-T-----	103
HIVNDK	-T-----C-----C-----G-T-----	103
HIVMAL	TG-----C-----AT-----G-T-----	103
HIVBLI	-T-----A-----C-----C-----A-TC-----	103
HIVZ226	-T-----A-----C-----AA-----C-----A-TC-----	103
HIVUG024	-T-----A-----T-----G-ACA-A-----C-A-C-----	106
HIVUG021	-C-T-----A-----T-A-----AC-----A-G-----A-T-----	99
HIVUG038	-C-----T-A-A-----AGC-----AA-----	100
HIVUG046	GTGGAGA-----T-----A-----T-----AA-----G-----	121
HIVUG266A	-A-CT-AA-----G-A-----A-A-A-TP-----AC-----C-----A-C-----	103
HIVUG269A	-CG-----T-----G-T-----A-----G-----G-G-----	103
HIVUG274A	-C-----A-----T-----A-----T-----C-G-----	103
HIVSE365A	-C-----A-----T-----T-----A-A-----	103
HIVCAR4020	-CG-----A-----C-----T-----A-A-----	103
CONSENSUS.E	TGGGGgActtTGaTCCCTGGGGTggtTGATPaATtTGtagtgcc...7TCAGac...AACTTgTGGG	103
HIVTH966	-T-----A-----	99
HIVTH975	-----T-----	99
HIVTH022	-----A-----	99
HIVTN243	-----	103
HIVTH006	-----	99
HIVTH011	-----NTAGTG-----C-A-----	104
HIVCARELO	-A-----G-----G-----A-----	103
HIVCAR4017	-C-GCA-C-----A-AT-----G-C-----T-----	106
HIVCAR4039	-A-----G-----	103
HIVCAR4071	-----G-----	103
HIVCARMBA	-----A-----A-----	103
CONSENSUS.F	TGGGGCCCTTTTATTCCTGGGGATATTTG?TAATCTGTAAATGCT.....GAA...aACTTgtGGG	97
HIVBRA7944	-----G-A-CG-----	13
HIVB2163A	-----T-----	100
HIVB2126A	-----A-----	100
CONSENSUS.G	TGGGGgAcTTTGATCCCTTGGGTTGGTgATAATtTGTAgtGCC...tCAaaT...AaCTTGTGGG	103
HIVVI525A	-----A-----	103
HIVLBV217	-A-T-----A-----A-----GG-----	103
HIVCAR4067	-----C-----A-----G-----	103
CONSENSUS.O	AFTCTgTaCtTegccaTGCCCTTCaTAAtCCCATGTTTG????AGc??ta??CAacTaTATG	101
HIVANT70	-C-----C-----CT-GA-G-----	106
HIVMVP5180	-G-A-A-----C-----AGTTAT--TAAA-----T-G-----	109
HIVVAU	-T-G-G-G-QATTT-----TG--AC-----	106
CONSENSUS.CPZ	??????????T??????T?AT???G?T?A????????GAC?..??TGA?..??????T??G	24
SIIVCPZGAB	...ATAATTACAA-CATAACAA-C-TTP-C-A-CCCATGTTT--C...TC--G.....TFA-GG-	103
CONSENSUS.U	tGGGgGcaTt??T?atctTtgGgaTGaTaAT?ATTtGTAAatGct...g?agaa???AaCTtGTGGG	97
HIVZAM184	-CT-AGG-----T-----T-----G-----G--TG-----TG-G-----	103
HIVCAR4081	-G--T-G--C-----T--G-G--A-----C-----TC--CAAC-----	106
HIVKI24A	G--A--A-G--G--G--A--C-----C-----	100
HIVZ3	-A-C-GA-A-T-A-A-----G-A-----A-----AT-----A-C-----	103

HIV1 ENV

CONSENSUS .A		TcAC7GTcTAcTATGGGGTACCTGTgTGGaaaga7gcagagAACACccTaTTcttGtGCATCagATGctAaa	166
HIVSF1703	-T-T-----T-----C-----T-----	174	
HIVU455	--G-----T--TT-----	171	
HIVZ321	-G-----T-----T-----	174	
HIVRW020	-TG-T-----C-----T-G-	170	
HIVUG031	-T-T-----A-----A-----	173	
HIVNT	-G-----GACC-----	171	
HIVRW009	-T-T-----C-----T-GG-	170	
HIVUG275A	-T-T-----A-G-C-----T-----G-	174	
HIVUG273A	-T-T-A-----G-C-----C-----	174	
HIVV1191A	--A-A-T-----T-C-----G-	171	
HIVDJ264A	-G-----G-T-A-----	171	
HIVDJ263A	-G-----G-A-CA-----	171	
HIVDJ258A	-G-----G-C-A-----	171	
HIVCARGAN	--A-T-----A-T-----C-----A---C-G-	174	
HIVCAR5AS	-A-T-----T-A-----	174	
HIVCAR4054	-A-A-T-----G-G-A-T-----T-----C---	174	
HIVCAR286A	-A-A-----T-----TCG-----T-----	174	
HIVCAR4023	-T-A-----T-----G-T-A-T-----T-----	180	
HIVCAR423A	-A-A-T-----CG-T-A-T-----T-----	177	
HIVKENYA	-T-T-----A-C-----T-----	174	
CONSENSUS .B		TcACAgTcTATTATGGgGTaCCTGTgTGGhaaGAagChaccCACActcTaTTttGtGCAtCaGatgctAAA	180
HIVJRC5F	-----A-----	174	
HIVJRFL	-----	174	
HIVALA1	-----	174	
HIVBRVA	-----A-----	177	
HIVJH3	-----G-----	174	
HIVBAL1	-----CG-----	174	
HIVYU2	-----	174	
HIVMN	-----	174	
HIVHXB2R	-----G-----	177	
HIVLA1	-----G-----	177	
HIVNL43	-----G-----	177	
HIVMFA	-----G-----	177	
HIVCAM1	-----A-----	174	
HIVNY5OG	-----C-G-----	174	
HIVADA	-----T-----	174	
HIVJPL	-----A-----C-----	174	
HIVS1M184	-T-----C-----	174	
HIVD31	-----A-----T-----C-----	174	
HIVSP162	-----	174	
HIVBCSG3C	-A-----A-----C-----G-	174	
HIVOYI	-----T-----T-----T-----	174	
HIVSP33	-----T-----T-----T-----	180	
HIVCD4	-----T-----A-----T-----G-	174	
HIVSF2	-----G-----T-----G-	174	
HIVSP2813	-----T-----	174	
HIVHAN	-----T-----	174	
HIVRF	-----A-----	174	
HIVMAJ2	-----G-----G-----	174	
HIVTB132	-----	174	
HIVBR020	-----A-----G-G-----	170	
HIVTH014	-----A-----G-----	170	
HIVTH026	-T-----A-----A-----T-----	170	
HIVUS711	-T-----A-----TT-----	170	
HIVUS712	-T-----A-----C-----	170	
HIVUS715	-T-----A-----	170	
HIVUS716	-----	167	
HIVHA593	-----	170	
HIVHA594	-----	170	
HIVHA596	-----C-----T-----	170	
HIVHA599	-----T-----G-----	167	
HIVRA51	-----T-----	174	
HIVRJ5	-----A-----	89	
HIVGUN	-----	174	
HIVSC	-----G-----T-----C-----	173	
HIVBR014	-----G-----	177	
HIVBR021	-----A-----G-----	54	
HIVACN9	-----A-----G-----	54	
HIVACP1	-----A-----G-----	54	
HIVJM	-----A-----G-----	54	
HIVWM	-----A-----G-----	54	
HIVMA208	-----A-----G-----	198	
HIVP896	-----A-----G-----	174	

CONSENSUS .C	TCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAAGCAAAAACtAaCtCtATTtGTGTCATCaGAtgCTAAA	173
HIVMA959	-----C-----C-----	170
HIVMA960	-----G-----C-----	170
HIVZAM20A	-----G-----C-----	174
HIVD1024	-----C-----	174
HIVBR025	-----C-----	84
HIVSM145A	-----G-----G-----G-----T-----	174
HIVZAM18A	-----T-----C-----	177
HIVDJ259A	-----T-----CC-C-C-----	168
HIVDJ373A	-----C-----CC-----	174
HIVSE364A	-----A-----G-----	174
HIVUG268A	-----C-----	174
CONSENSUS .D	TcACaGtTtATtATGGGGTtACCTgTgTGGAAAGAAAGCAaaccACCActcTATTtGtGCATCAGAtGCTAAA	169
HIVJY1	-----G-----	174
HIVMDK	-----G-A-A-----G-----T-----	174
HIVMAL	-T-----T-----	174
HIVEL1	-----G-A-G-----	174
HIV2226	-----G-G-A-G-----	174
HIVUG024	-----T-----C-----	177
HIVUG021	-----T-----	170
HIVUG038	-----C-----A-----	171
HIVUG046	-----T-----C-----	192
HIVUG266A	-T-T-C-C-----T-GAG-CT-G-G-----	174
HIVUG269A	-----T-----C-----	174
HIVUG274A	-----G-----T-----C-----	174
HIVSE365A	-----G-----	174
HIVCAR4030	-----GAA-----	174
CONSENSUS .E	TTACAGTCTATTATGGGGTTCCTGTGTGGAGAGATGCCAGATACCAACCTATtTtGTGTCATCAGATGCCAAA	174
HIVTH966	-----	170
HIVTH975	-----A-----	170
HIVTH022	-----A-----	170
HIVTH243	-----	174
HIVTH006	-----G-----G-----	170
HIVTH011	-----C-----	175
HIVCARELO	-----C-----C-----C-----A-----	174
HIVCAR4017	-----C-A-----A-----	177
HIVCAR4039	-----A-G-A-----	174
HIVCAR4071	-----	174
HIVCARMBA	-----A-----GC-----	174
CONSENSUS .F	TCACAGTCTATTATGGGGTACCTGTGTGGAAAGAAAGCAACcACTACTCTATTtTGTGTCATCAGATGCTAAA	168
HIVBRA7944	-----T-----	84
HIVB2163A	-----T-----	171
HIVB2126A	-----	171
CONSENSUS .G	TCACAGTCTATTATGGGGTtACCTGTGTGGGAAGAtgcAGAtaCCcCtCTATTtTGTGTCATCTGATGCTAaa	174
HIVVI525A	-----C-----A-----	174
HIVLBV217	-----C-----C-----A-G-----	174
HIVCAR4067	-----G-----CG-----G-----GG-----	174
CONSENSUS .O	CcACAGTCTATTtCTGGGGTtACCTGTATGGGAAGAtGCAacACCAGtAcTtTCTGTGCTTCAGATGCTAAC	172
HIVANT70	-A-----G-----G-----	177
HIVXVP5180	-----T-----G-G-----	180
HIVVAU	-----A-AC-T-G-----	177
CONSENSUS .CPZ	TAACAGTAT??TATGGAGT?CCTGT?TGG? ???ATGC???CC???CT?TTTGTGCC?CA?A?GC????	71
SIVCP2GAB	-----AT-----A-----T-----CATG-----TGAC-----GGTA-----C-----T-----G-C-----TAAG	174
CONSENSUS .U	TcACaGtTtATtATGGGGTACCTGTGTGGA?AGAtGCAaA?ACcActcTATTtTGTGTCATCaGAtGCTAAA	165
HIVZAM184	-----C-----G-A-A-A-T-----C-----	174
HIVCAR4081	-----T-----A-A-G-C-----T-----	177
HIVK124A	-T-T-T-C-----G-----G-----T-----G-----	171
HIVZ3	-----C-----A-----A-----G-----	174

HIV1 ENV

CONSENSUS.A	GCatAtgatacGAAat?CAtAATGTcTGG?GctACaCatgCCTGTGTACCCACAGACCCcaacCCCaAaAG	235
HIVSF1703	-----G-G-----G-G-----G-----G-----	244
HIVU455	-----G-G-----G-G-----	241
HIVZ321	-----AA-----C-----	244
HIVRW020	--T-----G-G-----A-----AA-----	240
HIVUG031	--A-G-----G-----G-----	243
HIVNI	-----G-A-----T-----	241
HIVRW009	-----C-----AG-----C-----	240
HIVUG275A	-----G-----G-----T-----	244
HIVUG273A	-----G-----G-----T-----	244
HIVV1191A	---AG-GA---G-AA-C-----G-----A---	241
HIVDJ264A	---C---GT---G-A-----T-----	241
HIVDJ263A	---G---G-A-----T-----	241
HIVDJ258A	---G---G-A-----T-----	241
HIVCARGAN	-----G-G-----T-----	244
HIVCAR5AS	-----G-----C-----	244
HIVCAR4054	--C-AG--T--AG-----C-----	244
HIVCAR286A	---AG-G-T---AA-----G-----T-----	244
HIVCAR4023	---AG--T---AA-----A-----G-----G---	250
HIVCAR423A	---AG--T---AA-----T-----	247
HIVKENYA	---A-----AA-----	244
CONSENSUS.B	gCaTAtgatacAGAggtacATAAagTtTGg.GccAcaCatgCCTGTGTaCCcACAgACCCcaacCCCaAaAG	250
HIVJRC5F	-----	244
HIVJRF1	-----	244
HIVLA1	-----	244
HIVBRVA	---A-----C-----	247
HIVJH3	-----	244
HIVBAL1	-----	244
HIVYU2	-----	244
HIVMN	-----A-----	244
HIVHKB2R	-----	247
HIVLAI	-----	247
HIVNL43	-----	247
HIVMPA	-----T-----	247
HIVCAM1	---C-----T-----	244
HIVNY5CG	---C-----T-----	244
HIVADA	-----	244
HIVJFL	---C-----T-----	244
HIVSIM184	---C-----T-----T-G-----	244
HIVD31	---C-A-----G-----	244
HIVSF162	--C---C-----C-----T-----	244
HIVBCSG3C	-----A-----	244
HIVQY1	---C-----	244
HIVSF33	-----T-----C-----	244
HIVCDC4	-----C-----A-----	250
HIVSF2	-----	244
HIVSF2B13	-----	244
HIVHAN	--T-----	244
HIVRF	---A-A-----C-----A-----T-----T-----	244
HIVW02	---AG-----C-----C-----T-----	244
HIVTB132	-----C-----G-----	244
HIVBR020	---CA-AG---G-----C-----	240
HIVTH014	-----	240
HIVTH026	-----	240
HIVUS711	-----G-----	240
HIVUS712	-----A-----	240
HIVUS715	-----G-T-----	240
HIVUS716	-----T-----	237
HIVHA593	---A---A---C-----	240
HIVHA594	---G---G---C-----	240
HIVHA596	---G---G---C-----	240
HIVHA599	---G---C---C-----	240
HIVHA651	---C-C---C-----C-----	237
HIVRJS	---C-----	244
HIVGUN	---C-----A-----T-----	159
HIVSC	-----A-----T-----	244
HIVBR014	T---C---C-----C---G-----T-----	243
HIVBR021	---C---A-G---C-----T-----	247
HIVACH9	---C-----G-----	124
HIVACP1	A-----G-----G-----	124
HIVJN	-----G-----C-----	124
HIVMM	-----AG-----T-----	124
HIVMA208	---C---A---C-----C-----	268
HIVP896	--C-----	244

CONSENSUS .C	GCATATgAga?aGAAGtgCATAATGTCTGG .GCTACACATGCCGTGTGTaCCCCACAGacCCCCaaccCACAaG	242
HIVMA959	-----G-----	240
HIVMA960	-----CAG-----	240
HIVD747	-----A-----	50
HIVD760	-----T-----	52
HIVZAM20A	--T-----C-----	244
HIVD1024	-----GA-----	244
HIVBR025	-----T-C-----	154
HIVSM145A	-----G-----	244
HIVZAM18A	-----G-----	247
HIVDJ259A	-----A-AGT-----A-----	238
HIVDJ373A	-----C-----	244
HIVSE364A	-----G-----G-----	244
HIVUG268A	-----C-----	244
CONSENSUS .D	tCATAtaaaacAGaggcaCATAAtaTcTGG .GCTACACATGCCGTGTaCCaACaGACCCCaACCCACgAG	239
HIV0Y1	-----G-C-----	244
HIVNRK	G-----A-----	244
HIVMAL	-----G-----A-T-----C-----	244
HIVELI	-----G-----	244
HIVZ226	-----C-----	244
HIVUG024	G-----A-GG-----C-----	247
HIVUG021	-----G-G-----	240
HIVUG038	-----T-A-----	241
HIVUG046	-----G-C-----	262
HIVUG266A	G-----G-T-----A-TG-----G-----	244
HIVUG269A	-----A-G-----	244
HIVUG274A	-----T-G-----	244
HIVSE365A	-----G-----AA-----C-----	244
HIVCAR4020	-----G-----A-----A-----G-----C-G-----A-----	244
CONSENSUS .E	GCACAtgaGACaGAAGTGCACAATgTcTGG .GCCACACATGCCGTGTaCCaACAGAccCcAAcCCacaag	244
HIVTH966	-----C-----	240
HIVTH975	-----A-----	240
HIVTH022	-----T-----	240
HIVTH243	-----G-----	244
HIVTH006	-----A-----	240
HIVTH011	-----T-----	241
HIVCARELO	-----G-----	244
HIVCAR4017	-----A-----	247
HIVCAR4039	-----A-----	244
HIVCAR4071	-----C-T-----A-----	244
HIVCARMBA	-----C-T-----A-----T-----	241
CONSENSUS .F	TCATATGAAaAgAGAGGTACATAATGTCTGG .GCTACACATGCCGTGTaCCaACAGATCCCaATCCACAaG	238
HIVBRA7944	-----G-----	154
HIVBZ163A	-----A-----	241
HIVBZ126A	-----A-----	241
CONSENSUS .G	tCATATAGTACTGAAAgcCATAATGTCTGG .GCTACACATGCCGTGTaCCaACAGATCCCaATCCACAaG	244
HIVVI525A	-----G-----	244
HIVLBV217	G-----AA-----	244
HIVCAR4067	-----C-----	244
CONSENSUS .O	cTaaCAAGCACTGAACAGCATAATATTGG .GCATCACAGCCTGTGTCCcACAGacCCCa?TCCA?ATG	240
HIVAMP70	-----A-----	247
HIVMP5180	-----C-----	250
HIVVAU	T-G-----A-----C-----G-----A-----	247
CONSENSUS .CPZ	?????AGTACAGAGG??CA?AAT?T?TGG .GC?AC??????TGTGT?CC?A?AGATCC??TCCCT???G	113
SIVCPZGAB	-----CT--T--A-T--C--ACAGGCA--A--T-C--CAG--CAG--	244
CONSENSUS .U	gCATAT?agAcaGAAatgCAPAATGTCTGG .GCTACACATGCCGTGTaCCCaCAGACCCCaACCCACAAG	234
HIVZAM184	-----G-G-G-----	244
HIVCAR4061	T-----AGT--T--AA-----	247
HIVK124A	-----A-----	241
HIVZ3	-----G--A--TC-----	244

HIV1 ENV

CONSENSUS.A	AaaTatcattTggaAAAT...GTgACAGAAaagTTtaAcATgtGGhAAATAaacATGGTAGA?CagATgcAt	302
HIVSF1703	---T---TC---GG---G---	312
HIV0455	---G---TC---G-A---	309
HIVZ321	---C-AG---GC---A-G---	312
HIVRW020	---G---A---G-T---G---	308
HIVUG031	---G---A---G---A---	311
HIVNI	---A---A---A---A---	309
HIVRW009	---G---G---G---G---	308
HIVUG275A	---A---G---A---G---	312
HIVUG273A	---G---G---G-A---G---	312
HIVVI191A	---A---TT---G-C-C---A---G---G	309
HIVDJ264A	---A---A---A---A---	309
HIVDJ263A	---A---A---A---A---	309
HIVDJ258A	---T---A---A---A---	309
HIVCARGAN	---A---TC---G---A---	312
HIVCARSA5	---GG---A---T---G-A---	312
HIVCAR4054	---A-AC-A---A---T---T---A---	312
HIVCAR286A	---G---C-C---A---T---A---GT---A---CG---	312
HIVCAR4023	---G---C-C-A---A---T---GT---A---G---	318
HIVCAR423A	---G---T-C---A---T---GT---T---	315
HIVKENYA	---G---G---G---G---	312
CONSENSUS.B	AagTagtAtTggaAAat???GTgACAGAAaAtTTtaAcATyTGGaaaAATaacATgGTgAaCagATGcAt	318
HIVJRC9F	---A---G---G---	312
HIVJRF1	---A---C---G---	312
HIVALA1	---A---A---A---	312
HIVBRVA	---T---A-G---G---	315
HIVJH3	---A---A---A---	312
HIVBAL1	---A---A---A---	312
HIVYU2	---AA---A---	312
HIVMN	---A---T---A---	312
HIVHXB2R	---T---T---G---	315
HIVLAI	---T---T---G---	315
HIVNL43	---T---T---G---	315
HIVMTA	---A---T---G---	315
HIVCAM1	---T---T---G---	312
HIVNY5CG	---C---A---	312
HIVADA	---A---A---	312
HIVJF1	---AA---G---G---	312
HIVSIMI84	---A---T---G---A---	312
HIVD31	---G---G---	312
HIVSF162	---A---A---	312
HIVBCSQ3C	---GA---A---G---A---G---	312
HIVQY1	---G---G---T---A---A---	312
HIVSF33	---G---G---C---	312
HIVCDC4	---G---G---	318
HIVSF2	---G---G---G---	312
HIVSF2B13	---G---G---G---	312
HIVHAN	---A-G---G---	312
HIVRF	---C---G---T---	312
HIVWMJ2	---A---G---T---	312
HIVTB132	---C---T---A---C---A---	312
HIVBR020	---A---A---	308
HIVTH014	---G---C---	308
HIVTH026	---G---C---G---A---G---	308
HIVUS711	---G---G---A---	308
HIVUS712	---A---A---	308
HIVUS715	---A---G---	308
HIVUS716	---G---G---	305
HIVHA593	---T---C---G---	308
HIVHA594	---C---G---T---	308
HIVHA596	---C---G---	308
HIVHA599	---A---G-C---C---A---	308
HIVHA651	---A-G-A---A---	305
HIVRJS	---AA---A---T---	312
HIVGUN	---T---T---G---	227
HIVSC	---G---GAAT---G---	312
HIVBR014	---G---G---G---	314
HIVBR021	---G---G-C---	315
HIVACH9	---C---A---	192
HIVACP1	---AC---A---T---	192
HIVJM	---A-CC---AT---C---A---	192
HIVMM	---T---G---	192
HIVHA209	---T-CC---G---A---	336
HIVP896	---G---T---	312

CONSENSUS.C	AaaT7gtTtTGGaAAAT...GTAACAGaAaatTTTAAcATgTGGaAAATGacATgGTggATCagATGCAT	309
HIVMA959	---A---G-----A---A-----	308
HIVMA960	---G-A---G-----A---A-----	308
HIVD747	---G-G---G-----G-----A-----	118
HIVD760	---A-----G-----A-----	120
HIVZAM20A	---G---C-----C-----	312
HIVD1024	---G-----A-----	312
HIVBR025	---G-----G-----A-----	222
HIVSM145A	---GC---G-----T-----	312
HIVZAM18A	---A---G-----	315
HIVDJ259A	---C-A-C-----G-----	306
HIVDJ373A	---C-A-A-----G-----	312
HIVSE364A	---G-G-----T-----A-----	312
HIVUG268A	---A-C-----G-----	312
CONSENSUS.D	AAaTagaacTggaaAAT...GTcaCaGaaacTTTaaCATgTGGaAAATaAcATGGTGGagCagATGCAT	307
HIVJY1	---A---G---T-----A---A-----	312
HIVNDK	---T---G-----A---A-----	312
HIVMAL	---GGO-----	312
HIVELI	---C-----G-----A-----	312
HIVZ226	---G-----G-----A-----	312
HIVUG024	---C---A-----A---A-----	315
HIVUG021	---T-A-----G---C-----	308
HIVUG038	---G-----	309
HIVUG046	---A-----G-----	330
HIVUG266A	---A---A-----G-----	312
HIVUG269A	---A-T-A-----C-----T-----	311
HIVUG274A	---AG---A-----G-----A-----	312
HIVSE365A	---G---C-----	312
HIVCAR4020	---G-G---A---T-----A-----	312
CONSENSUS.E	aaatatac?Tggaaaat...gtaacagaaaatTTTaaCATgTGGaAAATaAcATGGTGGagCagATGCAT	311
HIVTH966	---C-C-----A-----	308
HIVTH975	---C-C-----	308
HIVTH022	---C-----	308
HIVTH243	---C-----C-----	312
HIVTH006	---C-C-----	308
HIVTH011		241
HIVCARELO	---A-----C-----	312
HIVCAR4017	---A-----C-----	315
HIVCAR4039	---A-----	312
HIVCAR4071	---T-T-----	312
HIVCARMBB	---G---A---T-----A---T-----	309
CONSENSUS.F	AAGTAGctCTGgAAAT...GTAACAGAAAATTTTGAATGTGGAAAATaAcATGGTGGaACAAATGCAT	306
HIVBRA7944	---AG---A-----C-----	222
HIVB2163A	---G-----	309
HIVB2126A	---G-----	309
CONSENSUS.G	AgATAtcTCTGgAAAT...GTAACAGAAAATTTTAAcATGTGGAAAATaAcATGGTGGaACAGATGCAT	312
HIVV1525A	---A-----A-----	312
HIVLBV217	---A---A-C-----	312
HIVCAR4067	---GT-----T-----	312
CONSENSUS.O	AATaTccacTa?acaAT...GTGACAGaTaAcTTTaaATATATGGAAAATTAacATgGTaGacCAAATGCAT	306
HIVAM70	---T-GC-----G-----A-----G-----	315
HIVMVP5180	---T-----GG-----G-----G-----T-----C-----	318
HIVVAU	---GAG---A-A---G---A-C-----T---A-----C-----	315
CONSENSUS.CPZ	7AGT7??CT??TAAAT...77777AG??T??TTTAA??T??T??AAAAAT?ATATGGT?GA??TATG???	148
SIVCP2GAB	A---ATTT---TCCA---GTAAT---AA-CA---CATG-GG---A---G---CCAA---CAT	312
CONSENSUS.U	AaaTA?cTPTgGaaAAT...GTAACAGAAAATTTTAAcATGTGGAAAATaAcATGGTGGaACAGATGCAT	301
HIVZAM184	---G-----G-----G---T-----	312
HIVCAR4081	---C---A-----G-----	315
HIVKI24A	---CA-----G-----	309
HIV23	---GC---G---G---C-----G---A-----G-----	312

HIV1 ENV

LAI/IIIB		<-tev or tnv in	
		\3'sj in LAI/IIIB	
CONSENSUS . A	gaaGATaTaAtcAGtcTATGGGAccAaAGcCTAAAGCCATGTgTaaAgtTaaCCCCCTCTCTGcGTtACtTT	373	
HIVSF1703	AC-----G-T-----	383	
HIVU455	-----G-----	380	
HIVZ321	--G--G-----T-----T-----T-----C-----	383	
HIVRW020	AC-----T-G-----A-----A-----	379	
HIVUG031	AC-----T-----G-----	382	
HIVD687	-----T-----C-----	60	
HIVNI	-----T-----	380	
HIVRW009	AC-----C-----	379	
HIVUG275A	-----T-----C-----	383	
HIVUG273A	AC-----G-----G-----	383	
HIVV1191A	-----G-----	380	
HIVDJ264A	-----T-----G-----	380	
HIVDJ263A	-----T-----T-----	380	
HIVDJ258A	-----T-----	380	
HIVCARGAN	AC-----C-----C-----C-----C-----	383	
HIVCAR5AS	AC-----C-----C-----C-----T-----	383	
HIVCAR4054	--G-----T-----TG-----T-----A-----AC-----T-----	383	
HIVCAR286A	--G-----T-----TG-----A-----G-----T-----	383	
HIVCAR4023	--G-----T-----CT-----TG-----A-----C-----T-----	389	
HIVCAR423A	--G-----T-----TG-----C-----T-----	386	
HIVKENYA	AC-----T-----T-----C-----	383	
CONSENSUS . B	GAAGATaTaAtcAgttTATGGGAtcAAAGcCTAAAGCCATGTGTaAaatTaaCCcCaCtCtGTGTtActtT	389	
HIVJRC5F	-----G-----A-----T-----	383	
HIVJRFL	-----	383	
HIVALA1	-----	383	
HIVBRVA	-----	386	
HIVJH3	-----	383	
HIVBAL1	-----	383	
HIVYU2	-----T-----	383	
HIVMN	-----	383	
HIVHKB2R	-----G-----	386	
HIVLAI	-----G-----	386	
HIVNL43	-----G-----	386	
HIVMFA	-----A-----	386	
HIVCAM1	-----	383	
HIVNY5CG	-----A-----C-----	383	
HIVADA	-----	383	
HIVJFL	-----	383	
HIVSIM184	-----	383	
HIVD31	-----	383	
HIVSF162	-----T-----G-----C-----	383	
HIVBCSG3C	-----C-----	383	
HIVQY1	--A-----	383	
HIVSF33	-----G-----C-----	383	
HIVCDC4	-----C-----C-----	389	
HIVSF2	-----	383	
HIVSF2B13	-----	383	
HIVHAN	--A-----	383	
HIVRF	-----	383	
HIVWMJ2	-----A-----G-----T-----	383	
HIVTB132	-----A-----G-----T-----	383	
HIVBR020	-----T-----G-----	379	
HIVTH014	--A-----	379	
HIVTH026	--A-----	379	
HIVUS711	-----G-----T-----	379	
HIVUS712	--A-----T-----T-----	379	
HIVUS715	-----T-----C-----	379	
HIVUS716	--A-----T-----	376	
HIVHA593	-----	379	
HIVHA594	-----A-----	379	
HIVHA596	-----A-----G-----	379	
HIVHA599	-----C-A-----A-----G-----	379	
HIVHA651	-----	376	
HIVRJS	-----	383	
HIVGUN	-----G-----TT-----G-----	298	
HIVSC	-----	383	
HIVBR014	-----	385	
HIVBR021	-----C-----	386	
HIVACH9	-----	263	
HIVAC91	-----C-----	263	
HIVJM	-----	263	
HIVWM	-----A-----A-----	263	
HIVMA208	-----G-----A-----	407	
HIVP896	-----G-----	383	
HIVBWB11A	-----	2	

380

HIV1 ENV

CONSENSUS.A	aaacTGtaTc????????????????????????????????aa??ccac????????????????????????	368
HIVSF1703	-----GCCAT.....--CA---CACCCAT.....	411
HIVU455	-G-----CCA.....--CA---CATT.....	402
HIVZ321	-G-----CCA.....	393
HIVRW020	-G-----A.....GCCAC-B-CTCC.....	401
HIVUG031	---C---CC.....-CTG--GAC.....	401
HIVD687	G-----G.....G-CG---C.....	79
HIVNI	-G-----CA-AATKTC.....-CAA--GCTACAGC.....	411
HIVRW009	-G-G---AC.....-CA---CAATGTC.....	404
HIVUG275A	-----G-GATGTCAAT.....-TG---TGTCACC.....	417
HIVUG273A	-----CCAATGCC.....-CC---CAATGCC.....	414
HIVVI191A	-G-A---CA.....	390
HIVDJ264A	-G-----CA.....-C.....	393
HIVDJ263A	-G-----CA-AATGTC.....-CAG--GC.....	405
HIVDJ258A	-G-G---CA.....-TG--GCAATAGC.....	406
HIVCARGAN	-G-A---CCA.....-ACC.....	399
HIVCAR5AS	-G-C---C.....-TG-AG-T.....	402
HIVCAR4054	---C---C-GATGTAAAGAAG.....-TAG---T.....	414
HIVCAR286A	---C---C-GATGCAAGATAGACTGTACTAATGGAAGG-TAA--GTGCTGGAGGACAGACCTGTACTG	454
HIVCAR4023	---C---C-GATGTAAAA.....-TA---TGTC.....	420
HIVCAR423A	---C---C.....-TG-A--TGTAACC.....	411
HIVKENYA	---GC.....-TG--.....	399
CONSENSUS.B	AaatTgcactgaTtgg??g??	406
HIVJRCSF	-----AA--G.....	399
HIVJRFL	---G---G.....	399
HIVALA1	-C-----T---ATTT.....	402
HIVERVA	-----CA-----C.....	402
HIVJH3	-----T---G-GG.....AATGATACTAGCCCC.....	417
HIVBAL1	-----AG.....AATGCTACTAATGGG.....	417
HIVYU2	-----AAG.....	402
HIVMN	-----AG.....AATACTACTAATACCAAT.....	420
HIVMXB2R	---G-----AA.....	405
HIVLA1	---G-----GG.....AATGCTACTAATACC.....	420
HIVNL43	---G-----AA.....	405
HIVMFA	---G-----AA.....	405
HIVCAM1	--T-----A-G-A.....	399
HIVNY5CG	-----AC.....	402
HIVADA	-----AG.....	402
HIVJFL	-----T-A-----AA.....	402
HIVSIMB4	-----A--C-AA-AGTGCT.....	408
HIVD31	-----C-AA.....	402
HIVSF162	-C-----A-AA.....	402
HIVBCSG3C	-----GA-TT-AAG.....	405
HIVVOYI	-G-----G-T.....AATACCCTAGTAGTAGTTTGAGG.....	423
HIVSP33	-----ATTT-GGG.....	405
HIVCDC4	-----AATACTAAT.....	414
HIVSF2	-----GG.....	402
HIVSF2B13	-----GG.....	402
HIVHAN	---A-----ATAAT.....	402
HIVRF	-----GCTAACTTG.....	405
HIVAMJ2	-----T-AA.....	399
HIVTB132	-----A---AG.....	402
HIVBR020	-----T-----AATAATAGTACTAAC.....	410
HIVTH014	---C-----AG.....	398
HIVTH026	-----AG.....	398
HIVUS711	-----AA-TT.....	398
HIVUS712	-----C-AGT.....	398
HIVUS715	-----CA-C-AG.....	398
HIVUS716	---G-T-----C-GG.....	395
HIVHA593	-G-G-TCA---G-AATGTGAATGGCACT.....GCTAATAATGGCACTACTAATGTCACTGAG.....	440
HIVHA594	-----G-GA.....	398
HIVHA596	-----G-AA.....	398
HIVHA599	-C-----GC-AG.....	398
HIVHA651	-----C-GCAGA.....	395
HIVRJS	-----A--O-AA.....	402
HIVGUM	-----A---GT.....	317
HIVSC	-----A---AG.....AATGATACTAGCACC.....	417
HIVBR014	-----G.....	401
HIVBR021	---A-G-T.....AATACCAATAGTACT.....	417
HIVACH9	-----AATTT.....	282
HIVACP1	-----GATTT-AGG.....	285
HIVJM	-----A---GCCTC-AGG.....AATGTACTAAGAGG.....	300
HIVMM	-----GTGTAACCTCA.....	288
HIVHA208	-----GCCACTATTAGA.....ACTAATACC.....	441
HIVP896	-----A.....AATATCACTAAG.....	411
HIVBWB11A	-----A-GATTT-GGG.....GATGCTACTAATACC.....	39

CONSENSUS.C	AaaTGTGcaga????????????	aatgcTca????????????	396
HIVMA959	---T---C---	---C---AC---	398
HIVMA960	---T---GCT---	---	389
HIVD757	---T---C---	---CA---	76
HIVD747	-C-T---C---	---C---CCTA-AGT---	214
HIVD750	-G-G---G---AATGTT---	---C---CC---	216
HIVNOF	---T---C---GATGTPAAAAGTT---	---	50
HIVZAM20A	-G-A---G---	---TG---CTGT---CCCAAGGAT---	414
HIVD1044	-G-A---G---CAGGCT---	---TT---CC---	46
HIVD1034	-G-A---G---AAGGTT---	---C---CC---	408
HIVD744	---A---GAT---	---C---CCAG---	78
HIVD766	-G-A---G---	---T---CCAG---AATGCT---	39
HIVD808	-G-A---G---	---T---GCAG---AATGGT---	90
HIVD868	-G-A---G---GAGGTT---	---A---CTAGAATGGGGTT---	90
HIVBR025	-C-C---C---T---	---CAGA---CTAT---GACTACAAT---	324
HIVSM145A	---T---C---GAA---	---GC---AA---TT---	405
HIVZAM18A	-G-A---G---	---T---	402
HIVDJ259A	---C---T---	---T---CAGAG---	399
HIVDJ373A	-C---T---	---AAGGTTAC---GCTACCGGT---	414
HIVSE364A	---C---G---	---	393
HIVUG268A	---T---C---	---T---ATGT---	405
CONSENSUS.D	aaactGcactgatg?g????????	aa????act????????????	399
HIVJY1	---T---A---CTGG---GGG---	---	405
HIVNDK	---A---AATT---AGG---	---	405
HIVMAL	---A---T---	---TGGG---GCTGTG---	414
HIVELI	---T-G---A---AATT---AGGAAC---	---	408
HIVZ226	---TA---AAGT---ATGGAG---	---	408
HIVUG024	---T---ATG---AA---AATGAA---	---TGAA---AAGGGG---	426
HIVUG021	---ATG---AA---	---	398
HIVUG038	-C-----T---	---TAGT---	405
HIVUG046	-C-T-----ATATAA---GCCCT---	---TGCC---ATT---	438
HIVUG266A	---G---A---G---AG---GTT---	---CAGAT---A---	414
HIVUG269A	---T---A---CCAAATTC---	---TGTC---ACAGCCACTAATGHAAG---	431
HIVUG274A	---ATG---AA---	---	402
HIVSE365A	---GG---ATTCTAGT---	---G---CGCC---AGTAACACCATTAAGT---	429
HIVCAR4020	GC-----CGATT---GA---AATATTGCA---	---CAGT---GCCTTACAAAC---	439
CONSENSUS.E	AaatTGTAccaatgct????????	????????????????	398
HIVTH966	---T---	---AATTTGACCAATGTCAATAACATA---ATC---	422
HIVTH975	---T---CATTGACCAAAGTC---	---AACATAACC---	413
HIVTH032	---	---AATTTGATCAATGTCAAT---	413
HIVTN243	---T-----AAGTTGACCAATGCT---	---AATTTGACCAATGTCAAT---	432
HIVTH006	---AAGTGACCAAT---	---GTCAGT---	413
HIVTH011	---T---N-----NT---	---AATTTGACCAATGTCAAG---	341
HIVCARELO	---GG---	---	399
HIVCAR4017	---G---	---	402
HIVCAR4039	-C-----GG---	---	399
HIVCAR4071	-G-----A---GATTTT---	---	405
HIVCARMBR	---GG---AATCTT---	---	402
CONSENSUS.F	AaATtGTACTAaTGCcact?cT	AATGCCACT	407
HIVBR47944	---G---G---	---	324
HIVBZ163A	---AT---	---	408
HIVBZ126A	-C-----G---C---	---	411
CONSENSUS.G	AAACTGTACTAa?Gtaa?C?ac	AAT?g?Acta?g?C???	412
HIVV1525A	---G---C-T---	---A-C---GCA-C-AAC---	423
HIVLBV217	---T-C-T---GCT---	---GTA---C-ACAG---	420
HIVCAR4067	---G-A---G-A-A---	---T-T---GG-A---	420
CONSENSUS.O	GaA?TGTacagat?c?aa????	AACA?aaCa????????c?ta????	404
HIVANT70	-G-G---	---C-T-G-T---	405
HIVMWP5180	---C---GT---C-GC---CA---	---A---GG-C---TTA---	429
HIVVAU	---T---A-CA---ATAGTATT---	---CC---AACAGTCC-T---AACTCA---	441
CONSENSUS.CP2	G?A?TG?A? ?????????		196
SIVCP2GAB	-C-G---C-GTAAGGCTAACCTTAGCCAGGCCAAA		417
CONSENSUS.U	AAA?TGTAgTgatgtc????????	AATagcActaa????	395
HIVZAM184	---T-----C-CTAGTACCAATAGC---	---C-C---	423
HIVCAR4081	---CC---C---	---TA---C---A---	408
HIVK124A	---T---A---	---G---GGGTGGC---	411
HIVZJ	---C---T---AAG---	---T---C---	411

I-A-70
DEC 94

CONSENSUS -A	?????????aacac?ac?a? ??????????????????????aa?a?ac? ????????	400
HIVSF1703	---T---C-CCAACAACCCC	441
HIVU455	---A---C-AT	429
HIVZ321	---T---C-TCAGGAC	426
HIVRW020	---TGT---C-AT	413
HIVUG031	---C---G-AT	425
HIVD687	---A---CTTTAGC	103
HIVNI	---GT-GC---	429
HIVRW009	---A---TGTC	425
HIVUG275A	---TCAT-GT-TT---	429
HIVUG273A	---T-AA-ATGCCAAGGTCAACAGCGCCACG-TGGC-C-	453
HIVVI191A	---G---C-CCCCAGAT	420
HIVDJ264A	---GT---CTCTAAC	420
HIVDJ263A	---T-G---GC-CCTCTATC	435
HIVDJ258A	---T-C-T-T-GCTACAGG	432
HIVCARGAN	---C---AGATAGCACCAAA	441
HIVCAR9AS	---GA---GT-CA	414
HIVCAR4054		414
HIVCAR286A	GAGTACAG---T-G---TGCT	474
HIVCAR4023	---T-GT---T-GT---ACTGGGAATAGTGCT---T-CA-TA	456
HIVCAR423A	---G---GT-AT	438
HIVKENYA	---TGT---TGCCACG	426
CONSENSUS -B	...aatg?act? ??? ??????????????????aata?ta?k? ??????????????????gg	424
HIVJRC5F	---C---C	429
HIVJRFL	---C---C	429
HIVALA1	---C---C	432
HIVBRVA	---C---C	432
HIVJH3	---C---C	447
HIVBAL1	---AC---C	447
HIVYU2	---C---C	432
HIVNM	---AG---GC-AAT	453
HIVHXB2R	---A---AA-ACC	435
HIVLAI	---AG---G-AA-ACC	450
HIVNL43	---A---AA-ACC	435
HIVMFA	---A---AA-ACC	435
HIVCAM1	---AA---AGGACC	438
HIVNYSOG	---C---TA-GCC	426
HIVADA	---T---AA-ATC	432
HIVJFL	---T---AA-ATC	429
HIVSIMI84	---AG---CAG-AATACC	435
HIVD31	---C---AA-ACCAAT	438
HIVSF162	---C---C	429
HIVBCSG3C	---C---C	447
HIVQY1	---C---C	453
HIVSF33	---C---AA-ACC	435
HIVCDC4	---AC---C	456
HIVSF2	---G-C---AA-ACC	432
HIVSF2B13	---G-C---AACC	429
HIVHAN	---CC---C	423
HIVRF	---G---GTC-C-AGTAGTAGCGGG	438
HIVNMJ2	---CAT---GA-TGGAAG	429
HIVTB132	---A---A	411
HIVBR020	---CAA-A-	434
HIVTH014	---C---AA-ACC	422
HIVTH026	---C---C	422
HIVUS711	---T-AA-G-ACCTATAGCAAT	428
HIVUS712	---C---AA-ACGAAT	446
HIVUS715	---A---AA-ACCAAG	446
HIVUS716	---C---C	422
HIVHA593	---G-G-GTCAA-AGTAGTAGTCACTAGT	494
HIVHA594	---AG---AA-GCC	428
HIVHA596	---AG---AA-GTC	428
HIVHA599	---AG---AA-GTC	425
HIVHA651	---GGAA-GA-AC-GGG	422
HIVRJS	---A---GA-ACT	441
HIVGUN	---C---TA-AGT	344
HIVSC	---C---C	447
HIVBR014	---T---AA-ACTAAT	449
HIVBR021	---ACC---AA-TTGAGG	453
HIVACH9	---AG---AA-TTGGGG	309
HIVACP1	---A---A	294
HIVJM	---A---AA-GAGAACTCTACTGATAGT	360
HIVMM	---TGC---AA-TTGAGG	336
HIVMA208	---AC---AA-ACT	468
HIVP896	---AC---C	441
HIVBMB11A	---AG-G-	51

CONSENSUS.C	AAtg?taaccaataat?????????????aat???ac????????????????????	415
HIVMA959	CAC-G-T-CC-----GTA-TGCTACCAATAATGTAACC.....	440
HIVMA960	T-T-TC-G-----GAG-CTAT.....	416
HIVD757	T-T-G-AATGGT.....CTAC-T.....	106
HIVD747	AG-T-C-----AGT-CTACAAT.....	241
HIVD760	AT-----GGGGAAAT.....CCT-CAATATTACCAATAATAGG.....	267
HIVN0F	C-G-CAGTAAT.....CATG-T.....	80
HIVZAM20A	AG-T-C-GGTGATATCAATAGT-ACT-CATGAAA.....	459
HIVD1044	C-----GC-AAC-----GCT-G.....	73
HIVD1024	AT-----GGG.....	426
HIVD744	C-G-G-GGG.....	96
HIVD766	G-T-C-GAGACCTAC-----GAG-GCGTG.....	75
HIVD808	G-T-C-GAGACCTAC-----GAG-GCGTG.....	126
HIVD869	T-----GGG.....	108
HIVBR025	CAGA-TG-----CATGGGG.....	345
HIVSM145A	AT-----G-GT-GGC.....GTT-CAGTAATGACAATAAG.....	447
HIVZAM18A	T-C-G-G-----AGC-GAAG.....	429
HIVDJ259A	C-T-CG-G-----AGC-AAAA.....	426
HIVDJ373A	G-GTGT-CCTAC-----AGC-AGAA.....	444
HIVSE364A	T-----CC-GCAAT-----CCC-CAGT...AATGATAACCTAAT.....	438
HIVUG268A	AT-----GCT-----GCT-CAATAGCCCTACGAAATGGGAAA.....	456
CONSENSUS.D	AAt?cacta?ta?????????aat???act????????????????????	416
HIVJY1	AAG-C-ATGGGAAT.....ACT--AATCAGGAGGAACAGATGATGGAG.....	459
HIVNDR	CAG-AGGGC.....GG-AGGTAGAAGAGGAG.....	438
HIVMAL	GGG-A-GCTGGGAGT.....AGG--AATGCAGAAATGAAAATGGAAATT.....	468
HIVELI	GG-TGGGGAAC-----TC-ACAGAGGAG.....	444
HIVZ226	GT-A-TG-AGATAAT-----TC-AGAGGAA.....	444
HIVUG024	AAA-GTGGACCCCT-----AT-AGGAC.....	453
HIVUG021	GC-C-C-----CC-----	419
HIVUG038	ATA-G-CTAATGTG-----AAT-----	435
HIVUG046	GC-GA-GGGACATA-----GC-----	459
HIVUG266A	GG-T-GGA-----GC-----	438
HIVUG269A	CACTGT-GA-----CC-----	452
HIVUG274A	AC-GAACAGGAGAAG-----ACCG-----	432
HIVSE365A	CGT-GGA-TACCG.....	447
HIVCAR4020	GC-G-GCA-GAAGT-----AT-AGACAATAACCATACA.....	471
CONSENSUS.E	aat?t?accaatg?????????????aata??ac?????????????????7??7??7??7?	415
HIVTH966	C-GGCC-T-C-----ATAATAGGA.....	443
HIVTH975	TCT-TAA-----TTCT-TAA-----ATAGGA.....	431
HIVTH022	CA-A-----TCTTAACATAATAGGA-TA-A.....	452
HIVTH243	CA-A-----TCTTAACATAATAGGA-TA-A.....	471
HIVTH006	ACA-T-----CCCTGAC...ATAGGA-TA-A.....	449
HIVTH011	CA-A-----TCCCTAACAAAATAGGA-TA-A.....	380
HIVCARELO	T-T-----GCT-CGATGCT.....	423
HIVCAR4017	T-P-G-CT-----GT-CGAGGCTGCTTTG.....	438
HIVCAR4039	T-T-----TCT-TAT-CAAGGCT.....	429
HIVCAR4071	CACC-A-ATTT-----C-CA-AGAAAAG.....	435
HIVCARMBA	TCF-----CT-----TCT-C-----AAT.....	429
CONSENSUS.F	??????...	407
HIVBRA7944	ACCACC.....AATGACACT.....	339
HIVB2163A		408
HIVB2126A		411
CONSENSUS.G	AAtAa?ac?act??????...	426
HIVVI525A	GAT-AT-CACTGTC.....C-C-CCT...ACTAGTACTGTAAACC.....	465
HIVLBV217	G-T-G-----AAC-A-G-AGCCTTGGG.....	450
HIVCAR4067	C-T-GC-GTG.....	435
CONSENSUS.O	aa?ga?ACaa?a.....	413
HIVANT70	GA-C-----	414
HIVMVP5180	T-G-T-----	441
HIVVAU	CA-T-----	450
CONSENSUS.CPZ	AA2??AACA7?????????????????A?????ACA?????C????...???	207
SIVCPZGAB	CCT-----ACCAG--TCTT-T.....	441
CONSENSUS.U	aatagcactgaag?a???...	409
HIVZAM184	T-----T-----AACAAACAATAGCACTGTG.....	423
HIVCAR4081	C-----ACG-GTAC-----AACATCACT.....	441
HIVK124A		438
HIVZ3	A-----A-----	429

HIV1 ENV

CONSENSUS.A	?????gga?atg??a.....????????????????????gg	409
HIVSF1703	...AAAT-C---GA.....	455
HIVU455	ACAGAT-GAG--AGG.....	446
HIVZ321	GATACT--A---AA.....	443
HIVR020		413
HIVUG031	...AAG--C---AT.....	439
HIVD687	...TAT--C---AA.....	117
HIVNI	...AGT--C---AAT.....	443
HIVR009	...GAT--C---AA.....	439
HIVUG275A	AGCAGT--C---AA.....	446
HIVUG273A	ATCCAGAGC---GA.....	470
HIVV1191A	CAGTCCACT--CAT.....	437
HIVDJ264A	TCTAGG--C---CA.....	437
HIVDJ263A	GATAGTA-A---CA.....	452
HIVDJ258A	...ACT--A---CG.....	446
HIVCARGAN	ACCTTT--AGG-AT.....	458
HIVCAR5AS		416
HIVCAR4054	CTGGAA--	422
HIVCAR286A	GAAGTT--	482
HIVCAR4023	AATAGAACT.....GTGGATA-	473
HIVCAR423A	AACAGTAGC.....CAATCCTCTAAGTATAACATTGAT--	473
HIVKENYA	CACGGG-GAGC.....	440
CONSENSUS.B	g7aa?g???atggagaaa???.....gg	439
HIVJRCSP	-G--T-----G-----	446
HIVJREFL	-G--C-----G-----	446
HIVALA1	AC--A-----G-----	449
HIVBRVA	AA--T-----GG-----	449
HIVJH3	-A--A-----	464
HIVBAL1	-G--T...G---G-GG-----	464
HIVYU2	-A--C-----	449
HIVMN	-G--CA...--AA--GG-----	470
HIVHKB2R	AG--T-ATA-----	455
HIVLA1	-A--T-ATG-----	470
HIVNL43	AG--T-ATA-----	455
HIVMFA	AG--T-ATA-----	455
HIVCAM1	-A--GGAGAA-A-AT-----	458
HIVNY5CG	AGTGA-----G-----	440
HIVADA	-G--T-AGA-----	443
HIVJFL	-G--A-----G-----	446
HIVSIM184	TTGGA-----	449
HIVD31	AC--GGAGAA-----	455
HIVSF162	TGG--AAGAG-----C-G-----	449
HIVBOS03C	TGG--AAAAG-----A-G-----	467
HIVQY1	-A--C-----	470
HIVSF13	-G--C...G-----	452
HIVCDC4	-A--CAA...CG--GT-----	473
HIVSF2	AA--GAA...GAAATA-----	449
HIVSF2B13	AA--GAA...GAAATA-----	446
HIVHAN	-G--G-----	440
HIVRP	AC--T-----C-----	455
HIVM4J2	A--GAGG-----	440
HIVTB132		413
HIVBR020	AC--GG-----A--C--G-----	451
HIVTH014	GG-----	433
HIVTH026	GG-----	433
HIVUS711	-A--GGA...GAAATA-----	445
HIVUS712	ATG--T-----C--GG-----	463
HIVUS715	AC--T-----G--G-----	463
HIVUS716	AG--T-----AGGG-----	439
HIVHA593	-G--CA-----A--G-----	511
HIVHA594	-G--TA...--A--C-----	445
HIVHA596	-GG--TA...--A--GG-----	445
HIVHA599	-A--A...G--A-----	442
HIVHA651	AGTGGG...--A--T-----	439
HIVRJS	-A--AA-----G-----	458
HIVGUN	-A--A-----C--G-----	361
HIVSC	-G--A-----GG-----	464
HIVBR014	-A--GA-----	466
HIVBR021	-A-----ATAT--G-----	467
HIVACH9	GATATC-----	320
HIVACF1	-A--CA-----	311
HIVJM	ACT--T-----A--GTAT-----	380
HIVMN	-A--TA-----A-----	353
HIVMA208	-A--GT...--A-----G-----	485
HIVP896	-G--T-----	458
HIVBWB11A	-G--T-----A--T-----	71

CONSENSUS.C	????????????	ga	417
HIVMA959		AG	442
HIVMA960		A-	428
HIVD757		--	108
HIVD747		--	243
HIVD760		--	269
HIVNDF		-G	82
HIVZAM20A		--	461
HIVD1044		--	75
HIVD1024		--	438
HIVD744		--	98
HIVD766		A-	77
HIVD808		A-	128
HIVD868		--	110
HIVBR025		-G	347
HIVSM145A		--	449
HIVZAM18A		-G	431
HIVDJ259A		--	428
HIVDJ373A		A-	446
HIVSE364A		-G	440
HIVUG268A	CTCATGGAACAG	-G	470
CONSENSUS.D	a????????????????????????????????????	-G?	418
HIVJY1	-AA	-G	464
HIVNDK		-A	440
HIVMAL		-G	470
HIVELI		AA	446
HIVZ226		AT	446
HIVUG024		AT	455
HIVUG021	-ATGAGGGC	AT	430
HIVUG038	-GCATAGAGAAA	-G	449
HIVUG046		--	459
HIVUG266A	-TT	-G	443
HIVUG269A	-AGGGGGGC	CT	463
HIVUG274A	GTACCGGGGAGTACAACATCTGAC	AC	458
HIVSE365A	-TGATGGGGAAA	-G	461
HIVCAR4020	-ACAGTACCAATACTAGATGGGTATGACGGAAATGGAA	AA	512
CONSENSUS.E	??t??ta????????????????????????????	GA	419
HIVTH966	AA-ATA-CA	--	454
HIVTH975	AA-ATA-CA	--	442
HIVTH022		--	454
HIVTN243		--	473
HIVTH006		--	451
HIVTH011		--	382
HIVCARELO	AA-GTT-CCAATGCTACTGCTAAATTTAATGCTACAAAAGAAATOGAA	A-	473
HIVCAR4017	AA-GGC-GTTACAGTAATGAAACAGCAGAAATOGAA	A-	476
HIVCAR4039	AA-GAA-CTGTTGAAATGGAA	A-	452
HIVCAR4071	CCAGAA-TAGAAATATCAGAAATGCAA	A-	464
HIVCARHBA	GC-AAT-CCCTAAAAATGGAA	A-	452
CONSENSUS.F	gcca?tcAA?ACgE??TadCcTGAAGGAAGaaCCA	GG	439
HIVBR7944	C-A-C.....T.....G.....	--	368
HIVB2163A	T.....G.....GA.....G.....	--	443
HIVB2126A	C.....A.....AACA.....	--	449
CONSENSUS.G	AAtagta?tg?????tag?????	GA	438
HIVVI525A	--C-A--GCACTGTG--TGGCAAA	--	491
HIVLSV217	----C-CT-TGAGT--CATA	--	473
HIVCAR4067	---T--GAG-AA	--	449
CONSENSUS.O		AA	415
HIVANT70		--	416
HIVAMP5180		--	443
HIVVAU		--	452
CONSENSUS.CP2	????????????????P???	?T	209
HIVCP2GAB	CCGCC-	C-	449
CONSENSUS.U	??????????	GA	411
HIVZAMB4		--	425
HIVCAR4081	GATAATAAA	--	452
HIVK124A	AGGGAGGAC	AT	449
HIV23		-C	431

HIV1 ENV

		-> tev or tnv in LAI/IIIB	
	5' sj\		
CONSENSUS.A	ggaA?????ataAaaAACTGctCTTtCAATAtgACCaCagAa?TAAGagATAagaaacagaaagtatatatt		472
HIVSF1703	A-----G-----A-----C-----G-----		520
HIVU455	A-----G-----T-----C-----		511
HIVZ321	A-----T-----T-----A-----T-----C-----AG-----A-----		508
HIVRW020	-----G-G-----A-----T-----AG-----C-----G-----		477
HIVUG031	A-----A-----C-----Q-----A-----		504
HIVD687	G-----GT-----		180
HIVNI	G-----A-----G-----A-----GA-GC-G		508
HIVRW009	A-----T-----G-----A-----G-----		504
HIVUG275A	A-----T-----G-----C-----A-----C-----C-----		511
HIVUG273A	A-----C-----G-----C-----G-----G-----		535
HIVVI191A	T-----C-----C-----A-----A-----T-----GT-----A-----C-----		502
HIVDJ264A	G-----G-----T-----C-----C-----C-----G-C-G		502
HIVDJ263A	G-----C-----C-----G-C-G		517
HIVDJ258A	G-----C-----C-----A-G-C-G		511
HIVCAR0AN	A-----C-----GC-----		523
HIVCAR5AS	T-----C-----G-----A-----A-----C-----AT-----A-----G-----A-----G-----C-----		481
HIVCAR4054	CC--GTAGAA--G-----A-----A-----AT-----G-----GA-----A-----C-----CG		493
HIVCAR286A	CC--GTAGAA--G-----A-----T-----A-----G-----GA-----A-----CG		553
HIVCAR4023	CC--GTAGAA--G-----T-----A-----G-----GA-----A-----CG		544
HIVCAR423A	AA--GAAGAA--T-----C-----A-----T-----GGA--C-----A-----CG		544
HIVKENYA	GAG-----G-----A-----A-----GC-----G-----G-----G-----		505
CONSENSUS.B	aGaa.....aTaaaaAACTGcTCTTtCaAtatcaccacaagcataagagataaggtgcagaaagaataTG		504
HIVJRCSP	-----A-----		511
HIVJRF	-----G-----		511
HIVALA1	-----A-----C-----		514
HIVBRVA	-----G-----A-----		514
HIVJH3	-----G-----C-----		529
HIVBAL1	-----G-----T-----A-----G-----		529
HIVYU2	-----A-----		514
HIVMN	-----G-----A-----		535
HIVHKB2R	-----G-----G-----G-----		520
HIVLAI	-----G-----G-----		535
HIVNL43	-----G-----G-----		520
HIVMFA	-----G-----A-----G-----		520
HIVCAM1	-----G-----G-----A-----G-----		521
HIVNYSOG	-----G-----T-----G-----T-----A-----A-----A-----		505
HIVADA	-----A-----A-----C-----		508
HIVJFL	-----C-----A-----A-----A-----G-----		511
HIVSIM184	-----G-----T-----A-----A-----C-----		514
HIVD31	-----		520
HIVSF162	-----T-----GG-----A-----A-----		514
HIVBCSG3C	-----G-----G-----T-----G-----AA-----C-----		523
HIVOTI	T-----C-----A-----G-----C-----		535
HIVSP33	-----G-----G-----G-----A-----G-----C-----		517
HIVCDC4	-----G-G-----T-----A-----T-----A-----		538
HIVSP2	-----A-----T-----A-----		514
HIVSF2B13	-----T-----A-----T-----A-----		511
HIVHAM	-----C-----GG-----A-----A-----C-----		505
HIVRF	-----C-----AG-----T-----T-----G-----AC-----A-----		520
HIVWJ2	-----G-----C-----G-----G-----T-----		505
HIVTB132	T-----G-----G-----G-----T-----G-----T-----		478
HIVBR020	-----G-----C-----TG-----A-----A-----G-----		516
HIVTH014	-----G-----G-----A-----AC-----AA-----A-----C-----		495
HIVTH026	-----G-----CAC-----C-----		495
HIVUS711	-----C-----A-----A-----		510
HIVUS712	-----T-----G-----G-----A-----G-----		528
HIVUS715	-----T-----A-----A-----T-----		528
HIVUS716	-----T-----A-----		504
HIVHAS93	-----A-----A-----AC-----		576
HIVHAS94	G-----G-----A-----A-----ACAAC-----GC-----		510
HIVHAS96	G-----C-----T-----G-----TG-----A-----ACAAC-----GC-----		510
HIVHAS99	-----C-----A-----AT-----A-----G-----A-----A-----		507
HIVMA651	-----C-----C-----G-----T-----A-----AC-----C-----C-----		504
HIVRJS	-----T-----A-----T-----G-----		523
HIVGUN	-----C-----A-----A-----A-----C-----		426
HIVSC	-----G-C-----AG-----A-----		529
HIVBR014	-----G-----G-----A-----A-----		531
HIVBR021	-----A-----A-----A-----T-----		532
HIVACH9	-----G-G-----G-----A-----		385
HIVACP1	-----T-----A-----A-----A-----A-----C-----		376
HIVJM	-----A-----AT-----GG-----GA-----A-----		445
HIVMM	-----T-----A-----AAG-----		418
HIVMA208	-----T-----A-----A-----G-----A-----G-----C-----		550
HIVP896	-----T-----T-----A-----A-----AA-----		523
HIVBWB11A	-----A-----A-----A-----		136

			-> tev or tnv in LAI/IIIB	
			5' sij\	
CONSENSUS.C	agaa.....aTaAAAAATTGctCTTTcaAtgcAACcACaGaATTAagaGATaagaacacagaaagtgatG	481		
HIVMA959	T-C.....-G-----T-----T-----G-----G-----	507		
HIVMA960	-G.....G-----T-----A-----A-----A-G-T-----	483		
HIVD757	-----A-----T-----C-----C-----TGC-----	173		
HIVD747	-----A-----T-G-C-----C-----AC-A-----	308		
HIVD760	-----C-----A-----G-C-G-----	334		
HIVNOF	-----T-A-----A-G-----	147		
HIVZAM20A	-G-G-----T-A-----AT-----C-AGT-G-C-----	526		
HIVD1044	-G-----A-----A-----G-----G-----C-----	140		
HIVD1024	-----T-----A-----G-----G-----C-----	493		
HIVD744	-G-G-----A-----G-----G-----C-----	163		
HIVD766	G-G-----T-----G-----G-----G-----	142		
HIVD808	G-----G-----G-----G-----C-----	193		
HIVD868	-----A-----G-----G-----C-----	175		
HIVBR025	-----T-----ATG-----G-----G-----G-AC-----	412		
HIVSM145A	-----T-----T-----T-----	514		
HIVZAM18A	G-----G-----C-----A-----A-----G-----	496		
HIVDJ259A	-----G-----T-----T-----C-----G-----CA-----	493		
HIVDJ373A	-C-----G-G-----CAT-----C-----AAA-----	511		
HIVSE364A	-----AT-----C-----G-AG-----T-AA-----	505		
HIVUG268A	-----T-----A-----A-----C-C-C-----	535		
CONSENSUS.D	agaa.....atGaaAACTGCTcTTTCaatataaccacagttagtaAgaGATAagaagaagcagtagtAtaTg	483		
HIVJY1	-----G-----C-----A-----C-----	529		
HIVNDK	-A-----G-----G-----G-----G-G-----	493		
HIVMAL	-G-----G-----C-----G-----T-----A-G-----A-----	532		
HIVELI	-G-----C-----C-----A-----C-----C-----	511		
HIVZ226	-AG-----T-----C-----CA-A-----C-----	511		
HIVUG024	-G-----G-----A-----A-----	520		
HIVUG021	-G-----A-----C-----C-----	486		
HIVUG038	-----G-----A-----C-----G-A-A-----	514		
HIVUG046	-G-----G-----AA-----T-----A-C-----	523		
HIVUG266A	G-C-----G-----A-A-----A-A-----CG-----	508		
HIVUG269A	-----G-----T-A-----G-----A-A-----	528		
HIVUG374A	-C-----G-----AGT-----G-----	523		
HIVSE365A	-----C-----T-----A-A-----A-GAC-A-T-----T-----	526		
HIVCAR4020	-----C-----G-----A-----A-----C-----	577		
CONSENSUS.E	tgaa.....GTAAgaAACTGTCTTTTAAATaTgACCACagaAcTaagAGATAagaagcAGaAGBTCCaTG	484		
HIVTH966	-----A-----	519		
HIVTH975	-----A-----T-----	507		
HIVTH022	-----A-----A-----AC-----	519		
HIVTN243	-----	538		
HIVTH006	-----	516		
HIVTH011	-G-----A-----T-----A-----G-----	447		
HIVCARELO	-AC-----A-----G-----GO-----C-T-----TC-----	538		
HIVCAR4017	-----G-----G-----GA-----T-----	541		
HIVCAR4039	-----T-----A-----A-----C-A-G-----C-T-----	517		
HIVCAR4071	A-----C-----A-----G-----G-----T-----	529		
HIVCARMBA	-----A-----TC-----	517		
CONSENSUS.P	GGcA.....ATACAAAACGTCTCTTTCAATaTGaCCaCaGAAGTA?GAgATAAGCAGCaGAAAGTACATg	503		
HIVBRA7944	-A-----T-----A-----T-----T-----	433		
HIVBZ163A	-----G-----TG-----	507		
HIVBZ126A	-----T-----A-----	513		
CONSENSUS.G	AgAA.....ATGAaAAACTGCTCTTTCAATaTAACCACAGAA?TAAGAGATAAGag?AAGacAGAATACG	501		
HIVV1525A	-----C-----G-----C-----CAG-G-----	556		
HIVLBV217	-----G-----C-----T-----A-----	538		
HIVCAR4067	-A-----A-----A-----	514		
CONSENSUS.O	cGag?????aTGaAaaAgTGTp?TTTTAATGTAACTACaGT?cTCAaAGACAAAaAGGAGaAAAAACAGG	478		
HIVANT70	---AAACCTT---G-----G-----C-----T-----TA-----	487		
HIVMVP6180	-----G-----T-----AGT-----C-----C-----C-----	508		
HIVVAU	A-----G-----C-----C-----G-----C-----	517		
CONSENSUS.CPZ	??A?.....?T?TA?A?A?TGT???TTTAAAT?TGA?ACAG?ATT?A?AGNTAAGAA?AA?A?T?T?A??	247		
SIVCPZGAB	CG-A.....A-GA-A-C---TCA-----GT---CA-----A-G-----GA---C-GG-CT-TT	514		
CONSENSUS.U	agaa.....ATaAAAAAT?TGCTCTCTTCAa?ataaCCACAGAACTaA?AGATAaGA?GaaGacAGTatAtg	472		
HIVZAM184	-----C-----A-----GG-CT-----A-----C-CT-G---C---T	490		
HIVCAR4081	-----T-----T-----A-----G-----A-----C-----C-----	517		
HIVK124A	-G-----G-----C-----T-----G-----G-----G-A---CA-----	514		
HIVZ3	TACC.....-C-C---T---C---GG-C-----G-A-----C-G-----C-A	496		

HIV1 ENV

CONSENSUS .A	CaCTtTtTATAgACTtGATgTAGTaCaAAttaac????????????????aata??atc?????	513
HIVSF1703	-----A-----C-----G-----AAA-----ATGG-----	557
HIVU455	-----A-----C-----C-----AA-C-GAT...	555
HIVZ321	---G-----A-----C-----GGGGA...	546
HIVRW020	-----AG-----C-----GAAAAG...	518
HIVUG031	-----C-----G-----GAA-G-----	545
HIVD687	-----G-----G-----GAA-A-----	214
HIVN1	-----G-----A-----AGC-----GT-A-----	549
HIVRW009	-----G-----C-----GGC-GCAAC...	551
HIVUG275A	-----A-----A-----GT-G-GAG...	555
HIVUG273A	-----A-----GGAAT...	582
HIVV1191A	-----GGA-C-AGTAAT	558
HIVDJ264A	-----GAA-G-AGC...	546
HIVDJ263A	-----GAA-G-AGCAAT	564
HIVDJ258A	-----G-GAAAAT...	564
HIVCARGAN	-----A-----AG-GGAAT...	576
HIVCAR5AS	-----A-----A-----C-C-----AAAATTAATGATGCTAGTGCT--TT-G-GAGAAC	552
HIVCAR4054	---G---C-----AGTAATAGT	549
HIVCAR286A	---G---C-----C-CA--AATGATGAT	609
HIVCAR4023	---C---C-----A-----C-----AATAATAAT	597
HIVCAR423A	---G---C-----A-----C-----AATAAT	600
HIVKENYA	-----C-----GAAAAT	561
CONSENSUS .B	CacktttTtTaTaaacttgATgTagtaccAatAgat????????????????????????????	539
HIVJRC5F	---T---G-----AATAAG	552
HIVJHFL	---T---A-----AAT	549
HIVALA1	---G---C-----AAT	552
HIVBRVA	-----A-----AATGATAAACT	561
HIVJH3	-----A-----A-----AATAGTACCAAGGAT	579
HIVBAL1	-----A-----C-----	564
HIVYU2	-----C-----	549
HIVMN	---C---C-----A-----T-----	570
HIVHXB2R	---T---A-----A-----	555
HIVLAI	---T---A-----A-----	570
HIVNL43	---T---C-----A-----	555
HIVMFA	---T---C-----A-----AG-----	555
HIVCAM1	-----AAGGCT	564
HIVNY5CG	---G---A-----AAGGAT	546
HIVADA	---G---A-----AATGAT	549
HIVJFL	-----AAGGAT	552
KIVSINI84	-----A-G-----GGATGAT	555
HIVD31	---AC-----AATGAC	561
HIVSF162	-----AATGAT	555
KIVBCS33C	---C-----AATGATAAG	567
HIVQY1	-----T-----AAGAAT	576
HIVSF33	---TA-----GATGAT	564
HIVCDC4	---T---G-----A-----GATAATAAA	591
HIVSF2	---CG---C-----AATGCTAGTACTACT	564
HIVSF2B13	---CG---C-----AATGCTAGTACTACT	561
HIVHAN	-----AC-----AAT	558
KIVRF	-----G-----GAAGGCT	576
HIVWJ2	-----A-GGCTGAT	549
HIVTB132	---C-----A-----A-----AATGAT	519
HIVBR020	---T---C---C---T-----A-----G-----	551
HIVTH014	-----AAT	533
HIVTH026	-----TC-----	530
HIVUS711	-----GC---A-----A-----G-----	545
HIVUS712	-----A-----A-----A-----	563
HIVUS715	-----A-----	563
HIVUS716	---G---G-----A-----A-----C-----	539
HIVHA593	---AG-----A-----G-----	611
HIVHA594	-----G---G-----	545
HIVHA596	-----G---G-----	545
HIVHA599	---C-----T-----	542
HIVHA651	-----GG-----C-G-----	539
HIVRJS	-----	558
HIVGON	---C---A---G---A-----	461
HIVSC	-----C-----	564
HIVBR014	-----G-C-----	566
HIVBR021	---A---G-----	567
HIVACH9	---C-----A-----AATAGTACT	429
HIVACP1	---G---G-----C-----	411
HIVJM	---A---C-----G-----AATAATAAT	495
HIVNM	---AC-----A-----AATGATAAAGAT	455
HIVMA208	---AC-----A-----G-----AATGATAGTACTAGT	600
HIVP896	---A---G-----AATAACT	564
HIVBWB11A	---G-----G-----GGTAGT	177

CONSENSUS.C	CACCTTTTATAGactTGATaTAGTAccacTtaatt????????????.....??aaataa?atc???????	523
HIVMA959	-----A-----G-G-GGA-----	551
HIVMA960	-----G-----G-G-TGAGAA-----	530
HIVD757	-----A-----G-----G-GGT-----	217
HIVD747	-----A-----G-----CT-C-GAT-----	352
HIVD760	-----G-----G-----T-A-----	378
HIVNOF	-----A-C-----G-----G-G-GG-----	188
HIVZAM20A	-----GG-----GATCAGAAATAGT-----C-G-C-AATGCC-----	588
HIVD1044	-----A-----G-----G-G-G-AAG-----	187
HIVD1024	-----G-----G-AG-----G-C-----	534
HIVD744	-----N-----G-----G-GG-A-GAGAGA-----	213
HIVD766	-----C-----CC-----T-----G-G-ACTCT-----	189
HIVD808	-----G-----T-----G-G-----	234
HIVD868	-----G-----G-----G-AGAG-----	222
HIVBR025	-----G-----G-G-G-AGT-----	459
HIVSM145A	-----GAGGCTGAA-----C-CYC-AGGAAG-----	573
HIVZAM18A	-----A-----GAGAAAT-----AAC-CTCT-G-GAG-----	552
HIVDJ259A	-----C-----GAG-----T-G-GAA-----	543
HIVDJ373A	-----A-----GGG-----GGG-A-GAGGAA-----	561
HIVSR364A	-----G-GGGAGTAA-----C-GTTC-----	558
HIVUG268A	-----A-----G-----T-----G-AGGGAATAGT-----C-CT-----	585
CONSENSUS.D	CACCTTTTATAGactTGATgTAGTaCaaATagat?at????????????????aaata?tagt???????	528
HIVJY1	-----G-----G-GATAATAGT-----GCT-----CC-AATACC-----	594
HIVNDK	-----A-----G-C-----CA-AAT-----GG-CC-----	543
HIVMAL	-----AC-----C-----C-----G-AGTGAT-----	576
HIVELI	-----G-----A-----C-----C-----GA-AGTACC-----	561
HIV2226	-----G-----A-----CC-----CA-GAT-----G-CC-----	561
HIVUG024	-----G-----G-----G-CA-G-----A-C-----	567
HIVUG021	-----G-----G-----GA-G-GAT-----G-C-----	536
HIVUG038	-----A-----GG-----G-----GACA-----A-AATGAT-----	564
HIVUG046	-----G-----G-----G-----G-A-ACT-----	570
HIVUG266A	-----G-----C-----G-----AAT-G-C-----	558
HIVUG269A	-----G-----GA-A-AGTGAGAGT-----ATT-C-CC-C-----	587
HIVUG274A	-----G-----G-----G-----AAT-G-C-----	573
HIVSE365A	-----G-----T-----C-----GCA-AGTAATAAG-----GC-CC-----	582
HIVCAR4020	-----C-----A-----A-----C-----GG-GAGGATAATATTAAAGGT-----C-GATAAT-----	648
CONSENSUS.E	CACCTTTTATAGactTGAT?TAGTAccaTtga?g????????????????...??AAtagTAGtgag???	530
HIVTH966	-----C-----A-----A-ATAAA-----	572
HIVTH975	-----C-----A-----G-A-GG-----C-----	557
HIVTH022	-----G-----A-----A-----GC-TAATGAA-----	572
HIVTH243	-----A-----A-----GA-ATAAG-----	591
HIVTH006	-----A-----A-G-----	560
HIVTH011	-----C-----A-----GA-AT-----AAT-CG-----AGTGAG-----	503
HIVCARELO	-----T-----T-----T-GCAATATTGATGCC-----A-----	600
HIVCAR4017	-----AC-----A-----T-----T-----C-----AGTGAT-----	591
HIVCAR4039	-----C-----A-----C-----AA-G-AA-----AGTGAG-----	570
HIVCAR4071	-----C-----G-TC-----C-----GAGAAATAGTGGGGAA-----AAT-G-----AGT-----	594
HIVCARHBA	-----A-----GTATC-----C-----T-----GA-T-CA-----A-C-----	564
CONSENSUS.F	CACCTTTTATAGactTGATATAGTACCAATtAGcaat?g?ag?.....AATaaCAGT??gga	556
HIVBRA7944	-----G-----C-----T-----G-A-A-----GG-----GGA-AC-----	480
HIVB2163A	-----A-----A-----A-C-T-----A-----	565
HIVB2126A	-----A-----A-----A-C-T-----A-----	569
CONSENSUS.G	GgCTTTTCTATAGactTGATGTaGTACCAATtAATgataAT?aa.....AATagTAGT-----	553
HIVV1525A	-----A-----G-----A-----AG-----GA-----	609
HIVLBV217	-----T-----G-----G-----G-----G-----	591
HIVCAR4067	-----GG-----C-G-----	567
CONSENSUS.O	CTCTATTCTATGTatCAGAttTGatAaG?TtAAtGacaCaa??????a.....AATgaaACA??????	529
HIVANT70	-----GG-AC-G-----G-----GCAGCAC-----A-G-----AACAGC-----	552
HIVMVP5180	-----C-----G-----T-----ATGCAGT-----G-----	567
HIVVAU	-----GA-----G-----A-----C-C-----TCA-----	570
CONSENSUS.CP2	C??TATTTTAT????G??GAT??T??TAA??????G???.AA?????	267
SIVCP2GAB	-----AC-----GTG-AG-----G-GG-A-----CCTAGG-----TGAG-----	555
CONSENSUS.U	CaCTTTTtTATAaACTtGATtTAQTaC?AaTtAATgaaat?????.....??AATAactacTag?????	523
HIV2AM184	-----G-----A-----G-----CC-----AGT-----TG-----	543
HIVCAR4081	-----G-----C-----G-----A-----C-----T-----AAT-----A-----	570
HIVK124A	-----A-----A-----A-----G-----CAGGAT-----GC-G-----CGGT-----	576
HIV23	-----G-----G-----C-C-----TG-CA-----T-----TATA-----	552

HIV1 ENV

[illegible]

CONSENSUS.C	aac???agt?????.....?????????a?TatagATTAAaTaaATGtaaTAC	553
HIVMA959	---AGC---CTCT.....GGAG-A---C-G---	592
HIVMA960	---AAC---C.....C---	562
HIVD757	---TCT---.....C---	249
HIVD747	---TCT---.....C---	384
HIVD760	---TAGC---CC.....C---	410
HIVNDF	A-C---T---C---C---	214
HIVZAM20A	---TCT-C---AAGCCTACTA-G---T---	632
HIVD1044	---GACTC---AGTAGGA-T---	228
HIVD1024	---TAAG-C---A-G---	569
HIVD744	---TCT---AAGT-T---	251
HIVD756	---TGAT---T-T---	224
HIVD808	---TCT---GAGT-T---	272
HIVD868	---AATTC---AGTGAGT-T---	263
HIVR025	---ACTTC---GGTG-T---G---	497
HIVSM145A	C---	596
HIVZAM18A	---G-TCT---G-G---	587
HIVD7259A	C---	566
HIVD7373A	---C---	584
HIVSE364A	---TTCC---G-G---	593
HIVUG258A	C---	608
CONSENSUS.D	???????.....aala?ta??.....???TATAGatTaaTaaATTGtaaTaaC	557
HIVJY1	---TA---CC.....AAT---C---	629
HIVMDK	---G---CT.....AAT---G---	578
HIVMAL	---G---GT.....GC---	608
HIVELI	---G---CC.....AAT---G---	596
HIV2226	---G---CC.....AAT---	596
HIVUG024	---CC-GC---	599
HIVUG021	---CC-GT---	568
HIVUG038	C---C---	587
HIVUG046	---CC-AC---	602
HIVUG266A	---CC-GC---	590
HIVUG269A	T---	610
HIVUG274A	---G---CT.....AAC---G---C---	608
HIVSE365A	---TA---CC.....AGT---	617
HIVCAR4020	ACTAAA.....T---CT.....AGA---G---	689
CONSENSUS.E	?????????.....TataggtTAATAaaTtgtAaatac	553
HIVTH966		595
HIVTH975		580
HIVTH022		595
HIVTH243		614
HIVTH006	---C---	583
HIVTH011	TATAGTGAG.....	535
HIVCARELO	---TA---C---	623
HIVCAR4017	---T---A---	614
HIVCAR4039	---GT-C---CT-GTA-TACA	593
HIVCAR4071	---A---C---	617
HIVCARMBA		587
CONSENSUS.F	gaa.....TACAGgCTAATAaaATTGcaATAC	582
HIVRA7944	---T---	506
HIVB2163A	T-C.....	591
HIVB2126A	---A---C---G---	595
CONSENSUS.G	?????????cTATAGgcTAATAaaATTGTAATGc	577
HIVVI525A	C---	632
HIVLBSV217	ACAACTGA.....C	623
HIVCAR4067	AC---AT---	593
CONSENSUS.O	???.....AtgtATA?atTAACTAATTGTAACTC	554
HIVANT70	AAA.....C---	581
HIVMVP5180	---CA---TG---	593
HIVVAU	G---T---	596
CONSENSUS.CPZ	AA?A?AC?.....??????TAT????TA?A?A?TGCAA?AC	285
SIVCP2GAB	---CA-C---A.....AGGA---ATTA-T---T---	587
CONSENSUS.U	?????????.....?????????.....???TATAG??TAATAaaTTGTAATAC	544
HIVZAM184	AC---G---	566
HIVCAR4081	AGT---GC---	596
HIVCL24A	GAATATAGA.....AT---	608
HIV23	AGTAGTACC.....GT---	584

HIV1 ENV

CONSENSUS .A	ctCagCcatTAcAcAgGctTgtCCaAaggTatCCTTTGAGCCa.ATTCCCATaCATTATGTgGcccGgCT	614
HIVSF1703	-----A-----T-----	681
HIVU455	---A-----C---G-----	657
HIVZ321	T-----A-----G-----	666
HIVRW020	-----A-----A-----	626
HIVUG031	-----A-----A-----	662
HIVD687	-----A-----A-----	322
HIVNI	-----A-----A-----	654
HIVRW009	-----A-----A-----	662
HIVUG275A	T-----A-----C-----	657
HIVUG273A	-----A-----A-----	675
HIVVT191A	-----A-----A-----G-----	666
HIVDJ264A	-----A-A-A-----A-----C-----	651
HIVDJ263A	-----A-A-A-----A-----	663
HIVDJ258A	-----A-A-A-----A-----	657
HIVCARGAN	-----A-A-----A-----A-T-----	684
HIVCAR5AS	---A-----C-----A-----	678
HIVCAR4054	---A-TGC---A-----C-----T-----	657
HIVCAR286A	---A-----A-----GA-----T-----	717
HIVCAR4023	A---AG---A-----A-A-----T-----	714
HIVCAR423A	---A-TG---A-----T-G-----	693
HIVKENYA	-----A---C-----A-----	657
CONSENSUS .B	ctCagtcCattAcacAggCctTgtCCAAaggTatCgTTTgAgCCA.ATTCCCATaCAtTaTTGtgcccCggcT	641
HIVJRCSP	-----A-----A-----	657
HIVJRFPL	-----A-----A-----	654
HIVALA1	-----A-----C-----	657
HIVERVA	-----A-----A-----	663
HIVJH3	-----A-----A-----	699
HIVBAL1	-----A-----A-----	675
HIVYU2	-----A-----A-----	651
HIVMN	-----A-T-----A-----C-----	678
HIVHXB2R	-----A-----A-----	663
HIVLAI	-----A-----A-----	678
HIVNL43	-----A-----A-----	657
HIVNFA	-----A-----A-----	657
HIVCAN1	-----A-----A-----	666
HIVNY5CG	-----T-----A-----	648
HIVADA	---AC-----A-----A-----	651
HIVJFL	---AC-----A-----A-----	654
HIVSINT84	-----A-----A-----A-A-----	657
HIVD31	-----A-----A-----	663
HIVSF162	-----A-----A-----	657
HIVBCSG3C	---C-----A-----	672
HIVQYI	---AC-----A-----G-----A-T-----	678
HIVSF33	-----A-----T-----	666
HIVCDC4	-----A-----A-----	696
HIVSF2	A-----A-----A-----	672
HIVSF2B13	A-----A-----A-----	669
HIVHAN	A-T-----A-----A-----	669
HIVRF	-----A-----CA-----	702
HIVWMJ2	-----A-----A-----	654
HIVTB132	---A---C-----A-----A-----	621
HIVBR020	-----C-----A-----	659
HIVTH014	-----A-A-T-----A-----	636
HIVTH026	-----A-T-----A-----	638
HIVU6711	---C-----C---G-----A-----T---A-T-----	653
HIVU6712	-----A-----A-----	665
HIVU6715	---C-----C-----T-----	671
HIVU6716	-----A-----A-----	647
HIVHA593	-----A-----A-----	731
HIVHA594	---AC-----A-----	665
HIVHA596	---AC-----A-----	668
HIVHA599	---A---G---A-----A-----A-----	665
HIVMA651	-----A-----A-----	653
HIVRJS	-----A-----A-----	666
HIVGUN	-----A-----A-----	562
HIVSC	---C---A-----C-----	664
HIVBR014	-----A-----T---G-----A-----	668
HIVBR021	-----A-----A-----	675
HIVACH9	-----A-----A-----	540
HIVACP1	-----A-----A-----	519
HIVJN	-----A-----A-----	603
HIVWM	---C-----A-----T-----	567
HIVMA208	-----A-----A-----	723
HIVPB96	-----C-----T-----	669
HIVBNB11A	-----A-----A-----	279

CONSENSUS.C	GTCAgCCaTAACaCAaGCCTGTCCAAAGGtCtctTTtGacCCa.ATTCCTATACaTATtGtGtCtccGcCt	623
HIVMA959	---A-TG-----A-----	662
HIVMA960	---A-----AG-----G-----	632
HIVD757	-----AG-----C-----	319
HIVD747	-----A-----GG-----	454
HIVD760	-----A-----T-----C-----	480
HIVNOF	---A-----C-----	284
HIVZAM20A	---A-----AG-G-G-----	702
HIVD1044	-----A-C-----C-C-CA-----	298
HIVD1024	-----A-----T-----C-----	639
HIVD744	-----A-----T-----C-----A-----	321
HIVD766	-----A-----T-----C-----A-G-----	294
HIVD808	-----A-----T-----C-CA-----	342
HIVD868	-----C-A-----T-----C-----	333
HIVBR025	-----G-----	567
HIVSM145A	-----	666
HIVZAM18A	-----A-G-T-----	657
HIVDJ259A	-----G-----	636
HIVDJ373A	---T-----A-----	654
HIVSE364A	-----G-----C-----	663
HIVUG268A	---G-----A-G-----	678
CONSENSUS.D	CTCAgCCATTACACAgGctTGTCCaAAGgTaacCTTTGagCCA.ATTCCCATACATTaTTGTGCCCCAGCT	627
HIVJY1	---A-----A-T-----A-----T-----	699
HIVNDK	---TA-----T-----	648
HIVMAL	-----T-----	678
HIVELI	-----GT-----	666
HIV2226	-----AG-----	666
HIVUG024	---A-----G-----A-T-----	669
HIVUG021	---A-----A-----	638
HIVUG038	---T-----G-----T-----	657
HIVUG046	-----A-G-----T-----	672
HIVUG266A	-----G-----T-A-----	660
HIVUG269A	-----A-G-----	680
HIVUG274A	---T---A-----A-----	678
HIVSE365A	-----T-----A-----	687
HIVCAR4020	-----T-----A-----	759
CONSENSUS.E	TTCAgTCATTAAgCaGcCTTGTCCaAAAGgTatCCTTTGatCCa?ATTCCTATACATTATtGtGtCtCCaGCT	623
HIVTH966	-----A-----A-----	665
HIVTH975	-----A-----A-----	650
HIVTH022	-----A-----A-----	665
HIVTH243	-----A-----A-----	684
HIVTH006	-----A-----A-----	653
HIVTH011	-----G-----A-----	605
HIVCARELO	-----G-----A-C-----C-G-----	693
HIVCAR4017	-----A-----A-----G-----C-----C-----	684
HIVCAR4039	-----A-A-----C-----C-----	663
HIVCAR4071	-----A-A-----G-----C-----	687
HIVCARMBA	-----G-----C-----	657
CONSENSUS.F	CTCAACCATTTACACAGGCTTGTCCAAAGGTATCTTGGaTCCA.ATTCCCATACATTATtGtGtCtCCaGCT	652
HIVBRA7944	-----A-----G-----	576
HIVBZ163A	-----A-----	661
HIVBZ126A	-----A-----	665
CONSENSUS.G	CTCAACCATTTAAACAgGCTTGTCCaAAGGTgaCTTTTGA?CCA.ATTCCCATACATTATtGtGtCtCCaGCT	646
HIVVT525A	-----C-----AT-----A-----	702
HIVLBV217	---A-----C-----C-----	693
HIVCAR4067	-----G-----	663
CONSENSUS.O	CACAACCATCA?GCAgGCCTGTCCaAAGGTatCCTTTGagCCc.ATTCCCATACACTATtGtGtCtCCaGCa	623
HIVANT70	---C---A-----A-----A-----	651
HIVMVP5180	---TT---A-----G-----AG-----A-----	663
HIVVAU	---G---G-----G-----	666
CONSENSUS.CPZ	7ACA?C?AT?ACACAA?C?TGT?AAA?PC?PC?TTTGA?CC?.ATTCC?ATACA?TATtGtGtC?CCAGCA	337
SIVCP23AB	T---G-C-A---G-A---CC---AA-GT-C---G-T---A---C-----	657
CONSENSUS.U	CTCAgCCATTacACAgGcCTGTCCAAAGGTATCCTTTGAGCCA.ATTCC?ATACATTATtGtGtC?CCaGCT	612
HIVZAM184	---CG---A-----T-----A-C-----	636
HIVCAR4081	-----A-----C-----T-----	666
HIVK124A	-----G-C-----C-----T-----	678
HIVZ3	---A-----T-----C-T-----	654

HIV1 ENV

CONSENSUS.A	GGTTTTCGCGATTCTAAAGTGTAAAGTatagaagagTTCaatGGAAcAGGGccatgcaagAATGTcAGCaCaGT	684
HIVSF1703	-----T-----GCA-----	752
HIVU455	-----G-----CCT-----A-----G-----	728
HIVZ321	---A-----GG---G---TG-A---A---G-----	737
HIVRW020	---C-----G-----CA-----	697
HIVUG031	-----GA---T---GG-T---T-----T-----	733
HIVD687	-----G---G---A-A-----T---GC---A-----	390
HIVNI	-----G-----A---A-G-----A-----	725
HIVRW009	---C-----A-----AA-----	733
HIVUG275A	-----T---C-----AT---T---T-----	728
HIVUG273A	---C---A---T-----T-----T-----T-----	746
HIVVI191A	-----A---GA-----G-----T-----	737
HIVDJ264A	-----T---A-----T-----T-----	722
HIVDJ263A	---G---A-----A---T---A-----T-----T-----	734
HIVDJ258A	-----T---A-----T-----T-----	728
HIVCARBAN	-----T---G---AA---T-----G-----T-----	755
HIVCARBAS	-----G---C-----A-----T---T-----	749
HIVCAR4054	-----T---CAA-A-----	728
HIVCAR286A	---G-----	786
HIVCAR4023	-----T---C-A---G-----	785
HIVCAR423A	-----GG---G-A-----T-----C-----	764
HIVKENYA	---C-----C-----A-----A-----C-----	728
CONSENSUS.B	GGTTTTCGCGATTCTAAAGTGTAAAGTatagaagagTTCaatGGAAcAGGccatGtAcAAATGTcAGCACaGT	712
HIVJRCSP	-----A---CA-----A-----A-----	728
HIVJRPL	-----A---C-----A-----A-----A-----	725
HIVALA1	-----A-----A-----A-----A---C-----	728
HIVBRVA	-----A-----	734
HIVJH3	---A-----	770
HIVBAL1	-----A-----A-----A-----	746
HIVYU2	-----A-----A-----	722
HIVMN	-----A---C---A---G---A---T---A-----	749
HIVHXB2R	-----A---A---C-----	734
HIVLAI	-----A---A---C-----	749
HIVNL43	-----A---A---C-----	728
HIVMFA	-----A---A---C-----	728
HIVCAN1	-----A-----	737
HIVNY5CG	---C---A-----T---T-----	719
HIVADA	---A---C-----G---A-----	722
HIVJFL	-----A-----A-----	725
HIVSIM184	-----A-----	728
HIVD31	---G---A---G---A---A---C-----	734
HIVSF162	-----T-----	728
HIVBCSG3C	-----A-----T---A-----	743
HIVQYI	-----A-----G---A---T-A-----	749
HIVSF33	---AC-----C-----G-----	737
HIVCDC4	-----A---A-C---A-----	767
HIVSF2	-----A---A-C---A-----	743
HIVSF2B13	-----A---A-C---A-----	740
HIVHAN	---T-----A---A-----A-----	740
HIVRF	-----A-----A-----	773
HIVNMJ2	---A-----GA-----G---AC---A-----	725
HIVTB132	-----A---G-----G---AC---C-----	692
HIVBR020	---A-----C---CA---A-----T-----	730
HIVTH014	-----A---C---CA---A-----T-----	709
HIVTH026	-----A---C---CA---A-----T-----	709
HIVUS711	-----C---C---C---A-----	724
HIVUS712	-----C-----A-----	736
HIVUS715	---A-----A-----	742
HIVUS716	-----C-----T-----	718
HIVHA593	-----G---T-----A-----	802
HIVHA594	---A-----A---CA-----	736
HIVHA596	---A-----A---C-----	739
HIVHA599	---A-----G---T-----A-----	736
HIVHA651	-----GG---CCA-A---G-----	724
HIVRJS	---T-----G---T-----	737
HIVGUN	---C-----A-----A-----T-----	633
HIVSC	---T-----A-----	734
HIVBR014	---T-----A---T-----G-----	739
HIVBR021	---G---T-----A---T---C---A---A---A-----	746
HIVACH9	-----A-----A-----	611
HIVACP1	-----A-----A-----T-----	590
HIVJM	-----GA---G---C---A---A-----	674
HIVWM	-----CA---C---G---A-----	638
HIVMA208	---C-----C---C-----A-----	794
HIVP896	---G---G---CA---CA---T---C-----	740
HIVBWB11A	-----T-----G-----	350

CONSENSUS.C	GGTtTaTGCgATTcTaAAGTGTAAataATAagAcaTTcaATGGGACaGGaCCaTGccataATgTcAGcACAgT	694
HIVMA959	-----G-----A-----T-----CG	733
HIVMA960	-----G-----A-----T-----CG	703
HIVD757	-----G-----A-----T-----CG	390
HIVD747	-----G-----A-----T-----CG	525
HIVD760	-----G-----A-----T-----CG	551
HIVNOF	-----G-----A-----T-----CG	355
HIVZAM20A	---A-T---G-----CC-----G-----	773
HIVD1044	-----C-----	369
HIVD1024	-----C-----	710
HIVD744	-----G-----	392
HIVD766	-----G-----G-----T-----	365
HIVD808	-----G-----G-----G-----	413
HIVD868	---T---A-----AG-----	404
HIVBR025	-----A-G---T---A-----A-	638
HIVSH145A	-----G-----G-T-----	737
HIVZAM18A	---A-----	728
HIVDJ259A	---GGG---AC---A---G---A---	707
HIVDJ373A	---GGG---CA-A---A---A-G---T---	725
HIVSE364A	---T-----A-----A-----	734
HIVUG268A	-----A---T---A---TA-----	749
CONSENSUS.D	GGTtTtTGCaATTCTAAAgTGTAAagATAAgaagTTCAATGGaAcaGgcccataTGCaaaAatGTcAGcACAGT	697
HIVJY1	---T-----A---T---	770
HIVNDK	---T-----G-----G-----TC-----	719
HIVMAL	---T-----T-----G-AAAT---T-----T---	749
HIVELI	---T---G-----G-----C-----	737
HIVZ226	---T-----A---G---A-G-----C-----	737
HIVUG024	---A-----A---CA---A-T---G-G-T---T---C---C---	740
HIVUG021	---A-----A---T---A---G---T---T---T---	709
HIVUG038	---C-----A-----C-----C-----	728
HIVUG046	---A---G-----T-----G-G-T---C-----C-----	743
HIVUG266A	---A---A-----T---G-G-T---C-----C-----	731
HIVUG269A	---A---A---CC-----G-AG-T-----	751
HIVUG274A	---A-A---G---A---T---G-G-T---C---C---	749
HIVSE365A	---T-----A-----A-----	758
HIVCAR4020	---C-----A---C-----T---C-T---	830
CONSENSUS.E	GgtTtTGCgATTcTaAAGTGTAAatgatAgaatTTcaATGGgacAGGGcCATGTaaaaAatGTcAGctCAGT	694
HIVTH966	-----C-----C-----C-----	736
HIVTH975	-----C-----C-----C-----	721
HIVTH022	-----C-----T-C-----	736
HIVTN243	---T-----	755
HIVTH006	-----A-C-----	724
HIVTH011	-----	676
HIVCARELO	---A-T-----G-G---A---C-----	764
HIVCAR4017	---G-----A---C-----	755
HIVCAR4039	---T---C-----A---G-G---G---A-A---A---C-G---T---TA	734
HIVCAR4071	-----A-----T---A-----G-A-----	758
HIVCARMBA	-----AG---C---A-----A-----	728
CONSENSUS.F	GGTtTtTGCgATTcTAAAgTGTAAATGatAAaAGTTCaATGGGACgGGaCCaTGCaaAAGATGTcAGcACAGT	721
HIVBRA7944	-----A---G-----A-G-----	647
HIVBZ163A	-----T-----A-G-----	731
HIVBZ126A	-----A-----	736
CONSENSUS.G	GGTtTtTGCgATTcTAAAgTGTAGGgATaAGAAATTTCAATGGAAcAGGACcaTGTAAAAATGTcAGTACAGT	717
HIVV1525A	-----A---G-----C-----T-----	773
HIVLEV217	---G---C---CA-----G-----G-----	764
HIVCAR4067	---A-----A---G-----C-----	734
CONSENSUS.O	GGATaTGCcATCTTTAAGTGTAAAgacACAGaattTTAATGGAAcAGGccataTGCaaaaAcaTLaCaGTaGT	694
HIVANT70	-----T-----CAG-----AC-----G-----A-G-----	722
HIVMVP5180	-----G-----C-----C-C---T---G---	734
HIVVAU	---G-----A---G-----T---C---T---G---	737
CONSENSUS.CP2	GGTtTtTGCgATTcTAAAgTGTAAATGatAAaAGTTCaATGGGACgGGaCCaTGCaaAAGATGTcAGcACAGT	381
SIVCF2GAB	---G-T---A---TC-T-A---CA-T---A-A---C---T---AA---T-AA---CT---AAGTAC---	728
CONSENSUS.U	GGTtTtTGCgATTcTAAAgTGTAAAgacACAGaattTTAATGGAAcAGGccataTGCaaAAGATGTcAGcACAGT	679
HIVZAM184	-----C-----C-G-A-----GT-----G-----T-----	707
HIVCAR4081	---T-----A---A-----A---A---T---A-----T---	737
HIVK124A	-----C-----G-----G-----G-----T---	749
HIV23	---T-----T---A-----A-----A-----	725

HIV1 ENV

CONSENSUS .A	ACAATGcACacATGga?ATcAagCCAGtagTatCAACtCAaCTgcTGtTaAATGGcAGtCtTAGCAGaBg?a	753
HIVSF1703	-----G-----A-----C-----G-----	822
HIVU455	-----T-----G-----G-----AG-----	798
HIV2321	-----G-----C-----C-----G-----	807
HIVRM020	-----G-----A-----C-----T-----A-----	767
HIVUG031	-----CA-----A-----C-----AA-----	804
HIVD687	-----G-----G-----A-----A-----	460
HIVNI	-----G-----G-----A-----A-----	795
HIVRM009	-----G-----G-----A-----A-----	803
HIVUG275A	-----G-----G-----T-----G-----	798
HIVUG273A	-----G-----G-----AC-----	816
HIVVI191A	-----G-----G-----A-----A-----	807
HIVDJ264A	-----T-----A-----A-----A-----	792
HIVDJ263A	-----T-----A-----A-----AG-----	804
HIVDJ258A	-----C-----A-----A-----T-----G-----	798
HIVCARGAN	-----A-----A-----A-----A-----	825
HIVCARGAS	-----G-----A-----A-----C-----G-----	819
HIVCAR4054	-----T-----A-----A-----A-----A-----	798
HIVCAR286A	-----G-----G-----AT-----G-----	855
HIVCAR4023	-----G-----G-----A-----A-----	855
HIVCAR423A	-----A-----A-----C-----GG-----	834
HIVKENYA	-----G-----G-----A-----G-----	798
CONSENSUS .B	aCAaTGtACeCAtpGa .ATtAggCCagtaGT?TCaaCtCAaCTgcTGTraAATGGcAGtCtTagCAGaAagaa	781
HIVJRCSE	-----A-----C-----C-----	798
HIVJRFI	-----A-----C-----C-----	795
HIVALA1	-----A-----A-----C-----	798
HIVBRVA	-----A-----C-----T-----	804
HIVJH3	-----A-----A-----G-----	840
HIVBAL1	-----A-----A-----	816
HIVYU2	-----A-----A-----	792
HIVM4	-----A-----A-----	819
HIVHXB2R	-----A-----A-----	804
HIVLAI	-----A-----A-----G-----	819
HIVML43	-----C-----A-----	798
HIVMFA	-----A-----A-----	798
HIVCAM1	-----G-----G-----C-----AT-----A-----	807
HIVNY5CG	-----AA-----G-----	789
HIVADA	-----G-----G-----	792
HIVJFL	-----G-----G-----CT-----G-----	795
HIVSIHI84	-----C-----A-----G-----C-----T-----G-----	798
HIVD31	-----G-----G-----C-----	804
HIVSF162	-----G-----G-----T-----	798
HIVBCSG3C	-----A-----A-----C-----	813
HIVGYI	-----A-----G-----C-----	819
HIVSF33	-----A-----G-----	807
HIVCDC4	-----G-----G-----	837
HIVSF2	-----A-----G-----	813
HIVSF2B13	-----A-----G-----	810
HIVHAN	-----G-----G-----T-----A-----	810
HIVRF	-----G-----G-----	843
HIVM432	-----A-----G-----	795
HIVTB132	-----A-----A-----C-----	762
HIVBR020	-----A-----G-----C-----	800
HIVTH014	-----A-----A-----T-----	779
HIVTH026	-----A-----A-----T-----	779
HIVUS711	-----G-----G-----C-----	794
HIVUS712	-----A-----A-----T-----C-----	806
HIVUS715	-----C-----A-----A-----	812
HIVUS716	G-----C-----C-----A-----	788
HIVHA593	-----C-----A-----G-----C-----	872
HIVHA594	-----G-----G-----T-----	806
HIVHA596	-----G-----G-----T-----	809
HIVHA599	-----C-----G-----T-----	806
HIVHA651	-----G-----A-----G-----G-----	794
HIVRJ5	-----G-----G-----	807
HIVGUN	-----G-----G-----C-----G-----	703
HIVSC	-----A-----C-----	804
HIVSBA	-----A-----A-----T-----	47
HIVSEB	-----A-----A-----G-----	47
HIVSBC	-----A-----A-----C-----	47
HIVBR014	-----G-----G-----A-----	809
HIVBR021	-----G-----G-----T-----G-----	816
HIVACH9	-----G-----A-----	681
HIVACP1	-----C-----G-----	660
HIVJM	-----C-----A-----C-----A-----	744
HIVMM	G-----A-----A-----C-----	708
HIVMA208	-----A-----A-----A-----A-----	864
HIVP896	-----G-----G-----	810
HIVBM811A	-----T-----G-----	420

CONSENSUS.C	acaatgtacacatgga. attaagccagtggtatcaactcaacttactgtttaaatggtagcctagcagaaagga	764
HIVMA959	-----G-----	803
HIVMA960	T-----	773
HIVD757	TAC-----A-----	460
HIVD747	-----A-----	595
HIVD760	-----GC-----	621
HIVNOF	-----T-----G	425
HIVZAM20A	-----T-----A	843
HIVD1044	-----GT-----	439
HIVD1024	-----	780
HIVD744	-----G-----T-----G	462
HIVC766	-----A-----	435
HIVD808	---G-T---C-----A	483
HIVD868	-----G-----G-----A	474
HIVBR025	---G-----AA-G---C-----T-----A	708
HIVSN145A	-----C-B-----	807
HIVZAM18A	---C---C-----A	798
HIVDJ259A	-----A-----T-----T-----AT	777
HIVDJ373A	-----A-----T-----TAC---G-AG	795
HIVSE364A	-----T-----	804
HIVUG268A	-----A-----A-----T-----TA-----A	819
CONSENSUS.D	acaatgtacacatgga. attaggccagtggtatcaactcaacttactgtttaaatggtagcctagcagaaagga	767
HIVJY1	-----C-----A-----	840
HIVNDK	-----A-----G-----	789
HIVMAL	-----A-----G-----A-----	819
HIVELI	-----G-----	807
HIV2226	-----G-----	807
HIVUG024	---G-C---G-----T-----	810
HIVUG021	---G---G---A-----T-----	779
HIVUG038	-----G-----C-----C-----T-----T-----	798
HIVUG046	G-G-----G---C-----C-----T-----	813
HIVUG266A	G-G-----A-----T-----	801
HIVUG269A	---G---G---A-----T-----	821
HIVUG274A	---G---G---C-----T-----	819
HIVSE365A	-----C---A---G-----A-----	828
HIVCAR4030	---G---A---G-----C-----	900
CONSENSUS.E	ACAATGTACACATGGA. ATTAAAGCCAGTGgtatcaactcaacttactgtttaaatggtagcctagcagaaagga	764
HIVTH966	-----	806
HIVTH975	-----	791
HIVTH032	-----	806
HIVTH243	-----	825
HIVTH006	-----	794
HIVTH011	-----	746
HIVCARELO	-----T-----	834
HIVCAR4017	-----C-----	825
HIVCAR4039	---C---A-G---C-----G---A	804
HIVCAR4071	-----T-----	828
HIVCARMBA	-----	798
CONSENSUS.F	ACAATGTACACATGGA. ATTAAAGCCAGTGgtatcaactcaacttactgtttaaatggtagcctagcagaaagga	793
HIVBR7944	-----A-----	717
HIVB2163A	---G---	801
HIVB2126A	-----	806
CONSENSUS.G	acaatgtacacatgga. attaagccagtggtatcaactcaacttactgtttaaatggtagcctagcagaaagga	787
HIVVI525A	-----A-----	843
HIVL8V217	G-----T-----	834
HIVCAR4067	-----T-----G	804
CONSENSUS.H		
HIVCA13		?A?
HIVVI557		G-A
		A-C
CONSENSUS.O	TACTGTACACATGGC. ATcagccacagtaagtaactcaacttaacttactgtttaaatggtagcctagcagaaagga	764
HIVAM770	-----T-E-----G-----T-----	792
HIVMVF5180	-----A-----G-----G-A	804
HIVVAU	-----A-----C-----	807
CONSENSUS.CPZ	?CATGT?AC?CATGGA. AT?A?PCCA?TGGT?TC?ACT?TGTACT?T?AATGGAA?T?T?T?T?T?T?A?T?T?	424
SIIVCPZGAB	G---T-A---T-AA---G---TA-C---TCA---TA-T---G---TAGCTGA-GGA	798
CONSENSUS.U	acaatgtacacatgga. attaagccagtggtatcaactcaacttactgtttaaatggtagcctagcagaaagga	745
HIVZAM184	-----T-----T-E-----G-----G-----	777
HIVCAR4081	G-----C-A-----G-----C-G-----A	807
HIVK124A	C-----C-A-----A-----A-----G	819
HIVZ3	-----T-E-----G-G-----T-A-A-A-----T-T-----A	795

HIV1 ENV

CONSENSUS .A	gaggTAatgATTAgatCTgAaaAataTcaacAaAcAAATgccccaaAccataatagtagacagcttg77aagcctgt	822
HIVSF1703	AG--A-----A-----A-----A-----AAC--A-----	893
HIVU455	--AA--G-----T-----A-----TC--T-----	869
HIVZ321	-----G-----T-----T-----TC-----	878
HIVRW020	-G-A-CA-----A-----AT--G-----	838
HIVUG031	-G--GA-----G-----T-----A-----	850
HIVD587	-A-G-----G-----T-----T-G-TT-----	531
HIVNI	--A-G--C-----T-G-CT--T-----	866
HIVRW009	--A--A-----T-----C-C-----A--AACG--A--	874
HIVUG275A	A-----A-----AT--C-G--T--A--ACC-C-----	869
HIVUG273A	-----AA--A-----T-G-----T-----C-----A--ACC-C-----	887
HIVVI191A	A--A-----CT--G-T--GG-----TCG-----	878
HIVDJ264A	-----G-----A-----G--T-G-TT-C-----	863
HIVDJ263A	-----G-----C--A-----T-G-TTGC-----	875
HIVDJ258A	-----G-----A-----T-G-TT-C-----	869
HIVCARGAN	-----C-C-----A--A-----TC-----	896
HIVCARSA5	A--A-----T--T-----AACG--T-----	890
HIVCAR4054	-----CT--G-----T-----AAA--T-----	869
HIVCAR286A	-G--A-----C-----T-----AC-GT-----	926
HIVCAR4023	-G-----C--A-----T-----AC-GT-----	926
HIVCAR423A	-----A-----CT--T-----AC-GT-----	905
HIVKENYA	-----A-----AT-----AT--CCG-----	869
CONSENSUS .B	gaggtagtaattAGatctgaaaattTcaoggAcaaTgctAaaAaccATAatAgTaCagctgAA7gaatCtgT	851
HIVJRCSE	A--T-----C-----T-----T-----	869
HIVJREF	-----C-----A-----A-----	866
HIVLAI	-----C-----C-----	869
HIVBRVA	-----A-----T-----T-----	875
HIVJH3	-----G-----T-----A--C-----	911
HIVBAL1	-----C-CC-----G-----GT-----T-----	887
HIVYU2	--A-----AA-----T-----C-----	863
HIVBN	-----G-----T--T-----C-----T-----T-----	890
HIVHXB2R	-----TC-----A-----CAC-----	875
HIVLAI	-----CC-----A-----CC-----	890
HIVNL43	-T-----CC-----A-----CAC-----	869
HIVMFA	-G-----CC-----A-----CAC-----	869
HIVCAM1	-----AA-----A-G--C-----	878
HIVNY5CG	-G-----A-----T-----T-----	860
HIVADA	-----AGT-----A-----A--A-----T--A-----	863
HIVJFL	-----T-----TA-----A-----	866
HIVSIMB4	-----T-----TA-----	869
HIVD31	-----C-----G-----	875
HIVSF162	-G-----A-----T-----G-----	869
HIVBC9G3C	-----CC-----T-----A--G-----	884
HIVOI1	-----A-----AGC-----AA-----T-----TA-----	890
HIVSF33	-----C-----A-----T-----T--T-----	878
HIVCDC4	-----A-----T-----T--T-----	908
HIVSF2	-----C-----A-----T-----	884
HIVSF2B13	-----C-----T-----	881
HIVHAN	-A-----C-----A--T--A-----T--C-----	881
HIVRF	-----A-----T-----T--C-----	914
HIVWMJ2	--A-----A-----C-----T-----	866
HIVTB132	-----Y-----T--A-----A--C-----	833
HIVBR020	-TA-----C-----AA-----G-----A--G--CC-----	871
HIVTH014	-----AGC-----G-GT-----T-----	850
HIVTH026	-----A-AGC-----A-----T-----	850
HIVUS711	-----C-----T--AA-----G--G-----	865
HIVUS712	-G-----C-CC--C--T-----T-----A--T-----	877
HIVUS715	-----G-C-CC--C--T-----T-----	883
HIVUS716	--A-----A-----T--TA-----G-----	859
HIVHA593	-----C-----A-----T-----A-----T--A-----	943
HIVHA594	-----AA-----A-----C--G-----	877
HIVHA596	-----A-----A-----G-----	880
HIVHA599	--A-----G-----T--A-----A-----TA--G--A--	877
HIVHA651	--A-----A-----TAC--A-----	865
HIVRJS	-----A-----T-----T--A-----	878
HIVGUM	--A-----A-----T--T--A-----	774
HIVSC	-----C--G-----A--G-----	875
HIVSBA	-----C-----AA-----T--A--A--G-----	118
HIVSBB	-----G-----A-----C--A-----	118
HIVSBC	-----C-----C-----AA-----C-----	118
HIVBR014	-----C-----T--A-----T-----	880
HIVBR021	-----T--A-----T-----T--A-----	887
HIVACH9	-----AA-----G-----C-----	752
HIVACF1	-----A-----T-----TA-----	731
HIVJM	-----G-----C-----AA-----CA--A-----	815
HIVHM	-----A-----A--C--A-----	779
HIVMA208	-----A-----A--CA-----T-----A--C--A-----	935
HIVPB96	--CA-----A-----A-----T-----	881
HIVBWB11A	-----G--C-----A-----G-----C-----T-----	491

CONSENSUS.C	GagATAAaatTAGATCTgAAAAATCCgaCAAAcCAATgtcaAAAcaATAaTAGLaCATctTaaT?AAacCTgT	838
HIVMA959	--A--TAATaATLAGATCTGAAAAATCCGAACAACAAATGTCAAACAATAATAGLACATCTTAACTGT	874
HIVMA960	-----C-C-----CT---G---	844
HIVD757	-----G-----C---	531
HIVD747	-----C---	666
HIVD760	-----U-----C---	692
HIVN0F	--T-----A-----G-C-C-A-	496
HIVZAM20A	-G-----C-C-----G-A--A-	914
HIVD1044	-----C-----C---	510
HIVD1024	-----T-----T-----C-C---	851
HIVD744	-----A-----C---	533
HIVD766	-----C-----C---	506
HIVD808	-----C-----T-C---	55
HIVD868	-G-----C---P---G-C-C---	545
HIVBR025	-----A-----A-G-----CG---	779
HIVSM145A	-----G-C-----C-----CC---	878
HIVZAM18A	-----G-----A-----C-----G---	869
HIVDJ259A	--T-----C-----T-----G-----G---	848
HIVDJ373A	--T-----C-----T-----G-----C---	866
HIVSE364A	-A-----T-----C-----G-----A-	875
HIVUG268A	-A-----T-----C-----T-----G-----CA---	890
CONSENSUS.D	GAgATaATeTaactAGATCTcGAAaAtcTcaCAaAcAAATgcTAAAAccATAATAGtACAgcTtAatGagtCTgT	838
HIVJY1	-----C-----T-----CA--A-	911
HIVMDK	-----G-----G-----A-----AA---	860
HIVMAL	--G-C-----C-----A-----C-T-----A---	890
HIVELI	--C-----C-----T-----C-----C---	878
HIVZ226	-A-----A-----T-----A-----T-----T-----	881
HIVUG024	-----A-----T-----TT-----T-----	850
HIVUG021	-----A-----T-----T-----T-----	869
HIVUG038	-----T-----T-----T-----C-----	884
HIVUG046	-A--G-----T-----T-----C-----	872
HIVUG266A	-----T-----T-----C-----	892
HIVUG269A	-----A-----T-----T-----A-----	890
HIVUG274A	-A-----G-----T-----C-----T-----A-----	899
HIVSE365A	-----G-----C-----G-----G-----A-----	971
HIVCAR4020	-----G-----C-----G-----G-----A-----	971
CONSENSUS.E	GAgATaATaATcAGATCTGAAaATcTCaCAaAcAAATGCCaaAAccATAATAGTgcAcCTtaaT?AaatCTGT	835
HIVTH966	-----G-----C-----T-----	877
HIVTH975	-----G-----C-----T-----	862
HIVTH022	-----G-----C-----T-----	877
HIVTN243	-----G-----C-----T-----	896
HIVTH006	-----G-----C-----T-----	865
HIVTH011	-----G-----C-----T-----	817
HIVCARELO	-----T-----T-----T-----G-----	905
HIVCAR4017	-G-G-G-T-----A-----A--G--GGT--CGC--	896
HIVCAR4039	-A-----T-----G-----C-----	875
HIVCAR4071	-----G-----C-----C-----	899
HIVCARNBA	-----G-----C-----C-----	869
CONSENSUS.F	gATATAATaATcAGATCTgAAAAATCTcAgATaATacAAAAaccATAATAGTACAgcTtAATGaATCTgT	864
HIVBRA7944	-----C-----C-----G-----C-----	788
HIVBZ163A	-----C-----C-----G-----CT---	872
HIVBZ126A	-----C-----CC-----G-----CT---	877
CONSENSUS.G	GAAATaataATTAGATCTGAAaActTCACAgACAATgCCAAAgTCATAATAGTgCAGcTTAAATaaaactTsT	858
HIVI525A	--T-----G-----A-----AA-----AT-----G-----G-----	914
HIVLBV217	-----G-----G-----A-----AA-----AT-----G-----T-----	905
HIVCAR4067	-----G-----G-----C-T-A-----A-----T-----G-----	875
CONSENSUS.H	?A?GTAAT?AT?AGATCT?AAAAATATCACAGACAATACCAAAAACATAATAGTACAGCTTAAAA?TCCTGT	66
HIVCA13	A-G-----G-C-----G-----G-----A-----	74
HIVVI557	C-T-----A-T-----A-----G-----	74
CONSENSUS.O	AAaATAAgANT?ATGGGAbaaaATATPtcgGAc?gtGgggaAaAATATCaTAgTaACCCTAAAttctTactaT	833
HIVANT70	-----G-----C-----G-----T-----AG-----G-----C-----CC---	863
HIVMVP5180	--G-----T-----G-----A-A-ATCA-CA--G-----CA--C---	875
HIVVAU	--T--C--C-----G-----A-A-----G-G-C--C-A-----A-A---	878
CONSENSUS.CPZ	A????A???GTAA????PQ????AA????AA??T????TGT??G??T??G?A?A???APT?GAA????T	450
STVCPZGAB	-ACAT-ACT---GAGTG-AGAAT--GTCAA--AA-CTGA---CT-GA-A-T-C-GCT-G-A--GCAG-	869
CONSENSUS.U	GAGgTaALaATTAGATCTgAAaAtcTCACA?ACAAATGCCAAAaccATAATAGTACAgcTTaatgAgcCTGT	815
HIVZAM18A	-----G-----G-----T-----G-----A-----T-----C-----A-----	848
HIVCAR4081	--A-----C-----C-----G-----T-----T-----T-----T-----	878
HIVK124A	-----G-A-----C-----C-----A-----T-----A-----GTCA-----	890
HIVZ3	-----A-----A-----A-----A-----A-----A-----A-----	866

HIV ENV

CONSENSUS.A	aa?aat?attgtaccagacct...aacacacacatacaaga...aaaagt?????gtacgtata????	873
HIVSF1703	-GA-----	945
HIVU455	-A--C-----T-----T-----A-----A-----GGTATA	922
HIV2321	-AC--C-----TG-----A-----A-----TC-----	930
HIVRW020	-A-----G-----A-----A-----G-----	890
HIVU0031	-----C-----GG-----AT-----A-----A-----	850
HIVD687	-G-----C-----GG-----AT-----A-----A-----	583
HIVNI	-A-----G-----A-----A-----A-----	918
HIVRM009	-CT-----T-----A-----A-----A-----	926
HIVUG275A	-C-----C-----C-----C-----C-----	921
HIVUG273A	-A-----T-----GG-----C-----C-----	939
HIVVI191A	-A-----G-----G-----A-----A-----	930
HIVDJ264A	-GG--C-----G-----G-----A-----G-----	915
HIVDJ263A	-G--C-----G-----G-----A-----A-----	927
HIVDJ258A	-G--C-----G-----GG-----A-----A-----	921
HIVCARGAN	-AT--C-----T-----G--GG--C-----G-----A-----TGG--A-----	948
HIVCAR5A5	-GA-----T-----C-----T-----G--A-----A--G--A-----	942
HIVCAR4054	-G-----G-----GG-----GG-----A-----A-----	921
HIVCAR286A	-GG-----GG-----A-----A-----TATGC	982
HIVCAR4023	-T-----G-----C-----G--C-----A--G-----	978
HIVCAR423A	-G-----G-----T-----TG-----A--C--C-----	957
HIVKENYA	-A-----TG-----A-----A-----	921
CONSENSUS.B	AgaaaATCAATGTGACaAGaCcc???aacacacacatacaaga...aaaagt???...aTAcataata????	903
HIVJRC5F	-A-----G-----G-----A-----A-----	921
HIVJRF1	-----T--T--G--A--GG--G-----	918
HIVALA1	-----T--T--G--A--GG--G-----	921
HIVBRVA	-----G--G-----G--G-----AC--G-----	927
HIVJH3	-T-----G--G--C-----G--G-----	963
HIVBAL1	-----A-----A-----	939
HIVYU2	-T-----A-----A-----	915
HIVMN	-C-----T-----A-----G-----	942
HIVHXB2R	-----A-----C--G--CCAGA	931
HIVLAI	-----C--G--CCAGA	946
HIVNL43	-----C--G--CCAGA	925
HIVMFA	-----C--G--CCAGA	925
HIVCAM1	-----T-----GC-----	930
HIVNY5CG	-----G-----GC-----	912
HIVADA	-----G-----	915
HIVJFL	-C-----T-----CAC--C-----	918
HIVSINI84	-T-----G--A-----C-----C--G-----	921
HIVD31	-----TT-----C-----G-----	927
HIVSF162	-----T-----T-----AC-----	921
HIVBCSG3C	-----C-----A--A-----C-----AC--CG-----	936
HIVQY1	-----C--G-----TCA-----	942
HIVSF33	-----G-----G--G-----AC--GT-----	930
HIVCDC4	-----C-----G-----G--ACGC-----	960
HIVSF2	-C-----C-----CT-----	936
HIVSF2B13	-C-----G-----C-----	933
HIVHAN	-----G-----G-----	933
HIVRF	-C-----AC--AG-----	966
HIVM032	-----T-----GT-----G-----C--TC-----	918
HIVTB132	-----T-----T-----G-----C--r-----hC--G-----	885
HIVBR020	-C-----G-----G-----	923
HIVTH014	-----C-----C-----	902
HIVTH026	-----C-----C-----	902
HIVUS711	-----G-----T-----C-----	917
HIVUS712	-G-----G-----	929
HIVUS715	-C-----G-----G-----	935
HIVUS716	-G-----G-----G-----	911
HIVHA593	-G-----C-----A-----TC-----	995
HIVHA594	-T-----T--G--G-----G-----TC-----	929
HIVHA596	-T-----T-----G-----G-----TC-----	932
HIVHA599	-TT-----C-----T-----GG-----G-----CCGGAC	933
HIVHA651	-C-----A-----A-----	917
HIVRJS	-T-----TAC--CT--A--A--T-----G--CA-----	933
HIVGUN	-T-----G-----C-----G-----AC-----	826
HIVSC	-----G-----C-----G-----	927
HIVSBA	-----G-----G-----A-----	170
HIVSBB	-----C-----C-----G-----TC-----	170
HIVSBC	-C-----T-----AG-----	170
HIVBR014	-----T--CA-----G--GATA-----	935
HIVBR021	-----C--CA-----G-----G-----	939
HIVACH9	-----C-----C-----	804
HIVACP1	-A-----T-----G-----GG-----	783
HIVJM	-T-----G-----G-----	867
HIVM	-C-----T-----G-----	831
HIVMA208	-----C-----GG--C-----G-----C--G-----	987
HIVP896	-T-----G-----G-----T--TC-----	933
HIVBWB11A	-T-----G-----A--G-----	543

CONSENSUS.C	AgaaatTgtgtgtaCaAGaccc...aacATAAtACAAGa...aaAaGt...ataAGgaTA...	886
HIVMA959	-----G-----A-----	926
HIVMA960	-----C-----G-----	896
HIVD757	AG-----A-----G-----G-----	583
HIVD747	-----G-----G-----G-----	718
HIVD760	-----T-----G-----G-----	744
HIVKOF	-----G-----G-----T...GG-----	548
HIVZAM20A	-----G-----A-----C-----	966
HIVD1044	-----A-----G-----G-----	562
HIVD1024	-----A-----G-----G-----	903
HIVD744	-----A-----G-----G-----	585
HIVD766	-----C-----C-----C-----	558
HIVD808	-----C-----C-----C-----	606
HIVD858	-----C-----C-----A-----	597
HIVBR025	-----G-C-AAT-----G-----	831
HIVSM145A	-----AG-----TA...GC-----G-----A-----	930
HIVZAM18A	-----G-----G-----G-----	921
HIVDJ259A	-----C-----AAC-----G-----G-----	900
HIVDJ373A	-----AAT-----G-----T-----G-----C-----C-----	818
HIVSE364A	-----AAT-----G-----T-----G-----G-----	927
HIVUG268A	-----AAT-----G-----G-----G-----	942
CONSENSUS.D	AacaATTaatTGTaCaAGGCCc???tacaa?aatatAaga???caaagg...acacctata...	889
HIVJY1	-----GA-----GACA-----A-T-C-----T-----	966
HIVNDR	-----GT-----AT-----C-G-----T-G-----	912
HIVMAL	-----T...GGA-----C-----AG-G-----T-A-T-C-----	942
HIVELI	-----A-----CC-----G-----TC-A-----C-----A-----	930
HIVZ236	-----G-----C-----GA-----T-----	930
HIVUG024	-----GT-----C-----C-----C-----C-----C-----	933
HIVUG021	-----C-----C-----G-T-CG-----C-----T-C-----	902
HIVUG038	-----T-----G-----G-----T-T-TGC-----GA-----AG-CAA-----G-A-ACCT-----	921
HIVUG046	-----C-----CT-----G-A-----A-----G-C-T-----	936
HIVUG266A	-----T-----C-----C-----C-----A-ATA-----	927
HIVUG269A	-----C-----C-----CC-G-C-----C-----	935
HIVUG274A	-----G-----G-----CG-----G-A-----C-----	942
HIVSE365A	-----TG-----T-----C-----A-----	951
HIVCAR4020	-----AC-----G-----A-----CT-----A-----G-T-----	1023
CONSENSUS.E	AgAAATCAATTGTACCAGACCC...TcCAACaAtAcAAGa...aCAAGT...gtaagtaTa...	887
HIVTH966	-----ACC-C-----	929
HIVTH975	-----C-----	914
HIVTH022	-----A-----C-----	929
HIVTH243	-----G-----R...C-----A-----C-G-R-----	948
HIVTH006	-----A-----C-----	917
HIVTH011	-----A-----GC-----	869
HIVCAR017	-----A-----A-----C-----G-----	957
HIVCAR4017	-----G-----A-----T-----G-----	948
HIVCAR4039	-----C-----T-----T-----G-----	927
HIVCAR4071	-----T-----T-----A-----G-TG-----A-----	951
HIVCAR88A	-----C-----	921
CONSENSUS.F	AgAAATTAAATTGTACAGACCC...AaCAACaAtAcAAGa...AaAaGt...ATAcTaTa...	916
HIVBR7944	-----C-G-----G-----G-----T-----	840
HIVBZ163A	-----C-G-----G-----	924
HIVBZ126A	-----C-G-----T-----T-----T-T-----	929
CONSENSUS.G	AGaaATTaatTGTACAGACCC...AATACCAATACAAGa...AAAAGT...ATaa?a?Tc...	908
HIVV1525A	-----C-----C-----A-T-----	966
HIVLBV217	-----T-----GTG-----C-----CACA-A-----	957
HIVCAR4067	-----TA-----GCT-T-----	927
CONSENSUS.H	A?AATTAACTGTACCAGACCC...AACCAATAA?ACGAGA...A?AAGT?????..ATA?GTATA...	113
HIVCA13	-----GT-----C-----G-----ATGCGT-----G-----	132
HIVV1557	-----CC-----T-----A-----A-----	126
CONSENSUS.O	Aaa?ATgaCctGtgagAGAccA...ggaAatca?acagTA...CAagAg...ATaa?atA...	881
HIVANT70	-----C-----A-----CA-TA...GACA-----G-GA-----	912
HIVMVP5180	-----C-----CATA-----GA-----T-GCAGAG-----T-----TAT-C-----	927
HIVVAU	-----CA-AG-A-----G-----A-----A-----TGCC-----	930
CONSENSUS.CP2	TA?TTA?A?TGT???AGACCAGGAAA?TTA?TTAAG?..??TA?..??TA?ATA...	478
SIVCP2GAB	T-GTC-A-AT---CAT---C...-ATAC---G...GGAG-G...G-G-G-----	921
CONSENSUS.U	aaaAATTAAATTGTACCAGACCC???aaCaCAATAcAagA???aaAAGT...a?acGtaTa...	866
HIVZAM184	-----C-----G-----C-----G-----G-----C-CT-T-----	900
HIVCAR4081	-----G-----G-----G-----G-----G-----A-G-----	930
HIVK124A	G-G-----T-----T-----	942
HIVZ3	-----CGATC-G-----G-AG-T-AGAC-----	924

HIV1 ENV

CONSENSUS .A	?????ggaccaggacaa?????gcattctatgca...acagggtgacata.....ataggg	915
HIVSF1703		997
HIVU455	GTATA--T-----T-----A-A-----	969
HIVZ321	G-----T-----A-----	972
HIVRW020	GT-----	932
HIVUG031		850
HIVD687	G-----G-----G-----CG-----	625
HIVNI	G-----	960
HIVRW009	T-----T-----G-----	968
HIVUG275A	T-----	363
HIVUG273A	T-----A-----	991
HIVVI191A	G-----C-G-----C-----	972
HIVDJ264A	G-----A-----	957
HIVDJ263A	G-----A-----	969
HIVDJ258A	G-----A-----	963
HIVCARGAN	T-----C-----T-A-G-----C-A-----	990
HIVCAR5AS	G-----AT-----A-C-----G-----	984
HIVCAR4054	C-GGACAA-TA-T-----T-----	969
HIVCAR286A	ATATA-----G-----T-T-----A-----TA-----	1026
HIVCAR4023	A-T-----T-----C-----	1020
HIVCAR423A	GGACAA-TCA-----A-A-----	1005
HIVKENYA		963
CONSENSUS .B	?????ggaccaggagaga.....gcattctatgca...acaggagaaATA?????????.....ataGgA	945
HIVJRCSP		963
HIVJRFL	T-----	960
HIVALA1	C-----A-C-----A-----	963
HIVBRVA	T-A-----C-----	969
HIVJH3	AA-C-----GC-----	1005
HIVBAL1	C-----	981
HIVYU2	G-----	957
HIVMN	AA-A-T-----	984
HIVHXB2R	GA.....GT-----T-A-----	972
HIVLA1	GG.....GT-----T-A-----	987
HIVNL43	GG.....GT-----T-A-----	966
HIVNFA	GG.....GT-----T-A-----	966
HIVCAM1	G-----A-G-----G-----ACAG-----	972
HIVNY5CG	G-----A-C-C-----G-----G-A-A-----	954
HIVADA	A-----	957
HIVJFL	G-----C-----	960
HIVS1MI84	GAA-G-----G-----C-----GGAGGATAT-----C-----	972
HIVD31	G-A-----A-----A-----	969
HIVSF162	G-----G-----C-----	963
HIVBCSG3C	T-A-----C-----G-----	978
HIVQY1	T-C-----C-----AA-C-----	984
HIVSF33	A-----T-C-----C-----	972
HIVCDC4	T-GG-----C-----	1002
HIVSF2	C-----AG-----	978
HIVSF2B13	G-----G-----AG-----	975
HIVHAN	G-----G-----AG-----G-----	975
HIVRF	T-A-----G-----C-----	1008
HIVRAJ2	CGT-----A-----	957
HIVTB132	Tm-m-----	927
HIVBR020	G-----C-----	965
HIVTH014	C-----GG-----C-----	944
HIVTH026	CA-----GG-----C-----	944
HIVUS711	G-----	959
HIVUS712	C-----T-----	971
HIVUS715	C-----	977
HIVUS716	C-G-----C-----	953
HIVHA593	CGCG-----A-G-----	1034
HIVHA594	T-GG-----C-----	971
HIVHA596	T-GG-----C-----	974
HIVHA599	ATATA--CGGC-----G-----A-----G-T-----C-C-----G-----	977
HIVHA651	G-----GG-----G-----A-C-----	956
HIVRJS	G-----GT-----GC-----G-----	975
HIVGUN	T-----C-G-----T-A-A-----	868
HIVSC	G-----C-----	969
HIVSBA	G-----C-----	212
HIVSBB	GT-G-----G-A-----	212
HIVSBC		212
HIVBR014	C-G-----G-----A-----	977
HIVBR021	TGG-----G-----	981
HIVACH9	A-----C-----	846
HIVACP1	A-G-----G-A-A-----	825
HIVJM	C-----A-----T-----	909
HIVMM	A-----	873
HIVMA208	T-G-----C-----	1029
HIVP896	G-----G-A-A-C-----	975
HIVBWB11A	G-----C-----	585

CONSENSUS.C	GGACCAGGacAa.....	acatTctatTgca...AcaggaGacATA.....	ATAGGA	928
HIVMA959	GT.....	ATAAT.....		968
HIVMA960	AAT-A.....			938
HIVD757				625
HIVD747				760
HIVD760	T.....			786
HIVNCF	G.....	AAT-C.....		590
HIVZAM20A	C.....	T.....	T-C.....	1000
HIVD1044				604
HIVD1024				945
HIVD744			A.....	627
HIVD766			G.....	600
HIVD808				648
HIVD868				639
HIVBRD25	G.....	A.....		873
HIVSM145A		AAT.....		969
HIVZAM18A	G.....	T-G.....		963
HIVDJ259A				942
HIVDJ373A			T.....	960
HIVSE364A	C.....			969
HIVUG268A				984
CONSENSUS.D	...??GgactAgGGCaa.....	gcactctatataca??Aca??agaata??????...	a?Agga	927
HIVJY1		AGG.....	A.....	1005
HIVNDK	A.....	T.....	ATA-GGA-A-AGAAGAAA.....	963
HIVMAL	C-C.....		G-G.....	981
HIVELI		T.....	T.....	969
HIV2226	T.....	G.....	A-C.....	969
HIVUG024			G.....	972
HIVUG021	G.....	TC.....		941
HIVUG038	ATA-A.....	T.....	AG-A.....	963
HIVUG046		GTA.....	A-T.....	975
HIVUG266A	G.....	T.....	ACA-G.....	966
HIVUG269A	C-A.....	G.....	TGG-GG.....	974
HIVUG274A	AC.....		CAAG-G-AAAAAG.....	990
HIVSE365A	G.....	T-C.....	ACA-GG.....	990
HIVCAR4020			G.....	1062
CONSENSUS.E	GGACCAGGacAa.....	GTATTCTATagA...ACAGGagacata.....	AtAGGA	929
HIVTH966			T.....	971
HIVTH975				956
HIVTH022				971
HIVTH243				990
HIVTH006				959
HIVTH011				911
HIVCARELO		A.....	TC.....	999
HIVCAR4017		A.....	C.....	990
HIVCAR4039		A.....	AG.....	969
HIVCAR4071	G-G.....	A.....	TC.....	993
HIVCARMBA		G.....	A.....	963
CONSENSUS.F	GGaCcAGGAcgA.....	GcatTctatTgCA...acagGtga?ata.....	AtaGGA	957
HIVBRA7944	A.....		A-A.....	882
HIVB2163A			A-C.....	966
HIVB2126A		C-A.....	G-AA-A.....	971
CONSENSUS.G	GgAcCaGGACaA.....	GcgTTCtATGCA...ACAGGtGacATA.....	ATAGGA	950
HIVVI525A	A.....	T-A.....	C.....	1008
HIVLBV217	T.....	A-A.....	C-C.....	999
HIVCAR4067			G.....	969
CONSENSUS.H	GG7C7AGGACAA.....	TCATTC7ATGCA...A?AGGTG??AT?.....	ATAGGA	147
HIVCA13	A-G.....	A-C.....	T-CT-A.....	174
HIVVI557	G-C.....	G-T.....	C-AC-C.....	168
CONSENSUS.O	GGTCCAaTGgc?.....	TGGtacAGc...ATGggccTagag?ga?????aatacaaaaa?c		926
HIVANT70	C.....		AA-G-G.....	960
HIVHVP5180	AGA.....	CG-T.....	ACA-TA-AA-AGTAAC.....	981
HIVVAU	T.....		C-TAGT.....	975
CONSENSUS.CPZ	GG7CCAGGAATG.....	ACTTTTATATA?...?TAGAAA??GTA.....	??AGGA	510
SIUCP2GAB	G.....	C-T.....	A-AT-T.....	963
CONSENSUS.U	GGaCCAGGAcAA.....	gcacTTCtATgca...acaggtgacAtA.....	ataGGA	908
HIVZAM184		C.....	AA.....	939
HIVCAR4081	G.....	A.....		972
HIVK124A		T.....		984
HIV23	A.....	T.....	A.....	957

HIV1 ENV

CONSENSUS.A	gatataagacaagcacattgttaattgtcagtagA?cagaATGGaAtaaaactttacaa?aGGTa	976
HIVSF1703	-----T-----G-----C-----GG-----	1050
HIVU455	-----AGG-C-----G-A-----C-----	1032
HIVZ321	-----G-----C-----A-----G-C-----TCCA-----	1035
HIVRW020	-----G-----C-----G-A-----CG-C-----G-GG-----	995
HIVUG031	-----A-----T-----A-----GAC-C-----	884
HIVD687	-----A-----G-----GG-----C-----	688
HIVNI	-----A-----A-----TC-----	1023
HIVRW009	-----T-----C-----A-G-A-CA-----G-----A-----	1031
HIVUG275A	-----A-----T-----A-----G-----G-----A-----	1026
HIVUG273A	-----A-----T-----GG-----G-----A-----	1044
HIVVI191A	-----CA-----A-----CC-----	1035
HIVDJ264A	-----A-----T-----A-----C-----C-----G-----	1020
HIVDJ263A	-----T-----T-----G-----C-----G-----	1032
HIVDJ258A	-----C-----T-----A-----G-----A-----G-----	1026
HIVCARGAN	-----A-----T-----CG-G-----A-C-A-----C-----C-----	1053
HIVCAR5AS	-----G-----T-----G-----AC-----AT-----	1047
HIVCAR4054	-----A-----G-----G-----C-----C-----	1032
HIVCAR286A	-----CCA-T-A-A-G-A-----C-----A-----	1089
HIVCAR4023	-----T-----C-----A-----GA-----C-CC-----C-G-----A-----	1083
HIVCAR423A	-----A-----T-----CA-----A-G-----A-C-TG-----AG-A-----	1060
HIVKENYA	A-C-----G-----C-----G-----A-----	1026
CONSENSUS.B	gataTaAgacaagCacATTGtAaCaTTagtGagagcAaaATGGaataacaCttTaaaacagatA	1008
HIVJRC5F	-----C-----	1026
HIVJRF1	-----G-----	1023
HIVALA1	A-----G-GG-----	1026
HIVBRVA	-----G-----C-----T-----G-G-----	1032
HIVJH3	-----T-----A-----G-----GC-----	1068
HIVBAL1	-----C-----G-----TA-----	1044
HIVYU2	-----C-----A-A-C-----G-A-----G-----	1020
HIVMN	AC-----G-----G-----	1047
HIVHXB2R	A-----G-----GC-----	1035
HIVLAI	A-----G-----GC-----	1050
HIVNL43	A-----G-----GC-----	1029
HIVMFA	A-----G-----TGT-----	1029
HIVCAM1	-----C-----TA-----C-----	1035
HIVNY5CG	-----C-----G-----	1017
HIVADA	-----C-----A-----T-A-----	1020
HIVJFL	-----C-----T-C-----	1023
HIVSIM184	T-----G-----	1035
HIVD31	-----G-----G-----G-----	1032
HIVSF162	-----G-A-----	1026
HIVBCSG3C	-----C-----G-----G-----	1041
HIVVOY1	-----C-----G-G-A-----G-----	1047
HIVSF33	-----A-----T-----A-----G-----G-----	1035
HIVCDC4	A-----G-----C-----C-----	1065
HIVSF2	-----A-----C-----G-----	1041
HIVSF2B13	-----A-----G-----A-C-----G-----	1038
HIVHAM	-----T-----G-----A-C-----C-A-----	1038
HIVRF	-----A-----C-----C-----G-----	1071
HIVMDJ2	AT-----	1020
HIVTB132	-----G-----CA-----GG-----	990
HIVBR020	-----C-----C-----C-----G-----GCA-----	1028
HIVTH014	-----C-----CA-----G-----	1007
HIVTH026	A-----C-----CA-----G-A-----	1007
HIVUS711	-----G-----C-C-A-A-A-----G-AG-----G-----	1022
HIVUS712	-----A-----	1034
HIVUS715	-----C-----A-----G-A-----AG-----	1040
HIVUS716	-----C-----GC-----G-A-C-----	1016
HIVHA593	A-----A-----C-----C-G-----A-----	1097
HIVHA594	-----A-----C-----CA-C-----G-----C-GG-----	1034
HIVHA596	A-----A-----C-----CA-GC-----C-GG-----	1037
HIVHA599	-----A-----C-----T-----C-----G-----G-----	1040
HIVHA651	-----C-----G-C-----G-----GG-----C-----	1019
HIVR3S	-----C-----G-----G-G-G-----G-----	1038
HIVGUN	A-----A-----A-----T-----	931
HIVSC	-----	1032
HIVSBA	-----C-----GA-----G-----G-----	275
HIVSBB	-----A-----C-T-A-----G-T-----	275
HIVSBC	-----C-----C-T-----T-----GC-----A-----	275
HIVBR014	A-----C-T-----C-C-GA-----TAA-----	1037
HIVBR021	A-----C-----T-----T-----GC-----	1044
HIVACH9	-----G-----C-----GC-----	909
HIVACP1	-----A-----G-----GT-----C-----	880
HIVJM	-----A-----T-----A-A-----T-----T-----G-----G-----	972
HIVWM	A-----A-----T-----C-----G-----	936
HIVMA208	-----A-----CG-----G-----	1092
HIVP896	-----C-----	1038
HIVBWB1A	-----A-----GC-----	648

CONSENSUS.C	gACATAAGcaAGCacATGtAACATTAGtaagataaaTGGaattgaaaCttTAcAaagGTa.....	991
HIVMA959	-----C-----G-----GTC-----CACT-----	1031
HIVMA960	A-----G-----GG-AG-----CA-----C-----	1001
HIVD757	-----G-----G-----G-----	688
HIVD747	-----C-----G-----	823
HIVD760	-----G-----G-----	849
HIVNOF	-----A-GCT-----G-C-G-----	653
HIVZAM20A	-----G-AG-G-----A-----C-----T-T-----	1071
HIVD1044	-----T-----T-----	667
HIVD1024	-----G-----G-----C-----G-----	1008
HIVD744	-----G-----G-----C-----C-----G-----	690
HIVD766	-----G-----G-----	663
HIVD808	-----G-----G-----	711
HIVD868	-----G-----T-----G-A-C-----G-----G-----	702
HIVBR025	--T-----G-ACAGC-----CA-----C-----G-----	936
HIVSM145A	-----GG-----CAG-----C-----	1032
HIVZAM18A	A-----G-----G-----C-----A-----	1026
HIVDJ259A	--T-----G-A-AG-----AG-----	1005
HIVDJ373A	--T-----G-C-A-G-----A-----C-----	1023
HIVSE364A	--T-----GGT-----C-----C-----G-----	1029
HIVUG268A	--T-----T-C-----G-A-G-----AT-----TG-----	1047
CONSENSUS.D	?atAtA?GacAGCacATTGTaataTTAgT?gagcagaTGGaataAaaCTTTAcAAcaggTA.....	987
HIVJY1	G-----A-----T-----GC-----C-----	1068
HIVNDK	T-C-----G-----A-----CA-----G-----	1026
HIVMAL	G-----A-AG-----T-----C-----A-CA-A-----G-----	1044
HIVELI	ATA-----G-----A-----C-----G-----A-----	1032
HIV2226	ATA-----G-----T-----AA-AAT-----G-A-----	1032
HIVU3024	G-----A-AG-----G-----CA-----C-----G-----G-----	1035
HIVU3021	G-----A-----C-----G-----A-A-----	1004
HIVU3038	AGG-----G-----CAG-----T-----G-----	1026
HIVU3045	T-----A-----C-----G-----	1038
HIVU3266A	ATA-A-G-----A-----A-G-----	1029
HIVU3269A	G-----A-----T-----C-----AA-----A-----C-GG-----	1037
HIVU3274A	A-G-----C-----G-----AA-----G-----G-A-----	1053
HIVSE365A	G-----A-----C-----AA-----A-AG-----	1053
HIVCAR4020	G-----A-A-----A-AA-----T-----A-----	1125
CONSENSUS.E	gATATAAGcaAGCacATGtAGATTAatGaaCaaaATGGaattgaaacTTTAAacAGGTa.....	992
HIVTH966	A-----G-----G-----	1034
HIVTH975	-----C-----G-----	1019
HIVTH022	-----G-----A-G-----	1034
HIVTN243	-----R-G-----AG-GT-----	1053
HIVTH006	-----C-----A-----	1022
HIVTH011	A-----T-----GT-----	974
HIVCAR610	-----G-----GC-----C-----A-G-----	1062
HIVCAR4017	-----T-----A-----G-----G-----	1053
HIVCAR4039	-----TC-----G-A-G-----C-----A-----G-----	1032
HIVCAR4071	-----G-----T-----C-----A-----	1056
HIVCARMB4	-----G-----T-----G-----	1026
CONSENSUS.F	GACATAAGcaAGCacATGtTaaacATtAgTAgaaACaATGGaattgaaacTTTAAacAGGTa.....	1020
HIVBR4744	-----C-----G-----G-----T-----G-----	945
HIVB2163A	-----C-----C-----TG-----G-----A-----A-----	1029
HIVB2126A	-----C-----C-----G-----G-----A-----G-----	1034
CONSENSUS.G	gATATAAGcaAGCacATTGTaataTTAgT?AaaTGGaagTAgATGTTAcAGaATGT?.....	1009
HIVV1525A	A-----C-CA-----GGTA-----A-----	1071
HIVLBW217	-----GA-----C-G-C-----GAG-C-----G-A-----	1062
HIVCAR4067	-----G-----C-GCT-----TC-----G-----	1032
CONSENSUS.H	GATATAAGcaAAGCA?ATTGTaataTTA?T?G?G?A?T?GGA?T?T?A?T?A?A?AGGTA.....	192
HIVCA13	-----C-----C-----G-G-A-C-A-G-----TGAC-TG-----ACCA-----	237
HIVV1557	-----A-----T-----C-A-G-A-G-T-----AAGG-CT-----CATG-----	231
CONSENSUS.O	agctCAAGGg?AGCTTATTGtAa?TATAaTgccACTGatTGGgaasangcctTAAaAcAaAcs.....	987
HIVANP70	-----C-----C-----G-----G-----ATA-----	1023
HIVVNP5180	--A-----T-----CA-----AAG-----TA-----T-----C-----C-----	1044
HIVVAU	GATA-----C-----T-----G-----C-----A-C-----A-C-T-----	1038
CONSENSUS.CPZ	GA?AC?AG?T?GCT?T?CTGTa?T?T?AAT?GAC?T?ATGG?A?C?A?C?T?T?A?T?A?T?T?T?T?T?	543
SIVCP2GAB	--T-C-ATCT--C-A--AGA-A--GG--CAC--A-T-G-A-GGTGG-AG-GT--	1026
CONSENSUS.U	?atAtAaGACaAGCacATGtAaT?T?AaTgTaaCa?AATGGaattgaaACTTTACAAcAgT?TA.....	964
HIVZAM184	-----T-----G-A-C-----A-----CC-----A-G-----	1002
HIVCAR4081	G-----G-C-G-AG-----A-----C-G-----AA-----	1035
HIVK124A	A-----G-----G-----G-T-----G-----A-----A-----	1047
HIVZ3	ATA-C-G-----CA-T-CA-ATGGGG-----GGA-T-----G-----	1021

HIV1 ENV

CONSENSUS.A	...Gctta?acAatTAagaaa?.....?acTct/???????AAcnaaaca???.??????a	1007
HIVSF1703	...C-AC-----A-GT.....T-----AGT-----	1087
HIVU455	...GA-----AG-A.....A-G.....AAT-----	1069
HIV2321	...OC-----A.....C-----GTA-T-CC-GT.....ACAGAT-	1078
HIVRW020	...C-A-A-C-C-G-A.....C-----A-C-A.....ACA-	1032
HIVUG031	...T-A-G-----A.....T--GG.....	918
HIVD687	...GA--C-----G-G.....T-----AAG--T--C-----	725
HIVNI	...T-C-----AG-CG.....T-----AAG--CC-----	1060
HIVRW009	...GA-A-----TC-C.....T-----GAG.....TT-----ACA-	1071
HIVUG275A	...TC-OC-----CA.....C-----GOC.....	1063
HIVUG273A	...TC-A-----CA.....C--GG.....	1078
HIVV1191A	...T-----G-A.....C-----ATAACTAAG--C-----	1078
HIVDJ264A	...T-----G-G.....C--C-----C-----	1054
HIVDJ263A	...C-----G-G.....C-----	1066
HIVDJ258A	...C-----G-G.....C-----C-----	1060
HIVCARGAN	...TA-A-A-----CA.....T-T-GG.....	1087
HIVCAR5AS	...GA-G-----A.....C-G-----	1081
HIVCAR4054	...A-A-----TATG-G.....GT-----G	1066
HIVCAR286A	...GA-A-----GGA.....AT-----AAC-----	1126
HIVCAR4023	...A-A-----G-CT.....AT-----AGC-----	1120
HIVCAR423A	...C-A-C-GA-GC.....A-G-----T-----	1102
HIVKENYA	...TC-C-A-----GG-A.....T-----CGGAAC-----	1066
CONSENSUS.B	...gtta?aAaatTAagagaa?????caattf...ggg???AAtAaaACA.....a	1044
HIVJRC5F	...GA-----AAT-----	1063
HIVJRF1	...T-----A-----	1060
HIVALA1	...GA-----AAG-----	1063
HIVBRVA	...C-----T-----AAG-----	1069
HIVJH3	...GG-----A-----T-----	1105
HIVBAL1	...T-----	1081
HIVYU2	...C-T-----A-----AAT-----	1060
HIVM6	...GC-----A-----AAG-----	1084
HIVHXB2R	...C-GC-----AAAT-----	1075
HIVLA1	...C-GC-----AAAT-----	1090
HIVNL43	...C-GC-----AAAT-----	1069
HIVMFA	...CA-GC-----AAAT-----G	1069
HIVCAM1	...C-----A-----G-----	1072
HIVNYSOG	...C-----A-----A-----	1054
HIVADA	...C-C-----A-----AAT-----	1060
HIVJFL	...T-G-----G-----	1060
HIVSIME84	...A-----A-----AGG-----	1075
HIVD31	...A-----G-----C-----	1069
HIVSF162	...C-----CA-C-----	1063
HIVBCSG3C	...T-----A-----	1078
HIV0Y1	...C-C-----A-----A-----	1084
HIVSF33	...C-C-----SF-----	1072
HIVCDC4	...C-C-CC-----	1102
HIVSF2	...A-----G-----AAT-----	1081
HIVSF2B13	...A-----G-----AAT-----	1078
HIVHAN	...T-G-----ATTAGA-----AA-----	1081
HIVRF	...C-----AC-----	1108
HIVWJ2	...GAG-----AA-----	1057
HIVTB132	...A-----A-----AAC-----T-----	1027
HIVBR020	...C-----A-----	1065
HIVTH014	...AC-GA-----C-----	1044
HIVTH026	...AC-GA-----C-----	1044
HIVUS711	...TC-G-----G-----	1059
HIVUS712	...C-T-----A-----A-----	1071
HIVUS715	...T-----G-----A-----A-----	1077
HIVUS716	...AC-T-----GAT-----G-----	1053
HIVHA593	...GC-----	1134
HIVHA594	...GA-----C-----	1071
HIVHA596	...GA-----C-----	1074
HIVHA599	...C-A-----C-----GAT-----G-----	1080
HIVHA651	...C-----A-----	1056
HIVR79	...C-----	1075
HIVGUN	...C-----	968
HIVSC	...T-----C-----A-----	1069
HIVSRA	...G-GA-----AA-----	312
HIVSBB	...GC-----CAG-----AAG-----	315
HIVSBC	...GA-----AA-----	312
HIVBR014	...C-----A-----	1074
HIVBR021	...AA-C-----A-C-----G-----	1078
HIVACH9	...C-----G-----	946
HIVACPI	...A-----C-----AAT-----	928
HIVJM	...C-GA-----G-----AA-----G-----	1009
HIVWM	...A-----AA-----	973
HIVMA208	...A-----	1129
HIVP896	...T-----A-----A-----	1075
HIVWB11A	...C-----T-----	685

CONSENSUS.C	...ggtaaaaAaTAgcAgaa...cactTccct?????AATaaaaca...a	1028
HIVMA959	...-----AA-----	1068
HIVMA960	...-----T-----GT-T-----	1038
HIVD757	...-----T-----	725
HIVD747	...-----	860
HIVD760	...-----C-----	886
HIVNOF	...AAG-----A-----	690
HIVZAM20A	...A--GG-----G-----GG--C-----	1108
HIVD1044	...--A-----	704
HIVD1024	...-----G-----	1045
HIVD744	...A-----	727
HIVD766	...A-----AA-----	700
HIVD808	...A-----	748
HIVD868	...A-----T-----	739
HIVBR025	...-----G-----G-----	973
HIVSM145A	...--A-----	1069
HIVZAM18A	...--A-----G-----	1063
HIVDJ259A	...A-AG-----AA-ACG...T---AGAACTGCA---T---	1047
HIVDJ373A	...AAGG-----AA-----	1057
HIVSE364A	...AAGG-----CA-----GGTAAA---T-GC	1071
HIVUG268A	...A-AG-----AA-AGG-----	1084
CONSENSUS.D	...GctaaaaAaTTAagAgac?...CTtCTt...Aacaaacaa...acaa	1024
HIVJY1	...G-----C-G-----	1105
HIVNDK	...C--C--G--A--	1063
HIVMAL	...GT--C--G--AG--	1081
HIVBLT	...G--G--AC--	1069
HIVZ226	...T--G--A--	1069
HIVUG024	...-----A--T-----T---G-CGG---	1072
HIVUG021	...GGG-----C-----G---	1041
HIVUG038	...-----T-----G	1063
HIVUG046	...-----G-----T-C---C-G---	1075
HIVUG266A	...A-----A-----G---	1066
HIVUG269A	...-----G-----G---CA---	1074
HIVUG274A	...-----A--C---T-T-A---GGC-G-T---AT-	1089
HIVSE365A	...T-G-----A---CT---	1087
HIVCAR4020	...G-----A---T---GT---A---	1162
CONSENSUS.E	...actgaAAaTTAaaAgAg...CACITT...???AATAAGAc...???	1026
HIVTH966	...A-----AAT-----	1071
HIVTH975	...G-----	1053
HIVTH022	...-----AAT-----	1068
HIVTH243	...-----AAT-----T-AG...ACA-	1090
HIVTH006	...A-----ACT-----	1059
HIVTH011	...A--G-----	1011
HIVCARELO	...GT--G-----G-A--	1096
HIVCAR4017	...AG-----G-A--	1087
HIVCAR4039	...TA-G-----G-----	1066
HIVCAR4071	...T-AG-----G-----AAT-----ATAC	1096
HIVCARMB	...-----G--A-----	1060
CONSENSUS.F	...a?gGaaaagtTaaaa???...catttcc??...???AAta?aaCa...???	1050
HIVBR7944	...G--C--G-----GGTCT...CT...GC---	982
HIVBZ163A	...G--C-----GTCT...T-CT...GC---	1056
HIVBZ126A	...G--C-----GTCT...CT...C---	1071
CONSENSUS.G	...a??gcAcaaCTAaaa?aa...atcTat?????aa?AacAaaAcc...?????????	1042
HIVVI525A	...GCT--A-----C--A--GAA-T---AAT--A---T---	1110
HIVLBV217	...AA-----C--GG--C...AAC--A-G--T---	1101
HIVCAR4067	...CTA--AG-----G---C--T-CGGGGGG-G---T---TTAGAAAAC-	1084
CONSENSUS.H	...G?T???CAATTA?GA?AA...CAC?T?...??AA??A?A??...???	213
HIVCA13	...C-ACA-----G--A---C-G-----CA-G-GC...AAA-	274
HIVVI557	...T-CAG-----A--G---T-T...AAT--TC-A-CA...	268
CONSENSUS.O	...gCTgaAAGgTATtTAgAa...CTtgTA...aAc?AtAca?????aata?tgTtAccA	1031
HIVANT70	...-----A-----A-----GG--G-A---A--	1069
HIVMVP5180	...AT-----A--T-----C-A---G-G-A-----	1090
HIVVAU	...A-----A-----G-AT---AT...CAA-C-GA-----	1087
CONSENSUS.CPZ	????????AG????TT????AA??...??A...??...AA??AAC?...???	558
SIVCPZGAB	...AAAA--GC--TAGC--CC...TC...TCA...--TCG--A...--AGCC...	1066
CONSENSUS.U	...gCTAcacaAPTaa?Aaaa...?????c???...a??AA?AaaaCA...???	992
HIVZAM184	...--A-----CA--C...TACTT...CCT--C-----	1039
HIVCAR4081	...-----G-----TACTTT...GGG--T-C---ACA-	1075
HIVK124A	...AG-----A-----CA-TTT...TG--C--G---	1084
HIV23	...--TCGC--G--G---CAATT...AT--T--T---	1057

HIV1 ENV

CONSENSUS .A	TaAtcTtGctaac.?ccTca...GGaCGGGAT?TaGAAaTtAcAacAcATAgTtTTAaTTGTggAgGaGA	1072
HIVSF1703	-----G- T-----T-----	1154
HIVU455	-----G- T-----A-----C-----	1136
HIVZ321	-----T-----G-----	1145
HIVRW020	-----AG- A- T-----A-----C-----T-----	1099
HIVUG031	-----A-----T-----T-----	985
HIVD687	-----C- CA-----T-----A-----	792
HIVNI	-----C- T-----G-----	1127
HIVRW009	-----T- AAG-----G-----T-----	1138
HIVUG275A	-----G- T-----A-----A-----A-----	1130
HIVUG273A	-----A- A- T-----T-----	1145
HIVVI191A	-----AA- G- T-----T- G-----C-----	1145
HIVDJ264A	-----C-----A-----	1121
HIVDJ263A	-----T-----A-----	1133
HIVDJ258A	-----T-----A-----A-----	1127
HIVCARGAN	-----A- CA-----A- C-----G-----T-----	1154
HIVCARSA5	-----A-----C- T-----G-----	1148
HIVCAR4054	-----N- A-----CA-----A-----C-----A-----G-----	1133
HIVCAR286A	-----A-----CA-----T-----G-----C-----C-----	1193
HIVCAR4023	-----AA- G- AG-----T-----G-----C-----C-----A-----	1187
HIVCAR423A	-----T- C-----A-----A- C- T-----A-----T- T-----C-----AA-----	1169
HIVKENYA	-----AA-----T-----C-----C-----	1133
CONSENSUS .B	TAgTcTtTaatacaa?cCCTCa...GgaCGggAcCcAGAAaTtGtaantgCacAgTtTTAaTTGTggaGgggA	1111
HIVJRC5F	-----C- C-----T-----	1130
HIVJRFL	-----C-----A-----	1127
HIVALA1	-----G-----CA-----	1130
HIVBRVA	-----G-----	1136
HIVJH3	-----G-----C-----	1172
HIVBAL1	-----G- C-----G- C-----	1148
HIVYU2	-----A-----C-----CA-----	1127
HIVMN	-----	1151
HIVHXB2R	-----A-----G-----C-----	1142
HIVLAI	-----A-----G-----C-----	1157
HIVNL43	-----A-----G-----C-----	1136
HIVMFA	-----A-----G-----C-----	1136
HIVCAM1	-----A-----G-----	1139
HIVNY5CG	-----	1121
HIVADA	-----	1127
HIVJFL	-----G-----C-----	1127
HIVSIM184	-----G-----C-----	1142
HIVD31	-----G-----C-----	1136
HIVSF162	-----G-----	1130
HIVBCS03C	-----T-----	1145
HIV0Y1	-----C- G- G-----T-----	1151
HIVSF33	-----A-----A-----A-----A-----	1139
HIVCDC4	-----C-----	1169
HIVSF2	-----A-----	1148
HIVSF2B13	-----A-----A-----	1145
HIVHAN	-----G-----A-----C-----	1148
HIVRF	-----CGTC-----C- T-----	1175
HIVMA72	-----T-----C-----	1124
HIVTB132	-----XG-----T-----T- G-----	1094
HIVBR020	-----AAG- C- G- C- C-----T-----T- T-----	1132
HIVTH014	-----G-----G-----	1111
HIVTH026	-----G-----A-----G-----A-----	1111
HIVUS711	-----A-----A-----A- CA-----	1126
HIVUS712	-----A-----A-----A- G- CT- T- TG-----	1138
HIVUS715	-----A-----T-----C-----	1144
HIVUS716	-----G- CC-----G-----T-----C-----	1120
HIVHA593	-----C- A- C-----	1201
HIVHA594	-----AGA-----C-----A-----A-----T-----	1138
HIVHA596	-----AGA-----C-----A-----A-----T-----A-----C-----	1141
HIVHA599	-----A-----G-----CA-----C-----	1147
HIVHA651	-----TC-----T-----A-----	1123
HIVRJS	-----G-----C-----	1142
HIVGUN	-----A-----C-----	1035
HIVSC	-----A-----G-----A-----	1136
HIVSBA	-----C-----T-----CA-----C-----	379
HIVSBB	-----C-----T-----CA-----C-----	382
HIVSBC	-----C-----T-----A-----	379
HIVBR014	-----C- GA-----G-----C-----A-----	1141
HIVBR021	-----C-----C-----	1145
HIVACH9	-----C-----	1013
HIVACP1	-----A-----	995
HIVJM	-----C-----G-----	1074
HIVWM	-----A-----	1040
HIVMA208	-----A-----C-----	1196
HIVP896	-----C-----	1142
HIVBWB11A	-----A-----T-----	752

CONSENSUS.C	taaaaTtTgcacca.tccTCA..GGAGGgacctaGAAATPACAACacACAGcTTTAATTGTaGAGGAGA	1095
HIVMA959	---C---A---CA---..---A---T---T---	1135
HIVMA960	---G---T---..---G---T---	1106
HIVD757	---G---T---..---G---T---	792
HIVD747	---G---T---..---G---T---	927
HIVD760	---G---T---..---G---T---	953
HIVNOF	---GC---AACT---..---TT---G---	757
HIVZAM20A	---T---A---..---..---	1175
HIVD1044	---G---C---..---..---	771
HIVD1024	---C---T---..---..---	1112
HIVD744	---T---T---..---G---T---	794
HIVD766	---GT---..---G---T---	767
HIVD808	---T---T---..---..---	815
HIVD868	---C---T---..---..---	806
HIVBR025	---AAG.CA---..---..---	1040
HIVSM145A	---G---A---..---T---T---	1136
HIVZAM18A	---AC-A-.CA---..---..---	1130
HIVDJ259A	CA---..---G---T---	1109
HIVDJ373A	---A---..---A---A---G---	1124
HIVSE364A	---A---..---C---T---	1133
HIVUG268A	---C---A-T-A-.C---..---T---	1151
CONSENSUS.D	TaattTTTaAacCa.tccTca..GGAGGgGACCCAGAAATtacAacACACagctTTAATTGTggAGGpGa	1091
HIVJY1	---C---G---..---A---T---TGC---	1172
HIVNDK	---C---G---..---T---TGC---	1130
HIVMAL	---TT---..---T---A---	1148
HIVELI	---AG---..---T---T---	1136
HIV22Z6	---..---..---T---T---	1136
HIVUG024	---GA---..---G---T---	1139
HIVUG021	---A---..---G---T---	1108
HIVUG038	---A---..---C---T---	1130
HIVUG046	---C---CA---G---..---C---	1142
HIVUG266A	---G---G---..---G---A---	1133
HIVUG269A	---G---G---..---G---A---	1141
HIVUG274A	---C---G---..---G---T---	1156
HIVSE365A	---A---C---C.CAT---..---A---T---G---T---CA---	1154
HIVCAR4020	---A---C---C.CAT---..---A---T---G---T---CA---	1229
CONSENSUS.E	LAatctttCaaCca.ccCTCA..GGAGGgGATCTagAAATPACAAtgCatcatTTAATTGTaGAGGGGA	1093
HIVTH966	---..---..---A---T---	1138
HIVTH975	---..---..---A---T---	1120
HIVTH022	---..---..---G---T---	1135
HIVTH243	---..---..---G---T---	1157
HIVTH006	---C---..---T---C---	1126
HIVTH011	---..---..---C---T---	1078
HIVCARELO	---T---..---C---A---	1163
HIVCAR4017	---CA---G---T---..---G---C---C---AT---AT---	1154
HIVCAR4039	---G---G---T---..---C---C---T---ACC---G---	1133
HIVCAR4071	A-T-TAAA-C---TA---..---C---T---	1163
HIVCARMBA	---G-T---..---AG---	1127
CONSENSUS.F	TAA?atTTTaa??ca.tCctCA..GGAGGgGACCTAGAAATPACAAtgCATagctTTAATTGTaGAGGAGA	1114
HIVBRA7944	---A---CT---..---..---	1049
HIVBZ163A	---A---TT---..---..---	1133
HIVBZ126A	---A---CT---..---..---	1138
CONSENSUS.G	taa?cTTTaaCTCA.tCctCA..GGAGGgGACCTAGAAATPACAacACATAGTTTCAATTGTaGAGGAGA	1108
HIVVI525A	---CT---..---..---	1172
HIVLBV217	---CT---..---O---	1163
HIVCAR4067	---C---C---..---G---A---	1151
CONSENSUS.H	TA?TTT??AaCCA.TCCTCA..GGAGGgGACAT?GAAATTA?AA??CA?A??TTTAA?TGT??AGGAGA	264
HIVCA13	---CAA-FA---..---A---A---T---CA---C-TG---C---GC---	341
HIVVI557	---ATP-CG---..---G---G---C---TG---T---CT---T---AG---	335
CONSENSUS.O	TgA7ATTCaatca?.agcAgT??GgtGgAGATgcaGAGgtAAcccaTtT?catTTTAACtGtCATGGaGA	1095
HIVANT70	---C---C---..---CT---A---C---	1136
HIVMVP5180	---A-T---GCAGA.-CT---..---A---G---A---A---	1157
HIVVADU	---A---GG-A-T.CA---..---A---AA---CTT---	1154
CONSENSUS.CPZ	?????ACA?TC?A???.?????GGAGATCC?GA?GT?A?A??CAT?GT?TAATTG???AGGAGA	596
SIVCP22GAB	ACATA--C--A-CAGA.GCCTCTGGA--A--G--C-C-CAT--AT--TGG--	1136
CONSENSUS.U	TAatcTTTaata?a.TCctCa..GGAGGgGacctaGAAATPACAACACA?A?cTTTAACtGtGAGGAGA	1056
HIVZAM184	---G---GC---AC---T---..---T---G---	1106
HIVCAR4081	---A---G---A---..---A---TT---C---C---	1142
HIVK124A	---T---ACC---G---..---C---C---C---GC---A---G---	1151
HIVZ3	---A---CTC---..---A---G---T---C---C---	1124

HIV ENV

CONSENSUS.A	atTtTtTcTatTgcaatatacatcagggccTgTtTt.AATagcaccTGG?a????????????????	1116
HIVSF1703	-----T-----TG-----A-----GGGGCA.....	1203
HIVU455	-----C-----T-----	1179
HIVZ321	-----G-----TTA.....	1191
HIVRW020	-----T-----C-----C-----GTCA.....	1148
HIVUG031	G-----T-----T-----TC-----TGGGAA.....	1034
HIVD687	-----AA-----G-----T-----	838
HIVNI	-----AAA-----G-----TAATAGC.....	1179
HIVRW009	-----T-----T-----C-----AGTAAAGA.....	1190
HIVUG275A	-----TG-----C-----A-----TGGC.....	1179
HIVUG273A	-----T-----G-----T-----A-----TAAA.....	1194
HIVV1191A	-----G-----G-----A-----	1191
HIVDJ264A	-----G-----AA-----G-----TAACAGCTTA.....	1176
HIVDJ263A	-----G-----AA-----	1176
HIVDJ258A	-----A-----T-----	1170
HIVCARGAN	-----G-----A-----GAATGAG.....	1206
HIVCARSA5	-----C-----AA-----TGGGGGGCCA.....	1200
HIVCAR4054	-----A-----AT-----C-----T-----TAACATT.....	1185
HIVCAR286A	-G-----TCTA-TGC-ATACATCA-----G-----ACTGACACT.....	1245
HIVCAR4023	-C-----C-----G-----A-----A-----T-----A-----	1233
HIVCAR423A	-----AAA-----A-----A-----	1215
HIVKENYA	-----T-----T-----ACCAACATGCGAGAGTCA.....	1194
CONSENSUS.B	aTtTtTCTactTgaAt?CaacAcaacTgTtTt.aatagtactTGG????????????????????	1153
HIVJRCSP	-----T-----	1173
HIVJRPL	-----T-----AAT.....	1173
HIVALA1	-----T-----A-----	1173
HIVBRVA	-----T-----A-----TATAGG.....	1185
HIVJH3	-----T-----C-----TTGTCT.....	1221
HIVBAL1	-----T-----	1191
HIVYU2	-----T-----	1164
HIVM	-----A-----T-----C-----AATGGT.....	1200
HIVHXB2R	-----T-----T-----	1188
HIVLAI	-----T-----T-----	1203
HIVNL43	-----T-----T-----	1182
HIVMFA	-----T-----T-----	1182
HIVCAM1	-----A-----G-----C-----CTGTTT.....	1189
HEVNY5CG	-----GA-----G-----CTGTTT.....	1170
HIVADA	-----T-----AATTTT.....	1176
HIVJFL	-----T-----G-----CTGTCC.....	1176
HIVSIM184	-----T-----A-----G-----	1165
HIVD31	-----T-----G-----	1179
HIVSF162	-----T-----G-----T-----	1173
HIVBCSG3C	-----GT-----GCTGGG.....	1194
HIVOX1	-----A-----T-----	1194
HIVSF33	-----A-----A-----AATGAA.....	1188
HIVCDC4	-----T-----G-----CG-----	1212
HIVSF2	-----A-----A-----A-----AGGTTA.....	1197
HIVSF2B13	-----A-----G-----AATCCA.....	1194
HIVHAN	-----T-----A-----	1191
HIVRF	-----A-----	1210
HIVMA12	-----T-----	1167
HIVTB132	-----Y-----A-----A-----	1137
HIVBR020	-----A-----C-----ACA.....	1178
HIVTH014	-----A-----	1154
HIVTH026	-----A-----T-----AATAAT.....	1160
HIVUS711	-----A-----ACAACATTG.....	1178
HIVUS712	-----T-----AATTCACA.....	1190
HIVUS715	-----T-----A-----AATGGTACT.....	1196
HIVUS716	-----T-----TATAATAAT.....	1172
HIVHA593	-----C-----A-----AATAGTACT.....	1253
HIVHA594	-----A-----T-----A-----AATGATACT.....	1190
HIVHA596	G-----A-----T-----A-----AATGATACT.....	1193
HIVHA599	-----C-----A-----T-----A-----GAGTTTAAT.....	1199
HIVHA651	-----CA-----A-----C-----AATAGCACT.....	1175
HIVRJS	-----T-----G-----G-----ATGTCT.....	1191
HIVGUN	G-----A-----A-----G-----CTGGTT.....	1084
HIVSC	-----T-----G-----	1179
HIVSBA	-----A-----A-----ACTTGG.....	428
HIVSBB	-----CA-----A-----A-----AATGAT.....	431
HIVSBC	-----T-----TATAGT.....	428
HIVBR014	-----A-----G-----G-----ATGTTG.....	1190
HIVBR021	-----A-----G-----TTG.....	1191
HIVACH9	-----T-----A-----T-----	1056
HIVACP1	-----GA-----CG-----A-----TCGGAT.....	1044
HIVJM	-----G-----T-----G-----CAACTTAATAAAAGTGATGGTAATTGG.....	1143
HIVNM	-----T-----G-----TC-----	1083
HIVMA208	-----A-----C-----ACTTGGGAAT.....	1248
HIVP896	-----A-----G-----	1185
HIVBNB11A	-----T-----	795

CONSENSUS.C	ATTTCCTATGCAATACATCAaaccTGTTt.AaTagtAcatac?????????a??ccct???	1142
HIVMA959	-----GGG-----AC-TT-----	1178
HIVMA960	-----T-----GA-----C-----AT-----	1154
HIVD757	-----G-----G-----TG-----	841
HIVD747	-----G-----G-----TG-----	976
HIVD760	-----GG-----C-----G-----TG-----	1002
HIVNOF	-----A-----G-----TC-GCA-----	794
HIVZAM20A	-----C-----T-----AT--C--TTAGGC-----TTGTT--	1230
HIVD1044	-----T-----C-----AATAGTACATAC-TG-T-----	832
HIVD1024	-----C-----G-----TG-----	1161
HIVD744	-----T-----GG-----G-----TG-A-----	843
HIVD766	-----T-----GG-----G-----TG-A-----	816
HIVD808	-----T-----GG-----G-----TG-----	864
HIVD868	-----G-----G-----TG-----	852
HIVBRO25	-----G-----G-----CA-----	1089
HIVSM145A	-----GG-----G-----AT-----	1185
HIVZAM18A	-----C-----G-----T-----AG-----	1179
HIVDJ259A	-----C-----G-----G-----	1243
HIVDJ373A	-----T-----GC-T-----GA-T--ATTACAAAT....GCA--CAIT.	1182
HIVSE364A	-----G-----G-----G-----ACACATACT.	1182
HIVUG268A	-----G-----G-----GTG-TAATAAT.	1200
CONSENSUS.D	atTtTTCaCTGcaATaCATCAagAcTgTTT.AaTagtAcatac?????????a??ccct???	1134
HIVJY1	C--C-----T-----A-----CAG-----	1215
HIVMDK	-----T-----A-----CAG-----	1173
HIVMAL	-----C-----T-----G-----	1194
HIVELI	-----C-----T-----G-----	1179
HIVZ226	-----C-----T-----G-----	1179
HIVUG024	-----A-----T-----CA--GTGTAGGGGAG.	1180
HIVUG021	-----G-----G-----C-A--GTGTATGGACATCA.	1160
HIVUG038	-----CA-----G-----GGGAGT.	1179
HIVUG046	-----G-----G-----G-----T-AT.	1185
HIVUG266A	-----G-----G-----G-----GTGTATGGAGG.	1182
HIVUG269A	-----G-----G-----G-----GT-AT.	1184
HIVUG274A	-----A-----A-----A--GTG-ATGGGAAAATTAT.	1211
HIVSE365A	-----G-----G-----AAAACA.	1203
HIVCAR4020	-----CA-----CA--A-----	1272
CONSENSUS.E	ATTTCCTATTGCAATACAAaCAaaccTGTTt.AaTagtActtoga?a??a???	1139
HIVTH966	-----G-----C-----A-----C-TTGU-----	1187
HIVTH975	-----C-----G-----A-----T-GG-----	1169
HIVTH022	-----CG-----G-----A-----C-T-AA-----	1184
HIVTH243	-----CG-----G-----A-----C-T-GG-----	1206
HIVTH006	-----C-----T-----G-A--CAT-GAGGGGCCCT.	1178
HIVTH011	-----C-----A-----A-----C-T-GG-----	1127
HIVCAR010	-----C-----G-----A-----GA-AT-----	1212
HIVCAR4017	-----C-----G-----A-----G-AT-----	1203
HIVCAR4039	-----G-----G-----G-----GG-AC-----	1182
HIVCAR4071	-----G-----T-----GCACA-----	1206
HIVCARMBA	-----G-----G-----GA-ATG-----	1176
CONSENSUS.F	atTTTTCTACTGCAATACa?cagaaCTGTTT.AaTagAcac?g?at?????????	1155
HIVBRA7944	-----GT--A-----A-----C-TTGU-----	1088
HIVB2163A	-----T-----G-----A-----A-G--CC-----	1179
HIVB2126A	-----T-----G-----A-----A-G--CC-----	1184
CONSENSUS.G	atTTTTCTATTGtAATACATCAgacTGTTt.AATaata?t?t?????????	1148
HIVVI525A	-----T-----G-----AG-----G--C-TGGTTGTTT.	1221
HIVLBV217	-----G-----G-----G-----G-A-ATTATCT.	1212
HIVCAR4067	-----C-----G-----G-----	1168
CONSENSUS.H	ATTTCCTATTG?AATACA?CAAAA?TGTTT.AA?AGTAC?TGG?????????????	302
HIVCA13	-----T-----A-----C-----T-----C-----ACGACTACACC.	396
HIVVI557	-----C-----T-----T-----C-----T-----GTT-----	381
CONSENSUS.O	aTTCCTTTATTGtAACACatcT?ggatTGTTT.AAttATAccTTTcatgt?????????????????	1143
HIVANT70	-----G-----AA-----C-----T-----ATC-----	1185
HIVMVP5180	-----G-----G-----C-----T-----ATC-----	1203
HIVVAU	-----AA-C--C-----C-----CAAGAAGAAATGACCAATAAC	1224
CONSENSUS.CP2	ATT?TT?TATTGT?ATA??A??C?????TT?A?TG?C?????????????????	618
SIVCP2GAB	---T--C---A---CT-GC-AAATA--T--C--A-----	1173
CONSENSUS.U	atTTTTCTATGCAATACATCAaaccTGTTT.AaTaGtat?Tggaa????	1101
HIVZAM184	-----G-----G-----CA-AC--TCCT.	1155
HIVCAR4081	-----CG-T-----C-GT-----GGA.	1191
HIVK124A	-----C-----A-----G-A--GG-----	1197
HIVZJ	G-----C-----G-G-----C-G--T-----	1167

HTY1 ENV

CONSENSUS.A	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619</
-------------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

CONSENSUS.C	aatgggtaca?aa?c????????????????????.....aat????a?????.....?????????	1158
HIVMA959	-----A-----GA-GG-----TCA-CT-----	1202
HIVMA960	-----GT-G-----GAT-CAGAA-----	1181
HIVD757	-----AA-GG-----TCA-----	862
HIVD747	-----AA-G-----TCA-GC-----	1000
HIVD760	-----AA-G-----TTA-AC-----	1026
HIVNOF	-----C-CA-AA-TT-----GAA-GT-----	818
HIVZAM20A	-----GA-----AATPCA-----	1251
HIVD1044	-----AC-G-----TCA-CC-----	856
HIVD1024	-----AA-G-----TCA-AC-----	1185
HIVD744	-----TAT-TGTTTAAATGGTACATACATGTT-----AGT-CA-----	889
HIVD766	-----AA-G-----TCA-GC-----	840
HIVD808	-----AA-G-----TCA-GC-----	888
HIVD868	-----AA-G-----TCA-GC-----	876
HIVBR025	-----A-----GT-----	1100
HIVSM145A	-----AAC-GG-G-GAGTCAAAC-----	1209
HIVZAM18A	-----A-----AGT-CA-----	1197
HIVDJ259A	-----CAG-GTATC-CA-----	1158
HIVDJ373A	-----AA-T-GAGT-CA-----	1197
HIVSE364A	-----GCAAA-T-----	1191
HIVUG268A	-----A-----C-----	1209
CONSENSUS.D	aata?taca?ca????????????????????.....aatagtaca???.....?????????	1153
HIVJY1	-----G-GTA-TGGAAAT-----GA-----	1242
HIVNDK	-----CAG-T-----	1191
HIVMAL	-----A-GCTG-AGACTAAGT-----C-----	1224
HIVELI	-----T-GTG-----TGGAAT-----T-----GAGTCAAA	1214
HIVZ236	-----G-T-GTAA-TCAGAAATGGGCT-----GAGTCAGA	1220
HIVUG024	-----A-TCT-TCA-----T-C-AC-----	1212
HIVUG021	-----G-G-----AT	1171
HIVUG038	-----GA-T-----TAT-----	1203
HIVUG046	-----CAT-GGT-----	1203
HIVUG266A	-----GA-TATC-TCA-----	1206
HIVUG269A	-----CAT-G-TTAAATAGTACACAGGGA-----GAG-GT-----	1223
HIVUG274A	-----TAT-----	1229
HIVSE365A	-----ATCAGGT-----	1212
HIVCAR4020	-----GA-A-GGAGAAATGTCAGAACAG-----C-A-ATACA-----GAGTCA	1320
CONSENSUS.E	AATgaaACcatg.....g?gg?gt?c.....AATggcAct.....	1166
HIVTH966	-----A-----G-G-G-G-----A-----	1217
HIVTH975	-----G-----A-G-G-G-----A-GC-----	1199
HIVTH022	-----C-----C-G-G-G-----C-----	1214
HIVTN243	-----AAT-G-G-----	1236
HIVTH006	-----A-C-GGA-----G-CCTAA-----CACT-----	1208
HIVTH011	-----AA-G-G-----A-----	1157
HIVCARELO	-----A-----AA-C-----A-A-----	1239
HIVCAR4017	-----A-----AAAA-GT-----A-----	1233
HIVCAR4039	-----A-----A-C-----T-A-----	1209
HIVCAR4071	-----G-----ACA-----	1227
HIVCARMBA	-----A-----A-A-A-CT-----	1206
CONSENSUS.F	AAtaacact????????????????????.....	1164
HIVBZ163A	-----GG-----	1188
HIVBZ126A	-----GG-----	1193
CONSENSUS.G	...???AAtaatAct????????????????????.....aat???ac?????????	1163
HIVVI525A	-----G-----GGCTGGTTAATAGTACTTGG-----GAT-ACGA-----	1263
HIVLBV217	-----C-----A-GCACCATCT-----AGC-GT-----	1239
HIVCAR4067	-----GC-----AAGACTACAAATGATCTAAGACA-----GAT-T-----	1230
CONSENSUS.H	AA?GGA?T?P?????????.....AAT?A?A?P?A.....AATGG?AC?.....	320
HIVCA13	-----TG-C-GC-ATGTGAGC-----A-C-CAGA-----T-T-----	435
HIVVI557	-----CA-T-CA-CA-----G-T-TCTC-----C-A-----	414
CONSENSUS.O	???????AAt?gtAcc?????????tg?????????aag?ta?cAATgt?ActaAt??t.....	1171
HIVANT70	-----CG-A-----ACC-T-----AGT-T-GT-----T-G-C-AGG-----	1224
HIVMP5180	-----CT-----AAAGTCCGGA-CCAGGAGATCA-----GG-G-----AG-C-----	1254
HIVVAU	AAGATC-----T-----T-----A-T-GC-----AA-----	1254
CONSENSUS.CPZ	??CAT?ACA.....	624
SIVCPZGAB	-----AA-T-----	1182
CONSENSUS.U	AAttgtA?c?????????c?A?????????aag?cact?????????.....?????????	1119
HIVZAM184	-----A-----CA-----GA-AGTAATTCA-----	1176
HIVCAR4081	-----G-AC-G-GGGCATGAG-C-----A-----	1221
HIVK124A	-----CA-G-----TG-AAAGGTA-----AATGG	1226
HIVZ3	-----G-C-----GGGAT-----AAG-----TG-AGCACT...GAGAGTAA	1208

HIV1 ENV

CONSENSUS A	?????????aatgacactataact???????????????	ctccaa	TGcAgaATAAagcaaATT	1170
HIVSF1703	GAGTCA-----T-	---C-	-----	1284
HIVU455	GGGCCA---G-	---C-	---T-----	1248
HIV2321	TGACACGGTA-----T-	---C-	-----	1269
HIVRW020	ACGTCA-----	-TAC-	-----	1217
HIVUG031	GAGACA---G-	---C-	---A-----	1109
HIVD687	GCA---A----	-A----	-----	904
HIVNI	GGGTCA-----	---	-----	1245
HIVRW009	G-ATTA=A-----A	--C-	-----	1256
HIVUG275A	G-A-----	---	-----	1254
HIVUG273A	---T-----	-T---	-----	1266
HIVV1191A	-----	---C-	-----	1263
HIVDJ264A	-----A-----	---	-----	1239
HIVDJ263A	AAATCACACAG A-G-A-----	-----G-	-----	1263
HIVDJ258A	AATGACACA-----A-----	-----	-----	1248
HIVCARGAN	-----	---G	-----	1272
HIVCAR5AS	-----	---C-	-----	1269
HIVCAR4054	GAAACA---G-----C	---C-	-----	1254
HIVCAR286A	ACTGGG-----GAC-TCATAACT	---C-	-----	1317
HIVCAR4023	GGAGAATTTC---GC-TC	--A-C-	-----A G-	1320
HIVCAR423A	ACTGAAGAA-----GGAGT-AGAGGACCCTATAGTT	---C-	---G-----	1287
HIVKENYA	-----	---C-	---A-----	1245
CONSENSUS B	?????????aatgacac?atcaca	CTcCca	TGcAgaATAAaacAaatT	1205
HIVJRCSE	GAAGGA-----C---T-	---	-----	1245
HIVJRFL	GAAGGA---T-----	---	-----	1242
HIVLA1	GAAGGA-----AT-----	---	-----	1248
HIVBRVA	GAAGGA---AG-C A-----	---	-----	1251
HIVJH3	GGAGGA-----A-----C	---	-----	1305
HIVBAL1	GTAGAA---A---A---	---	-----	1263
HIVYU2	GAAGGA-----AT-----	---	---T-----	1230
HIVMN	GGGTCA---A---AT-----	-T-A-	---A-----	1269
HIVHXB2R	GAAGGA-G-----A---C	---	-----	1269
HIVLAI	GAAGGA-G-----A---	---	-----T-	1284
HIVNL43	GAAGGA-G-----A---	---	-----T-	1263
HIVMFA	GAAGGA-G-----A---	---	-----T-	1263
HIVCAM1	AAATAACACTG-AAGA-AT---T-	-T---	-----	1266
HIVNY5CG	A GATAAC-----A---T---T	---	-----	1245
HIVADA	T GAAGGA-----T-----	---	---T-----	1257
HIVJFL	AAAAGACACT-GA--AC-T---T-	---	-----	1254
HIVSIH104	AGGATCAAAAT---A GA-AT -	---T-	-----	1263
HIVD31	---GA---T-----	---	-----	1245
HIVSF162	---GA---T-----	---	-----	1239
HIVBCSG3C	AGATGACACTGGA-CA-AT-----	---	-----	1275
HIVQY1	G-A-TA -T-----	---	---T-----	1260
HIVSF33	GGA---A---T-----	---	---T-----	1251
HIVCDC4	AAAAGACACTGGA---TT-----	---	-----	1302
HIVSF2	-----A---T-----	---	---T-----	1260
HIVSF2B13	---C-A---T-----	---	---T-----	1254
HIVHAN	-----T---T-----	-G---	---T-----	1260
HIVRF	GGAGGA-----A-----	---	-----	1290
HIVJMJ2	CTC-----	---	-----	1239
HIVTB132	GAGG-AAGA-AT-----	---	-----	1206
HIVBR020	GGG A - A---	---	---T-----	1250
HIVTH014	GGA---G---T-----G	-P	-----	1226
HIVTH026	ACAGGAT-AG---T---G	-T	-----	1226
HIVUS711	ACTGCA GA AC-----	-A	-----G-	1244
HIVUS712	---GA-AT-----	-A	-----	1247
HIVUS715	---C-A - GT-	---	---G---	1253
HIVUS716	-CACA-TA-----	---A-	---G-----	1235
HIVHA593	G-A-G---A---T	---	---T-----	1313
HIVHA594	---T-----	---	-----	1244
HIVHA596	---T---T-	---	-----	1250
HIVHA599	TAAATCGCACTGGA-GA AC--	-	---T-G-	1286
HIVHA651	---AT-AC-	---	-----	1238
HIVRJS	ACAGGGTTA AT--A	-G	-----	1254
HIVGUN	AAAGGTCA -T--T-----	---AG	-----G-	1153
HIVSC	GGAGGA---A---C	---	-----G	1251
HIVSBA	GAAGGAAAT--CACT-AT-----	-T-	-----	497
HIVSBB	AGCTGAAGAA -T-TAC--T	-A	-----	512
HIVSBC	GGAGGA-AAT-----	---	-----	491
HIVBR014	A-AAT -G	---	---T---	1256
HIVBR021	-P-----	---	-----	1251
HIVACH9	CACCATAGGA - - -C--TGT	---	-----	1143
HIVACP1	GGAGCA T AC	-	---T-	1127
HIVJM	AGGAAATATC CA GAGAT	-A	-----G	1224
HIVMM	GGA T G	---	---G---	1158
HIVMA208	AACTGAAGGA AG A -	A	-----T	1329
HIVP896	GAAGGA - TA -	A	-----	1257
HIVBWR11A	GGAA A AT T	---	-----	876

CONSENSUS.C	?????????????caa?cATcAcA.	CTcCcA.	TgcAGaATAAAaCAaaTT	1192
HIVMA959	TT-G--A-C.			1238
HIVMA960	AGTAATT--C-----			1223
HIVD757	AACT--C-----	A--A-		901
HIVD747	T--C-----	A-T-		1036
HIVD760	T--C-----	A--		1062
HIVNOF	GATCCAAACG--C-----		G-----	863
HIVZAM20A	CCCCAA--A-----T.	A--	-----C-----	1290
HIVD1044	TT-A-----	A--		892
HIVD1024	T-C-C-----	A--T-	T-T-----	1221
HIVD744	AGTAATG--C-----	A--		930
HIVD766	T--T-----	A--	-----G-----	876
HIVD808	T--T-----	A--	-----G-----	924
HIVD868	T--C-----	A--T-	-----G-----	912
HIVSM145A	AAC-A-----	A--		1245
HIVZAM18A	TATAATCCTAATGAT-CA-			1245
HIVDJ259A	AATGA--A-----	T-A-		1197
HIVDJ373A	GATG--A-----	T-A-		1236
HIVSE364A	--A-----	A--A-	-----G-----	1224
HIVUG268A	T-----			1242
CONSENSUS.D	???????????aata??acaaTcAcA.	CTcCcA.	TGCAGaATAAAaCAaaATT	1190
HIVJY1	TTAAATAGTG-AGGA-----A-			1290
HIVNDK	GGGTTCaAT--GGC--G-----			1239
HIVMAL	GAGTCA-C-GGT-GT-----			1269
HIVBLI	TAATAGCACA--C-CA-AC-----	A--		1263
HIVZ236	TGATAAA--C-----	A--		1260
HIVUG024	GGGGTA--CT-AC-----	A--		1257
HIVUG021	AGGGCCA--GGG-----			1217
HIVUG038	AC-----A-----	A--	-----G-----	1242
HIVUG046	GGGGTAAT--C-GT-----A-			1251
HIVUG266A	GGAGTAAATG-G-GCGT-----	A--		1254
HIVUG269A	GAA-AC-----			1259
HIVUG274A	GGGGGA--GAG-----G-	A--		1274
HIVSE365A	GAT-AT--T-----		-----G-----	1248
HIVCAR4020	GAA-AT-----		-----G-----	1356
CONSENSUS.E	?????????????ATCacA.	CTTCCA.	TGCAGaATAAAaCAaaATT	1196
HIVTH966	---T-----			1247
HIVTH975	GATACACCT--GT-----			1238
HIVTH022	---T-----			1244
HIVTH243	---T-----			1266
HIVTH006	---T-----			1238
HIVTH011	---T-----			1187
HIVCARELO	AACCTTAAGCAAT-----	CA--		1281
HIVCAR4017	AACCTCAACATG-----	G--G		1275
HIVCAR4039	AACCCCAAGAAAT-----	G--G--A-		1251
HIVCAR4071	-----	G--G--G-		1257
HIVCARMEA	-----	G--		1236
CONSENSUS.F	?????????????????????	CTcc7a???	TGTAgAATAAaGACaa?Tt	1186
HIVBZ163A	ATCACT.	---T---	C-----A---A-G	1218
HIVBZ126A	ATCACT.	---T---	C-C-----A--	1223
CONSENSUS.G	???????aa????atta??ac???????	CTcCcA.	TGtAgAATAAAACAAATT	1194
HIVVI525A	--AGAT--CA-----	G-A-	C-A-----	1299
HIVLBV217	--GAAT--C-CA-----			1275
HIVCAR4067	---C-CA-----	A--		1260
CONSENSUS.H	A??A?A???????	CTGC?A.	TGCAGaATAAAACAAATT	346
HIVCA13	---TT-C-----	A--		465
HIVVI557	CA-T-ATTGA.	---C---		450
CONSENSUS.O	a??AATGgcACT?A?????	aTACCT.	TGCAGgtTaAgACAagTg	1206
HIVANT70	AC-----	C-----	AAC-G-----G--	1260
HIVMVP5180	AA---T---		A-----GC-A	1290
HIVVAU	GC-----C-GCCA.		-----G-----A	1296
CONSENSUS.CPZ	AA?GG?????T?ATA.	???C??	TGCAGaATTA??CAGATA	649
SIVCPZQAB	---T---CATTA-A---	CTG-CA.	GA-----	1221
CONSENSUS.U	?????????????ga?A??ataaca.	CTCCCa.	TGca?aATAAAACAAaTt	1151
HIVZAM184	GACTTAA-C-TC-C-----		G-----	1215
HIVCAR4081	-----T-----	G--	A--G-----	1251
HIVK124A	GACAAAGGTAAAA--G-CA--C--		T-A-----	1275
HIVZ3	TTGCACT.-----GA-AT--T--		GG-----G-G	1251

HIV1 ENV

CONSENSUS.A	gTaaatATgTGGCagagagtaGgacaA.....GCaAtgTATgCCcCtCCcATccaaA...gGagtaaT	1229
HIVSF1703	A-----	1343
HIVU455	A-----	1307
HIVZ321	-----A-----G-----	1328
HIVRW020	A-----A-----C-----T-----	1276
HIVUG031	A-----A-----C-----T-----A-----	1168
HIVD687	-----A-----G-----A-C-----	963
HIVNT	-----A-----A-----	1304
HIVRW009	A-----T-----GAC-----	1315
HIVUG275A	A-----C-----A-----T-----A-----	1313
HIVUG273A	A-----G-----A-----	1325
HIVV1191A	GA-----A-G-----	1323
HIVDJ264A	-----A-----G-----C-----A-----	1298
HIVDJ263A	-----A-----T-----TC-----A-----	1322
HIVDJ258A	-----A-----A-----TC-----A-----	1307
HIVCARGAN	-----C-----G-----AC-----A-----	1331
HIVCARSA5	GG-----A-----CG-----	1328
HIVCAR4054	A-----GG-----A-T-----	1313
HIVCAR286A	-----A-----A-----A-----	1376
HIVCAR4023	-----T-----C-----A-----	1379
HIVCAR423A	-----C-----A-----A-----	1346
HIVKENYA	A-----C-----C-----A-----	1304
CONSENSUS.B	aTAAAcacgTGGCaggaagTAgGaaA.....GCaATGtatGccCctCCCatcaga...GGAcacaaT	1264
HIVJRCSP	-----G-----T-----T-A-----	1304
HIVJRF1	-----	1301
HIVLA1	G-----GA-----	1307
HIVBRVA	-----	1310
HIVJH3	-----GA-----	1364
HIVBAL1	-----G-----	1322
HIVYU2	-----T-----	1289
HIVMN	-----TGA-----	1328
HIVHKB2R	-----A-----T-----	1328
HIVLAI	-----C-----	1343
HIVNL43	-----T-----	1322
HIVHFA	-----T-----	1322
HIVCAM1	-----G-----C-----AC-----	1325
HIVNY5CG	-----G-----T-----	1304
HIVADA	-----T-----	1316
HIVJFL	-----G-----AC-----	1313
HIVSIM184	-----G-----A-----GA-----T-----	1322
HIVD31	-----T-----	1304
HIVSF162	-----G-----	1298
HIVBCSG3C	-----C-----	1334
HIVQY1	G-----T-----	1319
HIVSF33	-----GA-----	1310
HIVCDC4	-----G-----TT-----T-----A-----T-----	1361
HIVSF2	-----TG-----	1319
HIVSF2B13	-----TG-----	1313
HIVHAN	-----TG-----T-----	1319
HIVRF	G-----T-----	1349
HIVWMJ2	-----G-----C-----CA-----	1298
HIVTB132	G-----G-----R-----T-----A-----G-----	1265
HIVBR020	G-----G-----C-----T-----AO-----A-----	1309
HIVTH014	G-----	1285
HIVTH026	G-----A-----	1285
HIVUS711	-----G-----AA-----TT-----	1303
HIVUS712	-----C-----	1306
HIVUS715	-----A-----	1312
HIVUS716	-----C-----	1294
HIVHA593	-----A-----TGA-----	1372
HIVHA594	G-----C-----G-----C-----	1303
HIVHA596	G-----T-----A-----T-----	1309
HIVHA599	G-----AG-----C-----GC-----C-----G-----	1345
HIVHA651	-----C-----T-----GAG-----	1297
HIVRJS	-----G-----G-----GA-----	1313
HIVGUN	-----GC-----CA-----	1212
HIVSC	-----A-----G-----	1310
HIVSBA	G-----	556
HIVSBB	O-----T-----	571
HIVSBC	-----	550
HIVBR014	-----GA-----GAC-----	1315
HIVBR021	-----C-----C-----	1310
HIVACH9	-----G-----	1202
HIVACP1	-----C-----AC-----	1166
HIVJM	O-----	1283
HIVNM	-----GC-----	1217
HIVMA208	-----C-----A-----CA-----A-T-----	1388
HIVP896	-----T-----A-----C-----	1316
HIVBWB11A	-----T-----	935

CONSENSUS.C	aTAAACaTGTCgCagGagGTaGdAcya???. gcaATgTATgCCCCCccATtTgaa gGaAacAT	1251
HIVMA959	-----A-----C-----	1297
HIVMA960	-----A-----C-----	1282
HIVD757	-----A-----C-----	960
HIVD747	-----A-----C-----	1095
HIVD760	G-G--C-----AT--CC--A-G--	1121
HIVNOF	-----GACG--AGC-----	925
HIVZAM20A	-----AA--T--C-----C-----	1349
HIVD1044	-----AA-----	951
HIVD1024	G-----	1280
HIVD744	-----T-----	989
HIVD766	-----	935
HIVD808	-----	983
HIVD868	G-----	971
HIVSN145A	-----A-----A-----	1304
HIVZAM18A	-----C--A-----C-----	1304
HIVD7259A	-----A-----TA--	1256
HIVD7373A	-----C-----TA--	1295
HIVS2364A	-----GA-----C-----	1283
HIVUG269A	-----G-----A-----A-----	1301
CONSENSUS.D	aTAAacaTGtggcaggagtagGaaaA GCaATgTATgCCCCCccATcgaA gGactaAT	1249
HIVJY1	-----T-----	1349
HIVMDK	G--T--A-----	1298
HIVMAL	-----T--AA-AC-----T-----C-----G-C--	1328
HIVELI	-----G--GT-GCA-C--A-G--	1319
HIV2226	-----A-----A-----	1319
HIVUG024	G-----C-----TC--	1316
HIVUG021	-----T-----A-----T-----A-----	1276
HIVUG038	-----C-----A-----A-----	1301
HIVUG046	-----C-----T-----	1310
HIVUG266A	-----C-----C-----T-----	1313
HIVUG269A	-----A-----T-----GA--	1318
HIVUG274A	-----A-----TAT--A-----	1333
HIVS2365A	-----C-----A-----T-G--G-G--	1307
HIVCAR4020	-----GA-----A-----C-----G-----	1415
CONSENSUS.E	atAAAcATGTGgCAGggAGtAGGAcAA GCaATGTatgcTCCTCCCaTcagt GgAA?AaT	1254
HIVTH266	-----C-----T-----G-----	1306
HIVTH975	-----C-----A-----	1297
HIVTH022	-----C-----T-G--G-----	1303
HIVTH243	-----C-----A-----	1325
HIVTH006	G-----T-----C-----G-----	1297
HIVTH011	-----G-----	1246
HIVCARELO	GG-----G-----A-CT-G--	1340
HIVCAR4017	G-----A-----C-----GT-G--	1334
HIVCAR4039	G-----A-----A-----GT--	1310
HIVCAR4071	-----G-----A-----GA--A-GC--	1316
HIVCARMBE	G-----AA-----G-----GT--	1295
CONSENSUS.F	GTAAACATGTGgCAG??aGT?GGAcaA GCaATGTATGCCcCTCCCaTtTgaa Ggaaa?AT	1241
HIVB2163A	-----GA--G-----CT--A--C--	1277
HIVB2126A	-----GA--A-----G-----A-----C--	1282
CONSENSUS.G	GTgAgBaTGTGGCAGAgagtgGGAcAA GCaATGTATgCCCCCccATtGCA??GGAACAT	1253
HIVVI525A	-----A-----C-----GCA--	1361
HIVLBV217	-----A-----G-----	1334
HIVCAR4067	-----A-----ACA-----AA--	1319
CONSENSUS.H	?TAA?CATGTGGCAGAGAG?AGGACAA GCaATGTATGCCcCTCCCATCAA? GGAaA?AT	400
HIVCA13	G--G-----C-----T-----G-----	524
HIVVI557	A--A-----T-----A-----A-----	509
CONSENSUS.O	GTAAAGatcaTGGATgagGGGAgggTCG gGactcTATGCaCCTCCCATCca? GGcAAccT	1264
HIVAM70	-----G-----A-----CA-----AA-----T--	1319
HIVMPS180	-----GAC-----A-----A-----A--A-----CC--C--T--	1349
HIVVAU	-----GAC-----A-----T-----C-----A-----	1355
CONSENSUS.CP2	GTAA?T?T?TGG?T?A?TGT?T?AA?A GG?AT?TA?T?A?T?CCCT?A?T?AGA GG?AA?T?	683
SIVCP2QAB	-----G-TCA--ATG-GG--AGG--G--	1280
CONSENSUS.U	?TAA?cAtgTGGCAG??AGTAGGAC?A GCaATGTATGCCcCTCCCATtTg?A GGagtaAT	1203
HIVZAM184	A--A-----AG--G-----T-C--	1274
HIVCAR4081	G--GG-----AA-----A-----CCC--	1310
HIVK124A	A--A-----GA-----G-----T-A-----C-C--	1334
HIVZ3	G--GA-CA-----GG-----A-----T--C-A-----GAC--	1310

HIV1 ENV

CONSENSUS .A	AAGgTgTgaacCAAAcATcACaGGacTtcTatTaACaAGaGATGG7Ggg??c???	1279
HIVSF1703	--A--AT--T--T--T--GA--	1396
HIVU455	-----G-TAC--	1360
HIV2321	--A--T-----A--T-TGG--	1381
HIVRW020	-----A--	1326
HIVUG031	-----C-----G--T-TGGG--	1221
HIVD687	-----A--T--	1013
HIVNI	--A--T-----A--	1354
HIVRW009	--C--T-----T-AAA-AAT--	1371
HIVUG275A	-----T-----A--T--GG-AAT--	1369
HIVUG273A	-----T-----T-AAA-GGG--	1381
HIVYI191A	--A-----T-A--T-TTGGGAAT--	1378
HIVDJ264A	-----A--AG-AAT--	1354
HIVDJ263A	-----A-----C--A--AG-AAT--	1373
HIVDJ258A	-----A--AG-AAT--	1363
HIVCARGAN	-----G-----C-G--T-TT-AAT--	1387
HIVCAR3AS	--AC--AG-----A--GGG--	1381
HIVCAR4054	-----C-----C--T-GTA--	1366
HIVCAR286A	-----T-----CTG--	1429
HIVCAR4023	--T--A-C-----A-N--T--GA-GAT--	1435
HIVCAR423A	--A-----C-----T--AAG--	1399
HIVKENYA	--A--T-----A--T--AAG-AAT--	1360
CONSENSUS .B	tAggTgttcaTCaAaAtATtACaGGgcTgcTatTaACaAGaGATgGcGGTaa??a??????????????...	1317
HIVJRC5F	-----A--	1357
HIVJRF1	-----TTA-TGAG--	1360
HIVLA1	--C-----	1357
HIVBRVA	--A-----TA-C--	1366
HIVJH3	--C-----G-TA-CCAG--	1423
HIVBAL1	-----CTTG-GGAC--	1381
HIVYU2	-----GG-CACG--	1348
HIVMN	-----A--GG-CACG--GACACG--	1393
HIVHXB2R	-----TAGCBAAC--	1387
HIVLAI	-----TA-CAAC--	1402
HIVNL43	-----T--TA-CAAC--	1381
HIVNFA	-----GA-CACG--	1381
HIVCAM1	--C-----GAGGCGAA--	1384
HIVNY5CG	--C-----G-TA-AGAG--	1363
HIVADA	-----AA--A-CT--	1369
HIVJFL	-----A--AG-G--	1369
HIVSIM184	-----T-C--G--	1375
HIVD31	-----AA-TAAGGAC--	1366
HIVSF162	--C-----A--AG-GATCAGT--	1360
HIVBC5G3C	--A-----CGAGCATAACC--	1396
HIVQY1	-----G--	1372
HIVSP33	-----GGTG-CAAG--	1369
HIVCDC4	-----GGTG-G--	1417
HIVSP2	--T-----CA--	1372
HIVSF2B13	--T-----TA--	1366
HIVHAN	-----A--TG-T--	1375
HIVRF	--A--AT-----A--G--G-AG-TACAAC--	1411
HIVMMJ2	-----CA-C-----A--CC-A--	1348
HIVTB132	--A-C-----C--A--A--G--G--	1359
HIVTH014	-----T-----	1335
HIVTH026	-----C-----T-----TA-TGAC--	1344
HIVUS711	--G-----GT--	1356
HIVUS712	-----TAA-C--	1362
HIVUS715	--C-----A--CAGT--	1368
HIVUS716	--A-----CA--	1347
HIVHA593	-----CA-T--	1428
HIVHA594	C-----G-----T--G-GA-T--	1359
HIVHA596	C-----G-----T--G-GA-T--	1365
HIVHA599	--T-----G--A--G--TGTT--	1401
HIVRA651	--G-----T--A--G--TA-T--	1353
HIVRJS	-----A--G--TA--	1366
HIVGON	-----A--C-----	1262
HIVSC	--A-----TAGCAAG--	1369
HIVSBA	-----A--G--AA--CACA--	615
HIVS8B	-----TA-CAATAACAATACC--	639
HIVSBC	-----A--TA--CACA--	609
HIVBR014	-----T--C-C-----CACCACG--	1374
HIVBR021	--C-----T-----	1360
HIVACH9	--AC-----G--A--C--	1255
HIVACP1	-----G-CG-G--	1222
HIVJM	-----T-----TAGTAATAGT--	1345
HIVMM	-----TTA-C--	1273
HIVMA208	--C-----TTA-T--	1444
HIVP896	-----C-----A--	1366
HIVHB11A	-----GCAAGCAGC--	999

[illegible]

HTV1 ENV

CONSENSUS .A	ata?ta???????	?????????aatgAgAcc	TTcAGaCctggaGGAgGagatAT	1316
HIVSF1703	---AC-GTCA	-----A---	---AC-----A---	1439
HIVU455	-C-A-CAAAA	-----	-----	1403
HIVZ321	---AC-CGTCA	-----	-----	1424
HIVRW020	---AC-CTACA	-----A-T-	-----A-	1369
HIVUG031	--GC-ACAGC	ACCAATCGA-----A-	-----A-	1273
HIVD687	-C-C-GTAAT	ACCAATGAG-C-----G	-----G	1065
HIVNI	---A-GTACA	-----A-	-----G	1397
HIVRW009	---C-CA-	-----C-A-	-----	1411
HIVUG275A	-C-G-GAA	-----A-	-----A-	1409
HIVUG273A	GC-A-CA-	-----A-	-----A-	1421
HIVV1191A	-C-G-CA-	-----A-	-----	1418
HIVDJ264A	--G-CA-	-----T-	-----G	1394
HIVDJ263A	--G-CA-	-----	-----C	1418
HIVDJ258A	--G-CA-	-----T-	-----G	1403
HIVCARGAN	-C-GG-CA-	AATGAT-C-----G	-----G	1433
HIVCARBAS	---AC-GGTCA	-----	-----A-----A-C	1424
HIVCAR4054	---AC-CA-	CAA-----T-	---AC-----A-	1409
HIVCAR286A	---GC-CA-	CAA-----	---AC-----	1472
HIVCAR4023	-CTCA-CA-	-----T-	---AC-----	1475
HIVCAR423A	-C-CA-CT-	---A-----T-	---AT-----	1439
HIVKENYA	-C-G-GAA	---G-A---	-----	1400
CONSENSUS .B	Acg??acc???	?????????ac?GAgatc	TTcAGAcctGgaGGAGgagAT	1354
HIVJRCFS	-T-AG-GT-	GAG-TC-----	-----C-	1400
HIVJRF	-T-GG-	-----	-----	1397
HIVALAI	-TAGC-G-	AACAAC-C-----	-----	1403
HIVBRVA	---AG-	ACCGAC-C-----	-----A-	1412
HIVJH3	---AG-	-----C-	-----A-	1460
HIVBAL1	---AAG-	-----G-	-----	1418
HIVYU2	---GG-T-	-----	-----	1385
HIVMN	---AC-	-----	-----	1430
HIVHXB2R	-T-AGT-	-----	-----	1424
HIVLAI	-T-GGT-	-----	-----	1439
HIVNL43	-T-GGT-	-----	-----C	1418
HIVNFA	-T-AGT-	-----G-	-----	1418
HIVCAM1	-T-AG-	-----C-	-----	1421
HIVNY5CG	---ACT-	-----C-	-----	1403
HIVADA	---AGT-GT-	GGGT-C-----	-----G	1412
HIVJFL	-T-GGG-	-----	-----A-	1406
HIVS1M184	---AG-G-	GAGACC-C-----C-	-----	1421
HIVD31	---AG-	-----C-	-----A-	1403
HIVSF162	---ACC-	-----	-----T	1397
HIVBC9G3C	-TACC-	-----C-	-----	1433
HIV0Y1	---ACC-G-	AACGGG-TC-A-----	---C-----C-C	1418
HIVSF33	---AOT-	-----C-	---T-----C-----A-	1409
HIVCD4	-T-CA-	-----C-	---T-----	1457
HIVSF2	-P-TA-T-	AATGAC-C-----G-	-----	1418
HIVSF2B13	-T-CA-T-	AATGAC-C-----G-	-----	1412
HIVHAN	---AGT-GT-	ACC-C-----	-----G-A-	1418
HIVRF	-TACT-A-	-----	---T-----A-	1448
HIVMMJ2	-TAGC-G-	AGCAAGGAA-----	-----A-	1394
HIVTB132	-T-GG-	AACGAG-T-A-----	-----A-	1367
HIVBR020	-TAATGAG-	ATGAACACG-C-----	-----	1408
HIVTH014	-TAGTGAG-	AACAAGACC-C-----C-	-----	1384
HIVTH026	---ACC-	-----	-----	1381
HIVUS711	---AAG-A-	GAG-C-----C-	-----	1399
HIVUS712	---AAGCGAG-	AACGGG-C-----	-----	1408
HIVUS715	---AG-	AATACC-C-----	-----G-A-	1414
HIVUS716	-T-CA-	---AC-C-----	-----A-	1387
HIVNA593	---AAA-A-	AACGGG-T-----	-----	1474
HIVNA594	---AA-T-	-----	---CT-----	1396
HIVNA596	---GA-TGAA	AATAATGAA-T-----CT-	-----	1417
HIVNA599	-T-AG-G-	AAAGAC-AT-----C-	-----	1447
HIVNA651	-T-AT-A-	ACCGAG-C-----C-	-----	1399
HIVRJS	---GG-AAC	GAC-C-----C-	-----	1412
HIVGUN	---GG-G-GAG	AAC-C-----	-----G	1308
HIVSC	-T-GT-G-AAG	ANTGAGAAC-A-----T-	---G-----	1421
HIVSBA	---AG-T-	-----	-----	652
HIVSBB	-TAAT-AT-	---T-A-----	---C-----	679
HIVSBC	---TG-T-	-----	-----	646
HIVBR014	---AGG-	GAA-----	-----A-	1414
HIVBR021	---ATT-G-	GCGGAT-AT-A-----C-	-----	1406
HIVACH9	-T-AG-ACC	GAC-C-----C-	-----	1301
HIVACP1	---AG-T-ACC	AAC-TC-----C-	---G-----A-	1268
HIVJM	-T-AC-TAAC	AAG-A-----C-	-----A-	1391
HIVMM	-TAAC-AAT	AATGAC-C-----	-----	1322
HIVNA208	---CG-G-AAG	GAC-AC-----C-	-----	1490
HIVP896	-TAGT-TGAG	ACT GAG-T-----	-----	1415
HIVBN11A	-T-AG-G-	-----	-----	1001

CONSENSUS.C	ata?caca?????????????????????????????gagaca.ttcagaccctggaggaggagatAT	1341
HIVMA959	-C-CA-AG.....ACA.....G.....	1393
HIVMA960	-C-CA.....AATAACACA.....	1381
HIVD757	-A.....	1056
HIVD747	-GA.....AAGACA.....G.....A.....	1197
HIVD760	-GA.....T.....T.....	1217
HIVN0P	-C-CA.....AATAATACACA.....T.....A.....	1030
HIVZAM20A	-G---AAAATCACACATGAAACAAATGGCACAAGGAAC---	1478
HIVD1044	-A.....T.....	1044
HIVD1024	-C.....	1376
HIVD744	-C.....T.....C.....	1091
HIVD766	-A.....GAGACA.....	1037
HIVD808	-A.....	1079
HIVD868	-CA.....T.....	1067
HIVSM145A	GA-C-GA-GAG.....CAAAATAACACA.....G.....	1418
HIVZAM18A	-GA---CACA.....AACACA.....	1403
HIVDJ259A	-CC.....ACAGAGGA.....T.....	1355
HIVDJ373A	-GC---GAC.....	1388
HIVSE364A	-A.....T.....	1379
HIVUG268A	-GC.....	1403
CONSENSUS.D	Aca?tagt?a???????...?????????????????aatgagacc.ctcagaccctggaggaggagATAT	1341
HIVJY1	-G-C.....	1445
HIVNDK	-G-TC.....C.....A.....	1394
HIVMAL	-G---G-CAATAGT.....GAC---A.....	1430
HIVELI	-G-C.....C.....T.....	1415
HIVZ226	-G-TC.....C.....	1415
HIVUG024	-C---C-G.....	1409
HIVUG021	-A---C-G.....C.....G.....	1375
HIVUG030	-A---C-G.....AAT---T.....T.....	1406
HIVUG046	-C---GC-G.....A---G.....G.....	1406
HIVUG266A	-A---C-G.....	1409
HIVUG269A	-C-A---AATGAGACTAGCAGTAAA---T.....	1414
HIVUG274A	-A---AGT.....AATGAGACTAGCAGTAAA---T.....	1447
HIVSE365A	-C---C.....AATGAGACC.....	1409
HIVCAR4020	-CGA.....AAGAATGAT-T---T.....G.....	1523
CONSENSUS.E	ataatag???????...aacgagacc.ttcagaccctggaggaggagATAT	1344
HIVTH966	-CG-GT.....T.....	1402
HIVTH975	-G---GACT.....T.....	1396
HIVTH022	C---ATGCG.....	1396
HIVTH243	-C---GACT.....	1421
HIVTH006	-GACT.....	1393
HIVTH011	T-G---AGACG.....	1342
HIVCAR010	-C-G-G.....T.....G.....	1433
HIVCAR4017	-G-G---ACACGAG.....	1430
HIVCAR4039	GC---GTGGTGAA.....CCT---A---G.....G.....	1406
HIVCAR4071	-G-G-C.....T---A.....	1415
HIVCAR018A	-ACAA.....T.....	1394
CONSENSUS.F	ATgagacc???????...???gagacc.ttcagaccctggaggaggagATAT	1330
HIVB2163A	-C---T.....T.....G---T---T.....	1373
HIVB2126A	-C---T.....T.....G---A.....	1378
CONSENSUS.G	ATagTACC?a?????...???AacGAGACC.ttcagaccctggaggaggagATAT	1346
HIVV1525A	-A.....	1457
HIVLBV217	-A-TGGA.....C.....	1430
HIVCAR4067	-AG-GAAT.....AGA-CA.....T.....	1421
CONSENSUS.H	AT???TCTGCAGA?????...?A?GT?ACC.ttcagaccctggaggaggagATAT	486
HIVCA13	-CGG.....G-G---G.....	620
HIVV1557	-AAGT.....A-T---T.....	605
CONSENSUS.O	Aca?cagc?????a?...???Aac???ACa.ctTAGACCagtAGGGGGAGATAT	1349
HIVAMT70	-G---AACA-C.....TGTA---T---A---C.....	1410
HIVMVP5180	-TTC---CAGGTG-A.....T.....	1445
HIVVAU	-T-AA-CACATCCT.....AAC-G-ACC-C.....G.....	1457
CONSENSUS.CPZ	A?A?A?A?T???...???T?A?A.?T???ACCT???G?AA???T	727
SIVCP2GAB	-CA-C-G-GGT.....AACG-G-C---T-TAG---ACA-GA-G---ACA-	1379
CONSENSUS.U	AtagtActca?????...?????A?TGAGACC.ttcagaccctggaggaggagATAT	1294
HIVZAM184	-C---C.....AATGGG-C.....	1379
HIVCAR4081	-AG---A-TAAC.....C---A---G.....	1409
HIVK124A	-G---G.....A.....	1433
HIVZ3	-GCA---A.....A.....T.....C.....	1410

HIV1 ENV

CONSENSUS . A	gAgggACAAATGGAGaAgTGAatTATATAagTAtAAaGtaGTaaasaTtgaaCCacTAGGagTAgcACCca	1387
HIVSF1703	-----A-----	1510
HIVU455	-----A-----	1474
HIV2321	-----G-----	1495
HIVRW020	-----T-----	1440
HIVUG031	-----G-----	1344
HIVD687	-----A-C-AT-----T-----T-----	1136
HIVNI	-----T-----	1468
HIVRW009	-----T-----C-----A-----	1482
HIVUG275A	-----T-----	1480
HIVUG273A	-----A-----C-----T-----T-----	1492
HIVV1191A	-----T-----	1489
HIVDJ264A	-----TA-----	1465
HIVDJ263A	-----T-----	1489
HIVDJ258A	-----T-----	1474
HIVCARGAN	-----A-----A-----T-G-----CG-----	1501
HIVCAR8AS	A-----A-----	1495
HIVCAR4054	AA-----A-----GG-G-----A-----A-----	1480
HIVCAR296A	A-----GG-----A-T-----	1543
HIVCAR4023	A-----C-----GG-----T-----	1546
HIVCAR423A	-----G-----GG-----T-----	1510
HIVKENYA	-----	1471
CONSENSUS . B	GAgggGacAAATGGAGaAgTGAatTATATAAATATcAAAGTAgTaaasaTtGAaccATaGGAGTaGCACCCA	1425
HIVJRCSE	-----	1471
HIVJREFL	-----	1468
HIVALA1	-----	1474
HIVBRVA	-----	1483
HIVJH3	-----TT-----	1531
HIVBAL1	-----T-----	1489
HIVYU2	-----	1456
HIVMM	-----C-----	1501
HIVHXB2R	-----	1495
HIVLAI	-----	1510
HIVNL43	-----	1489
HIVMFA	-----A-----	1489
HIVCAM1	-----	1492
HIVNYSCG	-----G-----	1474
HIVADA	-----	1483
HIVJPL	-----A-----C-----G-----	1477
HIVS1MI04	-----T-----C-----G-G-G-----G-----	1492
HIVD31	-----	1474
HIVSF162	-----G-----	1468
HIVBCSG3C	-----	1504
HIVGYI	-----	1489
HIVSF33	A-----	1480
HIVCDB4	-----C-----	1528
HIVSF2	-----A-----A-----A-----	1489
HIVSF2B13	-----A-----	1483
HIVHAN	-----A-----	1489
HIVRF	-----G-G-G-----G-----	1519
HIVNM12	-----G-----	1465
HIVTB132	-----C-----	1438
HIVBR020	-----G-----G-----A-----	1479
HIVTH014	-----G-----	1455
HIVTH026	A-----A-----	1452
HIVUS711	-----G-----A-----	1470
HIVUS712	A-----C-----	1479
HIVUS715	A-----G-----A-----	1485
HIVUS716	-----	1458
HIVHA593	-----	1545
HIVHA594	-----	1467
HIVHA596	-----	1488
HIVHA599	A-----G-----	1518
HIVHA651	A-----T-----C-----G-----G-----	1470
HIVRJS	-----G-----	1483
HIVGUN	-----G-A-----	1379
HIVSC	-----	1492
HIVSBA	-----G-----A-----	723
HIVSBB	-----A-C-----	750
HIVSBC	-----C-----	717
HIVBR014	-----	1485
HIVBR021	A-----G-----G-----	1477
HIVACH9	-----	1372
HIVACP1	-----	1339
HIVJM	A-----	1462
HIVNM	A-----	1393
HIVMA208	-----	1561
HIVP896	-----G-----A-----	1486

CONSENSUS.C	GAGggACAAtTGGAGAAgTgAAAtTaTATAAATATAAAgTpGTAGAAA*TTAAgCCAtTpGGAgTAGCaCCCA	1412
HIVMA959	-----G-----	1464
HIVMA960	-----	1452
HIVD757	---A-----G---	1127
HIVD747	-----G---	1268
HIVD760	-----C-----	1288
HIVNOF	--A-----A-----A-----	1101
HIVZAM20A	-----A-----C-A--A-----	1549
HIVD1044	---A-----G-----	1115
HIVD1024	---A-----	1447
HIVD744	---A-----C-----	1162
HIVD766	-----A-----	1108
HIVD808	---A-----	1150
HIVD868	---A-----C-----T-----	1138
HIVSM145A	---A-----G-----A-----A-----	1489
HIVZAM18A	-----C-----A-----CA-----	1474
HIVDJ259A	-----C-----G-A--A-----	1426
HIVDJ373A	-----C-----A--C-----	1459
HIVSE364A	-----A-----A-----A-----	1450
HIVUG268A	-----G-----C-A-----	1474
CONSENSUS.D	gAgagACAAtTGGAGAAgTGAATTAATACAAATATAAAgTAgTAagAaTtgAaCCActTAGGa?cagCACCCa	1411
HIVJY1	---A-----A-----A-----	1516
HIVNDK	---G-----T-----A-----A-----G-----	1465
HIVMAL	---G-----T-----T-----C-----G-----	1501
HIVELI	---G-----T-----G-----CA-----G-----	1486
HIVZ226	---G-----T-----G-----A-----G-----G-----T-----	1486
HIVUG024	-----A-----A-----TC-----	1480
HIVUG021	-----C-----A-----TC-----	1446
HIVUG038	-----A-----	1477
HIVUG046	-----TC-----	1477
HIVUG266A	---G-----A-----C-----TC-----T-----	1480
HIVUG269A	-----TC-----	1485
HIVUG274A	-----TCGG-----	1518
HIVSE365A	---G-----T-----C-C-----G-----	1480
HIVCAR4020	---G-----T-----A-----A-----	1594
CONSENSUS.E	aAaGGACAAATGGAGAAgTGAATTAATATAAAATATAAAgTAGTAcAAATTGAACCACTAGGAATAGCACCCA	1415
HIVTH966	-----	1473
HIVTH975	-----	1467
HIVTH022	-----	1467
HIVTN243	-----	1492
HIVTH006	-----	1464
HIVTH011	-----	1413
HIVCARELO	G-----G-----	1504
HIVCAR4017	-----C-----T-----G-----	1501
HIVCAR4039	G-G-----G-----G-----A-----G-----	1477
HIVCAR4071	G-----	1486
HIVCARMBA	---T-----A-----	1465
CONSENSUS.F	GA?AGACAAATGGAGAAgTGAATTAATATAAAATATAAAgTAGTAAaAATtGAACCACTAGGAGTAGCACCCA	1399
HIVB2163A	---A-----T-----T-----G-----A-----G-----	1444
HIVB2126A	---A-----T-----G-----A-----G-----	1449
CONSENSUS.G	GAGGGAcAAATGGAGAAgTGAATTAATATAAGTATAAAgTAAaAATtAAACCACTAGGAGTAGCACCCA	1417
HIVV1525A	---G-----G-----C-----T-----	1528
HIVLBV217	-----	1501
HIVCAR4067	---T-----A-----	1492
CONSENSUS.H	GAGGGACAAATGGAGAAgTGAATTAATATAAAATATAAAgTAGT?A?A?T?GAACCACTAGGA?TAGCACCCA	552
HIVCA13	-----G-A-A-T-----G-----	691
HIVV1557	-----A-G-G-C-----T-----	676
CONSENSUS.O	GAaAGATATATGGAGAACTGAATTTgTtCAAAATACAAAGTAGTaaaggTAAaACCTTTTAGTGTgGCACCLA	1420
HIVANT70	-----C-----C-----	1489
HIVAMP5180	-----A-----A-----C-A-A-----A-----	1516
HIVVAU	-----C-----T-----A-----A-----A-----	1528
CONSENSUS.CPZ	??TGA?T??TGGAGA??TGA?T??T?A?GPA??A?GT?GT??T?AT?G??CC?TT?TCAGT?GC?CCAA	768
SIVCP2QAB	GAA---TATT---AGT---C-C-AT-A---TA-A---A---TCGA---A-AG---T---A---G---A---	1450
CONSENSUS.U	GAg?GACAAATGGAGAAgTGAATTAATATAAAATATAAAgTAGTAAAAATtGAACCACTAGGAGTAGCACCCA	1363
HIVZAM184	---G-----G-----C-A-----	1450
HIVCAR4081	---G-----A-----	1480
HIVK124A	---A-----C-A-----TC-----	1504
HIVZ3	---AA--T-----G-----G-----	1481

HIV1 ENV

	gp120 \/ gp41	
CONSENSUS .A	cCggg.GCAAaaaGaAGAGTGGTG.....gagAGAGAAAAAGAGCa?????gTtGga??	1434
HIVSF1703	--C--G-----C-----	1557
HIVU455	--A--G-----C-----	1521
HIVZ321	--AA--G-----C-----A-----	1543
HIVRM020	--A--G-----C-----	1487
HIVUG031	--A--G-----A-----	1391
HIVD687	--T-----G-----	1183
HIVMI	--T-----G-----	1515
HIVRW009	--A--G--G-----G-----	1529
HIVUG275A	--A--GG--G-----	1527
HIVUG273A	--CA--G-----C-----	1539
HIVVI191A	G--A--G-----C-----	1536
HIVDJ264A	--AT--G-----G-----	1512
HIVDJ263A	--T-----	1536
HIVDJ258A	--T-----	1521
HIVCARGAN	-----	1507
HIVCAR5AS	-----	1501
HIVCAR405A	-----	1486
HIVCAR286A	-----	1549
HIVCAR4023	-----	1552
HIVCAR423A	-----	1516
HIVKENYA	--AA--G-----	1518
CONSENSUS .B	CcAaG.gCAAagAgaAGAGtgTG.....CagaGagAAAAAGAGCa.....gtggga??	1472
HIVJRCSP	-----A-----	1518
HIVJRFL	-----A-----	1515
HIVALA1	--G-----	1521
HIVBRVA	-----	1530
HIVJH3	-----	1578
HIVBAL1	-----	1536
HIVVU2	-----	1503
HIVMN	-----C-----	1545
HIVHB2R	-----	1542
HIVLAI	-----	1557
HIVNL43	-----	1536
HIVMFA	-----	1536
HIVCAM1	-----GC-----	1541
HIVNY5CG	-----GC-----	1523
HIVADA	-----A-----AC-----	1532
HIVJFL	-----A-----AC-----	1526
HIVSIM184	-----CT-----	1541
HIVD31	-----CT-----	1523
HIVSF162	-----AC-----	1514
HIVBCSG3C	--C--A-----	1551
HIVQY1	-----G-----AT-----	1538
HIVSF33	-----A-----GT-----	1529
HIVCDC4	-----AT-----	1577
HIVSF2	-----AT-----	1538
HIVSF2B13	-----AT-----	1532
HIVHAN	-----G-----AT-----	1538
HIVRF	-T-G--A-----AC-----	1568
HIVVMJ2	-----AC-----	1514
HIVTB132	-T-x--A-----	1485
HIVBR020	--G--A-----AC-----	1528
HIVTH014	--C--A-----AC-----	1504
HIVTH026	--G--A-----AC-----	1501
HIVHS711	-----G-----G-----	1517
HIVUS712	-----G-----	1526
HIVUS715	--G--C-----	1532
HIVUS716	--G-----	1505
HIVHA593	-----GAT-----	1594
HIVHA594	--G--GC-----	1516
HIVHA596	-----A-----GT-----	1537
HIVHA599	-----T-----GT-----	1567
HIVHA651	-----GAC-----	1519
HIVRJS	-----GT-----	1532
HIVGUN	-----AC-----	1428
HIVSC	-----AC-----	1541
HIVSBA	-----	770
HIVSBB	--A-----	797
HIVSBC	-----	764
HIVBR014	--G--A--G-----GC-----	1534
HIVBR021	--A--A-----AC-----	1526
HIVACH9	--G--A-----AC-----	1421
HIVACP1	--G--A--A--AC-----	1388
HIVJM	--G--AAC-----	1509
HIVVM	--G--A-----	1440
HIVNA208	-T-C--AA-----	1599
HIVP896	--G--ACA--A-----	1533

		gp120 \ / gp41	
CONSENSUS.C	CTa?g.gCAAAAAGgAGAGTGGTG??????...	gagagaGaaAAAAGAGCA.....Gtggga??	1458
HIVMA959	--A--		1511
HIVMA960	--CA--		1499
HIVD757	--CT--		1174
HIVD747	--CT--		1315
HIVD760	--CA--		1335
HIVNOF	--GGA.T--		1148
HIVZAM20A	--GG.C-C--	gAGAAG.....C--T--G--G--	1602
HIVD1044	--CA.C--		1162
HIVD1024	--GCA--		1494
HIVD744	--GCA--	G--	1209
HIVD766	--GCA--		1155
HIVD808	--CA--		1197
HIVD868	--CA--		1185
HIVSH145A	--GG--		1536
HIVZAM18A	--GC--	gAG.....AC--	1524
HIVDJ259A	--A--	G--	1470
HIVDJ173A	--A--	CA--	1503
HIVSE364A	--A--	CA--	1497
HIVUG268A	--A--		1521
CONSENSUS.D	Cca?g.gCAAAgAGaAGAGTGGTG.....	gaaAGAGAAAAAGAGCA.....ATAGGA..	1457
HIVJY1	--G--		1563
HIVNDK	--A--G--		1512
HIVMAL	--A--		1548
HIVELI	--G--		1533
HIVZ26	--G--		1533
HIVUG024	--G--	C--	1527
HIVUG021	--A--G--		1493
HIVUG038	--T--T--	G--	1524
HIVUG046	--TGA--		1524
HIVUG266A	--G--		1527
HIVUG269A	--C--		1532
HIVUG274A	--G--A--G--		1565
HIVSK365A	--CA--	A--	1527
HIVCAR4020	--CG--		1600
CONSENSUS.E	CCaGG.GCAAAAGAGaAGAGTGGTG.....	GAGAGAGAAAAAGAGCA.....GTGGGA..	1462
HIVTH966			1520
HIVTH975			1514
HIVTH022			1514
HIVTH243			1539
HIVTH006			1511
HIVTH011			1460
HIVCARELO	--C--		1510
HIVCAR4017	--C--		1507
HIVCAR4039	--C--		1483
HIVCAR4071	--C--		1492
HIVCARMBA	--C--		1471
CONSENSUS.F	CCAgg.GCAAAaAGAcAGTGGTG.....	aaGAGAGAAAAAGAGCA?????GTGGGA..	1446
HIVB2163A		G--	1491
HIVB2126A	A--	GTA..	1499
CONSENSUS.G	CCagG.GCAAGGAGaAGAGTGGTG.....	GTGAGAGAAAAAGAGCA.....ATT???	1460
HIVVI525A		A--GGA..	1575
HIVLBV217	--A--	G--	1545
HIVCAR4067	--C--		1498
CONSENSUS.H	C7A?G.GCAA7GAGaAGAGTGGTG.....	G??AGAGAAAAAGAGC?.....GT7GGA..	592
HIVCA13	--T-A--A--	GG--A--T--	738
HIVVI557	--C-G--G--	AA--T--G--	723
CONSENSUS.O	CAaaa.ATTgCAAGgCCAatcATAggcActTg?act??PCACAGaGAAAAAGAGCA.....	GcAGGa..	1477
HIVANT70	--CGT--	G--A--A--	1548
HIVMVP5180	--GT--A--A--A--T--CAC--COCT--C--G--		1578
HIVVAU	CT--A--A--AT--	G--C--T--	1587
CONSENSUS.CPZ	CAA?A.G??A?AAGGC?7777G??77AA?ACA?A?..??CAGACA?AAAAAGAG?7..??T?GG?..		802
SIUCP2GAB	--A--CA-G--ATACA-TGCC--G--GA-A..GA--G--CA..GCCT-C--T..		1512
CONSENSUS.U	CCagg.gCAAAgAGaAGAGTGGTG.....	GAGAGAGAAAAAGAGCA.....gT7GGA..	1409
HIVZAM184	--A.C--G--	T--	1497
HIVCAR4081	--C--		1486
HIVK124A	--C--	A--A--	1551
HIVZ3	A--	G--G..	1528

HIV1 ENV

CONSENSUS.A	?...cTgGGAGctgTctTCcTTgGGTTCTTAgGAgCaGCAGGAAGcACtATgGGCGCgGcGTCAaTAACGC	1501
HIVSF1703	...A-A-----A-A-----C-----	1624
HIVU455	...-A-----A-----T-----A-----	1588
HIVZ321	...A-----T-----	1609
HIVRW020	...A-A-----T-----	1554
HIVUG031	...A-----A-A-----G-----	1458
HIVD687	-----	1250
HIVNI	-----CG-----	1582
HIVRW009	-----A-----	1596
HIVUG275A	...G-A-----	1594
HIVUG273A	...A-----A-A-----	1606
HIVVI191A	...-----T-----A-----C-----	1603
HIVDJ264A	...-----T-----A-----	1579
HIVDJ263A	-----T-----	1603
HIVDJ258A	...-----T-----T-----	1588
HIVKENYA	...A-A-----	1585
CONSENSUS.B	?...aTaGGAGcTaeGttCctTCGgtTCttTgGGagCAGCaGGAAGCACTATCGgCgcagcgtCaatgaCgc	1539
HIVJRCSF	...-----T-----CG-----A-----	1585
HIVJRFL	...-----G-----A-----	1582
HIVLAI	...C-----T-----CG-----	1588
HIVBRVA	...C-----T-----A-----CC-----	1597
HIVJH3	...-----G-----G-----	1642
HIVBAL1	...-----G-----G-----	1603
HIVYU2	...C-----T-----A-----	1570
HIVMN	...-----C-----A-----G-----	1612
HIVHXB2R	...-----T-----	1609
HIVLAI	...-----T-----CG-----	1624
HIVNL43	...-----T-----TGCA-----	1603
HIVMPA	...-----T-----	1603
HIVCAM1	G...C-----C-----T-G-C-----	1609
HIVNYSOG	G...C-----C-----G-----	1591
HIVADA	G...-----A-----	1600
HIVJFL	G...-----G-----	1594
HIVSIM184	G...-----T-----A-----G-----	1609
HIVD31	G...C-G-----G-----CG-----G-----	1591
HIVSF162	G...C-----CG-----C-----	1582
HIVBCSG3C	-----C-----T-----	1618
HIVOI	G...C-----CG-----G-----	1606
HIVSF33	G...-----A-----	1597
HIVCDC4	G...C-----A-----G-----	1645
HIVSF2	A...G-----T-----T-----	1606
HIVSP2B13	A...-----CG-----	1600
HIVHAN	G...C-----CG-----C-----	1606
HIVRF	A...-----GC-----A-----	1636
HIVRMJ2	A...-----GC-----C-----	1582
HIVTB132	-----GC-C-C-----x-----C-C-----	1552
HIVBR020	A...T-----A-----G-A-----	1596
HIVTH014	G...-----A-----	1572
HIVTH026	G...-----A-----	1569
HIVUS711	-----G-----G-----	1584
HIVUS712	...T-T-----	1593
HIVUS715	...T-----T-----AG-----	1599
HIVUS716	-----C-----	1572
HIVHA593	A...G-----	1662
HIVHA594	A...C-----	1584
HIVHA596	A...C-----	1605
HIVHA599	A...C-----C-----	1635
HIVHA651	A...C-----	1587
HIVRJS	G...-----G-----G-----	1600
HIVGUN	A...-----T-----C-----G-A-----	1496
HIVSC	G...-----TA-----	1609
HIVSBA	-----A-----	837
HIVSBB	-----G-C-T-----T-----	864
HIVSBC	-----G-----A-----	831
HIVBR014	A...C-----C-----	1602
HIVBR021	G...-C-----A-----	1594
HIVACH9	G...-----A-----	1452
HIVACPI	G...-----T-----A-----	1419
HIVJM	...-G-----T-----	1539
HIVWM	...T-----	1470
HIVP896	...-----G-----CG-----G-----	1600

SEQUENCE	POSITION	SEQUENCE	POSITION
CONSENSUS.C	7	..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1525
HIVMA959		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1578
HIVMA960		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1566
HIVD757		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1241
HIVD747		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1382
HIVD760		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1402
HIVMOF		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1215
HIV2AM20A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1669
HIVD1044		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1229
HIVD1024		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1561
HIVD744		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1255
HIVD766		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1200
HIVD808		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1264
HIVD868		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1230
HIVSM145A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1603
HIV2AM18A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1591
HIVDJ259A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1537
HIVDJ373A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1570
HIVSE364A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1564
HIVUG268A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1588
CONSENSUS.D		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1522
HIVJY1		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1630
HIVMDK		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1579
HIVMAL		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1615
HIVELI		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1600
HIVZ26		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1600
HIVUG024		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1594
HIVUG021		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1560
HIVUG038		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1591
HIVUG046		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1591
HIVUG266A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1594
HIVUG269A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1599
HIVUG274A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1632
HIVSE365A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1594
CONSENSUS.E		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1529
HIVTH966		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1587
HIVTH975		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1581
HIVTH022		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1581
HIVTN243		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1606
HIVTH006		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1578
HIVTH011		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1527
CONSENSUS.F		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1512
HIVBZ163A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1558
HIVBZ126A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1566
CONSENSUS.G		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1520
HIVVI525A		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1643
HIVL8V217		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	1612
CONSENSUS.H		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCGCgTCAaTaaCgc	652
HIVCA13		..aTaGgAGCTgTgtTCCTTGGGTTCCTTGGGAgCAGCaGGAAGCACTATgGgCGCG	

HIV1 ENV

CONSENSUS .A	TGACGGTACAGGCCAGACaaTTATGTCTGGTATAGTgCAaCAGCAGAgCAAttTGCTgAggGCTATagAG	1572
HIVSF1703	-----T-----C-----A-----	1695
HIVU455	-----C-----	1659
HIV2321	-----AG-----G-----A-----C-----	1680
HIVRW020	-----C-----A-----	1625
HIVUG031	---T-----A-----	1529
HIVD687	-----C-----C-----T-----	1321
HIVNI	-----C-C-----AA-----	1653
HIVRW009	-----C-----A-----	1667
HIVUG275A	-----C-----A-----A-----	1665
HIVUG273A	-----C-----A-----A-----	1677
HIVVI191A	-----C-----G-----C-----A-----	1674
HIVDJ264A	-----	1650
HIVDJ263A	-----	1674
HIVDJ258A	-----C-----	1659
HIVKENYA	-----C-----A-----	1656
RRE primary binding site ->/		
CONSENSUS .B	TgACGGTACaGgcCAGacaatTattgtCtggTatAGTGCAaCagCagAacAAAttTGCTGAggGCTATTgAg	1610
HIVJRCSE	-----A-----	1656
HIVJRFL	-----T-----	1653
HIVALA1	-----	1659
HIVBRVA	-----T-----C-----T-----	1668
HIVJW3	-----T-----A-A-----	1713
HIVBAL1	-----T-----C-----	1674
HIVU2	-----C-----	1641
HIVMI	-----T-----C-----	1683
HIVHXB2R	-----G-----	1680
HIVLAI	-----G-----	1695
HIVNL43	-----A-----G-----	1674
HIVMFA	-----G-----	1674
HIVCAM1	-----A-----C-----	1680
HIVNY5CG	-----A-----A-----A-----	1662
HIVADA	-----T-----	1671
HIVJFL	-----T-----A-----	1665
HIVSIM184	-----G-----	1680
HIVD31	-----G-----	1662
HIVSF162	-----A-----	1653
HIVBCSG3C	-----	1689
HIVQY1	-----G-----C-----	1677
HIVSF33	-----A-C-----C-----A-----	1668
HIVCDC4	-----A-----A-----A-----	1716
HIVSP2	-----	1677
HIVSF2B13	-----A-----	1671
HIVHAN	-----A-----A-----A-----	1677
HIVRF	-----A-----C-----A-----	1707
HIVWHJ2	-----C-----T-----	1653
HIVT8132	-----C-----C-----	1623
HIVBR020	-----C-----A-----	1667
HIVTM014	-----G-----GA-----	1643
HIVTH026	---T-----G-----GA-----	1640
HIVUS711	-----T-----A-----C-----A-----	1655
HIVUS712	-----T-----G-----	1664
HIVUS715	-----A-----	1670
HIVUS716	-----T-----A-----	1643
HIVHA593	-----TT-----A-----	1733
HIVHA594	-----G-----	1655
HIVHA596	-----G-----	1676
HIVHA599	-----A-----	1706
HIVHA651	-----T-----	1658
HIVRJS	-----C-----	1671
HIVGUN	-----A-----A-----A-----	1567
HIVSC	-----T-----	1680
HIVSBA	-----T-----C-----	908
HIVSBB	-----C-----C-----A-----	935
HIVSBC	-----T-----	902
HIVBR014	-----C-----C-----AA-----	1673
HIVBR021	-----C-T-----GGA-CGGA--GT--CA	1628
HIVP896	-----G-T-----G-----C-----	1671

CONSENSUS.C	TGACGGTACAGGCAGACAA?TgtTGTtTGGtATAGTGCAaCAGcAAAgcAA?TTGCTGAggGCTATAGAG	1595
HIVMA959	-----A-----C-----T-----C-----	1649
HIVMA960	-----C-----G-----	1637
HIVD757	-----T-----	1312
HIVD747	-----T-----	1453
HIVD760	-----T-AC-----	1473
HIVNOF	-----C-----A-----	1286
HIVZAM20A	-----G-----A-----	1740
HIVD1044	-----C-CT-----	1245
HIVD1024	-----	1569
HIVD808	-----	1272
HIVSM145A	-----C-----T-----A-----	1674
HIVZAM18A	-----C-----G-----A-----C-----	1662
HIVDJ259A	-----T-A-----	1608
HIVDJ373A	-----T-A-----A-----	1641
HIVSE364A	-----T-----A-----	1635
HIVUG268A	-----T-A-----A-----	1659
CONSENSUS.D	TGACGGTACAGGCCAGACagTtATtGTCTGgtATAGTgCAaCAGCAAAACAAT?TTGCTGAggGCTATAgAG	1593
HIVJY1	-----G-----A-----	1701
HIVNDK	-----A-----C-----	1650
HIVMAL	-----C-----	1686
HIVELI	-----A-----A-----	1671
HIVZ226	-----	1671
HIVUG024	-----A-----	1665
HIVUG021	-----C-----	1631
HIVUG038	-----	1600
HIVUG046	-----	1600
HIVUG266A	-----T-----T-----	1665
HIVUG269A	-----AG-----C-----C-----	1670
HIVUG274A	-----G-----A-----	1703
HIVSE369A	-----C-----A-----	1665
CONSENSUS.E	TGACGGTACAGGCCAGaCAATtATtGTCTGgtATAGTGCAaCAGCAaAGCAAT?TTGCTGAGGGCTATaGAG	1600
HIVTH966	-----G-----G-----G-----	1658
HIVTH975	-----G-----	1652
HIVTH022	-----	1652
HIVTN243	-----R-----W-----	1677
HIVTH006	-----	1649
HIVTH011	-----	1536
CONSENSUS.F	TGACGGTACAGGCCAGaCAATtATtGTCTGgtATAGTGCAACAGCA?A?cAAT?TGCTGAaGGCTATaGAA	1581
HIVBZ163A	-----A-----G-CT-----C-----G-----T-----	1629
HIVBZ126A	-----AA-----G-A-----C-----G-----T-----	1637
CONSENSUS.G	TGACGGTACAGGTCAGACAATtATtGTCTGgtATAGTGCAACAGCAAA?CAAT?TTGCTGAGGGCTATAGAG	1590
HIVVI525A	-----G-----	1713
HIVLBV217	-----A-----	1683
CONSENSUS.H	TGACGGTACAGGC?AGACAATtATtGTCTGgtATAGTGCAACAGCAAA?AA?TTGCTG?P?GCTATa?AG	716
HIVCA13	-----T-----C-----T-----AGG-----G-----	876
HIVVI557	-----C-----T-----C-----CAA-----C-----	861
CONSENSUS.O	TGACGGTACAGaCCCAcacttTgaTgAAGGGTATAGTGCAACAGCAGGACAACTGCTaAGAGCaATACAG	1615
HIVANT70	-----G-----C-----	1686
HIVMVP5180	-----G-----G-G-AC-----G-----G-----	1716
HIVVAU	-----G-----GCA-C-----A-----T-----	1725
CONSENSUS.CPZ	TGACGG?ACAG?CC?G?A?T?T?G?T?T?G?G?T?T?G?T?CA?CA?CAG??AATCTG?TG?AAGC?ATAAGAG	897
SIVCP2GAB	-----T-----G-----C-AC-G-----C-GTCG-----A-G-----A-----AAT-----T-A-----A-----	1650
CONSENSUS.U	TGACGGTACAGGCCAGACAatTatTGTCTGgtATAGTGCAaCAGCAAAaCAAT?TTGCTGAggGCTATAGAg	1546
HIVZAM184	-----C-G-----G-----A-----	1635
HIVK124A	-----G-----A-----	1689
HIVZ3	-----C-----G-----	1666

HIV1 ENV

CONSENSUS .A	GcTCAACAaCatcTgTTgaaaCTCACgGTCTGGGGCATTTAAACAgCTCCAGGCAaGAgTCCTGGCT?TGG	1642
HIVSF1703	-----G-----	1766
HIVU455	-----G-----T-----G-----	1730
HIVZ321	-----G-----A-----A-----G-----	1751
HIVRW020	-----G-----A-----G-----C-----G-----	1696
HIVUG031	-----G-----G-----C-----C-----	1600
HIVD687	-----A-----C-----	1392
HIVNI	-----C-----	1724
HIVRW009	-----G-----A-----A-----A-----	1709
HIVUG275A	-----G-----A-----G-----	1736
HIVUG273A	-----G-----A-----A-----C-----G-----	1748
HIVVI191A	-----T-----C-----G-----T-----C-----	1745
HIVDJ264A	-----C-----	1721
HIVDJ263A	-----C-----	1745
HIVDJ258A	-----C-----	1730
HIVKENYA	-----G-----	1727
CONSENSUS .B	GcgCAACAgCatcTgTtGcAACTcACaGTCTGGGGCATcAAgCAGCTCCAGGCAAGagTCcTgGCTgTgGA	1681
HIVJRCSF	-----A-----	1727
HIVJRFL	-----G-----A-----	1724
HIVALA1	-----A-----G-----T-----A-----	1730
HIVBRVA	-----A-----G-----T-----A-----	1739
HIVJH3	-GC-----A-----	1784
HIVBAL1	-----C-----	1745
HIVYU2	-----A-----	1712
HIVMN	-----A-----	1754
HIVHX82R	-----A-----	1751
HIVLAI	-----A-----A-----	1766
HIVNL43	-----A-----A-----	1745
HIVMFA	-----A-----A-----	1745
HIVCAM1	-----A-----	1751
HIVNY5CG	-----C-----	1733
HIVADA	-----C-----	1742
HIVJFL	-----C-----	1736
HIVSIMI84	-----A-----	1751
HIVD31	-----A-----	1733
HIVSF162	-----A-----	1724
HIVBCSG3C	-----C-----	1760
HIVQYI	-----C-----	1748
HIVSF33	-----A-----	1739
HIVCDC4	-----A-----	1787
HIVSF2	-----A-----	1748
HIVSF2B13	-----A-----	1742
HIVHAN	-----A-----G-----A-----	1748
HIVRF	-----G-----A-----	1778
HIVRAJ2	-----G-----A-----A-----	1724
HIVTB132	-----CA-----G-----A-----T-----A-----	1694
HIVBR020	-----CA-----G-----A-----A-----	1738
HIVTH014	-----A-----	1714
HIVTH026	-----A-----	1682
HIVUS711	-----A-----G-----	1726
HIVUS712	-----A-----	1735
HIVUS715	-----A-----	1741
HIVUS716	-----A-----	1714
HIVHA593	-----G-----A-----	1804
HIVHA594	-----A-----A-----	1726
HIVHA596	-----A-----A-----A-----	1747
HIVHA599	-----C-----G-----T-----A-----	1777
HIVHA651	-----G-----A-----	1729
HIVR35	-----A-----	1742
HIVGUN	-----A-----A-----	1638
HIVSC	-----A-----	1751
HIVSBA	-----A-----	979
HIVSBB	-----C-----	1006
HIVSBC	-----C-----	973
HIVBR014	-----G-----	1715
HIVP896	-----A-----C-----	1742

CONSENSUS.C	GcGCaacAgCATaTgtTGCAACTCACggtCTGGGGcATTAAAGCAGCTCCaGaCAAGAGTCcTGGCTATaGA	1666
HIVMA959	-----T--C--C-----	1720
HIVMA960	-----G-----G-----	1708
HIVD757	-----C-----	1383
HIVD747	-----C-----A-----G-----G-----	1524
HIVD760	-----C-----	1544
HIVNOF	-----T-----	1357
HIVZAM30A	-----A-----G-----C-----	1811
HIVSM145A	-----G-G-----T-----	1745
HIVZAM18A	-----C-----T-----G-----	1733
HIVDJ259A	-----G-----	1679
HIVDJ373A	-T--T-----A-----	1712
HIVSE364A	-----A-----	1706
HIVUG268A	-----	1730
CONSENSUS.D	GCGCAACAGCATcTgTTGCAACTCAC?gTcTGGGGcATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTgGA	1663
HIVJY1	-----A-----A-----	1772
HIVNDK	-----A-----G-----	1721
HIVMAL	-----G-----	1757
HIVELI	-----G-----A-----	1742
HIVZ236	-----G-----A-----A-----	1742
HIVUG024	-----A-----	1736
HIVUG021	-----T-----A--T-----	1673
HIVUG266A	-----A-----A-----	1736
HIVUG269A	-----A-----A-----	1741
HIVUG274A	-----A-----A-----	1774
HIVSB365A	-----G-----A-----	1736
CONSENSUS.E	GCGCAGCAGCATcTGTGCAACTCACAGTCTGGGGcATTAAACAGCTCCAGGCAAGAGTCcTgCTGTgGA	1671
HIVTH966	-----	1729
HIVTH975	-----	1723
HIVTH022	-----A-----	1723
HIVTN243	-----A-----Y--R-----Y--R-----	1748
HIVTH006	-----	1720
CONSENSUS.F	GCGCAACAGCATcTGTGCAAGCTCACAGTCTGGGGcATTAAACAGCTCCAGGCAAGAGTCCTGGC?GTGGA	1651
HIVB2163A	-----C-----	1700
HIVB2126A	-----T-----	1708
CONSENSUS.G	GCGCA?CA?CATcTGTGCAACTCACAGTCTGGGGcATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTAGA	1659
HIVVI525A	-----G--G-----	1784
HIVLBV217	-----A--A-----	1754
CONSENSUS.H	GC?CAACAGCAT?TGTGAA?CTCAC?	739
HIVCA13	--T-----C-----A-----A-----	903
HIVVI557	--A-----A-----G-----G-----	888
CONSENSUS.O	GCCAGCAacAaTTGCTGAGGctATCTgTAAGGGTATLAGACAACCTCCAGAGCTCGCCTGCTAGCCTTAGA	1686
HIVANT70	-----G-----R-----C-----	1757
HIVMP5180	-----C-----T-----A-----	1787
HIVVAU	-----C-----C-----	1796
CONSENSUS.CP2	?C?CA?CA?CA?CT??T?CA?CT?TC??TCTGGGAGTAAACAACT?CA?GC?AGA?TGCTTGC?GT?GA	950
SIVCPZGAB	G-T--A--G--C--AT-G--A--A--AA-----A--A--C--T-----T--A--	1721
CONSENSUS.U	GCGCAACAGCATcTGTGCAACTCACaGTCTGGGGcATTAAACAGCTCCAGGCAAGAGTCCTGGCTGTgGA	1617
HIVZAM184	-----T-----A-----G-----C-----A-----	1706
HIVK124A	-----	1760
HIVZ3	-----	1737

HIV1 ENV

CONSENSUS .A	AAGaTACCTaagGGATCAACAGCTCCTAGGAATTTGGGGcTGCTCTGGaAAaCTCATCTGCaCCACTAaTG	1713
HIVSF1701	-----A-----	1837
HIVU455	-----CA-----G-----C-----	1801
HIVZ321	-----A-----A-----A-----C-----	1822
HIVRW020	-----A-----	1767
HIVUG031	-----T-----	1671
HIVD687	-----A-----C-----	1463
HIVNI	-----	1735
HIVUG275A	-----	1807
HIVUG273A	-----C-----	1819
HIVV1191A	-----T-----A-----C-----C-----	1816
HIVDJ264A	-----C-----	1792
HIVDJ263A	-----T-----	1816
HIVDJ258A	-----T-----C-----	1801
HIVKENYA	-----A-----	1798
RRE or CAR ->/		
CONSENSUS .B	AAGATaCcTaaagGATCaACAGCTCcTagggaTtTGGGGtTGCTCTGGAAaCtcaTtTGCaCCACTgctG	1752
HIVJRCSE	-----A-G-----	1798
HIVJRFL	-----GG-----G-----	1795
HIVALA1	-----G-----G-A-----A-----	1801
HIVBRVA	-----C-----G-----	1810
HIVJH3	-----G-----	1855
HIVBAL1	-----G-----G-----C-----	1816
HIVYU2	-----G-----A-----	1783
HIVMN	-----G-T-----A-----	1825
HIVHXB2R	-----G-----A-----	1822
HIVLA1	-----G-----	1837
HIVNL43	-----G-----	1816
HIVMFA	-----G-----	1816
HIVCAM1	-----C-----	1822
HIVNY5CG	-----A-----C-----A-----	1804
HIVADA	-----G-----C-----	1813
HIVJFL	-----C-----G-----A-----	1807
HIVSIMI84	-----C-----G-----A-----	1823
HIVD31	-----G-----T-----	1804
HIVSF162	-----	1795
HIVBCSG3C	-----G-----	1831
HIVOI1	-----G-----A-----	1819
HIVSF33	-----GA-----A-----	1810
HIVCDC4	-----T-----	1858
HIVSF2	-----G-----	1819
HIVSF2B13	-----G-----C-----	1813
HIVHAN	-----G-----C-----A-----	1819
HIVRF	-----G-----A-----A-----A-----	1849
HIVRAJ2	-----G-----A-----	1795
HIVTB132	-----G-A-----C-----A-----	1765
HIVBR020	-----GG-----G-----C-----A-----	1809
HIVTH014	-----	1785
HIVUS711	-----A-----	1797
HIVUS712	-----G-----AA-----	1806
HIVUS715	-----GA-----G-----C-----A-----	1812
HIVUS716	-----G-----A-----	1785
HIVHA593	-----A-----	1875
HIVHA594	-----A-----	1797
HIVHA596	-----A-----	1818
HIVHA599	-----C-----C-----G-----	1848
HIVHA651	-----G-----C-----T-----	1800
HIVRJS	-----T-----	1813
HIVGUN	-----T-----G-----C-----A-----	1709
HIVSC	-----G-----G-----A-----	1822
HIVSBA	-----G-G-----G-----A-----	1050
HIVSBB	-----G-----	1077
HIVSBC	-----G-G-----G-----A-----C-----	1044
HIVP896	-----G-----A-G-A-----T-----	1813

CONSENSUS.C	AAGATACCTAaAGGATCAACAGCTcCTAGGgATtTGGGcTcTCTGGAAAACtCATCTGCACCACtgcTg	1737
HIVMA959	-----	1791
HIVMA960	-----C-----T-G-----CA-----	1779
HIVD757	-----C-----	1454
HIVD747	-----A-----A-----	1595
HIVD760	-----G-----	1615
HIVNCF	-----	1377
HIVZAM20A	-----	1882
HIVSM145A	-----G-----	1816
HIVZAM18A	-----C-----	1804
HIVDJ259A	-----T-----T-----AA-----	1750
HIVDJ373A	-----G-----	1703
HIVSE364A	-----	1777
HIVUG268A	-----C-----	1801
CONSENSUS.D	aAGaTACCTAaaAGGATCAaCgGCTCCTaGGaATtTGGGGTtGCTCTGGAAAaCaCATTtGCaCCACtaatG	1754
HIVJY1	-----C-----	1843
HIVNDK	-----G-----GG-----	1792
HIVMAL	-----C-----G-----G-----ATT-----	1828
HIVELI	-----	1813
HIV2226	G-----T-----C-----	1813
HIVUG024	-----C-----G-----G-----C-----C-G-----	1807
HIVUG266A	-----C-----	1807
HIVUG269A	-----C-----GG-----G-----G-----	1812
HIVUG274A	-----A-----C-----	1845
HIVSE365A	-----G-----G-----	1807
CONSENSUS.E	aAGATACCTAAAGGATCAAAAGtTcCTaGGAcTtTGGGGcTgCTCTGGAAAATcATCTGCACCACtGCTG	1742
HIVTH966	-----	1800
HIVTH975	-----	1794
HIVTH022	-----	1794
HIVTN243	-----C-----R-----Y-----	1819
HIVTH086	G-----	1791
CONSENSUS.F	AAGATACCTACAGGATCAACAGCTcCTAGGACTtTGGGcTcTCTGGAAAACtCATCTGCACCACtAAATG	1722
HIVBZ163A	-----	1771
HIVBZ126A	-----	1779
CONSENSUS.G	AAGAT?CCTA?AGGATCAACAGC?CCTAGGATtTGGGcTcTCTGGAAAACtCATCTGCACCACtAAATG	1727
HIVVI525A	-----T-----A-----C-----	1855
HIVLBV217	-----A-----C-----T-----	1825
CONSENSUS.O	AACcTtLaTACAGAATCAGCAACtCCTaAaCCTaTGGGGcTgTAAaggaaAaaCTaATCTGcTACACATPCAG	1757
HIVANT70	-----AC-----G-----G-----	1820
HIVMVP5180	-----C-----G-----T-----G-----C-----GAAT-G-----	1858
HIVVAD	-----	1867
CONSENSUS.CPZ	TA?GTACCT????GATCA?CA?TT?TTG?GCC?TGGG?TG?CTG?AAGG?TTCTGT?A?AC?ACGG	996
SIVCP2GAB	A-G-----GCAG-----G-GA-TC-G-----C-CT-----GA-----CTGT-----T-T-C-----	1792
CONSENSUS.U	AAGaTACCTA?AGgaTCAACAGCTCCTaGgaaTTTGGGGcTgCTCTGGAAAACtCATcTGCACCACtAcTG	1687
HIVZAM184	-----C-----G-----	1777
HIVL124A	-----C-----A-----T-----A-----T-----A-----	1831
HIV23	-----G-AG-----GC-----	1808

HIV1 ENV

CONSENSUS.A	TGCCCTGGAACTctAGtTGG	AGtAATAaatCt?a?agtGA?ATA	1754
HIVSF1703	-----	-----G-A-C-G-----G-----	1801
HIVU456	-----	-----C-GAG-C-----	1845
HIVZ321	-----T-----C-C-----	-----C-G-----T-----	1866
HIVRW020	-----	-----ATG-A-----G-----	1811
HIVUG031	-----	-----T-C-----G-----	1715
HIVD687	-A-----	-----A-CT-T-----C-----	1507
HIVUG275A	-----	-----CC-G-----T-----	1851
HIVUG273A	-----	-----C-A-----G-----	1863
HIVV1191A	-----A-----C-----	-----T-TGA-----T-----	1860
HIVDJ264A	-A-----G-----	-----A-T-T-A-----C-----	1836
HIVDJ263A	-A-----	-----A-T-T-A-----C-----	1860
HIVDJ258A	-A-----	-----A-TTT-A-----C-----	1845
HIVKENYA	-----	-----C-----C-G-----G-----	1842
CONSENSUS.B	TGCCCTGGAACTgctAgTTGG	AGtAATAaatCtctggat?agaTT	1795
HIVJRCSP	-----A-----	-----AGT-----	1842
HIVJREFL	-----	-----AG-----	1839
HIVALA1	-----	-----A-C-----	1845
HIVBRVA	-----	-----AG-G-C-----	1854
HIVJH3	-----	-----AG-A-----	1899
HIVBAL1	-----	-----A-A-----	1860
HIVYU2	-A-----	-----A-G-A-----	1827
HIVMN	-----	-----G-T-----	1869
HIVHXB2R	-----	-----AC-----	1866
HIVLAI	-----	-----AC-----	1801
HIVML43	-----	-----AC-----	1860
HIVMFA	-----	-----AC-T-----	1860
HIVCAM1	-----	-----A-----	1866
HIVNY5CG	-----	-----A-----	1848
HIVADA	-----	-----A-----	1857
HIVJFL	-----	-----G-----	1851
HIVSIMI84	-----	-----C-A-C-----	1866
HIVD31	-----	-----A-----AT-----	1848
HIVSF162	-----	-----C-----	1839
HIVBCSG3C	-----	-----A-C-----	1875
HIVQY1	-----	-----A-G-----	1863
HIVSF33	-A-----	-----A-----	1854
HIVCDC4	-----	-----A-----C-----	1902
HIVSF2	-----	-----AG-C-----	1863
HIVSF2B13	-A-----	-----A-----AG-C-----	1857
HIVHAN	-----	-----A-C-----C-----	1863
HIVRF	-----	-----A-----AT-----	1893
HIVMAJ2	-----C-----	-----A-A-C-A-----	1839
HIVTB132	-----	-----G-----	1809
HIVBR020	-----CA-----	-----C-----G-C-----	1853
HIVTH014	-----	-----A-----	1829
HIVUS711	-A-----	-----T-----A-----	1841
HIVUS712	-AAG-C-----	-----C-----	1850
HIVUS715	-----	-----AT-----C-----	1856
HIVUS716	-A-----	-----C-----	1829
HIVHA593	-A-----	-----AG-G-----	1919
HIVHA594	-A-----	-----G-----AA-A-----	1841
HIVHA596	-A-----	-----G-----AA-A-----	1862
HIVHA599	-C-----	-----G-----C-GT-T-----	1892
HIVHA651	-T-----	-----C-----GC-----	1844
HIVRJS	-A-----	-----A-GG-----	1857
HIVGUN	-A-----	-----A-AC-A-----	1753
HIVSC	-A-----	-----T-----CA-----	1866
HIVSBA	-----	-----C-----	1094
HIVSBB	-----	-----A-G-T-----	1121
HIVSBC	-T-----	-----C-----	1088
HIVB896	-T-----	-----G-----G-T-----	1857

CONSENSUS.C	TgcCTTGGAACTCCAGTTGG	AgTAATAAATCTCaacagATAT	1781
HIVMA959	--G--T-----T-----	-----T-----	1835
HIVMA960	-----	-----A-----	1823
HIVD757	-A-----C-----	-----C-G-----	1498
HIVD747	-A-----	-----C-G-----	1639
HIVD760	-A-----	-----C-G-----	1659
HIVZAM20A	-----	-C-----A-T-CGA-T----	1926
HIVSN145A	-A-----	-----G-A-----GA--A--	1860
HIVZAM18A	--T-----	-----G-----	1848
HIVDJ259A	-----G-T-----	-----GG-----AG--G--	1794
HIVDJ373A	-----T-----	-----GAG--G--	1827
HIVSE364A	-----T-----	-----T-CAG-----	1821
HIVUG268A	-----T-----	-----T-GGG-----	1845
CONSENSUS.D	TgCCCTGGAACTCTAGTTGG	AGTAATAgATCTCTAgatgagATT	1778
HIVJY1	-----	-----A-----G-----	1887
HIVNDK	-A-----	-----	1836
HIVMAL	-----T-----	-----C-----	1872
HIVELI	-----	-----A-----	1857
HIVZ226	-----	-----A--C-----	1857
HIVUG024	-----C-----	-----A-----ACA--	1851
HIVUG266A	-----C-----	-----A-----A--	1851
HIVUG269A	-----C-----	-----G-----	1856
HIVUG274A	-----C-----	-----G-----	1889
HIVSE365A	-----	-----A-----AC--	1851
CONSENSUS.E	TGCCCTGGAACTCCACTTGG	AGTAATAgATCTCTGAAGagATT	1786
HIVTH966	-----T-----	-----	1844
HIVTH975	-----	-----	1838
HIVTH022	-----	-----A-----A--	1838
HIVTH243	-----	-----	1863
HIVTH006	-----	-----C-----	1835
CONSENSUS.F	TGCCCTGGAACTCTAGTTGG	AGTAATAAAATCTCAGGAGGAGATT	1766
HIVBZ163A	-----	-----	1815
HIVBZ126A	-----	-----	1823
CONSENSUS.G	TGCCCTGGAACTACTAGTTGG	AGTAATAAAATCTTAAATGA?AT?	1764
HIVVI525A	-----C-----T-----	-----T-----C-A-----C-T	1899
HIVLBU217	-----T-----C-----	-----C-----T-T-----G-A	1869
CONSENSUS.O	TAAAAATGCAATCTAaCATGGCTAGGA	???????ataATGAag?ATT	1793
HIVANT70	-----G-----AT-----	-----C-----	1666
HIVMVP5180	-----C-C-T-----TC-----	AGATATA--G-----C-T--	1905
HIVVAU	-----G-----A-----GG-----	-----G-----TCA-----	1908
CONSENSUS.CPZ	TGCCCTGGAACTAA?TC?TGG?????????????????A?CA?????A?A?A?A?A?A?A?	-----G-AT...TCC-C-G-TGAC	1022
SIVCPZGAB	-----C-C-T-----CCTGGG	-----	1839
CONSENSUS.U	TGCCCTGGAACTCTAGTTGG	AGTAATAAatCtctgga?gagatT	1730
HIVZAM184	-----T-----	-----GA--A--T-----	1821
HIVK124A	-----	-----AC-G-----C--	1875
HIVZ3	-----	-----CT--TAACAC--	1852

HIV1 ENV

CONSENSUS.A	TGGGAtAacATGACcTGGcTgcAATGGGAtAAaGAAaTtagcAaTTACACAgacaTAATATATaatCTAaT	1825
HIVSF1703	---A-----	1952
HIVU455	---A-----G-----T-G-----G-----C-A-----	1916
HIVZ321	---C-A-----AG-----G-----C-AG-----C-----	1937
HIVRM020	---T---T-----	1882
HIVUG031	-----A-----AC-----G-----	1786
HIVD687	-----A-A-----GC-----	1578
HIVUG275A	---C-----G-----A-A-----GC-----GG-----C-----	1922
HIVUG273A	-----T-----A-----G-----	1934
HIVV1191A	---A-C-----GA-----C-AT-----GG-----	1931
HIVDJ264A	---G-----	1907
HIVDJ263A	-----G-----G-----	1931
HIVDJ258A	-----C-----	1916
HIVKENYA	---G-----C-----	1913
CONSENSUS.B	TgggaataAcacGACcTGGcTggAgTGGGaaAgAGAAATtgacAaTTacACaagccttaATAtacacccTTaAT	1866
HIVJRCSF	-----A-----G-----A-AC-----C-----	1913
HIVJRF1	-----GA-----C-----	1910
HIVLAL1	---G-----	1916
HIVBRVA	---G-----T-----A-----T-----	1925
HIVJH3	---G-----	1970
HIVBAL1	---G-----A-----C-----A-----A-----G-----	1931
HIVYU2	---G-----T-----A-----CA-A-----T-----	1898
HIVMN	-----C-----T-A-----C-----	1940
HIVHXB2R	---C---C-----C-----A-----C-----T-----	1937
HIVLAI	-----C-----A-----C-TT-----	1952
HIVNL43	-----C-----A-----C-----T-----	1931
HIVMFA	-----C-----A-----C-----C-----T-----	1931
HIVCAM1	---T-----G-----A-----	1937
HIVNY5CG	---G-----G-----G-----	1919
HIVADA	---G-----C-----A-----G-----	1928
HIVJFL	---AA-----C-----G-----AG-----C-----	1922
HIVSINI84	---G-----G-----G-----GA-----	1937
HIV031	-----	1919
HIVSF162	-----G-----A-----	1910
HIVBCSG3C	---G-----T-----C-----A-----GA-----A-----	1946
HIV0Y1	---G-----C-----CA-----	1934
HIVSF13	-----T-----C-----	1925
HIVCDC4	-----C-----CA-----T-----	1973
HIVSF2	---G-----C-----A-AC-----C-----	1934
HIVSF2B13	---G-----C-----A-----A-AC-----C-----	1928
HIVHAN	-----A-----T-----	1934
HIVRF	-----C-----G-A-----A-----C-----	1964
HIVRMJ2	---G---C-----A-----G-----	1910
HIVTB132	-----A-----A-----G-----C-----	1880
HIVBR020	---CA-----T-----A-----T-----	1924
HIVTH014	---T---T-----AGA-----	1900
HIVUS711	---GG-----G-----	1912
HIVUS712	---C-G-----C-----T-G-----GA-G-----	1921
HIVUS715	---C-G-----C-----A-----T-----A-----	1927
HIVUS716	---GGG-----C-----T-----G-----	1900
HIVHAS93	---G-----C-----	1990
HIVHA594	-----G-----T-----	1912
HIVHA596	-----G-----T-----	1933
HIVHA599	-----A-----GA-A-----G-----	1963
HIVHA651	---T-A-----T-----	1915
HIVRJS	---G-----G-----A-----G-----	1928
HIVGUN	---G---T-----G-----G-A-----T-A-----	1824
HIVSC	---GG-----	1937
HIVSBA	---G-----C-----A-----A-----	1165
HIVSBB	---G-----C-----C-----A-A-----GAT-C-----CAA-----	1192
HIVSBC	-----C-----A-----A-----	1159
HIVP896	-----GA-AT-----TGA-----C-----	1928

CONSENSUS . C	TGGGAtAACaTGACcTGGATGcAgTGGGATAgAGAAATTAgTAATTAcacAgacacAATATAcagcTTGCT	1853
HIVMA959	-----G-----T-----A-----	1906
HIVMA960	-----C-----A-----	1894
HIVD757	-----A-----	1569
HIVD747	-----A-----	1710
HIVD760	-----G-----A-T-----	1730
HIVZAM20A	---A---G-A---T---T-A---	1997
HIVSM145A	---A---C---A---A---	1931
HIVZAM18A	-----A-----G-----A-----T-----GACT-----G-----	1919
HIVDJ259A	-----C-----G-----A-----T-----A-----	1865
HIVDJ373A	-----T-----A-----AGT-T-----C-----	1898
HIVSE364A	-----T-----A-----T-----GA-----	1892
HIVUG268A	-----A-----A-----C-----	1916
CONSENSUS . D	TGGcatAACaTGACCTGGATggAGTGGGAAAGAGAAATTgaCAATACACAGGctTAATATAcagcTTaAT	1849
HIVJY1	---A---A---C---TG---	1958
HIVNDK	---G---A---A---A---	1907
HIVMAL	---A---T---C---A---AG---A---A---	1943
HIVELI	---G---T---A---T---	1928
HIVZ26	---G---T---A---T---	1928
HIVUG024	---GG---AG---G---TGAT---G---	1922
HIVUG266A	---A---G---G---T---	1922
HIVUG269A	---A---G---G---T---	1927
HIVUG274A	---A---G---G---C---	1960
HIVSE365A	---A---A---T---	1922
CONSENSUS . E	TGGGaaCAaCATGACaTGGATAGAATGGGagAGAGAAATTAGCAATTAcACAAACCAaATATATGAGATACT	1857
HIVTH966	---G---C---T---	1915
HIVTH975	---C---T---	1909
HIVTH022	---C---T---	1909
HIVTN243	---M---R---	1934
HIVTH006	---G---	1906
CONSENSUS . F	TGGGAGAACATGAACCTGGATGGA?TGG?AAAA?GAGATTAGCAATTACTCA?A?GA?TTATACAGGTTAAT	1830
HIVBZ163A	---A---C---A---A-T---AG---	1886
HIVBZ126A	---G---G---G---T-C---GA---	1894
CONSENSUS . G	TGGGA?AACATGAACCTGGAT?GAATGGGAAAGGGAAAT?A?CAA?TACACA?A?CAAAATATAC??CCT??T	1824
HIVV1525A	---T---G---C---T---A-A---AG---AA---	1970
HIVLBV217	---C---A---A---T-A---C---C-C---TC---GC---	1940
CONSENSUS . O	TGGGACaaactTaACA?TGGCAGcAaTGGGATCaaCagATAAaCAAcgTAAGCTOCatcATATATGA?gAAAT	1863
HIVANT70	---C---G---GG---G---A---C---G---	1937
HIVMVP5180	---C-T---C---C---T---T---T---T---	1976
HIVVAU	---TG-G---G---T---T---AA---	1979
CONSENSUS . CPZ	TGGG??AAT?T?ACATGGCA??AATGGGA?A?ATTAGT?T??AA?T??ACAGG??A?AT?T?T?T?T?T?	1069
SIVCP26AB	---GG---C-A---GC---T-A---GTCT---C-AC---GA-A---T-T-GG-C-G---	1910
CONSENSUS . U	TGGGACAA?ATGACCTGGATGgAgTGGGA?AgAGAAATTagCAATTAcACAgaa?TAATATAc?ggtTact	1797
HIVZAM184	---A---T---C-A---C-A---	1892
HIVK124A	---T---C---A---A---GA---T---GGT---A-C---A---	1946
HIVZ3	---TC---T---C---G---TG---G---	1923

HIV1 ENV

CONSENSUS .A	TGAAGAacCgCagAACCCAGCAGGAAAagAATGAACaaGAcTTatTGGCATTGGACaAgTGGgcAaaTCTgT	1896
HIVSF1703	-----T-----A-----	2023
HIVU465	---G-----T-----	1987
HIV2321	-----C-----TA-----AGG-----	2008
HIVRW020	-----C-----T-----G-----	1953
HIVUG031	---G---AGC-----A---G---	1857
HIVD687	-----G-----A-----	1649
HIVUG275A	-----G-----G-----	1993
HIVUG273A	-----G-----G-----	2005
HIVVI191A	-----A-----A---A---G---	2002
HIVDU264A	-----A-----C---A---	1978
HIVDU263A	-----A-----A---GG---	2002
HIVDJ258A	-----A-----A---	1987
HIVKENYA	-----A-----A---	1984
CONSENSUS .B	tgaagaatCgCagAACCAaGaaAaAgAATGAACaaGAcTTatTggaaTTggAaaagTGGGcaagTtTgT	1937
HIVJRC5F	-----T-----A---A---	1984
HIVJRFL	-----A-----A---A---	1981
HIVALA1	---G---A-----A---A---	1987
HIVBRVA	---G---C---A---T-----A---A---C---	1996
HIVJH3	-----A-----GG---A---A---	2041
HIVBAL1	-----A-----A---A---	2002
HIVYU2	---C---A---C-----C---A---A---	1969
HIVMN	A---A---A---C-----A---A---	2011
HIVHXB2R	-----A---G-----A---A---	2008
HIVLAI	-----A---G-----A---A---	2023
HIVNL43	-----A---G-----A---A---	2002
HIVNFA	---C---A---G-----A---A---	2002
HIVCAM1	-----A---C-----C-----	2008
HIVNY5CG	-----T-----A---A---	1990
HIVADA	-----C---A---C---A---	1999
HIVJFL	-----T-----C---A---C---	1993
HIVSIM184	-----G---C-----C---A---C---	2008
HIVD31	-----G-----A---A---A---A---	1990
HIVSF162	-----A-----A---A---	1981
HIVBCSG3C	---G---G-----A---A---	2017
HIVOYI	-----G-----G-----G---	2005
HIVSF33	-----A-----G-----G---	1996
HIVCDC4	-----A-----C---G---C---C---A---	2044
HIVSF2	-----A-----A---A---	2005
HIVSF2B13	-----A-----A---A---	1999
HIVHAN	---C---A-----A---A---	2005
HIVRF	-----G-----A---A---	2035
HIVWMJ2	-----G---G-----A---A---C---	1981
HIVTB132	---C---A---T-----T---T---C---A---A---	1951
HIVBR020	-----GG---G-----G---AC---	1995
HIVTH014	-----T-----A---A---	1971
HIVUS711	-----G-----A---A---A---	1983
HIVUS712	---G---A-----A---A---	1992
HIVUS715	---T---A-----A---A---	1998
HIVUS716	-----A---G-----A---A---	1971
HIVHA593	-----A---G-----G---G---A---	2061
HIVHA594	-----G-----C-----	1983
HIVHA596	---G---G-----C-----	2004
HIVHA599	---A---G-----G---A---A---C---	2034
HIVHA651	---A---G-----A---	1986
HIVRJS	---C---G-----C---C---	1999
HIVGUN	---G-----T---G---A---G---	1895
HIVSC	-----A-----A---A---	2008
HIVSBA	-----A-----C---G---A---	1236
HIVSBB	-----A-----C---A---G---G---	1263
HIVSBC	-----A-----A---G---A---	1230
HIVP896	---A---A---C-----A---	1999

CONSENSUS . C	TGAAGacTCgCAAAacCAAGCAGGAAaaAAATGAaaAaGAtTATAGCATTGGACAgTGGGaaAaTCTgT	1923
HIVMA959	-----G-----C-----	1977
HIVMA960	-----T-----	1965
HIVD757	-----G-----	1640
HIVD747	-----G-----	1781
HIVD760	-----G-----	1801
HIVZAM20A	-----C-----C-----G-----AG-----T-T-----	2068
HIVSM145A	---T---A-----C-----GG---G-----	2002
HIVZAM18A	---A---TT---G-----C-----A-----	1990
HIVDJ259A	-----G-----C-----C-----AA---C-G-----	1936
HIVDJ373A	---T---C-----GC-----C-----AA---C-----	1969
HIVSE364A	---A---G-----C-----C-----A-----C-----	1963
HIVUG268A	---A---A-----G-----C-----AA---C-----	1987
CONSENSUS . D	TGAagAaTCgCAaatCCAGCAGAGAAAgAATGAacAaGAatTaTTGgaATPGGACAAaTGGGCAAGTCTGT	1920
HIVJY1	---A-T---A-----A-----CC-----C-----	2029
HIVNDK	-----G-----A-----	1978
HIVNAL	-----A-G-----G-----	2014
HIVELI	---G---G-C-----A-----G-----G-----	1999
HIVZ226	-----C-----G-----	1999
HIVUG024	-----A---TC-----	1993
HIVUG266A	-----G-----C-----	1993
HIVUG269A	-----C-----A-----C-----C-----	1998
HIVUG274A	---A---C-----C-----C-----T-----G-----	2031
HIVSE365A	---G---C-----G-----C-----G-----	1993
CONSENSUS . E	tACagAATCGCagaAACCAGCAGGACAGgaATGAaaaGgAtTTGTTAGAaATGGATAAATGGGCAAGTCTGT	1928
HIVTH966	---G---A-----	1986
HIVTH975	-----A-----G-----	1980
HIVTH022	-----C-----	1980
HIVTN243	-----C-----	2005
HIVTH006	G--GA-----G-----A-----C-----G-----	1977
CONSENSUS . F	TGAA?AATCACAGAACCCAGCAGGAAAAGAAATGAACAAG?ATTATGGCATTGGACAAATGGGCAAGTCTGT	1899
HIVB2163A	---A---G-----	1957
HIVB2126A	---C-----A-----	1965
CONSENSUS . G	?GAAGAATCGCAGATCCAGCAGGAAAAGAAATGAACAAGACTTATTTGGCATTTGGACAAATGGG?AAG?TTGT	1692
HIVV1525A	T-----C-----T-----	2041
HIVLBU217	C-----A---C-----	2011
CONSENSUS . O	ACAAGagGCACAAGaaCAaCAGGAaaAaaATGagAAAgaATTGcTGGAGrTAGATGAATGGGCTCTaTTT	1934
HIVANT70	---A---T---G---C---A---G---	2008
HIVMVF5180	---CA---C---G---TA---C---T---C---	2047
HIVVAU	-----G-----	2050
CONSENSUS . CPZ	A7A??AGCACA???CAACAAGAGA?AAAT?AAA?GGA?TAT?TGAA?T?GA?AATGG?TC?TTTPT	1115
STVCPZGAB	-G-GGA-----ATCA-----A---G---G---CC---TG---T-G---TC---GC-AGC--G-	1981
CONSENSUS . U	TGAagAcTCgCAaaAcCAgCAGGAAAGAAATGAacAAGA?LTactGgaATTgGACAAATGGGCAAGTCTGT	1867
HIVZAM194	---G---G-----C-----A-----	1963
HIVK124A	---A-A---C-----A-----GC---AC-----	2017
HIVZ3	---A---A-----G---A---T---C-----T---G-----C---	1994

HIV1 ENV

CONSENSUS . A	GgAaTTGGTTTgacATatCAAA?TGGCTgTGGTATATaAgaaTATTtaTAATgATAGTAGGAGGcTTaATA	1966
HIVSF1703	-----G-----G-----	2094
HIVU455	-A-----A--C--T-----T--C--G--A-----	2050
HIVZJ21	-----T--A-----A-----G--	2079
HIVRW020	-----GG-----A-----	2024
HIVUG031	-----C-----A-----	1928
HIVD687	---G-----A--A--A-----T--G--	1720
HIVUG275A	-----G-----C-----C-----	2064
HIVUG273A	---G-----T-----C-----	2076
HIVVI191A	---G-----T-----T--G--	2073
HIVDJ264A	-----A--A--A--A-----A--C--T-----	2049
HIVDJ263A	-----G--A--A--A--A--A--C--T-----	2073
HIVDJ258A	---G-----G--A--C-----A-----T-----	2058
HIVKENYA	-----C-----A-----	2055
CONSENSUS . B	GgAaTTGGTTtaacATAaCAAA?TGgCTgTGGTatATAAaaaTaTTCATAATgATAgTAGGaGGcTTggTA	2007
HIVJRCFS	-----GG-----A-----A-----	2055
HIVJRPL	-----G-----A-----A-----	2052
HIVALA1	-----T-----C-----Q-----	2058
HIVBRVA	-----C-----C-----G-----A-----	2067
HIVJH3	-----CT-----C-----G-----A-----	2112
HIVBAL1	-----G-----A-----A-----	2073
HIVU02	-----G-----A-----A-----	2040
HIVMN	-----G-----T-----	2082
HIVHXB2R	-----T-----T-----	2079
HIVLAI	-----T-----T-----	2094
HIVNL43	-----T-----T-----	2073
HIVMPA	-----T-----T-----	2073
HIVCAM1	-----G-----T-----A-----T--AA--	2079
HIVNY5CG	-----G-----G-----A-----T--AA--	2061
HIVADA	-----G-----T--T-----C-----A-----	2070
HIVJFL	-A-----G--C--C--C-----C--T--A--	2064
HIVSIM184	-----G-----A-----T--AA--	2079
HIVD31	---G-----G--T--C-----	2061
HIVSF162	-----G--T--A-----T--A--	2052
HIVBCSG3C	-----CA-----C-----A-----	2088
HIVQY1	---G-----G--C-----G-----A-----	2076
HIVSF33	-----G-----C-----G-----A-----	2067
HIVCDC4	---C--C--G-----A-----A-----	2115
HIVSF2	-----G-----C-----G-----	2076
HIVSF2B13	-----C-----C-----	2070
HIVMAN	---G-----A--G--T--T-----A-----	2076
HIVRP	-----G-----C--A-----G-----	2106
HIVMMJ2	-----G-----T-----A-----	2052
HIVTB132	-----G-----C--B-----A-----	2022
HIVBR020	-----C--T--A-----C-----	2066
HIVTH014	-----G-----A-----	2042
HIVUS711	-----C--C-----C-----C-----A-----	2054
HIVUS712	-----G-----C--A-----	2063
HIVUS715	-----G-----C-----A-----	2069
HIVUS716	-----C-----C-----C-----	2042
HIVHA593	-----G--G-----C-----	2132
HIVHA594	-----T-----G--A-----A-----	2054
HIVHA596	-----T-----A-----	2075
HIVHA599	-----GT-----A-----T-----A-----	2105
HIVHA651	-----T-----G--A-----	2057
HIVRJS	-----G--T--C-----G-----T--A--	2070
HIVGUN	-----G-----C-----C-----T--AA--	1966
HIVSC	-----T-----T-----A-----	2079
HIVSBA	---G-----G-----T-----G-----A-----	1307
HIVSBB	-----G--T--A-----	1334
HIVSBC	-----G-----T-----A-----A-----	1301
HIVPB96	-----G-----C-----G--T-----A-----	2070

CONSENSUS. C	GGAAATTGGTTT?aCATAaCAAA?TGGCTgTGGTATATAAAaATATTCATAATgATAGTAGGAGGCTTgATA	1992
HIVMA959	-----G-----T-----	2048
HIVMA960	-----A-----A-----	2036
HIVD757	-----AG-----T-----	1711
HIVD747	-----AG-----T-----A-----A-----	1852
HIVD760	-----AG-----T-----	1872
HIVZAM20A	-----G-----T-----A-----G-----	2139
HIVSM145A	-----G-----A-----A-----A-----	2073
HIVZAM18A	-----G-----T-----	2061
HIVDJ259A	-----AG-----G-----	2007
HIVDJ373A	---G-----G-----G-----G-----	2040
HIVSE364A	-----A-----G-----A-----	2034
HIVUG268A	---G-----G-----A-----A-----	2058
CONSENSUS. D	GGAAATTGGTTTtagCATaCaAAaTGGCTGTGGTATATAAAaATATTCATAATgATAGTAGGAGGCTTgATA	1991
HIVJY1	-----T-----T-----C-----	2100
HIVNDK	-----T-----G-----AG-----A-----	2049
HIVNAL	-----A-----C-----A-----	2085
HIVELI	-----A-----C-----	2070
HIVZ226	-----GA-----C-----G-----A-----	2070
HIVUG024	-----A-----C-----	2064
HIVUG266A	-----A-----C-----G-----CT-----	2064
HIVUG269A	-----GA-----T-----G-----A-----A-----	2069
HIVUG274A	-----G-----	2102
HIVSE365A	-----G-----A-----A-----	2064
CONSENSUS. E	GGAAATTGGTTTtGACATAACAAaTGGCTGTGGTATATAAAaATATTTATAATgATAGTAGGAGGTTTaATA	1999
HIVTH966	---G-----	2057
HIVTH975	-----A-----	2051
HIVTH022	-----G-----	2051
HIVTN243	---G-----	2076
HIVTH006	-----	2048
CONSENSUS. F	GGAAATTGGTTT??CAT?TCAAA?TGGCTGTGGTATATAAAaATATTCATAATgATAGTAGGAGGCTTgATA	1966
HIVBZ163A	---AG---A---C---	2028
HIVBZ126A	---GA---G---T---	2036
CONSENSUS. G	G?A?TTGGTTT??CATA?CA??TGGCTATGGTATATAAAaATATTTATAATG?TAGTAGGAGGTTTaATA	1954
HIVV1525A	-A-A-----GA---T---AGC-----G-----	2112
HIVLBV217	-G-G-----AG---A---CAA-----A-----	2082
CONSENSUS. O	GGAAATTGGCTTGACATAAaTAAATGGTTGTGGTATATAAAaATAGCtATAATCA?TAGTaGGAGCACTAaTA	2005
HIVANT70	-----T-----	2079
HIVNVP5180	-----T-----G-----	2118
HIVVAU	-----	2121
CONSENSUS. CPZ	GGAAATTGGTTTGACATAACA?AATGGCT?TGGTATATAAAaAT?TT?TTATG??AGTAGGAG??AT??TA	1175
STVCPZQAB	-----A-----G-----T---CC---GC---GA---CA---	2052
CONSENSUS. U	GGAAATTGGTTTgGCATAACAAaTGGCTgTGGTATATAAAaATATTCaTAATgATAGTAGGAGGCTTgATA	1938
HIVZAM184	-----A-----G-----A-----T-----A-----	2034
HIVK124A	-----A-----G-----	2088
HIVZ3	---C-----T---T---	2065

HIV1 ENV

	\ / 3'sj (minor)		
	\ / 3'sj (minor)		
CONSENSUS.A	GGATTAAgaaTAgTTTtTgCtGTgcTTtCtATaATaAATAGAGTTAGGcAGGGATACTCACcttTgTC?TT		2036
HIVSF1703	-----C-----A--		2165
HIVU455	-----A-----C-----C--G--		2129
HIVZ321	-----A-----		2150
HIVRW020	-----G-----C--A--		2095
HIVUG031	-----A-----T-----C--A--		1999
HIVD687	--T-----A--AA-----G--		1791
HIVUG275A	-----G-----G--		2135
HIVUG273A	-----C-----T--G-----A-----G--A--		2147
HIVV1191A	--T-----		2144
HIVDJ264A	--G-----C-----A--G-----G-----A--A--		2120
HIVDJ263A	--T--G-----A-----A--G--		2144
HIVDJ258A	--T-----A-----A--G--		2129
HIVKENYA	-----A-----C--G--		2126
CONSENSUS.B	GGtTTAAgaaTAgTTTTTgCtGTaCTtTCTaTAGTgAATAgAGtTAGGCAGGGaTAcTCACCctTaTCgTT		2078
HIVJRC5F	-----T-----C-----		2126
HIVJRF1	-----C-----A-----		2123
HIVALA1	-----T-----		2129
HIVBRVA	-----		2138
HIVJH3	-----A-----C-----		2183
HIVBAL1	-----T-----		2144
HIVYU2	-----T-----		2111
HIVMN	-----G-----		2153
HIVKX2R	-----T-----		2150
HIVLAI	-----T-----		2165
HIVNL43	-----T-----		2144
HIVMFA	-----T-----		2144
HIVCAN1	-----A--A-----T-----G--A--		2150
HIVNY5CG	-----A-----		2132
HIVADA	-----A-----A-----G--A--		2141
HIVJPL	-----A--G-----G--A--		2135
HIVSEN1B4	-----A-----		2150
HIVD31	-----T-----A--		2132
HIVSF162	-----G-----A--G-----A--		2123
HIVBC3G3C	-----A-----G-----		2159
HIVOF1	-----		2147
HIVSF33	-----A-----G-----A-----A--		2138
HIVCDC4	-----G-----		2186
HIVSF2	-----G-----G--A--		2147
HIVSF2B13	-----G-----G-----A--		2141
HIVHAM	-----G-----G-----		2147
HIVRF	--C--A-----G-----A--		2177
HIVWJ2	-----A-----A-----A--		2123
HIVTB132	-----A--A-----G-----		2093
HIVBR020	-----A--AG-----G--A--		2137
HIVTH014	-----A-----A-----G--A--		2113
HIVUS711	-----T-----A--		2125
HIVUS712	--C-----AG-----		2134
HIVUS715	-----AT-----T-----T-----		2140
HIVUS716	-----G-----AT-----C-----G--A--		2113
HIVHA593	-----G-----A-----G--A--		2203
HIVHA594	-----G-----A-----A--		2125
HIVHA596	-----G--A-----A--		2146
HIVHA599	--A-----A-----C-----		2176
HIVHA651	-----A-----A-----A--		2128
HIVRJS	-----G-----G--A--		2141
HIVGUM	-----A-----T-----		2037
HIVSC	-----A-----		2150
HIVSBA	-----		1335
HIVSBB	-----		1362
HIVSBC	--G-----		1329
HIVP896	-----A-----T-----		2141

	\\ 3'sj (minor)	
	\\ 3'sj (minor)	
CONSENSUS.C	GGCTTAAGaATAATTTTGTCTGTCTCTCTATAGTgAATAGAGTtAgGcagGGaTACTCACCTtTgTCgTT	2063
HIVMA959	-----T-----A-----	2119
HIVMA960	---G-----A-----	2107
HIVD757	---G-A-----A-----	1782
HIVD747	---C-G-----T-G-----A-----A-GCA-----	1923
HIVD760	---G-----A-----	1943
HIVZAM20A	-----A-----	2210
HIVSM145A	-----T-----	2144
HIVZAM18A	---C-----G-----T-----	2132
HIVDJ259A	-----G-----GC-A-----	2078
HIVDJ373A	-----T-----A-A-A-----	2111
HIVSE164A	-----C-----	2105
HIVUG266A	-----A-----A-A-----	2129
CONSENSUS.D	GGCTTAAGaATAgTtTTTGTCTGTCTTTCCTTAGTAATAGAGTtTAGGCAGGGATtTCACCTCTgTCtT	2061
HIVJY1	-----A-----A-----	2171
HIVNDK	-----G-----C-----A-----	2120
HIVMAL	-----A-----C-----G-----	2156
HIVELI	-----C-----G-----	2141
HIV2226	-----A-----	2141
HIVUG024	-----A-----	2135
HIVUG266A	---G---C-----GC-----	2135
HIVUG269A	---G---T-----C-G-----	2140
HIVUG274A	-----A-----A-----	2173
HIVSE365A	---G---A---C-----G-----	2135
CONSENSUS.E	GGCTTAAGaATAATTTTGTCTGTCTTTCCTATAGTgAATAGAGTtTAGGCAGGGATtTCACCTTTGTCTtT	2070
HIVTH966	-----A-----	2128
HIVTH975	-----A-----	2122
HIVTH022	-----A-----	2122
HIVTN243	-----K-----C-----	2147
HIVTH006	-----A-----	2119
CONSENSUS.F	GGCTTAAGaATAGTtTTTGTCTGTCTTTCCTATAGTAAATAGAGTtTAGGAAGG?ATACTCACCTTTGTCAATT	2036
HIVB2163A	-----C-----	2099
HIVB2126A	-----G-----	2107
CONSENSUS.G	GGTTTAA?AATAGTtTTTGTCTGTCTTTCCTATAGT?AATAGAGT?AGGCAGGGATtTCACCTCT?TCATT	2021
HIVV1525A	---G---A---C-----G-----	2183
HIVLBV217	---A---G---T-----A-----	2153
CONSENSUS.O	GGtGTaAGAGTtAtCATGATAgTACTTAACTTAGTgAAgAACATTAGGCAGGGATATCAACCCCTCTCGTT	2076
HIVANT70	---G---C-----A-----A-----	2150
HIVMVP5180	---A---A-----A-----	2189
HIVVAD	---AG-----A-----	2192
CONSENSUS.CP2	GG??TAAGaATT?T??T??A?T??T??T??T??T?AG?A?GTTAGGCAGGG?TA??TTCCTCT?TC?TT	1216
SIVCP2GAB	--GT-----A-TA-GAC-G-TT-CTCACTAG-C-G-GA-----C-CTCA-----C-G--	2123
CONSENSUS.U	GG?TTAAGaATAGTtTTTGTCTgTgTTTCT?TAgTAAATAGAGTtTAGGCAGGGATtTCACCTtTgTC?TT	2005
HIVZAM184	---A-----A-----AA-----G-----A-----	2105
HIVK124A	---T---G-----C-----T-----T-----C-----G-----	2159
HIV23	---C-----A-G-T-----A-A-----T-----	2136

HIV1 ENV

	1/ 3'sj (major)	
CONSENSUS.A	CCAGACCCttaCCCCa?aCCcaagGGaa??CCCGaCAGgCCCgaAaGAATCGaaGaAGaAGGTGGa...g	2100
HIVSF1703	T---T-A---A---GT	2230
HIVU455	T---G---AT---GA---GT...T-G---A-G---	2194
HIVZ321	---ACC---AG---	2215
HIVRW020	---T-G---A---A---T-G---C-G---	2160
HIVUG031	T---A---A---GT...T---G---	2064
HIVD687	---CC---AG---G---G---C---	1856
HIVUG275A	---A---A---GT...T---G---A	2200
HIVUG273A	T---A---A---GT...T---C---	2212
HIVVI191A	---CC---ACC---AG---G---G---A---	2209
HIVDJ264A	---ACC---AG---	2185
HIVDJ263A	---ACC---AG---G---A---G---C---	2209
HIVDJ258A	---ACC---TGG---CCA---A---G---C---	2197
HIVKENYA	---A---GC---GT...T---	2191
CONSENSUS.B	tCAGACCC?CctCCCCagccccgaGGGga...CcGacagGCCCGaaGgAatcGaAGaAGaAGGTGGA...G	2142
HIVJRC5F	---T---AA---	2191
HIVJRF1	---T---	2188
HIVALA1	---G---AG---A---	2194
HIVBRVA	---G---G---	2203
HIVJH3	---G---	2248
HIVBAL1	---A---T---T---G---	2209
HIVYU2	---A---T---A---C---	2176
HIVMN	G---G---C---TT---	2218
HIVHXB2R	---A---A---A---	2215
HIVLA1	---A---A---A---	2230
HIVML43	---A---AT---	2209
HIVMFA	---A---AA---G---A---	2209
HIVCAM1	---G---T---TT---A---	2215
HIVNYS6G	C---G---A---	2197
HIVADA	---A---	2206
HIVJFL	C---G---A---G---	2200
HIVSIH184	---T---	2215
HIVD31	---G---	2197
HIVSP162	---G---T---A---	2188
HIVBCSG3C	---A---A---	2224
HIVQY1	---G---A---A---	2212
HIVSF33	---T---A---A---G---	2203
HIVCDC4	---T---AA---C---G---	2251
HIVSF2	---G---T---C---	2212
HIVSP2B13	---G---A---C---	2206
HIVHAN	---T---AA---CGACA---A---	2212
HIVRF	---A---A---G---	2242
HIVRMJ2	---A---A---A---A---	2188
HIVTB132	---A---V---G---	2158
HIVBR020	---G---A---A---A---A---	2202
HIVTH014	A---G---T---A---GAG---A---	2178
HIVUS711	---G---	2190
HIVUS712	C---A---G---G---	2199
HIVUS715	C---A---AG---	2205
HIVUS716	C---A---A---T---	2178
HIVHA593	---A---T---	2260
HIVHA594	---T---G---A---G---	2190
HIVHA596	---T---G---A---G---	2211
HIVHA599	---T---T---G---	2241
HIVHA651	---G---A---C---	2193
HIVRJ5	C---G---A---	2206
HIVGUN	C---A---A---A---C---G---	2102
HIVSC	---G---AG---A---	2215
HIVP896	---T---T---CA---	2206

		V 3'sj (major)	
CONSENSUS.C	tCAGACCCCTTACCCCaAaCccgaGGGga???	CCGACAGGCTCggAagAATCGAAGAAGaAGGTGGA...	G 2128
HIVMA959	-----	-----	2184
HIVMA960	-----	-----T-----	2172
HIVD757	-----G-----AG-----	-----A-----G-----	1847
HIVD747	-----G-----	-----	1941
HIVD760	-----	-----	1950
HIVZAM20A	-----G-----ATC-----GGAA-----	-----	2278
HIVSM145A	-----	-----	2209
HIVZAM18A	-----G-----A-----	-----	2197
HIVDJ259A	-----	-----C-----G-----	2143
HIVDJ373A	-----T-----G-----A-----	-----C-----A-----G-----	2176
HIVSE364A	C-----T-----	-----	2170
HIVUG268A	-----C-----	-----C-----GA-----	2194
CONSENSUS.D	tCAGACCCCTCCCTCCCaGccCCGAGGGga???	CCCCGACAGGCCCGaAGgAATAGAAGAAGaAGGTGGA...	G 2126
HIVJY1	-----	-----	2236
HIVNDK	-----T-----	-----A-----	2185
HIVMAL	G-----A-----A-----CCA-----	-----	2224
HIVELI	-----	-----C-----	2206
HIVZ226	-----	-----	2206
HIVUG024	-----A-----	-----G-----A-----C-----	2200
HIVUG265A	-----	-----C-----	2200
HIVUG269A	-----	-----	2205
HIVUG274A	-----A-----	-----A-----	2238
HIVSE365A	-----A-----	-----	2200
CONSENSUS.E	CCAGACCCcTaoCCATCATCAGaGGGAA...	CcCGACAGACCCCGAAGgAATCGAAGAAGgAGGTGGC...	G 2135
HIVTH966	-----T-----G-----	-----	2193
HIVTH975	-----TA-----	-----A-----	2187
HIVTH022	-----T-----C-----	-----	2187
HIVTN243	-----	-----T-----	2212
HIVTH006	-----T-----G-----	-----	2184
CONSENSUS.F	?CAGACCCATATCCCAAGCCCCGAGGGAA...	CCCGACAGGCCCGAAG?AATCGAAGAAGGAGGTGGA...	G 2099
HIVB2163A	T-----	-----A-----	2164
HIVB2126A	A-----	-----G-----	2172
CONSENSUS.G	CCAGACCCCTT?CCCA?CA?CA??GGGAA...	CCCAGACAGGC?GAAAGAATCGAAGAAGGAGGTGGC...	G 2079
HIVV1525A	-----A-----T-----C-----AG-----	-----TC-----	2248
HIVLBV217	-----G-----C-----T-----GA-----	-----CT-----	2218
CONSENSUS.O	acCAGATCCcCatCCaAcCaagaGGAA...	GcAGgAACGCCAGGAaGAACAGGAaGAGgAGGTGGA...	G 2141
HIVANT70	-----A-----T-----	-----G-----	2215
HIVMVP5180	G-----TC-----C-----GGC-----	-----A-----A-----	2254
HIVVAU	-----A-----C-----	-----T-----G-----	2257
CONSENSUS.CPZ	?CAGA?CC?TA?CC?A?CCAG?GGA?...	C?AG???G?G?G?A?AAT???G?AAG?AGGTGG?...	? 1256
SIVCP2GAB	G---C---T---T---C---GT---AG---A...	A---GAAG---TC---G---CGAC---G---C...	G 2188
CONSENSUS.U	tCAGACCCtccctCCCA?cCCCGagGGGa...	CccGACAGGCCcGgAaGAATcGAAGAAGaAGGTGGA...	G 2069
HIVZAM1B4	-----T-----CTAC-----CA-----AGA-----T-----	-----T-----G-----	2170
HIVK124A	-----G-----	-----A-----A-----G-----A-----	2224
HIV23	C-----	-----A-----A-----C-----	2201

HIV1 ENV

CONSENSUS .A	AGCAAGaCagAGaCAGaTCgattCgattAgTgAgCGGatTCCTTAgCacTTGCCTGGGACGAcCTgCGGAgc	2171
HIVSP1703	-----G-----	2301
HIVU455	-----G-A-----C-----GA-----A-----	2365
HIV2321	-----C-----C-G-----C-----A-----	2286
HIVRW020	-----	2232
HIVUG031	-----G-----G-----T-----	2135
HIVD687	-----G-CG-G-C-----T-----A-----	1927
HIVUG275A	-----AC-----	2271
HIVUG273A	-----C-----C-----	2283
HIVV1191A	-----G-----C-----T-----A-----	2280
HIVDJ264A	-----A-----CG-G-----C-----T-----	2256
HIVDJ263A	-----AGGG-----T-----	2280
HIVDJ258A	-----A-----A-A-C-----T-----	2268
HIVKENYA	-----C-----T-A-----	2262
CONSENSUS .B	agagAgacAGAgacagATCCggtcgattAGtggaatGGaTtcTtagcacTtaTCTGGgacGAtcTgcGgAgc	2213
HIVJRCSE	-----A-A-----C-----T-C-----	2262
HIVJREF	-----A-----C-----T-C-----	2259
HIVALA1	-----G-----A-----T-C-A-----	2265
HIVBRVA	-----C-----A-CC-----G-----T-----	2274
HIVJH3	-----C-----A-----G-----A-----T-C-----	2319
HIVBAL1	-----C-----C-----G-G-----T-----	2280
HIVYU2	-----C-----G-C-----A-----T-C-A-----	2247
HIVSN	-----C-----C-----A-----T-C-----	2289
HIVHXB2R	-----AT-----C-C-G-----	2286
HIVLAI	-----AT-----C-C-----	2301
HIVNL43	-----AT-----C-C-----	2280
HIVMFA	-----T-----C-C-----	2280
HIVCAN1	G-----C-----G-----A-C-----C-----C-----	2286
HIVWY5CG	-----A-C-----C-----T-----	2268
HIVADA	-C-----TG-----G-----T-----C-----	2277
HIVJFL	-----TAC-----C-----A-C-T-----	2271
HIVSINI84	-----A-A-----C-----C-A-C-----G-C-A-----	2286
HIVD31	-C-----AA-----A-----C-----A-C-----	2268
HIVSF162	-----A-C-----C-----A-----C-----A-----	2259
HIVBCSG3C	-----C-----T-----	2295
HIVQY1	-----AA-----G-----	2283
HIVSF33	-----AC-----C-----GT-----T-----	2274
HIVCDC4	-----G-----G-AC-----C-C-----G-----	2322
HIVSF2	-----T-----G-----A-----	2283
HIVSF2B13	-----T-----G-----A-----	2277
HIVHAN	-----G-----T-----GC-----T-----C-----	2283
HIVRF	-----CG-TGC-----GA-----T-C-----	2313
HIVMMJ2	-----T-----G-C-----	2259
HIVTB132	-----C-----A-C-----G-----A-T-----T-----A-----	2229
HIVBR020	-C-----G-----G-----A-C-----C-----A-A-A-C-----C-A-----	2273
HIVTH014	-CA-----G-----AGA-----GC-----A-GA-----G-----A-----	2249
HIVUS711	-----C-G-----A-----C-----A-----T-----A-----	2259
HIVUS712	-----A-----G-----A-C-----T-----C-----	2270
HIVUS715	-----A-A-----C-----A-C-----T-----C-----	2276
HIVUS716	-A-----A-----G-----A-A-C-----T-----C-A-----	2249
HIVHA593	-----G-----TGA-----C-----C-----	2339
HIVHA594	-----G-----AG-CA-----	2261
HIVHA596	-----G-----AG-CA-----	2282
HIVHA599	-----G-C-----GA-----C-----A-----	2312
HIVHA651	-----A-C-----CAC-----C-----	2264
HIVRJS	-----C-----C-----C-----C-A-----	2277
HIVGUN	-----C-----C-----C-----	2173
HIVSC	-----G-----A-----T-----C-----	2277
HIVP896	-----C-----C-----G-----T-----T-----T-A-A-----	2277

CONSENSUS .C	AGCAAGACA?AGaCagaTccattcGaTTAgTga?CGGATTCTTgCacTTgcCTGGGACGAcCTgCGGAGC	2197
HIVMA959	-----A-----G-----	2255
HIVMA960	-----A-----C-----A-----TT-----	2243
HIVD757	-----A-----C-----A-----	1917
HIV2AM20A	-----G-----C-----A-----T-----	2349
HIVSM145A	-----A-----A-----A-----A-----	2280
HIV2AM18A	-----G-----A-----C-----GA-----TT-----	2268
HIVDJ259A	-----G-----G-----G-----G-----G-----	2214
HIVDJ373A	-----A-----G-----C-----G-----	2247
HIVSE364A	-----G-----G-----G-----GT-CA-----	2241
HIVUG268A	-----G-----G-GA-----A-C-G-----T-----	2265
CONSENSUS .D	AGCaAGGCAGAG?cAGATCcaTtCGATTggTGAaCGGATTCTcagCaCTcatCTGGGACGacCTgaGGAaC	2196
HIVJY1	-----A-----A-----A-----CT-----T-----CC-----	2307
HIVMDK	-----G-----A-----A-----TT-----T-----	2256
HIVMAL	-----G-----A-----A-----	2295
HIVELI	-----G-----A-----G-GA-----C-----G-----C-----G-----	2277
HIV22E6	-----G-----A-----A-----T-----	2277
HIVUG024	-----G-----C-----G-----T-----	2271
HIVUG266A	-----GT-----A-----G-----T-----GC-----C-----G-----	2271
HIVUG269A	-----G-----A-----G-----	2276
HIVUG274A	-----G-----T-----	2309
HIVBE365A	-----A-----A-----GT-----C-----	2271
CONSENSUS .E	agCAAGGCAGAGAcAGATCCGTGcGaTTAGTGAgCGGATTCTTAGCaCTTgcCTGGGACGATCTACGGAGC	2206
HIVTH966	-----A-----C-----	2264
HIVTH975	-----A-----A-----AT-----A-----	2258
HIVTH022	-----G-----	2258
HIVTN243	-A-----A-----C-----	2283
HIVTH006	C-----	2255
CONSENSUS .F	AGCAAGGCA??GA?AGATCCGTG?G?TTAGT??ACGGATTCTTAGCTCTTGTCGGGACGACCTGAGGAAC	2163
HIVB2163A	-----GA-----C-----A-----A-----AG-----	2235
HIVB2126A	-----AC-----G-----C-----C-----GA-----	2243
CONSENSUS .G	AGCA?GACAGA?cAGATCC??GCG?T?AG?GA?CGGATTCTT?c?c?TTG?CTGGGACGAC?T??GGAGC	2133
HIVVI525A	-----C-----AC-----AC-----A-C-----C-----A-----AC-AC-----T-----C-CT-----	2319
HIVLBV217	-----A-----GA-----GT-----C-T-----T-----G-----GT-GA-----C-----T-GA-----	2289
CONSENSUS .O	AaGaAGaCAGGccCagGTGGAcacCcTtCCCGCaAGGgTTCTTGCaactGTTGTACACGGAcCTCAGgACA	2212
HIVANT70	-----G-----T-----C-----C-----	2286
HIVMP5180	-----G-----A-----G-----A-C-----A-----A-----T-----	2325
HIVVAU	-C-----G-----T-----A-----T-----	2328
CONSENSUS .CPZ	??AAGACAG?A?CAG?T??G??T??G??A?GGGT?CTP??CACT??T?TGGGACGACCTCCGAAAC	1303
STVCP2GAB	AGC-----A-G-----A-CAGT-AGA--AGT-G-A-----G-----GC-----TA-C-----	2259
CONSENSUS .U	AgcAAGgCagagACAGaTccaTTCGATTaGTGAaCGGATTCTTAgCaCTTgcCTGGGAcGAcCTGaGGagC	2140
HIV2AM184	-----G-----G-----G-----G-----T-----	2241
HIVK124A	-----G-----G-----CC-----AT-----GA-----	2295
HIV23	-AG---A---ACA-----G-----A-----C-----	2272

HIV1 ENV

CONSENSUS . A	CTGTGCTCCTTCaGCTACCACCG?PTGAGAGACTTcctcTcGATTGca . gcgaggactGTGgAactTC TGg	2240
HIVSF1703	-----C-----C-----A	2371
HIVU455	-----A-----GC-----T-----AG-----	2335
HIVZ321	-----T-----T-----A-----G-GCA-----TC-----GAC-----	2355
HIVRM020	-----C-----C-----C-----	2301
HIVUG031	-----C-----C-----	2205
HIVD687	-----	1941
HIVUG275A	-----A-----	2341
HIVUG273A	-----C-----TC-----A-----	2353
HIVVI191A	-----T-----A-----A-----T-----GAC-----	2350
HIVDJ264A	-----AC-----TG-----T-----G-----	2326
HIVDJ263A	-----A-----TG-----T-----G-----	2350
HIVDJ258A	-----AA-----TG-----T-----G-----	2338
HIVKENYA	-----C-----C A-TGA-----	2332
CONSENSUS . B	CTGTgCcctcTTCagCTACCcCcgcttgAgAgACTTAcTcTTGattgta . gcgAgGATTCGTGGAactcCTGg	2283
HIVJRCSP	-----T-----C-----C-----	2332
HIVJRFL	-----C-----A-----	2329
HIVALA1	-----T-----A-----	2335
HIVBRVA	-----T-----A-----	2344
HIVJH3	-----A-----	2389
HIVBAL1	-----T-----G AT-----	2350
HIVYU2	-----T-----A-----	2317
HIVNM	-----T-----A-----C-----	2356
HIVHXB2R	-----A-----	2356
HIVLAI	-----A-----	2371
HIVNL43	-----A-----	2350
HIVNEA	-----G-----A-----	2350
HIVCAM1	-----A-----	2356
HIVHY5CG	-----A-----	2338
HIVADA	-----	2347
HIVJFL	-----A-----C-----	2341
HIVSIMI84	-----T-----C-----	2356
HIVD31	-----CT-----C-----	2338
HIVSF162	-----A-----C-----	2329
HIVBCSG3C	-----A-----G-----A-----	2365
HIVOYI	-----A-----	2353
HIVSF33	-----C-----	2344
HIVCDC4	-----	2392
HIVSF2	-----G-----C-----C-----A-----	2353
HIVSF2B13	-----G-----C-----T-----C-----A-----	2347
HIVHAN	-----G-----A-----A-----AC-----	2353
HIVRF	-----AG-----A-----T-----	2383
HIVNM32	-----G AA-----	2329
HIVTB132	-----G-----A-----	2299
HIVBR020	-----T-----G-----G-----	2343
HIVTH014	-----A-----	2319
HIVUS711	-----A-----	2329
HIVUS712	-----T-----A-----	2340
HIVUS715	-----	2346
HIVUS716	-----A-----	2319
HIVHA593	-----A-----T-----AA-----A-----	2409
HIVHA594	-----T-----	2331
HIVHA596	-----C-----	2352
HIVHA599	-----A-----	2382
HIVHA651	-----A-G-----AA-----A-----	2334
HIVRJS	-----	2347
HIVGUN	-----T-----	2243
HIVP096	-----CT-----T-----A-----A-----	2347

CONSENSUS.C	CTGTgcCtctTCAGcTACCACCGaTTGAGAGACTTcsttATTGaTtGcA.GCGAGAGcgGtGGAACTtCTGg	2267
HIVMA959	-----T-----G-----A-----	2325
HIVMA960	-----A-----AC-----G-----T-----C-----	2313
HIVD757	-----	1929
HIVZAM20A	-----GT-----A-----C-----	2419
HIVSM145A	-----T-----G-----G-----	2350
HIVZAM18A	-----T-----A-----C-----	2330
HIVDJ259A	-----G-----A-----C-----G-----	2284
HIVDJ373A	-----T-----T-----A-----C-----T-----T-----	2317
HIVSE364A	-----T-----A-----C-----A-----	2311
HIVUG268A	-----G-----C-----AC-----A-----	2335
CONSENSUS.D	CTGTGCCTCTTCAGcTACCACCGcTTGAGAGACTtAaTCTTAATTGCA.gcGAGgATTGTgGaacTtCTGG	2266
HIVJY1	-----A-----	2377
HIVNDK	-----C-----	2326
HIVMAL	-----T-----C-----A-----	2365
HIVELI	-----T-----A-----	2347
HIVZ226	-----G-----	2347
HIVUG024	-----A-----A-----C-----	2341
HIVUG265A	-----C-----A-----	2341
HIVUG269A	-----A-----A-----	2346
HIVUG274A	-----G-----C-----A-----C-----	2379
HIVSE365A	-----G-----	2341
CONSENSUS.E	CTGTGCCTcTTCagCTACCACCGCTTGAGAGACTTCATCTcGATTGCA.GCGAGGAcTgTGGAACTTCTGG	2276
HIVTH966	-----C-----	2314
HIVTH975	-----CT-----C-----	2328
HIVTH022	-----	2328
HIVTN243	-----T-----C-----G-----	2353
HIVTH006	-----G-----	2325
CONSENSUS.F	CTGTGCCTCTTCAGCTACC?CCAC?TGAGAGACTTA?TATTAATA?CA.GCGAG?ATTGTG.....	2218
HIVB2153A	-----A-----C-----C-----A-----A-----	2295
HIVB2126A	-----G-----T-----A-----G-----G-----	2303
CONSENSUS.G	CTGTGCCT?TTCAGCTACC?CC?TTGAGAGACTT??TCTT?ATTGTA.?CGAGGAC?GTGGAAC?TCTGG	2194
HIVV1525A	-----T-----G-----TC-----AC-----A-----G-----T-----	2389
HIVLBV217	-----C-----A-----AA-----CA-----G-----A-----A-----	2359
CONSENSUS.O	ATAATcTTGTGGAcTTACCAcCTCTTGAGCAACTTAgcatCAGgGATC.....	2240
HIVANT70	-----A-----	2334
HIVHVP5180	-----AT-----	2373
HIVVAU	-----T-----C-----A-----	2376
CONSENSUS.CPZ	CTGGGGATTTCGAGCTACCAGAGCTTGACAAcCTTAGCCTGCAATGTG.TGGAGACAAC?CAAGACCTTGG	1373
SIVCP2GAB	-----	2329
CONSENSUS.U	CTGTGCcTCTTCAGCTACCACCGCTTGAGAGAcTtAATCTTgATTGCA.?cgagga?tGTggAacttCTGg	2208
HIVZAM184	-----T-----G-----C-----	2311
HIVK124A	-----A-----A-----T-----	2365
HIVZ3	-----A-----T-----AA-GGGA-----	2332

HIV1 ENV

CONSENSUS .A	gacacagcag?ctcaaggg?ctgagagctgggg.....cgggaagggccTCaAgTAc	2288
HIVSF1703	A--G-----T---G--G-----C	2421
HIVU455	---G-----C-----A-----G-----A---	2385
HIV2321	A-CC-CAQ-----GACT-----A--C	2385
HIVRW020	-----T-----A-----	2351
HIVUG031	A---A---T-----GT-----A-----	2255
HIVUG275A	---G-----	2358
HIVUG273A	---G-----AA-----	2382
HIVVE191A	--G--G-----G-CT---A---	2379
XIVDJ264A	-----T-----A-----T-----G-C-----A---	2376
HIVDJ263A	-----C-----A-----C-----C-A---	2400
HIVDJ258A	-----C-----A-----C-----C-----	2388
XIVKENYA	-----T---G--GT-----	2382
CONSENSUS .B	ga.....???cgcaGGGGG.....TGGGaAgccCTcAaaTAc	2312
HIVJRC3F	--.....AT---G---C	2361
HIVJRF1	--.....T---G-----	2358
HIVALA1	--.....T---A-----C	2364
HIVBRVA	--.....T-----	2371
HIVJH3	--.....	2418
HIVBAL1	CTAGCAG-----T-----	2382
HIVYU2	--.....G--T-----	2346
HIVMN	-----T-----	2385
HIVHXB2R	--.....	2385
HIVLA1	--.....	2400
HIVNL43	--.....	2379
HIVNPA	--.....	2379
HIVCAM1	--.....	2385
HIVNY5CG	--.....	2367
HIVADA	-----T-----G---	2376
HIVJFL	--.....	2370
HIVSIMI84	-----AT-----C	2385
HIVDJ1	--.....T-----	2367
HIVSF162	--.....G---	2358
HIVBCSG3C	--.....G---	2394
HIVQY1	--.....T---G---	2382
HIVSF33	--.....T-----	2373
HIVCDC4	--.....T-----	2421
HIVSF2	-G-----A-----	2382
HIVSF2B13	--.....A-----T-----	2376
HIVHAN	--.....T---G---	2382
HIVRF	--.....T---G---	2412
HIVVMJ2	-----G---	2358
HIVTB132	r-----	2328
HIVBR020	--.....	2372
HIVTH014	-----G---	2348
HIVUS711	--.....CT-----	2358
HIVUS712	--.....	2369
HIVUS715	-----AT-----	2375
HIVUS716	--.....T-----	2348
HIVHA593	--.....T-----	2438
HIVHA594	-----G---	2360
HIVHA596	--.....G---	2381
HIVHA599	T-----G---C	2411
HIVHA651	-----G---C	2363
HIVRJS	--.....G---	2376
HIVGUN	-----TT-----	2272
HIVP896	-----	2376

CONSENSUS.C	gAcgcagcagtcctcaggggactaCAGAggGGG.....TGGGAagccCTTAAGTAT	2317
HIVNA959	-----A-----C-----	2375
HIVMA960	-----C-----	2383
HIVZAM20A	---G-----A---	2469
HIVSM145A	-----A-----A---	2400
HIVZAM18A	-----	2388
HIVDJ259A	-----A-----	2313
HIVDJ373A	-----A-----	2346
HIVSE364A	T-----AT-----C	2340
HIVUG268A	-----	2385
CONSENSUS.D	GAcgc.....??AGgGGG.....TGGGAagccCTCAAAATAT	2295
HIVJY1	-----A-----	2406
HIVNDK	-----C-----	2355
HIVNAL	-----	2394
HIVELI	-----CAT-----	2376
HIVZ225	-----	2376
HIVUG024	-----AA-----C	2370
HIVUG266A	---CGC-----C	2370
HIVUG269A	---CGC-----A	2375
HIVUG274A	---CGC-----A-----A-----	2408
HIVSE365A	---CGC-A-----T-----G---	2370
CONSENSUS.E	GACaCAGCAGTCCTCAAGGGACTgAGACGGGGG.....TGGGAAGCCCTCAAAATAT	2326
HIVTH966	---G-----	2384
HIVTH975	-----	2378
HIVTH022	-----	2378
HIVTN243	-----A-----	2403
HIVTH006	-----	2346
CONSENSUS.F	AACAGGGGGCT?AGGAGGGGG.....TGGGAAGCCCTCAAACTT	2256
HIVBZ163A	-----G-----	2334
HIVBZ126A	-----A-----	2342
CONSENSUS.G	G??GTA???G? ??????????????????.....TGGGA?G?CCTCAAAATA?	2213
HIVV1525A	-GA-A-GGG-G.....A-C-----T	2418
HIVLBV217	-AC-C-ACA-CCTCAAGGGACTGAGACTGGGG.....G-G-----C	2409
CONSENSUS.O	..CeGAaG?TGATCagctAcCTGgGactTcGGACTgTGGATCcTaGGgCagAaGAcAATTgAaGcTTGcAGA	2328
HIVAMP70	---G-----T---A-G-----T---A-T-T---	2403
HIVMVP5180	---G-G-C---GA-----G-----G-A-A-----T---	2442
HIVVAU	---T---AC-----A-----A-----G-----	2445
CONSENSUS.CP2	GACATCTAATACTTCACAGCCTCAGACTGCTA.....CGGGAGAGACTGTGCCTG	1423
BIVCPZGAB	-----	2379
CONSENSUS.U	gac?????????????????cgGGGG.....TGGGAAGccCTCAAAatAt	2236
HIVZAM184	---ACAGCAGTCCTCAAGGGACTGA-A-T---	2361
HIVK124A	---C-CA-----	2394
HIV23	---GA-----A-----C	2359

HIV1 ENV

		<- rev	
CONSENSUS.A	CTG...TGGaAT.CTtCTg?taTatTGGGGGcGgGAaCTaAAAA?TAGTGTattAaTTTGTCTTGatCCa		2353
HIVSF1703	-----C-----A-----		2488
HIVU455	-----C-G-----T-----C-----CG-TG		2452
HIVZ321	..G-A-..-G-A-C-C-..-A-..-G-A-..-AG		2452
HIVRW020	-----T-----C-----G-----C-----G-----G-----		2418
HIVUG031	..G-----C-T-G-----C-----T-----		2322
HIVUG275A	-----A-AGC-----T-----G-----T-----C-----		2425
HIVUG273A	-----C-..T-G-T-..-C-----T-----		2449
HIVVI191A	-----C-T-----G-----T-----C-----G-T-----		2446
HIVDJ264A	..G-----AC-C-----G-A-----G-A-----A-----		2443
HIVDJ263A	..G-----AC-C-----A-----G-A-----A-----		2467
HIVDJ258A	..G-----AC-C-----A-----G-A-----G-----A-----		2455
HIVKENYA	-----C-T-G-C-----A-----T-----G-----G-----		2449
CONSENSUS.B	egg...TGGaAT.CTcCTgCagTATTGGcagtcAGGAacTAAaGaaTagTGetgTtagcTtGcTcaatgccA		2379
HIVJRC6F	-----A-----T-----		2428
HIVJRFL	-----A-----		2425
HIVALA1	-----C-----		2431
HIVBRVA	-----A-----C-----C-----		2440
HIVJH3	-----A-A-----T-----C-----		2485
HIVBAL1	-----A-----		2449
HIVYU2	-----C-----T-----C-----		2413
HIVMM	-----A-----G-----T-----		2452
HIVHXB2R	-----A-----		2452
HIVLAI	-----A-----		2467
HIVNL43	-----A-----A-----		2446
HIVMFA	-----A-A-----G-----		2446
HIVCAM1	-----G-----T-----T-TG-A-----		2452
HIVNY5CG	..T...-G-----A-----T-----T-----		2434
HIVADA	-----G-----G-----T-----T-----		2443
HIVJPL	-----A-----T-----T-----C-----T-----		2437
HIVSINI84	-----T-----A-A-----		2452
HIV031	-----		2434
HIVSF162	..G-----T-----T-----T-TG-----		2425
HIVBCSG3C	..-G-----A-----C-----A-----T-----T-----		2461
HIV0Y1	-----C-----TCA-----C-----		2449
HIVSF33	-----T-----C-----		2440
HIVCDC4	-----G-----T-----G-T-----T-----		2488
HIVSF2	..-G-----T-----G-----C-----		2449
HIVSF2B13	G-..-G-----T-----G-----C-----		2443
HIVHAN	-----C-----C-----T-----A-T-----		2449
HIVRF	-----G-----G-----T-----A-----		2479
HIVMAJ2	-----CA-----G-G-----T-----		2425
HIVTB132	-----A-----		2395
HIVBR020	-----A-A-----G-----G-----T-----C-----		2439
HIVTH014	-----T-----T-----		2415
HIVUS711	-----C-----T-----C-----		2425
HIVUS712	-----A-----G-----T-----		2436
HIVUS715	-----A-----G-G-----C-----		2442
HIVUS716	-----A-----G-----		2415
HIVHA593	-----A-----TG-----		2505
HIVHA594	..-G-----A-----A-----A-A-----T-G-----		2427
HIVHA596	..-G-----A-----A-----A-----T-----		2448
HIVHA599	G-..-----T-----		2478
HIVHA661	-----T-----		2430
HIVRJ8	-----TC-----		2403
HIVGUN	CT-..-T-----		2289
HIVP896	-----A-----C-----C-----		2443

		<- rev	
CONSENSUS.C	CTG...GGAAgt.CTTGtgCagTATGCGGTCtgGAgcTAAAAAGAGTGCTaTTAGTcTgCTTgATaCCa		2384
HIVMA959	---A---GA-T---A---A---A-AG---C		2442
HIVMA960	---A---T---A---A---AG---		2430
HIVZAM20A	---C---A---C---T---		2536
HIVSM145A	---A---GC---		2467
HIVZAM18A	---A---G---G---		2455
HIVDJ259A	---C---A---A---		2380
HIVDJ373A	---C---A---G---		2413
HIVSE364A	---CA---T---AA---		2407
HIVUG268A	---C---AA-A---T---		2452
CONSENSUS.D	ctG...TGGAAst.CTCCTgcAgTATTGGattCaGgAACTgBagAAtAGTgcTATTAgtCTTG?TTgAtaCCA		2361
HIVJY1	---G---C---C---T---C---A---G---		2473
HIVNDK	---C---G---G---G---C---C---		2422
HIVMAL	---A---GG---C---A---A---		2461
HIVELI	---A---G---G---C---G---T---G---		2443
HIVZ226	---A---G---G---C---G---C---		2443
HIVUG024	---C---A---T---A---C---		2437
HIVUG366A	TG---A---A---A---A---A---		2437
HIVUG269A	---AA-A---A---T---		2442
HIVUG274A	---A---A---T---		2475
HIVSE365A	---G---A---A---A---C---		2437
CONSENSUS.E	CTG...GGGAAT.CTTCTGCTATATTTGGgGcCAGGAACtAAAAATtAGTGCTATTTCTTTGCTTAAAGCTA		2393
HIVTH966	---C---G---C---		2451
HIVTH975	---G---A---C---		2445
HIVTH022	---T---		2445
HIVTN243	---		2470
CONSENSUS.F	CTG...GGGAAT.CTCACGCTGTATTGGGGTCAGGAACtAAAGAATAGTGCTATT?GCTTG?TTAACACCA		2321
HIVBZ163A	---G---T---		2401
HIVBZ126A	---A---C---		2409
CONSENSUS.G	?TG...TGGAAst.CT?CTGC?GTATTGGGG?CAGGAACtAAAGAATAGTGC?ATT?ATTGCTT?A?AC??		2270
HIVVI525A	C---T---A---A---C---C---A---CA		2485
HIVLBV217	T---C---T---T---T---A---G---T---AG		2476
CONSENSUS.O	CTtTgTg?AGCT.gTAAtaCAATACtGGCTACAGAAATTGCAaAaTAGTGCTACAAaccTgCTaGAtACT?		2396
HIVANT70	A---C---C---G---T---C---AC		2473
HIVMVP5180	---G---G---T---A---T---A		2512
HIVVAU	---C-T-AA---A---C---T---A---G		2515
CONSENSUS.CP2	CTG...GGGGGA.ATTATTCAATATTGGGGAAAGGAGCTAAAAATtAGTGCTATTAAGCCCTTGATGCCAA		1490
SIVCP2QAB	---		2446
CONSENSUS.U	CTg...gGGAAst.CTtCTGctgTATTGGggTCaGGAAcTAAAGAATAGTGCTATTAGCTTtactTgATaCCA		2303
HIVZAM184	---A---T---T---A---T---G---		2428
HIVK124A	---C---A---AC---G---C---G---		2461
HIVZ3	---G---G---A---G---		2426

HIV1 ENV

CONSENSUS . A	cAGCAaTAGcAgTAGCTggctGGAcAGAcAGggTTATAGAAaTAGgACaagAattggTAGaGcTaTtctc	2424
HIVSF1703	-----G-----G-----T-----C-----	2559
HIVU455	-----G-----G-----T-----C-----	2523
HIV2321	-----A-----A-----G-----T-----GC-----T-----	2523
HIVRW020	-----A-----G-----TT-----T-----	2489
HIVUG031	-----A-----A-----A-----C-----A-----	2393
HIVUG275A	-----G-----T-----A-----G-----A-----T-----C-----T-----T-----A-----	2496
HIVUG273A	-----A-----A-----G-----T-----T-----	2520
HIVV1191A	C-----A-----G-----A-----T-----	2517
HIVDJ264A	-----AA-----C-----A-----G-----G-----GT-----	2514
HIVDJ263A	-----T-----AAT-----C-----A-----G-----	2538
HIVDJ258A	-----T-----AA-----C-----A-----GA-----G-----	2526
HIVKENYA	-----T-----C-----G-----A-----	2520
CONSENSUS . B	caGCTaTAGcAGTAgcLGAggGgACACATAGGgtTaTAgAAgtAgtaCaagAgcctt?TAGaGcTaTtctc	2449
HIVJRC5F	-----C-----A-----C-----T-----A-----A-----G-----	2499
HIVJREFL	-----C-----A-----C-----T-----A-----A-----	2496
HIVALA1	-----C-----G-----A-----C-----T-----G-----T-----C-----T-----	2502
HIVBRVA	-----C-----G-----A-----C-----T-----T-----C-----G-----	2511
HIVJH3	T-----C-----A-----A-----A-----T-----T-----	2556
HIVBAL1	-----G-----A-----T-----G-----GP-----	2520
HIVYU2	-----A-----A-----T-----T-----G-----T-----	2484
HIVNN	-----C-----G-----GG-----	2523
HIVHXB2R	-----C-----G-----G-----G-----	2523
HIVLAI	-----C-----G-----G-----G-----	2538
HIVNL43	-----C-----T-----GC-----A-----G-----	2517
HIVMFA	-----G-----A-----G-----	2517
HIVCAM1	T-----A-----G-----G-----	2523
HIVNY5CG	-----G-----C-----AT-----G-----G-----C-----	2505
HIVADA	-----AT-----A-----G-----	2514
HIVJFL	-----A-----AT-----GG-----C-----	2508
HIVSIM184	-T-----T-----AT-----GG-----G-----	2523
HIVD31	-----C-----GC-----	2505
HIVSF162	T-----A-----C-----AT-----GG-----T-----	2496
HIVBCSG1C	-----C-----A-----A-----A-----A-----	2532
HIVQY1	-----A-----A-----T-----	2520
HIVSF33	-----A-----TGGG-----	2511
HIVCDC4	-----A-----AT-----A-----T-----	2559
HIVSF2	-----A-----C-----A-----	2520
HIVSF2B13	-----A-----A-----	2514
HIVHAN	T-----A-----G-----A-----	2520
HIVRF	-----A-----A-----A-----C-----AT-----CT-----T-----T-----	2550
HIVMAJ2	TT-----A-----ATC-----G-----A-----	2496
HIVTB132	-----A-----T-----A-----	2466
HIVBR020	-----AG-----A-----A-----A-----C-----TT-----T-----G-----T-----	2510
HIVTH014	-----A-----C-----	2486
HIVUS711	-----A-----A-----A-----	2496
HIVUS712	-----C-----T-----AT-----GG-----G-----	2507
HIVUS715	-----C-----T-----GG-----C-----	2513
HIVUS716	-----A-----T-----T-----GG-----G-----C-----	2486
HIVHA593	-----A-----A-----G-----T-----C-----	2576
HIVHA594	-----A-----A-----A-----AT-----GG-----	2498
HIVHA596	-----A-----A-----AT-----GG-----	2519
HIVHA599	-----A-----A-----G-----G-----T-----G-----T-----	2549
HIVHA651	-----A-----A-----A-----AT-----T-----G-----T-----	2501
HIVF896	-----C-----A-----A-----G-----	2514

CONSENSUS.C	tAGCAATAgCAGTAgCTGAAAGGAACAGATAGgATTaTAgAAtta?tACAAAGaaTtGtAGaGcTatCcgC	2454
HIVMA959	-----C-----T-----C-----G-----	2513
HIVMA960	-----C-----A-----C-----G-----TA-----	2501
HIVZAM20A	-----G-----G-----G-----T-----	2607
HIVSM145A	-----A-----GC-----G-G-----	2538
HIVZAM18A	-----A-----CA-----	2526
HIVDJ259A	C-----A-----A-----CA-----G-GC-----G-----T-----	2451
HIVDJ373A	C-----A-----A-----A-----G-----G-----T-----T-----	2484
HIVSE364A	-----A-----A-----G-----C-----G-----A-----	2478
HIVUG268A	-----T-----A-----G-GGG-----G-----G-G-----T-----	2523
CONSENSUS.D	?AGCAATAGCAGTAgCTGAgGgGACAGATAGggTtTATAGAAaTAgTACaAAGAgcT?gtAGAGCT?TTCTCt	2429
HIVJY1	C-----A-----C-----A-----G-----TT-----G-----	2544
HIVNDK	T-----A-----G-----T-----C-----A-----	2493
HIVNAL	C-----AT-C-----G-----TT-G-----A-----C-----	2532
HIVSLI	T-----A-----A-----A-----T-----C-----G-----	2514
HIVZ226	T-----G-----T-----C-----G-----	2514
HIVUG024	C-----T-----A-----C-----AT-G-----A-----	2508
HIVUG266A	C-----C-----C-----GT-----A-----	2508
HIVUG269A	T-----T-----G-----A-----GT-----G-----	2513
HIVUG274A	C-----A-----A-----G-----GT-----G-----	2546
HIVSE365A	T-----A-----A-----CG-----T-----A-----	2508
CONSENSUS.E	CAGCAATAGCAGTAgCGGGGTGGACAGATAgggTtTATAGAAgGTAGCACAAgGAgcTTGGAGAGCCATTCTC	2464
HIVTH966	-----G-----AT-----	2522
HIVTH975	-----A-----	2516
HIVTH022	-----A-----A-----	2516
HIVTN243	-----A-----	2541
CONSENSUS.F	CAGCAATAGTAGTA?CTGAGGGGAACAGATAGAAATTATAGCAGCTTTGCAAGAATTTGGTAGAGCTATT?T?	2389
HIVB2163A	-----G-----T-----G-----	2472
HIVB2126A	-----A-----C-T-----	2480
CONSENSUS.G	?AGCAATAGCA??AGCT?A??G?ACAGATAGGGT?ATAGAA?TAG?GCAAA?AGCTTGCAGAGCT?TTCT?	2328
HIVV1525A	C-----GT-----G-GG-A-----T-----A-----T-----T-----G-----T-----	2556
HIVLBV217	T-----AC-----A-CT-G-----C-----G-----C-----G-----A-----C-----	2547
CONSENSUS.O	TTGCAGTGgCAGTtGCCaATTGGACTGAGgGCAtaATCtTAGG?aTACAAAGaATAGGaa?AGGaaTtCtT	2465
HIVANT70	-----A-----GC-----G-----C-----G-----	2544
HIVMVP5180	-----T-----C-----TC-----CA-----T-----C-----	2583
HIVVAU	-----T-----T-----A-----C-----C-----C-----G-G-----G-----	2586
CONSENSUS.CPZ	CAGCTATTGCAGTAGCAGAAGGTACAGATAGAAATAAGAGCTTTTCAAGTTACACTTAGAATTATTAGA	1561
SIVCPZGAB	-----	2517
CONSENSUS.U	tAGCAaTAGtAGTAGCTGggtGgACAGATAG?ATTATAgAAaTaGGACAAAGA?tTtGtAGaGCTaTTCTC	2372
HIVZAM184	-----C-----G-----GC-----C-----A-----C-----G-----T-----	2499
HIVK124A	-----G-----GC-----C-----T-----A-----C-----G-----T-----	2532
HIVZ3	C-----G-A-----A-----G-G-----A-----G-----	2497

HIV1 ENV

CONSENSUS . A	aACaTACcTagAAGAATCaCaCagGGcctTgAAAggGctTTgctATAA	2471
HIVSF1703	-----C-----T-	2603
HIVU455	-----A-----A-----	2568
HIVZ321	-----A-----A-----A-----	2568
HIVRW020	-----C-T-----GC-----	2534
HIVUG031	-----G-----T-----A-----A-----	2438
HIVUG275A	C-----A-----C-----	2544
HIVUG273A	-----G-----C-----A-----A-----	2568
HIVV1191A	-----C-----AGCG-----AA-----A-----	2565
HIVDJ264A	-----GT-----A-----TC-C-----	2562
HIVDJ263A	-----TC-T-----A-----	2586
HIVDJ258A	-----GC-T-----T-----	2574
HIVKENYA	C-----T-----A-----	2568
CONSENSUS . B	cAcaTaCcTagAAGAAaTaAGACagGGcctTggAAaGggtTTgCtaTAA	2497
HIVJRCSE	-----C-----	2544
HIVJRFL	-----C-----	2541
HIVALA1	-----C-----	2547
HIVBRVA	-----A-----	2556
HIVJH3	-----C-----	2601
HIVBAL1	-----	2565
HIVYU2	-----GT-----A-----	2529
HIVRW	-----C-----	2568
HIVHXB2R	-----AT-----	2568
HIVLAI	-----AT-----	2583
HIVNL43	-----AT-----	2562
HIVMFA	-----AT-----	2562
HIVCAM1	-----A-----A-----CT-----	2569
HIVNY5CG	--T-C-----A-----A-----CT-----	2550
HIVADA	-----C-----CT-----	2559
HIVJFL	-----GT-----A-----	2553
HIVSIM184	-----C-----	2568
HIVD31	-----GT-----A-----	2550
HIVSF162	-----T-----	2541
HIVBCSG3C	-----CT-----	2577
HIVQY1	A-T-----	2565
HIVSF33	-----C-----T-----	2556
HIVCDC4	-----C-----	2604
HIVSF2	-----A-----T-----CT-----	2565
HIVSF2B13	-----T-----A-----CT-----	2559
HIVHAN	-----G-G-----C-----	2565
HIVRF	-----G-G-----A-----G-----	2595
HIVNMJ2	-----	2541
HIVTB132	A-----C-----C-----A-G-TCGA	2505
HIVR020	-----A-A-----	2555
HIVTH014	-----C-C-----A-----	2531
HIVUS711	-----C-----	2541
HIVUS712	-----C-----	2552
HIVUS715	--T-----T-----	2558
HIVUS716	-----C-----	2531
HIVHA593	-----C-----T-----	2621
HIVHA594	---G---C-----A-----	2543
HIVHA596	---G---C-----A-----	2564
HIVHA599	-----A-----	2594
HIVHA651	A-----C-----	2546
HIVP096	A-----C-----	2562

CONSENSUS.C	aaCATACCTagAAGAATAagacagggcttgaagcagcttggc?ataa	2501
HIVMA959	-----GC-----T-	2558
HIVMA960	-----T-----	2546
HIVZAM20A	--T-----	2625
HIVSM145A	-----C-----T---	2586
HIVZAM18A	-----	2544
HIVDJ259A	-----T-----A---	2499
HIVDJ373A	-----A-----	2532
HIVSE364A	-----A-----G---A---	2526
HIVUG268A	C-----A-----A---	2571
CONSENSUS.D	7ACaTACCCa7A7GAaTaAGACAGGGcTTgGAAAGGgcTTT?CTataa	2473
HIVJY1	C-----G-C-G-----A---	2589
HIVNDK	A--G-----G-A-----CT--G---	2538
HIVMAL	C-----T-G-A---T-----C-----G---	2577
HIVELI	A-----G-A-----A-----T---A--T	2559
HIVZ226	C-----C-A-----A-----CT--G--T	2559
HIVUG024	A-----C-C-----CT--G---	2553
HIVUG266A	A-----C-A-----C-----A---	2556
HIVUG269A	C-----G-C-----T-----A---	2561
HIVUG274A	A-----C-C-----G-----	2594
HIVSE365A	C-----C-C-----G-----A-----	2556
CONSENSUS.E	CACATACCTAGaAGAATCAGACAGGGCTTAGAAAGGGCTTTGCTA	2509
HIVTH966	-----	2567
HIVTH975	-----	2561
HIVTH022	-----G-----	2561
CONSENSUS.F	AACATACCTACAAGAATAAGACAGGGCTTCGAAAGGGCTTTGCT?TAA	2436
HIVB2163A	-----A---	2520
HIVB2126A	-----G---	2528
CONSENSUS.G	?A?ATACCTAGAAg?ATAAGACA?GGCTT?GAAAG?GCTTTGCTATAA	2370
HIVV1525A	C-C-----A-----G---T---G-----	2604
HIVLBV217	A-T-----G-----A---G---A-----	2595
CONSENSUS.O	aaCATcCCaAGaAGaATTAGACAgGGc?taGAAaGAatctTatTg	2509
HIVANT70	-----G-----T-----GT-----	2589
HIVMVP5180	C-----A---TGC-----G---	2628
HIVVAU	-----A-----G-----C-T---C---C-G---A	2631
CONSENSUS.CPZ	AACATCCCTAGGCGCATaAGACaAGGGCTGGAAAGAGCTTTACTT	1606
SIVCPZGAB	-----	2562
CONSENSUS.U	aaCATACCTagAAGAATaAGACAGGGcTTcGAAAGGGCTTTaCTgTAA	2420
HIVZAM104	-----C-----T-----	2547
HIVK124A	C-C-----C-----G-----G---A---	2580
HIV23	-----T-----	2545

This page intentionally left blank.

LOCUS HIVAMK101 947 bp ss-RNA VRL 09-MAY-1994

DEFINITION Human immunodeficiency virus type 1 (HIV-1) gag sequence, sample AMK10, partial cds.

ACCESSION U08192

SOURCE Human immunodeficiency virus type 1 uncultured PCR-amplified PBMC DNA derived from a 40 year old male student originally from Zaire who had moved to Alabama, U.S.A. and was suffering with AIDS.

REFERENCE 1 (bases 1 to 947)

AUTHORS Gao,F., Yue,L., Hill,S.C., Robertson,D.L., Graves,A.H., Saag,M.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE HIV-1 Sequence Subtype D in the United States

JOURNAL AIDS Res. Hum. Retroviruses 10, 625-627 (1994)

COMMENT The viral strain derived from patient AMK produced the first identified case of a non-subtype B infection in the United States [1]. This sequence was derived from a 40 year old male student who was originally from Zaire and had subsequently moved to Alabama, U.S.A. In 1992 he was identified as HIV positive when he sought medical care for Staphylococcus aureus endocarditis, tuberculosis, and acute renal failure. Gao et al. report that the patient's clinical history suggests that he was infected before his move to the United States. Although a wife and three children who still lived in Zaire were healthy, two brothers who both lived in Zaire had died of AIDS. This gag sequence clusters with HIV-1 subtype D gag sequences. Also see accession number U08193 <HIVAMK82>, another sample from this individual.

NCBI gi: 483777

BASE COUNT 362 a 183 c 213 g 189 t

ORIGIN 83 bp downstream from beginning of gag cds.

1 aatatagact aaaacatata gtatgggcaa gcagagagct agaacgattt gcaattaatc
gag cds ->

61 ctggcctttt agaaacatca gaaggctgta gacaaataat aggacagcta caaccatccc
121 tccagacagg gtcagaagaa cttagatcat tgtttaatac agtagcaacc ctctattgtg
181 tacatgaaag aatagaggta aaggacacca aggaagcctt agacaaagtg gaggaagaac
241 aaaacaaaag taagaaaaag gtacagcaag aaacagctga cacaagaaac agcagccagg
301 tcagccagaa ttatcctata gtgcaaaatc tgcaggggca aatggtacac cagccctat
361 caoctagaac tttgaatgca tgggtaaaag taatagaaga caaggctttc agtccagaag
421 taatacccat gttcacagta ttatcagaag gagccacccc acaagattta accaccgtgt
481 taaacacagt ggggggacat caagcagcca tgcaaatgtt aaaagagacc atcaatgagg
541 aagcggcaga atgggacagg atgcatccag tgcattgcagg acctgttgca ccaggtcagg
601 tgagagaacc aaggggaagt gatatagcag gaactactag caaccttcag gaacaaatag
661 gatggatgac aagcaatcca cctatcccag taggagaaat ctataaaaaga tggataatcc
721 tgggattaaa taaaatagta agaattgata gccctgtcag cattttggac ataagacaag
781 gaccaaagga acctttttaga gactatgtag atcggttcta taaaactcta agagcogagc
841 aagcttcaca ggatgtaaaa aattggatga ctgaaacctt gttggtccaa aatgcgaacc
901 cagattgtaa aactatctta aaagcattgg gaccaggggc ttcat

gag ->

HIVAMK82

LOCUS HIVAMK82 612 bp ss-RNA VRL 09-MAY-1994

DEFINITION Human immunodeficiency virus type 1 (HIV-1) env sequence, sample AMK8, partial cds.

ACCESSION U08193

SOURCE Human immunodeficiency virus type 1 uncultured PCR-amplified PBMC DNA derived from a 40 year old male student originally from Zaire who had moved to Alabama, U.S.A. and was suffering with AIDS.

REFERENCE 1 (bases 1 to 612)

AUTHORS Gao,F., Yue,L., Hill,S.C., Robertson,D.L., Graves,A.H., Saag,M.S., Shaw,G.M., Sharp,P.M. and Nahn,B.H.

TITLE HIV-1 Sequence Subtype D in the United States

JOURNAL AIDS Res. Hum. Retroviruses 10, 625-627 (1994)

COMMENT The viral strain derived from patient AMK produced the first identified case of a non-subtype B infection in the United States [1]. This sequence was derived from a 40 year old male student who was originally from Zaire and had subsequently moved to Alabama, U.S.A. In 1992 he was identified as HIV positive when he sought medical care for Staphylococcus aureus endocarditis, tuberculosis, and acute renal failure. Gao et al. report that the patient's clinical history suggests that he was infected before his move to the United States. Although a wife and three children who still lived in Zaire were healthy, two brothers who both lived in Zaire had died of AIDS. Interestingly, this envelope amino acid sequence contains an extremely rare GQQQ V3 loop crown. The env and gag sequences cluster with HIV-1 subtype D sequences (U08192, <HIVAMK101>).

NCBI gi: 483779

BASE COUNT 252 a 101 c 121 g 138 t

ORIGIN 805 bp downstream from beginning of env cds.

1 gaagagataa taattagatc tgaaaatatc acaataaatg ctaaaaccat aatagtacag
env cds ->

61 cttaatgagt ctgtatcaat taattgtaca aggccctaca gcaaggaaag actaaagaca
V3 loop ->

121 tctataggac aagggcgaagc actctacaca acaagggtaa aggtacagg agatataaga

181 caagcacatt gtaattgttag tgaaataaaa tggataaaa cattacaaca ggtagtgcga
<- V3 loop

241 aaattaaaag gccttcttaa tattagtagt acaatacatt ttaaaccatc ctcaggaggg

301 gacccagaaa ttacaacaca cagttttaat tgtggagggg aattcttcta ctgcaataca

361 tcaaaactgt ttaatatatc aataggtaat agtagtgcaa cgagtagtac aacaaatagt

421 acagggggta ctataacaat cccatgcagg ataaaacaaa ttataaacat gtggcagggg

481 gtaggaaaag caatgtatgc cctcccatc gaaggacaaa ttagatgtac atcaaacatt

541 acaggactat tactgacaag agatgggggt gcaaacagta ctaacgagac cttcagacct

601 ggaggaggaa at
env ->

LOCUS HIVCA1 897 bp ss-RNA VRL 16-SEP-1994

DEFINITION Human immunodeficiency virus type 1 (HIV-1) env sequence, isolate CA1, partial cds.

ACCESSION X80438

SOURCE Human immunodeficiency virus type 1 cocultured PCR-amplified PBMC DNA derived from an asymptomatic patient from Yaounde, Cameroon.

REFERENCE 1 (bases 1 to 897)

AUTHORS Nkengasong, J.N., Janssens, W., Heyndrickx, L., Franssen, K., Ndumbe, P.M., Motte, J., Leonaers, A., Ngolle, M., Ayuk, J., Piot, P. and van der Groen, G.

TITLE Genotypic subtypes of HIV-1 in Cameroon

JOURNAL AIDS 8, 1405-1412 (1994)

COMMENT Nkengasong et al. conducted a survey designed to determine the prevalence of different HIV-1 subtypes in Cameroon. As a part of this study, viruses were isolated from nine symptomatic (CA5, CA7, CA11, CA15, CA16, CA17, CA18, CA20, and CA21) and nine asymptomatic (CA1, CA2, CA4, CA6, CA9, CA10, CA13, CA19 AND CA22) HIV-1 infected individuals. Isolates CA21 and CA22 were derived from patients in Douala, while all other samples were from patients in Yaounde. All of the strains except for isolate CA9 were sequenced over the envelope region (V3, V4, V5 and a segment of gp41). Isolate CA9 was sequenced over a 280 bp region of pol. From phylogenetic analysis, the following subtype designations were assigned [1]:

A	CA7, CA6, CA1, CA17, CA22, CA21, CA11, CA2, CA15, CA19, CA18
B	CA5
E	CA10
F	CA4, CA20, CA16
H	CA13
O	CA9 (based on analysis of pol)

Sequence analysis also revealed significant diversity at the tip of the V3 loop. An alignment of envelope sequences in this set is shown below (accession numbers X80438-X80454).

BASE COUNT 334 a 154 c 209 g 200 t

ORIGIN 796 bp downstream from the beginning of env cds

1 gaaaaggttag taattcgatc tgaaaatc acaacaatg caaaaagcat actagtacag
env cds ->

61 ttggctaagc ctgtaaaaat taattgtacc agacctaaac acaatacaag aacaagtgtg
V3 loop ->

121 cgtatagggc cagggcaagc attctatgca acaggtgaca taatagggga tacaagacaa
181 gcacattgta atgtcagtag aacagaatgg aatacaactt tacaacaggt agttacacaa
<- V3 loop

241 ttaagggaagt actttgggaa tgacacaaac ataactcttg ctaaaccctc aggaggggat
301 ttagaaatta caacacatag ttttaattgt ggaggagaat ttttctattg caatacatca
361 aaactattta atagcacttg gaatagcaat accactgtca gcccaaatga cacagagtca
421 aatgacacta taactctcca atgcagaata aagcaaattg taatatgtg gcagagagta
481 ggacaagcaa tgtatgcccc tcccatccaa ggagtaataa ggtgtgaatc aaacattaca
541 ggactaatat taacaagaga tggaggggct aataatagta caaatgagac attcaggcct
601 ggaggaggag atatgaggga caattggaga agtgagttat ataagtataa agtagtagaa
661 attgaaccgc taggtgtagc acccaccctt gcaaaaagaa gagtgggtgga gagagaaaaa
721 agagcagttg gactggagc tgtcttctt gggttcttag gagcagcagg aagcactatg
781 ggcgcggtgt caataacgt gacggtagag gccagacaat tattgtccgg tatagtgcga
841 cagcaggata atttgctgaa ggctatagag gctcaacaac atctgttgaa actcacg
env cds ->

HIVCA1

CONSENSUS	GAAGAGATAATGATTAGATCTGAAAATATCACAAACAATGCCAAAACCATAATAGTACAGTTTAATAAGC	70
CA1	--GA-----CT---G---A-T---A-----C-GTT	70
CA10	-----C-A-A-----C---C-----G-CC-C---AT	70
CA11	-G-A-G-G-----T-----GGC-G--	70
CA13	---A-G---C-----G---A---A-----C---A-T-	70
CA15	-----G-G-----G-G-----AGT-G--	70
CA16	---T---A-----G---A-A---A-----C---GAT	70
CA17	-G-A-----G-----GG---C--	70
CA18	A---GA-G-----GGT---	70
CA19	-----G-G-----T-----A-AGT-T-	70
CA2	---G---G-----T-----G---GGC-GC--	70
CA20	---T---A-----T-G-T---A-A-----GAT	70
CA21	---A-G-G-A---C-----A-G-C-----GGC---	70
CA22	---A-G-GCA-----C-----C-T---	70
CA4	---A-T---CA-----G---A---ATA---TT-----C---GAT	70
CA5	---G---G-AC-----T---GTG---T-----C-GG---AT	70
CA6	-G-GA---GA-----CT---G---T---C---AC---C-GTA	70
CA7	-G-GAG---G-G-----CT---G---T---C---C-GTA	70
CONSENSUS	CTGTAAAAATTAATTGTACCAGACCTAACAACAATACAAGAA?AAGTATACGTATAGGACCAGGA.....	134
CA1	-----G-----A-G---A-----	135
CA10	---GG---C-----CTC---CA-C-T---C---TC-A-G-----	135
CA11	---C-----C-G---A-----	135
CA13	---GT---C-----C---T-C-G---G---G---TAT---CGAGG	140
CA15	-----G-----	135
CA16	-A-G-----A---C-G---A-C-T---G-G---C-T-----	135
CA17	---T---C---T---C---A-G-----G....	135
CA18	-----GG---G-G-----	135
CA19	---G-----T-G---A-G-----G....	135
CA2	---G-----G---A-G-----	135
CA20	---G-----A---C---A-C---A-----	135
CA21	-----C---G-G---G---G....	135
CA22	---T---GG---G-G---G....	135
CA4	---G-----A---C---G-A-----	135
CA5	---G---C---A---C---A-G---A-----G....	135
CA6	---G-----T-----G-G---C-A-T-C-----	135
CA7	---G-----G-G---C---T-----	135
CONSENSUS	.CAAGCATTCCTATGCAACAGGTGACATAATAGGGGATATAAGACAAGCACATTGTAATGTCAGTAGAACA	203
CA1	.AGT--TA-----T-----C-G--	204
CA10	.-----T---AA--T---AA-G---TC---AA-----GC---T---G-GA-T-A-G---	204
CA11	.---A-----A-----G---	204
CA13	A---A---C---T---CT-----A-----A-T---G-G-	210
CA15	.---A---C-----A---	204
CA16	.-G-T-----TA...-AT-A-----A-G---T---GTCA-T-A---	201
CA17	.---A-----G-----A---CA-	204
CA18	.-----A-----C---A---C-----	204
CA19	.---A---T-----A-----G-----GA---	204
CA2	.---A-----G-----G---	204
CA20	.AG-----A-----A-G---T-C---GTCA-T-A---	204
CA21	.-----C-----	204
CA22	.-----A-----A---C---	204
CA4	.---T-----T---A-----A-G---T---GCA-T-A-T---	204
CA5	.AGG-----T---A-----ACATA...-A-----C-CCC-T-A-G---	201
CA6	.---A-TC-----A-T-----C-----A-	204
CA7	.---A-T-C-----G-T-----A---C-G-----C-----	204

CONSENSUS	GAATGGGAATGA?ACTTTACAACAGGTAGCTACACAATTAAGAAAGTACTTT...AATAA?ACAACAATA.	267
CA1	--T-----A-C-----Aa-----AG-----G-AT-AC.....C-A-----.	267
CA10	A-----A-----TA-----A-A-----G-----T-AG-----.	270
CA11	-----GAG-----T-----G-----GGG-----C-A-----.	273
CA13	A-G-----C-TG--ACCA-----G-----AC-C-G...-C-G-GC-A-----.	276
CA15	-G-----G-AA-A-----G-----TC-AAG--G-T-AC---C-A	274
CA16	CTG-----C--C--GG-A-----T-GA-A--G-CGA--GACT-C-----T-T-T--CC-A	265
CA17	A-----CAA-A-AA-----G-----A-----GG-C-----C-----.	267
CA18	-----GA-----AG-----T-----T-A-----AAG--GTT-----.	273
CA19	-----GACA-----G-----C-----.	267
CA2	-----A-C-----A-----G-----G-----GGG...-T-A-----.	270
CA20	CTG-----T-----A-CA-----T-GAGGCG--CCA-CGCA-AAGC.....-CCTT--TG--.	267
CA21	-----ACA-----T-----G-C-----GGG--G-C-----.	273
CA22	--T-----A-C-----G-OC-----G-----A-----TGG-----GAG-----A-----T--C.	273
CA4	CTG-----A-----AC-----T-GA-G-G--C-A--ACTTGACCAC---TA--C-AT--.	273
CA5	C-----C-----A-AT-A--TGGG-A-----G--G-AC-A--GGG-----T-A-----.	270
CA6	-----G-AA-C-TG-----A-----C-----A-G-AA-G-----C-A--A-----.	270
CA7	-----G-A-C-TG-----C-----G--GGAA-G-----C-CA--G-----.	270
CONSENSUS	..ATCTTTACTAAGTCTCAGGAGGGGATTAGAAAATTACAACACATAGTTTAAATTGTGGAGGAGAATT	335
CA1	..--ACA-----G-----C-----.	335
CA10	..-AA--CAACCA--C-----A--CC-----TG-----A--G-----.	338
CA11	..-----G-----T-----.	341
CA13	..CAA--AACCA-----A-CA-----T-----C-TG-----C--C-----.	344
CA15	TACA--G-----.	344
CA16	TA-----CAACCA--A-----CC--G-----C-----CA-----G--.	335
CA17	..-----G--G-----.	335
CA18	..G-----G-----A--G-----.	341
CA19	..-----G-----.	335
CA2	..-----A-----.	338
CA20	..-C-----A--GA--AG-----CC-----A-----.	335
CA21	..-----G--AC-----.	341
CA22	..-----CAT-----A-----.	338
CA4	..-C-----GCCCC-----CCC-----G-----A--A-----.	341
CA5	..G-A--C-A-C-A-----CC-----GTG--C-----G-----.	338
CA6	..-C-----AG-G-----C-----A--G-----.	338
CA7	..-----AA-G-C-----C--C-----A-----.	338
CONSENSUS	TTTCTATTGCAATACATCAAACTGTTTAATAGCACTTGG.....A...??AATA??	380
CA1	-----G-T-----C-AG--A-----GAC-AGAAA--GAC	390
CA10	-----A-----T-----GA-TA	384
CA11	-----TC--A-----AT	384
CA13	-----T-----A-----T-C-----ACG-CTAAC-CC-AT	399
CA15	-----G--A-----T-----ACA-ATAAC---CC	399
CA16	---T-C-----C--AGT--G-----AC-CA-GT	371
CA17	---T-----G-----AATACATCAACACTGTCTAAT-GC-CT	402
CA18	-----A--G--A-----ACGTA-C-GCC	393
CA19	-----GG-T-A-----GCT-TATTT--GGC	390
CA2	-----T-C-----A--G-----T--AC	384
CA20	-----C-----C--GATG-G-----A--GC-T-----.	375
CA21	-----A-----ATAGC---CC	393
CA22	-----CC-----TATAAT-GCACTTGGGGT	396
CA4	-----C-A--C--A--G-T-----A--TGGA--.	381
CA5	-----C--T--T--A--C--T--G--T-----TATTTTAGTAATAGT-CTTGGT--TTT	405
CA6	-----G-----TT-----AT	381
CA7	-----GGG-----GATGTGGTA-ATAAC-C-GCC	399

HIYCA1

CONSENSUS	AATAA?AA?GAAA?TAACACAGAGGCAAAT.....GACACTATAACTCTCCCATGCAGAATAA	435
CA1	-C-TGG--TP-CCCAGGGT--A-T-AC--	448
CA10	----A-CT--CC-TGGAG--TT-A.....-C-TG--T-----AG--	442
CA11	--C--T-CCA-T-GC-CTTGG--T--.....-A-----C--A-----	442
CA13	GGC-GCT-T-TG-GC--T-ACACA-A-----GT-----T-A--G-A-----	457
CA15	---GTC-GACC--A-G-----G-T-----AAT.....-C-----A-----	460
CA16	T--CTG-TGTTG-A-----T-CCA.....CTA--ATC-CTCT--TGG--T-A-----	424
CA17	TGGG-G-GCA-T-GC-GTGTGAGCT--T--.....-A-A-----A--A-----	460
CA18	-GCTCA--CT-C-CAG-GT--A-CT-----G-----	451
CA19	----T--CACC-TC-G-T--A-T-----A-----A-----C--	448
CA2	---C-CT-GC-TAGTGG--G-----GACACTATAAAT-----T-----	454
CA20	TTG--AACACAAAAAAA--A-A--C--T--T-----	421
CA21	-C-GTC-GCCC--A-G-----T-----A-----	451
CA22	G-C--T-GCTC--CATC--G-TT--AAT.....-A-----A-----	457
CA4	-A-ACTAAC--.....A-----C-T-----T-A-----	421
CA5	---GT-CT-TG-A-GGT--TC-A-G--.....-A-C-A-----	463
CA6	GGC-CT-TG--A-----TT-CA-A-----A--C--TA--A-----A--	439
CA7	-GC-TCC-TA--C--TA-ATA-TA-T-CCAGCAGCAATAAT--A-----A--A-----A--	469
CONSENSUS	AGCAAATGTGAAATATGTGGCAGAGAGTAGGACAAGCAATGTATGCCCTCCCATCCAAGGAGAAATAAG	505
CA1	-----C-----A-----GA	518
CA10	-----C-----G-----T-----AGT--G-T--T-A	512
CA11	-----C-----G-----A-----T-----	512
CA13	-A-----GC-----C-----A-T--A-G--T-C	527
CA15	-----A-----A-----T--A--T-----	530
CA16	GA--T--T--C-----GAG-----A-----TG--GAC--TCA	494
CA17	-----A-----G-----T-----A	530
CA18	-A-----G--T-A--A-----	521
CA19	-----GG--AG-----T-----	518
CA2	-----GA-----T-C--C-----	524
CA20	-A--T-----C-----A-G-----T--AGCC--C--TCA	491
CA21	-----T-----	521
CA22	-----A-----A-----G-----C-----	527
CA4	GA--T-----C-----A--G--G--C-----TGCC--C--TCA	491
CA5	-A-----A--C-----GAG-----A--C-----A--CA--T--	533
CA6	---G-----GGC-----T-----	509
CA7	---G--A--C-----	539
CONSENSUS	GTGTGAATCAAACATTACAGGACTACTATTAACAAGAGATGGTGGG.....AATAATAATACAAATAAT	569
CA1	---A-----T-----A-CACA-----	579
CA10	T---T-----T-----T.....GG--G--TG-C-CA	576
CA11	-----G-----A-----G--CGGC-----	573
CA13	A---T-----T-----A--G-----T-CCA--TT.....CGGTC-G--G-GG-G	591
CA15	-----A-----A-----GTG-G--	594
CA16	A--A-C--GT--C--T--A--G-----T.....T--CC-CC-A--G-A	558
CA17	-----A-----T-----GGTCCA--	594
CA18	-----T-----A-----G-----	582
CA19	-----T-G--C-----A--AGATGGAGGG--A--G--G-GCA	588
CA2	-----C-----G-----A-----G--C-----GT-CA-C-	588
CA20	---CA-C--GT--C--T-----G-----AAC.....A--GGGT-C.....	549
CA21	-----A-----A-----GC-----GT-CA--	585
CA22	---T-----G-----A-----G-----G--ATGCA--	591
CA4	---A-T--G--C--T-----G-----A-----G-G--CA-ATGGGTCC	555
CA5	A--C-T--T-----G--G-----T.....G--GG--C-GC-CC-C-	597
CA6	---C-----G-----T-----C--G-GG-G	573
CA7	---C--T-----G-----G-----C-----	600

HIVCA1

CONSENSUS	GAGACCTTCAGACCTGGAGGAGGAGATATGAGGGACAAATTGGAGAGAAGTGAATTATATAAGTATA	633
CA1	-A-----AC-----C-----	643
CA10	-A-----A-----A-----A-----	640
CA11	-A-----C-----	637
CA13	-T-----A-----	655
CA15	AATAAT---A---G-----	664
CA16	ACC...-C--C-----G--A-----AA-----C-G-----A---	625
CA17	-T---G-----	658
CA18	-A---G-----	646
CA19	AAT...-A---G-----	655
CA2	-A---G-----	652
CA20	-C-----A-G-----A-----C-----A---	613
CA21	-A---G-----G-----	649
CA22	-G-----A-----	655
CA4	-C--G-----A-----C-----A---	619
CA5	-T-----A-----A-----G--A---	661
CA6	AGC...-C-----	640
CA7	-AT-----	664
CONSENSUS	AAGTAGTAAAAATTGAACCACTAGGAGTAGCACCCACCAAGGCCAAAAAGAAGACTGGTGGAGAGA...GA	700
CA1	-----GG-----AG-----C-----	710
CA10	-----A-C-----A-----GG-----	707
CA11	-G-----C-----T-G-----CGT-----A...-G	704
CA13	-----G-----T-----G-----G-----	722
CA15	-G-----C-----T-G-----CGT-----G-----	731
CA16	-G-----C-----C-----C-----A-----	692
CA17	-----T-----CGT-----A-----	725
CA18	-G-----C-----T-----T-G-----C-----	713
CA19	-G-----C-----T-G-----CGT-----G-----	722
CA2	-G-----C-----T-G-----C-T-----	719
CA20	-G-----C-----G-----CA-----A-----	680
CA21	-----G-----G-----T-----CCT-----	716
CA22	-----G-T-----CGT-----	722
CA4	-G-----G--A-----T-----CA-----C-----G	686
CA5	-----G-----T-----G-----C-----GAGAG	731
CA6	-----G--GG-G--A-----TG-----A-----	707
CA7	-----G--GG-G--A-----G-----G-----	731
CONSENSUS	AAAAAGAGCAGTG...GGACTAGGAGCTGTCTTC?TTGGGTTCTTAGGAGCAGCAGGAAGCACTATGGGC	766
CA1	-----A-A-----GC-T-----G-----	777
CA10	-----A-----C-GA-T-----	774
CA11	-----AGTT-----G-----C-----	774
CA13	-----T...-----C-----	789
CA15	-----T...-----G-----C-----	798
CA16	-----G-----G--A-----G-----	759
CA17	-----T...-----G-----C-----	792
CA18	-----T...-----A-----C-----	780
CA19	-----C...-----A-G-----TC-T-----	789
CA2	-----ACTT-----A-----T-----	789
CA20	-G-----A-----G--C-----G-----	747
CA21	-----T...-----G-----C-----	783
CA22	-----T...-----G-----T-----	789
CA4	-----A-----AC-T-----	753
CA5	-G-----AG--CAGTG--A-G-----A-G--C-----G-C-----	801
CA6	-----GC-T-----G-----	774
CA7	-----GC-T-----G-----	798

HIVCA1

CONSENSUS	GCGGCGTCAATAACGCTGACGGTACAGGCCAGACAATTATTGTCCTGGTATAGTGCAACAGCAAAGCAATT	836
CA1	-----A-----A-----C-----G-----	847
CA10	--A-----T-----A-----	844
CA11	-----C-----G-----C	844
CA13	-----T-----C-----T---	859
CA15	-----C-----G-----	868
CA16	-----A-----A-----A-----	829
CA17	-----C-----G-A-----	862
CA18	-----C-----G-----C	850
CA19	--A-----G-----	859
CA2	-----	859
CA20	-----	817
CA21	-----C-----G-----GGAT---	853
CA22	-----C-----AT---	859
CA4	-----G-AT---	823
CA5	--A-----C-G-----A-----	871
CA6	--A-----C-----	844
CA7	--A-----T-----C-----A-----	868
CONSENSUS	TGCTGAAGGCTATAGAGGCTCAACAGCATCTGTTGAAACTCAG	880
CA1	-----C-----G-----A-----T	891
CA10	-----G-----G-G-----C-C-----A	888
CA11	-----G-----A-----	888
CA13	-----G-----A	903
CA15	-----GA-C-----A-----	912
CA16	-----C-----A-G-----C-G-----A	873
CA17	-----G-----A-A-----	906
CA18	-----G-----A-----	894
CA19	-----GA-----A-----	903
CA2	-----A-----	903
CA20	-----A-G-----C-----A	861
CA21	-----A-----	897
CA22	-----A-----A-GA-----G-----	903
CA4	-----A-G-----C-G-----A	867
CA5	-----G-----T-----G-----C-----A	915
CA6	-----C-T-A-G-----A-----T	888
CA7	-----C-----A-----A-----T	912

LOCUS HIVCAR286A 1549 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 1286CL1, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), isolate 1286CL1.
 REFERENCE 1 (bases 1 to 1546)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env A subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 579 a 258 c 336 g 376 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggggacaca gaggaattgg caaaacttgt ggagatgggg cctcatgac
 env start ->
 61 ttggaatgt tgatgatttg taatgcaaca ascaacttgt gggtaacagt ctattatggg
 121 gtacctgtgt ggaagatcg agagaccact ctattttgtg catcagatgc taagcatat
 181 agtgcgtgaa acataatgtc tggggctaca catgcctgtg taccacaga cccaatcca
 241 caagagatac ctctggaaaa tgtaacagaa aattttaaca tgaggaaaaa tagtatggta
 301 gaacagatcg atgaggatat aatcagttta tgggatgaaa gcctaaagcc atgtatagag
 361 ttaacccctc tctgtgttac tttaactgt actgatgcaa agatagactg tactaatgga
 421 aggaataaca gtgctggagg acagacctgt actggagtac agaatagcac tgctgaggtt
 481 ggccaagtag aaatgaaaaa ctgctctttc aatataacta cagaaataag aaataggaag
 541 aaaaaagaat acggcctttt ctatagaact gatgtagtac caacaaataa tgatgataat
 601 agtacagata attatacaca taattatagg ttaataaatt gtaatgcctc aaccattaaa
 661 caggcttgtc caaagggtgac ctttgagoca attcccatc attattgtgc tccagctggt
 721 ttgcgattct aaagtgtagg gataaggagt tcaatggaac agggccatgc aagaatgtag
 781 cacagtacaa tgcacacatg gaataagcca gtggtatcaa ctcagctatt gttaatggc
 841 agtctagcag aaggaggggt aataattaga tctgaaaaca tcacaaaca tgctaaaacc
 901 ataatagtac agcttgacag tctgttagga attaattgta ccagacctg caacaataca
 v3 loop ->
 961 agaaaaagta taogtattat gcatatagga ccaggacgat ctttctatgc aaaagggtga
 1021 ataggggata taagacaagc acattgtacc attaataaag caaatggaa tcaacttta
 <- v3 loop
 1081 caaaaggtag ctgaaaaatt aagaggaatc tttaacaaca aaacaataat ctttactaac
 1141 cactcaggag gggattttaga agttacaaca catagcttta attgtggagg agaagtlttt
 1201 ctattgcaat acatcagcct gttaataagc agttggactg acactaacag cactaacagc
 1261 agtaggactg ggaatgacac tgacatcata actctcccat gcagaataaa gcaaattgta
 1321 aatatatggc agagagtagg acaagcaatg tatgccccac ccctccaaag agtaataagg
 1381 tgtgaatcaa acattacagg actactatta acaagagatg gtgggctgaa tagcacaca
 1441 aatgagacct tcagacctac aggaggagat atgaggaaca attggagaag tgaattatat
 1501 aagtataaag tagtggaaat taatccacta ggagtgcac ccacccggg
 env cds ->

HIVCAR4017

LOCUS HIVCAR4017 1507 bp ss-DNA VRL 01-OCT-1994

DEFINITION Human immunodeficiency virus type 1, CAR isolate 4017-1, env glycoprotein gp120, partial cds.

ACCESSION

SOURCE Human immunodeficiency virus type 1 (HIV-1), isolate 4017-1.

REFERENCE 1 (bases 1 to 1507)

AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P..

TITLE ;

JOURNAL Unpublished

COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env E subtype sequences. It is one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.

BASE COUNT 565 a 240 c 322 g 380 t

ORIGIN Beginning of env cds

1 ATGagagtga aggagaaata tcagcacttg tggagatggg ggtggaaatg gggcagcatg

env start ->

61 ctcccttggga tattgatgat ctgtagtgtt tcagacaact tgtgggttac agtttattat

121 ggggttcctg cgtggaaaga tcagataacc accctatatt gtgcatcaga tgccaaagca

181 cataagacag aagtgcacaa tgtctgggct acacatgctt gtgtgcccac agaccccaac

241 ccacaagaaa tatacatgga aaatgtaaca gaaaatttta acatgtggaa aaataacatg

301 gcagagcaga tgcaggagga tataatcagt ttatgggatg aaaggctcaa gccatgtgta

361 aagttaactc ctctctgctt tacttttaaa tgtaccaagg ctaattttac caaggctaata

421 agtacogagg ctgcttttga tggcagttac agtaatgaaa caggagaaat ggaaaatgaa

481 gtaagaaact gttcttttaa tgtgaccaca gaactaagag atagaaagca gaaggctctat

541 gcactttttt ataagactga tatagtatca attgataaca gtagtagtga ttatatgtta

601 ataaattgta ataattcagt cattaagcag gcttggtcaa aggtatcctt tgatccaatt

661 cctatacatt attgtgctcc agctgggtat gcgattttaa agtgtaataa taagaatttc

721 aatggaacag ggcctatgca aaatgtcagc tcagtacaat gcacacatgg aattaagcca

781 gtggtatcaa ctcaactgct gttaaatggc agtctagcag aagaagagat aataattaga

841 tctgaaaatc tcacaaataa tgccaaaacc ataatagtgc accttaataa agctgtagaa

901 atcaattgta ccagaccctc caagaaaata agaacaagtg taaggatagg accaggacaa

v3 loop ->

961 gtattctata aaacaggaga cctaatagga gatataagaa aagcattttg tgaaattaat

<- v3 loop

1021 ggaacaaaat ggaatgaac tttaaaagag gtaactagaa aattaagaaa gcactttaat

1081 aagacaatac actttcgacc atcctcagga ggggatccag aaactacaaa tcataattttt

1141 aattgtagag ggaatttttt ctattgcaat acaacacaa tgtttaatag tacttgga

1201 ataaatgaaa ccatggaaaa ggtaaatggc acaaaactca acatgatcac acttccatgc

1261 aggataaggc aaattgtaaa catgtggcag agagtaggac aagcaatgta cgctcctccc

1321 atcagtggag tagttaattg tgtatcaaat attacaggaa tactattaat aagagatggg

1381 ggtaatgaga cacacgagaa cgagaccttc agacctggag gaggaatat aaaggacaat

1441 tggagaagtg aattatataa atataaagta gtccaaattg aaccattagg agtagcacc

1501 acccggg

env cds ->

LOCUS HIVCAR4020 1600 bp ss-DNA VRL 01-OCT-1994

DEFINITION Human immunodeficiency virus type 1, CAR isolate 4020-19, env glycoprotein gp120, partial cds.

ACCESSION

SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4020-19.

REFERENCE 1 (bases 1 to 1600)

AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P..

TITLE ;

JOURNAL Unpublished

COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic virus, clusters with env D subtype sequences. It is one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.

BASE COUNT 629 a 266 c 327 g 378 t

ORIGIN Beginning of env cds

1 ATGagagtga ggggataga gaggaattat caacacttat ggagatgggg cagcatgctc

env start ->

61 ctctgggatat tgatgacctg tagtgttgc gaaaaattgt gggtcacagt ttattatggg

121 gtacctgtgt ggaaagaagc agaaccact ctattttgtg catcagatgc taatcatat

181 aaagcagagg cacataatat atgggcgaca catgctgtg taccacgga ccccaacca

241 caagaaatgg agctgaaaaa tgtctcagaa aactttaaca tgtggaaaaa taacatgggtg

301 gaacagatgc atgaggatat aatcagttta tgggatcaaa gcctaaaacc atgtgtaaaa

361 ttaaacacctc totgtgtcac tttgcactgc actogattgg agaattattg aaacagtagc

421 cttacaaaca atgccagtgc aagaagtaat gatacagaca ataaccatac aaacagtacc

481 aatactagat gggatatga cggaatggaa aaagaatga caaactgctc tttaaatgta

541 accacagtta taagagataa gaagaagcaa gtacatgcac tcttttataa actagatata

601 gtaccaatag aggatgagga taatattaaa ggtaatacta gtgataatac taaaaatatt

661 actagatata gattgataaa ttgtaatacc tcagccatta cacagtcttg tccaaaggta

721 aactttgagc caattcccat acattattgt gccccagctg gctttgcaat tctaaaatgt

781 aaagataaga acttcaatgg aacaggccta tgcataatg tcagcacagt acagtgtaca

841 catggaatta agccagtggg gtcaactcaa ctgctgtgta atggcagctc agcagaagaa

901 gagataatag ttagatctga aaacctcaca gacaatgcta aaaccataat agtacagctt

961 aaggaatctg taaacattag ttgtacaagg cccaacaact ataaaagaca aggtacacct

v3 loop ->

1021 ataggactag ggcaagcact ctatacaaca agagtaatag gggatataag aaaagcacat

1081 tgtaatatga gtagaaaaga atggaataat actttacaaa aggtagctaa gaaattaaga

<- v3 loop

1141 gaacttttta acaagtcaaa aataaatttt caaccccat caggaggaga cctagaagtt

1201 acacacacaca gttttaattg tcaaggggaa tttttctact gcaatacatc acaactgttt

1261 aatcataaat ggaatgatac aacaggagaa atgtcagaac agaacaataa tacagagtca

1321 gaaatataca cactcccatg cagaataaga caattataa gaatgtggca gagagtagga

1381 caagcaatgt atgcccctcc catcgaagga gtaatcaatt gtacatcaa tattacaggg

1441 ctactattga caagagatgg tggtaatata aagaacggaa gtaagaatga tattgagatc

1501 ttcagaccgg gaggaggaga tatgagggac aattggagaa gtgaattata taaatataaa

1561 gtagtaagaa ttaaacact aggaatagca cccaccggg

env cds ->

HIVCAR4023

LOCUS HIVCAR4023 1552 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4023-2, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4023-2.
 REFERENCE 1 (bases 1 to 1552)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P..
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic virus, clusters with env A subtype sequences. It represents one of 14 CAR isolates that have been integrated into a 'hybrid' gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 590 a 255 c 324 g 382 t 1 others
 ORIGIN Beginning of env cds
 1 ATGagagtga aggagaaata tcagcacttg tggagatggg ggtggaaatg gggcagcatg
 env start ->
 61 ctcccttggga tatttgatgat ctgtaatgca gacaaagaca aattgtgggt tacagtctat
 121 tatgggggtac ctgtgtggag agatgcaaat accactctat tttgtgcac tgatgctaaa
 181 gcatatagta ctgaaaaaca taatgtcttg gctacacata cctgtgtacc cacagacccc
 241 gacccacgag agatacctct gaaaaatgta acagaaattt ttaacatgtg gaaaaatagt
 301 atggtagaac agatgcagga ggataattac agcttatggg atgaaagcct aaaaccatgt
 361 gtaaaagctaa cccctctctg tgttacttta aactgtactg atgtaaaaaa tatcactgtg
 421 aatagtacta gtactgggaa tagtctaat scaataaata gaactgtgga tagccaagta
 481 ggaataaaaa actgctcttt caatatgacc tcagaaataa gagataggaa gaagaaagaa
 541 tacgcacttt tctatagact tgatatagta ccaattaata ataataataa taatacaaat
 601 aatccaaatg attataatga ttatatgkta ataaattgta atgtatcagc agttaaacag
 661 gcttgtccaa agataacctt tgagccaatt cccatacatt attgtgctcc agctgggttt
 721 gcgattctaa agtgtaatga tacgaagttc agtggaaacag ggccatgcaa gaatgtcagc
 781 acagtacaat gcacgcatgg aatcaagcca gtggtatcaa ctcaactact gttaaatggc
 841 agtctagcag aagaaggggt aatgattaga tctgaaaaca taacaaacaa tgctaaaacc
 901 ataatagtag agcttgacag tcctgtaata attaattgta ccagacctaa caacaatata
 v3 loop ->
 961 agaagaagca tgcgtatagg accaggacaa actttctatg caacagggtga tataacaggg
 1021 gatataagac aagcatattg caacgtcagt aaagaagaat ggaacaccac tctgcaaaag
 <- v3 loop
 1081 gtagctaaaa aattaggaac tatcttttagc aacaaaacaa taatctttaa tagcagctca
 1141 ggaggggatt tagaagttac aacacacagc ttttaattgta gaggagaatt cttctactgc
 1201 gatacatcaa gactgtttta tagcatttgg aaaaatgaca ctggcaacat aactggcaac
 1261 ataactggct gggagaattt caatgacagc ataatectac catgcagaat aaaacagatt
 1321 gtaaatatgt ggcagagagt aggacaagca atgtatgcoct ctccatccc aggagaaata
 1381 agttgttaact caaacattac aggaatncta ttaacaagag atgggtgggga tgataactca
 1441 acaaatgaga tcttcagacc tacaggagga gatatgagga acaactggag aagtgaatta
 1501 tataagtata aagtagtgga aattgaacca ttaggagtag ccccccccg gg
 env cds ->

LOCUS HIVCAR4039 1483 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4039-TG2, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4039-TG2.
 REFERENCE 1 (bases 1 to 1483)
 AUTHORS Schmitt,D., Nathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env E subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 548 a 240 c 328 g 367 t
 ORIGIN Beginning of env cds
 1 ATGagagtga gggggacaca gaagaattgg ccaaacttat ggaaatgggg aactttgatc
 env start ->
 61 ctgggggttg tgatgatttg tagtgcctca gacaacttgt gggttacagt ttattatggg
 121 gttcctgtgt ggagagatgc aaagaccaac ctattttgtg catcagatgc caaagcacat
 181 aagacagaag tgcacaatgt ctgggctaca catgcctgtg taccacaga ccccaaccca
 241 caagaaatat acatggaaaa tgtaacagaa aattttaaca tgtggaaaaa taacatggta
 301 gagcagatgc aggaggatgt aatcagttta tgggatcaaa gtctaaagcc atgcgtaaag
 361 ttaactcttc tctgcgttat ttacattgt accagggtca attttacca ttctaattat
 421 accaagggtg atgaaactgt tgaaatggaa aatgaagtaa gtaactgttc ttttaataa
 481 accacaaaac taagagataa gacgaaggag gcctatgcac ctttttataa gcatgatcta
 541 gtaccaataa aggaaaatag tagtagtgag tatgtgctaa tacttgtaac acattcagtc
 601 attaaacagg ctgtgtccaa ggtaaccttt gagccaattc ccatacataa ttgtgoccca
 661 gctggttttg cgattctaaa gtgtaaatg aaggagttca gtggaaaaag gacatgcaag
 721 aatgttagta cagtacaatg cacacatgga atcaagccag tagtgtcaac tcaactgtg
 781 ttaaatggca gtctggcaga aaaagaggta gtgattagat ctgaaaatat cacaacaac
 841 gccaaaacca taatagtaca gctggttacg cctgtacaaa ttaattgtac cagaccctcc
 v3 loop ->
 901 aacaatataa gaacaagtgt aaggatagga ccaggacaag tattctataa aacaggaagc
 961 ataacaggag atataagaaa agcattctgt gagattaagg aagcacaatg gaataaaact
 <- v3 loop
 1021 ttaaaaggag taataggaaa attagagag cactttaata agacaataat cttccaacga
 1081 tctcaggag gagatccaga aattacaatg cttaccttta attgtggagg ggaatttttc
 1141 tattgcaata caacaagact gttaaatagt acttggggaa caaatgaaac catagagtct
 1201 aatgtcacia acccaagaaa tatcacactt ccattgcaaa taaggcaaaa tgtaaacatg
 1261 tggcagagag taggaaaagc aatgtatgct cctcccatca gtggagtaac taattgtgta
 1321 tcaaatatta caggactact attgacaaga gatggaggga gcaatagtgg tgaacctgag
 1381 acattcaggc ctggaggagg agatatgagg gacaattgga gaagtgaatt atataagtat
 1441 aaggtagtac aaattgaacc aataggagta gcacccccc ggg
 env cds ->

HIVCAR4054

LOCUS HIVCAR4054 1486 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4054-TG2, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4054-TG2.
 REFERENCE 1 (bases 1 to 1486)
 AUTHORS Kieny,M.P., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Schmitt, D.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env A subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 573 a 249 c 299 g 364 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggagacaca gaggaattgg cacaacttgt ggaatggcg cctcatgac
 env start ->
 61 ttggaatgt taatgatctg taatgctaaa gaaaacttgt gggcacagt ctattatggg
 121 gtacctgtgt ggagagaggg aaataccact ctattttgtg catctgatgc caaagcacat
 181 agtactgaaa agcataatgt ctgggctaca cagcctgtg taccacaga cccaaccca
 241 caagaaataa aactagaaaa tgaacagaa aattttaata tgtggaaaa taacatggta
 301 gaacagatgc atgaggatat aatcagttta tgggatgaaa gtctaaaacc atgtgtaaaa
 361 ctaacccctc tctgtgttac tttaaattgt actgatgtaa agaagaatag cactctggaa
 421 ggccaagtag aaatgaaaaa ctgctcttcc aatataacca cagaaataaa tgataggaag
 481 aagaagaac acgcgctttt ctatagactt gatgtagtac aaattaatag taatagtaat
 541 attagtaatg caaattataa tgattatagg ctaataaatt gtaatgtctc aactgctaaa
 601 caggcttgcc caaagggtat ctttgagcca attcccatc attattgtgc tccagctggg
 661 ttgcgattc taaagtgtaa tgatacaaaa ttcaatggaa cagggccatg caagaatgtc
 721 agcacagtac aatgcacata tggaaatcaag ccagtagtat caactcaact actgttaaat
 781 ggcagtctag cagaagaaga ggtaatgatt agatctgaaa acttcacaga caatgctaaa
 841 accataatag tacagcttaa aaatcctgta agaattaatt gtaccagacc tggcaacaat
 v3 loop ->
 901 acaagaaagg gtatacatat aggaccagga ccaggacaag ctattttatgc aacagggtgat
 961 ataatagggg atataagaca agcacattgt aatgtcagta aagcagaatg ggataacact
 <- v3 loop
 1021 ttacaacagg tagctaaaaa attatatgag gtctttaaca aaacagtaan ctttactaac
 1081 cactcaggag gggatataga aattacaaca catagtttta attgtagagg ggaatttttc
 1141 tattgcaata caacagatct gtttaatagc acctggata acattaatgg cactgggaat
 1201 aacactcagg aaacaaatgg cactataacc ctccaatgca gaataaagca aattataagg
 1261 atgtggcaga gaggtaggaca agcaatgtat gcccctccca tccaaggaga attaagggtg
 1321 gactcaaca ttacaggcct actattaaca agagatgggt ggttaataa cacacaaaat
 1381 gataccttca gacctacagg aggaatatg aaagacaatt ggagaagtga attatataaa
 1441 tataaagtag tggagattaa accactagga atagaccca cccggg
 env cds ->

LOCUS HIVCAR4067 1498 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4067-68, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4067-68.
 REFERENCE 1 (bases 1 to 1498)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env G subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 563 a 239 c 320 g 376 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggggatata gaggaattgg cagcacttat ggaaatgggg gactttgatc
 env start ->
 61 cttggggttg tgataatttg tagcgccaca aatagcttgt gggtcacagt ctattatggg
 121 gtgectgtgt ggaagatcg agatgcccct ctattttgtg catctgatgc taggtcatat
 181 agtactgaaa aacataatgt ctgggccaca catgectgtg taccacgga ccccaacca
 241 caagagatag ttctggaaaa tgtaacagaa aattttaata tgtggaaaa taacatggta
 301 gaacagatgc atgagatat aattagttta tgggatgaaa gcttagagcc atgtgtaaag
 361 ctaacccctc tctgtgttac tttaactgt actgaagtga acaacaattg tactagggac
 421 aacaatagca ctgtgaattg tgaggaagaa aaaatgaaa actgctcttt caatataacc
 481 acagaaataa gagataagag aaagacagaa tacgcgcttt tctatagact tgatgtagta
 541 ccaattaatg aggatcagaa tagtagtact tatagattaa taaattgtaa tgtctcaacc
 601 attaaacagg cttgtccaaa ggtgactttt gagccaattc ccatacatta ttgtgctcca
 661 gctgggttat cgattttaaa gtgtagagat aagaatttca atggaagagg accatgtaa
 721 aatgtcagta cagtacaatg tacacatgga attagccag tggatatcaac tcaattactg
 781 ctgaatggca gtttagcaga agggaaaata atgattagat ctgaaaacct tacaacaat
 841 gccaaaatca taatagtga gcttaataaa tctgtagaaa ttatatgtac cagacccaat
 v3 loop ->
 901 aacaatacaa gaaaaagtat aagcctttga ccaggacaag cgttctatgc aacaggggac
 961 ataataggag atataagaca agcacattgt aatgttagta gaacagcttg gaatcagatg
 <- v3 loop
 1021 ttacagaatg tgactacaag actaaaagaa acctttcggg gggagaacat aaccttagaa
 1081 aacataacct ttaactcacc tgcaggaggg gaactagaag ttacaaaaca tagttttaat
 1141 tgtggaggag aatttttcta ttgtaataca tcaggactgt tcaatagtaa tagcactaag
 1201 actacaaatg atactaagac aatgatact atcacactcc aatgtagaat aaaacaaatt
 1261 gtgaaaatgt ggcagagaac aggacaagca atgtataacc ctcccattgc aggaacatt
 1321 acctgcataa caaacattac aggtactctc ttaacacgtg acgggggaaa taatagtaca
 1381 gagaatagaa cagagacott cagacctgga ggaggagata tgagggataa ttggagaagt
 1441 gaattatata agtataaagt aataaaaatt aaaccactag gaatagcacc caccggg
 env cds ->

HIVCAR4071

LOCUS HIVCAR4071 1492 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4071-TG2, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4071-TG2.
 REFERENCE 1 (bases 1 to 1492)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env E subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 572 a 235 c 306 g 379 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggggacaca gatgaattgg ccaaacttgt ggaaatgggg gactttgac
 env start ->
 61 cttgggttgg tgataatgtg tagtgccctca gacaacttgt gggttacagt ttattatggg
 121 gttcctgtgt ggagagatgc agataccacc ctattttgtg catcagatgc caaagcacat
 181 gagacagaag tgcacaatat ctgggctaca catgcctgtg taccaacaga ccccaaccca
 241 caagaaatat tcttggeaaa tgtaacagaa aatttttaaca tgtggaaaaa taacatggta
 301 gagcagatgc aggaggatgt aatcagttta tgggatcaaa gtctaaagcc atgtgtaaag
 361 ttaactcctc tctgtgttac tttagattgt accaaagctg attttaacac cacaataatt
 421 aacacaacag aaaagccaga aatagaaata tcagaaatgc aaaaagaagt aagcaactgt
 481 tcttttaata taaccacaga actaagagat aaggagcagg aggtctatgc actcttttat
 541 aggtctgata tagtaccat tgagagaaat agtggggaaa ataatggtag tagttataga
 601 ttaatacatl gtaatacttc agtcatttaag caggcttgtc aaaaagtatc ctttgatcca
 661 attcccatac attattgcgc tccagctggc tatgcatatt taaaatgtaa tgataagaat
 721 ttcattggaa caggggccatg taaaaggta agctcagtac aatgcacaca tggaaattaag
 781 ccagtgggat caactcaatt gctgttaaat ggtagtctag cagaagaaga aataataatt
 841 agatctgaaa atctcacaaa caatgccaaa accataatag tgcaccttaa taaatctgta
 901 gaaatcaatt gtaccagacc ttccaaaag atgagaacaa gtgtaagaat aggaccaggg
 v3 loop ->
 961 cgagtattct ataaaacagg atccataaca ggagatataa gaaaagcata ttgtgagatt
 <- v3 loop
 1021 aatggaacaa aatggaatga aactttacaa cagataatta gaaaattaga agagcacttt
 1081 aataataaga caatacaatt taaaccacca tactcaggag gagatctaga aattacaatg
 1141 caccatthta attgtagagg ggaatttttc tattgcaata caacaagact gtttaatat
 1201 agcacaaatg gaaccacaaa tggcactatc acacttccat gcaggataaa gcaagttata
 1261 aagatgtggc aggaagtagg acaagcaatg tatgctctc ccatcgatga agcaattaat
 1321 tgtgtatcaa atattacagg aatactattg gtaagagatg gtggtaaaat tgaatatgag
 1381 acaatogaaa ccttcagacc tggaggagga aatatgaagg acaattggag aagtgaatta
 1441 tataaatata aagtagtaca aattgaacca ctaggaatag caccaccccg gg
 env cds ->

LOCUS HIVCAR4081 1486 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate 4081, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 4081.
 REFERENCE 1 (bases 1 to 1486)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env G subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 577 a 245 c 298 g 366 t
 ORIGIN Beginning of env cds
 1 ATGaaagtga tggagacaca gaggaattgt ccacgtcat ggagatgggg gattttgatc
 env start ->
 61 ctgggggttg tgataatttg taatgcctca gacaacaact tgtgggtcac agtttattat
 121 ggggtacctg tgtggaaga tgcagagacc accctatatt gtgcatttga tgctaaatca
 181 tatagtactg aaaaacataa tgtctgggct acacatgcoo gtgtacccac agaccccaac
 241 ccacaagaaa tactttttag aaatgtaaca gaaaagttaa acatgtggaa aaataacatg
 301 gtagaacaga tgcattgagg tataatcagt ttatgggatg aaagcctaaa accatgtgta
 361 aagctaacc ctttttgtgt tactttaacc tgtactaata taaccaaaaa tagtactgaa
 421 gtaacaaca atagcactgt ggataataaa gaagaaataa aaaattgctc ttcaatata
 481 accacagaaa taagagataa gaagaagaca gcatacgcgc tttctatag actagatgta
 541 gtaaccaatta atgataataa taataataat agttataggc taataaattg taatacctca
 601 gccattaaac aggtttgtcc aaaggtatcc tttagagcaa ttccatata ttattgtgct
 661 ccagctggtt ttgcgattct aaagtgtatg gataagaagt tcaatggaaa aggaccatgt
 721 aaaaatgtca gtacagtcca atgcacacat ggaatcaagc cagttagtat aactcaactg
 781 ctgttgaatg gcagcctggc agaagaagag ataataatta gatctgaaa catcacagac
 841 aatgccanaa tcataatagt acatcttaat gactctgtag aaattaattg taccaggcoo
 v3 loop ->
 901 aacaacaata caagaaaaag tgtaaggata gggccaggac aaacattcta tgcaacaggt
 961 gacataatag gagatataag acaagcacat tgtaattgtc gtagaacaaa atggaatcag
 <- v3 loop
 1021 actttacaac aaatagctac acaattaaga aaatactttg ggaatacaac aacaataaac
 1081 tttagtaaat cctcaggagg agatttagaa attacaacac acacttttaa ttgtggagga
 1141 gaatttttct attgcaatac atcagcttg tttaatagca gttggaaagg aaagtacagt
 1201 gggcatgagt caaatgacac tataactctc cctgcaaaa taagacaaat tgtaaggatg
 1261 tggcagaaag taggacaagc aatgtatgcc cctcccatcc caggagtaat aaggtgtgaa
 1321 tcaaacatta caggactaat actaacaaga gatgggtgaa ataataagac taataacact
 1381 gaaaccttca gacctggggg aggagatag agggacaatt ggagaagtga attatataaa
 1441 tataaagtag taaaattga accactagga gtagcacca cccggg
 env cds ->

HIVCAR423A

LOCUS HIVCAR423A 1516 bp ss-DNA VRL 01-OCT-1994

DEFINITION Human immunodeficiency virus type 1, CAR isolate 11423-21, env glycoprotein gp120, partial cds.

ACCESSION

SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate 11423-21.

REFERENCE 1 (bases 1 to 1516)

AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,P., Deslandres,A., and Kieny, M.F.

TITLE ;

JOURNAL Unpublished

COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env A subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.

BASE COUNT 577 a 255 c 313 g 371 t

ORIGIN Beginning of env cds

1 ATGagagtga aggagaata tcagcacttg tggagatggg ggtggaaatg gggcagcatg
env start ->

61 ctccctggga tattgatgat ctgtaatgct tcagacaaat tgtgggtcac agtctattat
121 ggggtacctg tgtggcgaga tgcaaatacc actctatttt gtgcacttga tgctaaagca
181 tatagtactg aaaaacataa tgtctgggct acacattcct gtgtaccac agaccocaac
241 ccacaagaga tacttctgga aaatgtaaca gaaaatttta acatgtggaa aaatagtatg
301 gtagatcaga tgcagtggga tataatcagt ttatgggatg aaagcctaaa gccatgtgta
361 aagctaacc cctctctgtg tactttaaac tgtactaatg taactgtaac caacagcagt
421 aataatgtaa ccaaaaccaa cagtagccaa tctcttaagt ataacattga tggaaaagaa
481 gaaataaaaa actgttcttt caatacgacc acagaaataa tagataagag gaagcaagaa
541 taaggccttt tctatagact tgatatagta ccaattaata ataataataa tacaatgag
601 tataggttaa taattgtat tgctcaact gttaaacagg ctgtgtccaa ggtatccctt
661 gagccaattc ccatacatta ttgtgtctcg gctgggtttg cgattctaaa gtgtagggat
721 gagaagttca atggaacagg gccatgcaat aatgtcagca ccgtacaatg cacacatgga
781 atcaagccag tagtatcaac tcaactactg ttaaatggca gcctagcaga aggggaggta
841 aagattagat ctgaaaactt cacaacaat gctaaaacca taatagtaca gcttgacagt
901 cctgtaagaa ttaattgtac cagacctaac aactatacaa gaaaatgat acctacagga
V3 loop ->

961 ccaggacaag gacaagtcac ctatgcaaca ggtaaaataa taggggatat aagaaaagca
1021 tattgtaaca tcagttaagc agaattgaaa aacatgttaa gaaaggtagc tacaatacta
<- V3 loop

1081 gaaagcaagt ttaacaaatc aattaccttt gataaacctt caggagggga tatagaattt
1141 acaatatata gctttaattg taaaggagaa tttttctatt gcaatacacc aaaactgttt
1201 aatagcactt ggaaaaatgg cactactgaa gaaaatgaca cggagtcaga ggaacctata
1261 gttctcccat gcaggataaa gcaaatgtga catatgtggc agagagtagg acaagcaatg
1321 tatacccttc ccatacaagg agaaataaag tgtgaatcaa acatcacagg actactatta
1381 acaagagatg gtgggaagaa cacaactaat aagactttca gacctatagg aggagatatg
1441 agggacaatt ggagaagtga attatatagg tataaagtag tggaaattga accattagga
1501 gtagcaccac ccggg
env cds ->

LOCUS HIVCARELO 1510 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate ELO, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate ELO.
 REFERENCE 1 (bases 1 to 1510)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env E subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 550 a 257 c 314 g 389 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggggacaca gatgaattgg ccacacttgt ggaaatgggg aactttgac
 env start ->
 61 cttgggttgg tgataatttg tagggcctca gaaaacttgt gggttacagt ctattatggg
 121 gttctctggt ggagagagcg agaaaccacc ctattttgtg catcagatgc caaagcacat
 181 gagacggaag tgcacaatgt ctgggctaca cagcctgtg taccacaga ccctaacca
 241 caagaaatat acatggaaaa tgtaacagaa cattttaaca tgtggaaaa taacatggta
 301 gagcagatgc agggagatgt aatcagttta tgggatcaaa gtctaaagcc atgtgtaaaa
 361 ttaactcctc tctgcgttac tttaagggtg accaatgcta attttacc aa tgcctaccgat
 421 gctaatgtta ccaatgctac tgcataattt aatgctacaa aagaatgga aaatacagta
 481 aaaaactggt cttttaatat gaccacagaa ctgggagata agcatcagaa ggtctctgca
 541 cttttttata agcttgattt agtaccaatt gatggcaata ttgatggcaa taatagttag
 601 tatatattaa tacattgtaa taattcagtc attaagcagg cttgtccaaa ggtgtccctt
 661 gatccaattc ccatacatta ttgtgccccg gctgattttg cgattttaaa gtgtaatgat
 721 aaggagtcca atggaacagg gccatgcaaa aatgtcagct cagtacaatg cacacatgga
 781 attaaagccg tggatcaaac tcaattgctg ttaaatggca gtttagcaga agaagagata
 841 ataactcagat ctgaaaatct cacaacaact gccagaacca taatagtgea ccttaataaa
 901 tctgtagaaa tcaattgtac cagaccctac aaaaatacaa gaacaagtgc aaggatagga
 v3 loop ->
 961 ccaggacaag tattctataa aacaggatcc ataacaggag atataaggaa agcatattgt
 <- v3 loop
 1021 gagattaatg gaacagcatg gcattgaaact ttanaacaga tgggttgaaa attaagaang
 1081 cactttaata agacaataat ctttcaacca tcttcaggag gagatccaga aattacaatg
 1141 cataatttta attgttagagg ggaatttttc tattgcaata caacaaacct gtttaatagt
 1201 acttgggaaa taataaaac catgaagtct aatgacacaa acttaagcaa tatcacactt
 1261 ccatgcagaa taaagcaaat tggaaacatg tggcagggag taggacgagc aatgtatgct
 1321 cctcccatca gtgaactagt taattgtgta tcaaatatta caggaatact attgacaaga
 1381 gatgggtggc ctaatactgc gaatgagacc ttcagacctg gaggaggaa tatgaaggac
 1441 aattggagaa gtgaattgta taatatataa gtagtacaaa ttgaaccact aggaatagca
 1501 cccacccggg
 env cds ->

HIVCARGAN

LOCUS HIVCARGAN 1507 bp ss-DNA VRL 01-OCT-1994

DEFINITION Human immunodeficiency virus type 1, CAR isolate GAN, env glycoprotein gp120, partial cds.

ACCESSION

SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate GAN.

REFERENCE 1 (bases 1 to 1507)

AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.

TITLE ;

JOURNAL Unpublished

COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env D subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.

BASE COUNT 574 a 264 c 316 g 353 t

ORIGIN Beginning of env cds

1 ATGagagtga tggggatata gaagaattat caacacttgt tgagatgggg gactatgac

env start ->

61 ttggggctaa taataattag ttgtgttgca gaccaactgt gggtcacagt ctactatggg

121 gtacctgtgt ggaaaaatgc agacaccacc ctattttgtg catcaaatgc cagagcatat

181 gatacagaag tgcataatgt ctgggctaca catgcctgtg taccacacaga cccaaccca

241 caagaataaa atttgggtcaa tgtgacagaa gagtttaaca tgtggaaaaa taacatggta

301 gaacagatgc atacagatat aaccagtcta tgggacccaa gectacagcc atgtgtaaag

361 ttaacccctc totgctttac cttagaatgc catgaaccca acaccacaga tagcaccaaa

421 aacaacagcg cagaaaacat gacctttgaa gggataggag aaataacaaa ctgctctttc

481 aatatgacca cagagctaag agataagaaa cagaaagtat attcactttt ttataaactt

541 gatgtagtac aaatagatga gaataatggt actagaaata acagtgaagt gtatagacta

601 ataaattgta atacctcagc cattacacaa gcttgtccaa aggttaacct ttgagccaatt

661 ccatacatt attgtactcc agctgggttt gogattctaa agtgaatga gaaaaagttt

721 aatggaacag ggccatgcag gaatgtcagc tcagtacaat gcacacatgg aatcaaacca

781 gtagtaacaa ctcaactgct gttaaatggc agtctagcag aagaagaggt aatcatcaga

841 tctgaaaata tcacaaacaa tgcanaaaca ataatagtac agcttgtcaa gcttgtaaat

901 attacttgta tcagacctag caggactaca aggaagaagat ggcataatag atcaggacaa

v3 loop ->

961 gccttctatg caatagatgg cataacagga gatataagaa aagcatattg cyaggtcaat

<- v3 loop

1021 acaacagaat ggaatcaaac tttaacaacag gtagtaaaaa aattaagaac atattggaac

1081 aaacaataaa acttcactaa accctcggga ggggatttag aaattacaac acatagtttt

1141 aattgtggag gagaattttt ctattgcaat acatcaggcc tgtttaatag cagttggaag

1201 aatgagaatg aactaaaag catgaaggga ttaaaaggta ctataactct ccagtgcaga

1261 ataaagcaaa ttgtaaatat gtggcagaga gcagggcaag caatgtatgc cctcccatc

1321 acaggaataa taaggtgtga gtcaaacatt acaggactac tactgacaag agatgggtggg

1381 ttttaataaa ggacaaatga tactgagacc tttagacctg ggggaggaga tatgagagac

1441 aattggagaa gtgaattata taaatataaa gtagtaaaat tgaccactag gatacagacc

1501 acccggg

env cds ->

LOCUS HIVCARMBA 1471 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate MBA, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate MBA.
 REFERENCE 1 (bases 1 to 1471)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,P., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env E subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 552 a 239 c 298 g 382 t
 ORIGIN Beginning of env cds
 1 ATGagagtga aggggacaca gatgacttgg ccaaactggt ggaaatgggg gactttgac
 env start ->
 61 cttgggttgg tgataatttg taatgacctc aacaacttgt gggttacagt ttattatggg
 121 gttcctgtgt ggaagatgc agataccacc ctatgctgtg catcagatgc caaagcacac
 181 gtgacagaaa tgcacaatgt ctgggctaca catgctgtg taccacatc ccaaccacaa
 241 gaagtataca tggaaaatgt aacagaaaat tttacatgt ggaataataa tatggtagaa
 301 cagatgcatg aggcataat cagtttatgg gatgaaagcc taaaaccatg tgtaaagcta
 361 actcctctct gtgttacctt aaattgtacc agggctaata ttaattctac caatgcta
 421 tctaccaatg ctaataccct aaaaatggaa aatgaagtaa aaactgttc ttttaatatg
 481 accacagaa taagagataa gaagcagaag gtctctgcaa ttttttatgt atctgatcta
 541 gtatcaatga atgcaataa tagctatagg ttaataaatt gtaatacttc agtcattaag
 601 caggcttgtc caaagggtgc ctttgatcca attcccatc attattgtgc tccagctggg
 661 tatggcattt taaagtgtaa tagtaagact ttcaatggaa cagggccatg taaaatgta
 721 agtcagtagc aatgcacaca tggaaattaag ccagtgggat caactcaatt gctgttaaat
 781 ggagctctag cagaagaaga gataataatc agatctgaaa atctcgcaaa caatgccaaa
 841 accataatag tgcaccttaa taaacctgta gaatacaatt gtaccagacc ctccaacaa
 v3 loop ->
 901 acaagaacaa gtgtacgtat aggaccagga caagtattct atggaacagg agaaataata
 961 ggagatataa gaagagcatt ttgtgagatt aatggaacaa aatggaatga aactttaaag
 <- v3 loop
 1021 caggttaactg aaaaattaag agaacacttt aataagacaa tagtttttca accaccctca
 1081 ggaggagatc tagaaattac aatgcatagt ttttaattga gaggggaatt tttctattgc
 1141 aatacaacaa aactgtttta tagtacttgg gaattgaata aaaccataga ggagcttaat
 1201 ggcactatca cacttccatg caggataaag caaattgtta acatgtggca gggagttaga
 1261 caagcaatgt ataactctcc cgtcagtgga gtaattaatt gtgtatcaaa tattacagga
 1321 atactattaa caagagatgg tgggatgctg aataatacac aaatgagac cttcagacct
 1381 ggaggaggaa atataaagga taattggaga aatgaattat ataataataa agtagtacia
 1441 attgaaccac taggaatagc acccaaccgg g
 env cds ->

HIVCARSAS

LOCUS HIVCARSAS 1501 bp ss-DNA VRL 01-OCT-1994
 DEFINITION Human immunodeficiency virus type 1, CAR isolate SAS, env glycoprotein gp120, partial cds.
 ACCESSION
 SOURCE Human immunodeficiency virus type 1 (HIV-1), CAR isolate SAS.
 REFERENCE 1 (bases 1 to 1501)
 AUTHORS Schmitt,D., Mathiot,C., Levy,J.-P., Girard,M., You,B., Barre-Sinoussi,F., Deslandres,A., and Kieny,M.P.
 TITLE ;
 JOURNAL Unpublished
 COMMENT Kindly provided prior to publication by Dr. Marie Paule Kieny of Transgene in Strasbourg, France. This sequence, from a Central African Republic isolate, clusters with env A subtype sequences. It represents one of 14 CAR isolates that have been integrated into a "hybrid" gp160 sequence for expression from a vaccinia virus expression system for study by Transgene and Pasteur Merieux.
 BASE COUNT 569 a 263 c 307 g 362 t
 ORIGIN Beginning of env cds
 1 ATGagagtga agggggacaca gaagaattac totcccttgt ggagatgggg gattatgata
 env start ->
 61 ctggggctga taataatctg tagtgctaca gaaaaattgt gggtcacagt ttactatggg
 121 gtacctgtgt ggaagatgc aaagaccacc ctatthttgt catcagatgc taagcatat
 181 gatacagaaa tgcataatgt ctggggccaca catgctctgt taccacaga ccccaaccca
 241 caagaaatgg atttgaaaaa tgtgacagaa aatttttaaca tgtggaaaaa taacatggta
 301 gagcaaatgc atacagatat aatcagccta tgggacaaa gcctaaagcc atgtgtacag
 361 ttaacccctc tctgtgttac tttagactgt actaatgtag ctaacgacag tacaatgaa
 421 ctaagaaaact gctcttaca tatgaccaca aaactaaatg ataaaagaca aaaggtacat
 481 tcaacttttt ataaacttga tatagtccca attaataaaa ttaatgatgc tagtggtaat
 541 attagtgaga acaataatag tagtagtggc aatattacta aacaatatag actaatagat
 601 tgtaatacct caaccattac acaggcttgt ccaaaggat cctttgagcc aattcccata
 661 cattattgtg cccagctgg ttttggcatt ctaaagtgt aggatagga gttcaaagga
 721 acagggccat gcaagaatgt tagtacagta caatgcacac atggaatcag gccagtagta
 781 tcaactcaac tactgtttaa tggcagccta gcagaaggaa agataatgat tagatctgaa
 841 aatttctcaa acaatgccaa aaccataata gtacagcta accagtctgt agaattaat
 901 tgtactagac ccaataacaa tacaagaaga agaatgcata taggaccagg acgagcattc
 v3 loop ->
 961 attgcaacag atgoccatagt aggggatata aggcaagcat attgtaattgt cagtagagca
 <- v3 loop
 1021 acatggaata aaactttaca aatggtagct gaagaattaa gaaaacagtt taacaaaaca
 1081 ataactcttg ataacccttc aggggggat gtagaaatta caacacatag ttttaattgt
 1141 ggaggagaat ttttctattg caacacatca ggcctgttta ataaaacttg gtgggggcca
 1201 aatagcacca ccaacaatca gaccattggc aatgacacta taactctccc atgcagaata
 1261 aagcaaatg taaggatgtg gcagagagta ggaaaagcaa tgtatgccc tcccatcaa
 1321 ggagcgataa actgttagatc aaacattaca ggactactat taacaagaga tggagggggg
 1381 aataacaggt caaatgagac ctccagacct gaaggaggaa atacaagga caattggaga
 1441 agtgaattat ataagtataa agtagtaaaa attgaaccac taggaatagc accaccogg
 1501 g

LOCUS HIVCI47 900 bp ss-RNA VRL 31-MAR-1994
 DEFINITION Human immunodeficiency virus type 1 (HIV-1) V3-V5 region envelope sequence, Ivory Coast sample CI-47, clone 1.
 ACCESSION X72049
 SOURCE Human immunodeficiency virus type 1 cocultivated PCR-amplified PBMC DNA derived from a patient from Cote d'Ivoire
 REFERENCE 1 (bases 1 to 900)
 AUTHORS Janssens,W., Heyndrickx,L., Van de Peer,Y., Bouckaert,A., Franssen,K., Motte,J., Gershy-Damet,G.M., Peeters,M., Piot,P. and Van der Groen,G.
 TITLE Molecular phylogeny of part of the env gene of human immunodeficiency virus type 1 isolated in Cote d'Ivoire
 JOURNAL AIDS 8, 21-26 (1994)
 COMMENT Janssens et al. conducted a survey designed to determine the genetic HIV-1 variation present in Cote d'Ivoire. As a part of this study, viruses were isolated from eleven symptomatic (CI-3, CI-13, CI-21, CI-22, CI-42, CI-45, CI-47, CI-326, CI-327, CI-329, and CI-330) and two asymptomatic (CI-14 and CI-20) HIV-1 infected individuals. Isolates CI-13, CI-14, CI-45 and CI-47 were derived from patients whose sera were dually reactive to both HIV-1 and HIV-2 antibodies. These thirteen isolates were sequenced over the envelope region including V3, V4, V5 and the beginning of gp41. In phylogenetic analysis, isolates CI-22 and CI-13 cluster with subtype B and subtype D respectively. The remaining eleven of the thirteen HIV-1 isolates cluster with HIV-1 subtype A, with significant variation among these sequences. Janssens et al. propose this variation suggests either a local burst of particular variants, or a recent spread of the epidemic in Abidjan. The envelope sequences in this set are shown in the alignment below (accession numbers X72024-X72065).
 NCBI gi: 468672
 BASE COUNT 350 a 142 c 202 g 206 t
 ORIGIN 790 bp downstream from the beginning of env cds
 1 gaaggaaatag tgattagatc tgaatatc acaacaatg ccaaaaccat aatagtagc
 env cds ->
 61 ttaagtaagc ctgtaaaaat taattgtacc agacctggca actatacaag aaaaagtgt
 V3 loop ->
 121 cgtacaggta taggaccagg gcaaacattc tatgcaacag gtaaaataat aggggatata
 181 agacaagcac attgtaattgt cagtagaaga gactggaata acactttaca acaggtagct
 <- V3 loop
 241 acaaaattea gggagtactt taataaaaca ataatttttg ataatacttc agggagggat
 301 gtagaaatta caacacatag ttttaattgt ggaggagaat ttttctattg ctatacatca
 361 aaactgttta atagcacttg gaataacaat gacactgcca gctcaataaa catagagtca
 421 aatgataata taactctcca atgcagaata aagcaaatg taaatatgtg gcagaaagtg
 481 ggacaagcaa tgtatgcccc tcccatccaa ggagaataaa agtgtgtatc aaacatcaca
 541 ggactactat taacaagaga tggagggaat aataataata gtatgaatga gactttcagg
 601 cctggaggag gagatatgag ggacaattgg agaagtgaat tatataagta taaagtagta
 661 aaaattgaac cactaggtgt agcaccacac catgcaaaaa gaagagtggg ggagagagaa
 721 aaaagagcag ttggactggg agcgtttctt cttgggttct taggagcagc aggaagcact
 781 atgggcgcgg cgtcaataac gctgacggta caggccagac aattattgtc cggatatagt
 841 caacagcaaa gcaatttgtt gagggctata gaggtcaca aacatctgtt gaaactcag
 env ->

CONSENSUS	GGAGA?...GTAGTGATTAGATCTGAAAATATCACAAACAATGCCAAACCATAATAGTACAGTT?G?TA	64
CI-3-1	--A-T...--CA-----G-----G-----TAC-G	67
CI-3-2	--A-T...--ACA-----G-----G-----TAC-G	67
CI-3-3	--A-T...--ACA-----G-----G-----TAC-G	67
CI-45-3	-A-A-G...--C--A--A--C-----C-T-C--	67
CI-13-3	-A--A...A--A-A-----G--C-T--T--A--T-----C-TAA-G	67
CI-13-4	-A--A...A--A-A-----G--C-T--T--A--T-----C-TAA-G	67
CI-14-5	-A--G...A-----G-T--	67
CI-14-13	-A--G...-----T-----G-T--	67
CI-14-21	-A--G...A-----G-T--	67
CI-20-1	-AGA-G...A--C-----G-----A-----G-T--	67
CI-20-2	-AGA-G...A--C-----G-----A-----G-T--	67
CI-20-3	-AGA-G...A--C-----G-----A-----G-T--	67
CI-20-5	-AGA-G...A--C-----G-----A-----G-T--	67
CI-21-1	----G...--AGA-----G-T-----A-----AC-T-CC-	67
CI-21-2	----G...--AGA-----G-T-----A-----AC-T-CC-	67
CI-21-4	----G...--AGA-----G-T-----A-----AC-T-CC-	67
CI-22-3	-A--A...--A-----C--T--GG-----T-----C-GAACG	67
CI-22-6	-A--A...--A-----C--T--GG-----T-----C-GAACG	67
CI-22-10	-A--A...--A-----C--T--GG-----T-----C-GAACG	67
CI-42-1	----TGAT--A-----G-C-G-G--A-----C-T-T-G	70
CI-42-2	----TGAT--A-----G-C-G-G--A-----C-T-T-G	70
CI-42-3	----TGAT--A-----G-C-G-G--A-----C-T-T-G	70
CI-42-4	----TGAT--A-----G-C-G-G--A-----C-T-T-G	70
CI-45-1	-A-A-G...--C--A--A--C-----C-T-C--	67
CI-45-2	-A-A-G...--C--A--A--C-----C-T-C--	67
CI-47-1	-A-GA...A-----AAG--	67
CI-47-3	-A-GA...A-----AAG--	67
CI-47-8	-A-GA...A-----AAG--	67
CI-326-2	----A...--C-----GTC--	67
CI-326-3	----A...--C-----GTC--	67
CI-326-4	----A...--C-----GTC--	67
CI-327-1	----G...A-----G-----AAA-G	67
CI-327-2	----G...A-----G-----AAA-G	67
CI-327-3	----G...A--C-----G-----AAA-G	67
CI-327-4	----G...A-----G-----AAA-G	67
CI-329-1	----A-T...--ACA-----T-A-----AAA--	67
CI-329-3	----A-T...--ACA-----G-----AAA--	67
CI-329-5	----A-T...--ACA-----T-A-----AAA--	67
CI-330-1	----T...--A-----G--C--A-----C-T-A-G	67
CI-330-2	----T...--A-----G--C--A-----C-T-A-G	67
CI-330-3	----T...--A-----G--C--A-----C-T-A-G	67
CI-330-4	----T...--A-----G--C--A-----C-T-A-G	67

CONSENSUS	AGCCTGTAA?AATTAATTGTACCAGACCTAGCAACAATACAAGAAAAAGTGTGTA.....C?TATAGG	123
CI-3-1	-----G-----A-----G-----	128
CI-3-2	-----G-----A-----C-----	128
CI-3-3	-----G-----AT-----G-----	128
CI-45-3	-----A-----C-A-----G-----	128
CI-13-3	-AT-----C-----C-----G-C-A-----C-G-AC-----A-----	128
CI-13-4	-AT-----C-----C-----G-C-A-----C-G-AC-----A-----	128
CI-14-5	-----G-----CA-----	128
CI-14-13	-----G-----CA-----	128
CI-14-21	-----G-----CA-----	128
CI-20-1	-----G-----C-----A-----G-----	128
CI-20-2	-----G-----C-----A-----G-----	128
CI-20-3	-----G-----C-----A-----G-----	128
CI-20-5	-----G-----C-----A-----G-----	128
CI-21-1	-AAA-----C-----G-----T-----AA-----	128
CI-21-2	-AAA-----C-----G-----T-----AA-----	128
CI-21-4	-AAA-----C-----G-----T-----AA-----	128
CI-22-3	-T-----GT-----A-----C-A-----T-----A-----A-----G-----	128
CI-22-6	-T-----GT-----A-----C-A-----T-----A-----A-----G-----	128
CI-22-10	-T-----GT-----A-----C-A-----T-----A-----A-----G-----	128
CI-42-1	-----A-----T-----C-AT-----G-----A-----	131
CI-42-2	-----A-----T-----C-AT-----G-----A-----	131
CI-42-3	-----A-----T-----C-AT-----G-----A-----	131
CI-42-4	-----A-----T-----C-AT-----G-----A-----	131
CI-45-1	-----A-----C-A-----G-----	128
CI-45-2	-----A-----C-A-----G-----	128
CI-47-1	-----A-----G-----T-----CGTACA...GG-----	134
CI-47-3	-----A-----G-----T-----CGTACA...GG-----	134
CI-47-8	-----A-----G-----T-----CGTACA...GG-----	134
CI-326-2	-----G-----G-T-----G-----TGGAGAAAGTTGG-A-----	137
CI-326-3	-----G-----G-T-----G-----TGGAGAAAGTTGG-A-----	137
CI-326-4	-----G-----G-T-----G-----TGGAGAAAGTTGG-A-----	137
CI-327-1	-A-----A-----T-----G-----G-----A-----	128
CI-327-2	-A-----A-----T-----G-----G-----A-----	128
CI-327-3	-A-----A-----T-----G-----G-----A-----	128
CI-327-4	-A-----A-----T-----G-----G-----A-----	128
CI-329-1	-----A-----G-----T-----G-----	128
CI-329-3	-----A-----G-----T-----G-----	128
CI-329-5	-----A-----G-----T-----G-----	128
CI-330-1	-----C-----C-A-----C-----G-----	128
CI-330-2	-----C-----C-A-----C-----G-----	128
CI-330-3	-----C-----C-A-----C-----G-----	128
CI-330-4	-----C-----C-A-----C-----G-----	128

HIVCM7

CONSENSUS	ACCAGGGCAAGCATTCTATGCAACAGGTGACATAATAGGGGATATAAGACAAGCACATTGTAATGTCAGT	193
CI-3-1	-----	198
CI-3-2	-----A-----	198
CI-3-3	-----	198
CI-45-3	-----A-----G-G-----	198
CI-13-3	-----G-C-A-----GGG-----A-----A-----A-T-----	195
CI-13-4	-----G-C-A-----GGG-----A-----A-----A-T-----	195
CI-14-5	-----A-----	198
CI-14-13	-----G-----A-----	198
CI-14-21	-----A-----	198
CI-20-1	-----A-T-----CCG-----	198
CI-20-2	-----A-T-----CCG-----	198
CI-20-3	-----A-T-----CCG-----	198
CI-20-5	-----A-T-----CCG-----	198
CI-21-1	-----A-----T-----	198
CI-21-2	-----A-----T-----	198
CI-21-4	-----A-----T-----	198
CI-22-3	-----AGC-A-T-----A-A-----A-----A-----T-C-CA-T-----	198
CI-22-6	-----AGC-A-T-----A-A-----A-----A-----T-C-CA-T-----	198
CI-22-10	-----AGC-A-T-----A-A-----A-----A-----T-C-CA-T-----	198
CI-42-1	-----A-T-----T-----A-----	201
CI-42-2	-----A-T-----T-----A-----	201
CI-42-3	-----A-T-----T-----A-----	201
CI-42-4	-----A-T-----T-----A-----	201
CI-45-1	-----A-----G-----	198
CI-45-2	-----A-----A-----A-----G-----	198
CI-47-1	-----A-----A-A-----	204
CI-47-3	-----A-----A-A-----	204
CI-47-8	-----A-----A-A-----	204
CI-326-2	-----G-----A-A-G-----T-----A-----T-----	207
CI-326-3	-----G-----A-A-G-----T-----A-----T-----	207
CI-326-4	-----G-----A-A-G-----T-----A-----T-----	207
CI-327-1	-----G-----	198
CI-327-2	-----G-----	198
CI-327-3	-----G-----	198
CI-327-4	-----G-----	198
CI-329-1	-----A-----C-----	198
CI-329-3	-----A-----C-----	198
CI-329-5	-----A-----C-----	198
CI-330-1	-----A-----A-----	198
CI-330-2	-----A-----A-----	198
CI-330-3	-----A-----A-----	198
CI-330-4	-----A-----A-----	198

CONSENSUS	AGAACAGAAATGGAATAAACTTTACAACAGGTAGCTA?ACAATTAAGG?AGCACTTT...AA?AA?ACAA	256
CI-3-1	-A-GAG-----T-A--GG-A-----T-A-----G-AT---AAG-T-C---	268
CI-3-2	-A-A-----A--GG-A-----T-A-----G-T---AAG-T-C---	268
CI-3-3	-A-A-----A--GG-A-----T-A-----G-T---AAG-T-C---	268
CI-45-3	G-----G---G-----G-GA--G--GCG-----A-----T-A--	262
CI-13-3	G-G-A-C-----A-----T-A-----AA-C-TTC-----C-G---	262
CI-13-4	G-G-A-C-----A-----T-A-----AA-C-TTC-----C-G---	262
CI-14-5	-A-----G-----G-----C-----A-A-----G-TG---	265
CI-14-13	-A-----G-----C-----A-A-----G-TG---	265
CI-14-21	-A-----G-----G-----C-----A-A-----G-TG---	265
CI-20-1	-----G-----C-----AGG-T-----A-TG---	265
CI-20-2	-----G-----C-----AGG-T-----A-TG---	265
CI-20-3	-----G-----C-----AGG-T-----A-TG---	265
CI-20-5	-----G-----C-----AGG-T-----A-TG---	265
CI-21-1	---G---C-----C---G---G---C-AC-----AAATA-----G-G-C-A--	265
CI-21-2	---G---C-----C---G---G---C-AC-----AAATA-----G-G-C-A--	265
CI-21-4	---G---C-----C---G---G---C-AC-----AAATA-----G-G-C-A--	265
CI-22-3	G-G-A-----C---A---A---T-GG-A-----AG-A-A---GGA-T-T-A--	268
CI-22-6	G-G-A-----C---A---A---T-GG-A-----AG-A-A---GGA-T-T-A--	268
CI-22-10	G-G-A-----C---A---A---T-GG-A-----AG-A-A---GGA-T-T-A--	268
CI-42-1	GA-----G-G---G-A---GG-----TAC-A-----C-A---	268
CI-42-2	GA-----G-G---G-A---GG-----TAC-A-----C-A---	268
CI-42-3	GA-----G-G---G-A---GG-----TAC-A-----C-A---	268
CI-42-4	GA-----G-G---G-A---GG-----TAC-A-----C-A---	268
CI-45-1	G-----G---G---G-GA--G--GCG-----A-----T-A--	262
CI-45-2	GA---A-G-A-----GA---GA-GCG-----A-----T-A--	262
CI-47-1	---G---C-----C-----C-A-----G-T-----T-A--	268
CI-47-3	---G---C-----C-----C-A-----G-T-----T-A--	268
CI-47-8	---G---C-----C-----C-A-----G-T-----T-GT-A--	271
CI-326-2	-A---TG---GG-C---A---G---G---A-T-G-----T-T-AC-	274
CI-326-3	-A---TG---GG-C---A---G---G---A-T-T-----T-T-AC-	274
CI-326-4	-A---TG---GG-C---A---G---G---A-T-T-----T-T-AC-	274
CI-327-1	-A-G-A-----GC-----G---C-----A-T---AGG-TG-C---	268
CI-327-2	-A-G-A-----GC-----G---C-----A-T---AGG-TG-C---	268
CI-327-3	---G-A-----GC-----G---C-A-----A-T---AGG-TG-C---	268
CI-327-4	-A-G-A-----GC-----G---C-A-----G-T---AGG-TG-C---	268
CI-329-1	---A-----G-----G---G---C-----GG-T---AGG-TG-C---	268
CI-329-3	---A-----G-----G---G---C-----A-----AGG-TG-C---	268
CI-329-5	---A-----G-----G---G---C-----A-----AGG-TG-C---	268
CI-330-1	-----C-----G-----A---GA-----G-AC-A-----C-A--	262
CI-330-2	-----C-----G-----A---GA-----G-AC-A-----C-A--	262
CI-330-3	-----C-----G-----A---GA-----G-AC-A-----C-A--	262
CI-330-4	-----C-----G-----A---GA-----G-AC-A-----C-A--	262

CONSENSUS	CAATAATCTTTGCTAAC?CCTCAGGAGGGGAT?TAGAAATTACAACACATAGTTTAAATTGTGGAGGAGA	324
CI-3-1	-----A-----CA-----G-----	338
CI-3-2	-----A-----CA-----G-----	338
CI-3-3	-----A-----CA-----G-----	338
CI-45-3	-----G-C-----G-----	332
CI-13-3	-----T---AACTCAT---G-----GCCC-----T-----C-----G---	332
CI-13-4	-----T---AACTCAT---G-----GCCC-----T-----C-----G---	332
CI-14-5	-----C-----T-----G-GT-----	335
CI-14-13	-----C-----T-----G-GT-----	335
CI-14-21	-----C-----T-----G-GT-----	335
CI-20-1	GT---C-----G-G-----T-----C-----	335
CI-20-2	GT---C-----G-G-----T-----C-----	335
CI-20-3	GT---C-----G-G-----T-----C-----	335
CI-20-5	GT---C-----G-G-----T-----C-----	335
CI-21-1	-----A-----T-A-----C-----	335
CI-21-2	-----A-----T-A-----C-----	335
CI-21-4	-----A-----T-A-----C-----	335
CI-22-3	----G-----AACC-TT-----CCC-----GT-TG-C-C-----G---	338
CI-22-6	----G-----AACC-TT-----CCC-----GT-TG-C-C-----G---	338
CI-22-10	----G-----AACC-TT-----CCC-----GT-TG-C-C-----G---	338
CI-42-1	T---A---A---AT-----T-----C-----	338
CI-42-2	T---A---A---AT-----T-----C-----	338
CI-42-3	T---A---AA---AT-----T-----C-----A---	338
CI-42-4	T---A---A---AT-----T-----C-----	338
CI-45-1	-----G-C-----G-----	332
CI-45-2	-----G-C-----G-----	332
CI-47-1	-----A---AT-----G-----	338
CI-47-3	-----A---AT-----G-----	338
CI-47-8	-----A---AT-----G-----	341
CI-326-2	-----CA-----G-----	344
CI-326-3	-----CA-----G-----	344
CI-326-4	-----CA-----G-----	344
CI-327-1	-----TG-----T-----C-----	338
CI-327-2	-----TG-----T-----C-----	338
CI-327-3	-----G-----T-----C-----	338
CI-327-4	-----TG-----T-----C-----	338
CI-329-1	A-----AC-----T-----	338
CI-329-3	A-----A---AC-----T-----	338
CI-329-5	A-----AC-----T-----	338
CI-330-1	-----C---A---T-----A-----AC-----	332
CI-330-2	-----C---A---T-----A-----AC-----	332
CI-330-3	-----C---A---T-----A-----AC-----	332
CI-330-4	-----C---A---T-----A-----AC-----	332

CONSENSUS	ATTTTCTATTGCAATACATCAGAACTGTTTAATAGCACTTGGAAT...AACAAATA?CA?TGCC.....	383
CI-3-1	-----G-----A-A-----C-----...G-----GA-C-----	399
CI-3-2	-----G-----A-A-----C-----...G-----GA-C-----	399
CI-3-3	-----G-----A-A-----C-----...G-----GA-C-----	399
CI-45-3	-----A-----C-----T-----	393
CI-13-3	-----C-----A-A-----T-A-----G-ATAG-G-G-AT-T-A-A.....	396
CI-13-4	-----C-----A-A-----T-A-----G-ATAG-G-G-AT-T-A-A.....	396
CI-14-5	-----G-----T-----C-T-A-----	396
CI-14-13	-----G-----T-----C-T-A-----	396
CI-14-21	-----G-----T-----C-T-A-----	396
CI-20-1	-----AC-----GG-----GA-C-----	396
CI-20-2	-----AC-----GG-----GA-C-----	396
CI-20-3	-----AC-----GG-----GA-C-----	396
CI-20-5	-----AC-----GG-----GA-C-----	396
CI-21-1	-----T-C-----GC-A-----GGC-CTGTCGG-A-T-----	399
CI-21-2	-----T-C-----GC-A-----GGC-CTGTCGG-A-T-----	399
CI-21-4	-----T-C-----GC-A-----GGC-CTGTCGG-A-T-----	399
CI-22-3	G-----T-T-----C-----T-----T-ACG-TGG-CT-A-AA.....	402
CI-22-6	G-----T-T-----C-----T-----T-ACG-TGG-CT-A-AA.....	402
CI-22-10	G-----T-T-----C-----T-----T-ACG-TGG-CT-A-AA.....	402
CI-42-1	-----T-C-T-----	399
CI-42-2	-----T-C-T-----	399
CI-42-3	-----T-C-T-----	399
CI-42-4	-----T-C-T-----	399
CI-45-1	-----A-----C-A-----	393
CI-45-2	-----A-----A-AA-----C-A-----	393
CI-47-1	-----T-----A-----GA-C-----	399
CI-47-3	-----T-----A-----GA-C-----	399
CI-47-8	-----T-----A-----GA-C-----	402
CI-326-2	-----A-----T-----GAT-G-C-TTG-A-A-----	408
CI-326-3	-----A-----T-----GAT-G-C-TTG-A-A-----	408
CI-326-4	-----A-----T-----GAT-G-C-TTG-A-A-----	408
CI-327-1	-----T-----G-G...G-----G-C-----	399
CI-327-2	-----T-----G-G...G-----G-C-----	399
CI-327-3	-----T-----G-G...G-----G-C-----	399
CI-327-4	-----T-----G-G...G-----G-C-----	399
CI-329-1	-----GC-----G-A-AG-----	399
CI-329-3	-----C-----G-A-AG-----	399
CI-329-5	-----C-----GC-----G-G-A-AG-----	399
CI-330-1	-----GC-----T-----TGG--GG-C---ATCACG	399
CI-330-2	-----GC-----T-----TGG--GG-C---ATCACG	399
CI-330-3	-----GC-----T-----TGG--GG-C---ATCACG	399
CI-330-4	-----GC-----T-----TGG--GG-C---ATCACG	399

HIVC147

CONSENSUS	...AGCTCAAAT?ACACAGAGTCAAAT...GA?ACTATAACTCTCCAATGCAGATAAAGCAAATTGTAA	445
CI-3-1	...-A-----A-T--A-----C-----	463
CI-3-2	...-A-----A-----G-----GACACT-T-----	466
CI-3-3	...-A-----A-T--A-----C-----	463
CI-45-3	AG-T-----..-GC-----A-----	448
CI-13-3	GATAGT--A--C-TA--T-C-----A--G-----	448
CI-13-4	GATAGT--A--C-TA--T-C-----A--G-----	448
CI-14-5	G---CC-----..AGC-----	460
CI-14-13	G---CC-----..AGC-----	460
CI-14-21	G---CC-----..AGC-----	460
CI-20-1	T-----GA-----..T-----	460
CI-20-2	T-----GA-----..T-----	460
CI-20-3	T-----GA-----..T-----	460
CI-20-5	T-----GA-----..T-----	460
CI-21-1	AGGC-GG-GT--.....-A-A-----	454
CI-21-2	AGGC-GG-GT--.....-A-A-----	454
CI-21-4	AGGC-GG-GT--.....-A-A-----	454
CI-22-3	ATTGAAAG---G-----..C-C-TA---C-----A-----A-----	454
CI-22-6	ATTGAAAG---G-----..C-C-TA---C-----A-----A-----	454
CI-22-10	ATTGAAAG---G-----..C-C-TA---C-----A-----A-----	454
CI-42-1	G---G-G-----..GT-A-----C-----G-----	463
CI-42-2	G---G-G-----..GT-A-----C-----G-----	463
CI-42-3	G---G-G-----..GT-A-----C-----G-----	463
CI-42-4	G---G-G-----..GT-A-----C-----G-----	463
CI-45-1	G-G.....-GC-----A-----	448
CI-45-2	G.....-GC-----A-----	448
CI-47-1	A--T-----..T-A-----	463
CI-47-3	A--GT-----..T-A-----	463
CI-47-8	A--T-----..T-A-----	466
CI-326-2	AGT---AGTGTGAG-T-----..CC-----T-T-----A-----	475
CI-326-3	AGT---AGTGTGAG-T-----..CC-----T-T-----A-----	475
CI-326-4	AGT---AGTGTGAG-T-----..CC-----T-T-----A-----	475
CI-327-1	CC-----A-----..C-----A-----A-----	463
CI-327-2	CC-----A-----..C-----A-----A-----	463
CI-327-3	T---CC-----A-----..C-----A-----A-----	463
CI-327-4	CC-----A-----..C-----A-----A-----	463
CI-329-1	G-----A-----..C-----T-----	460
CI-329-3	A---TGACA-T-----..C-----T-----	460
CI-329-5	G-----A-----..C-----T-----	460
CI-330-1	CAG-A--T-GC-G---G---T---..AGT-----C-----A-----GC	466
CI-330-2	CAG-A--T-GC-G---G---T---..AGT-----C-----A-----GC	466
CI-330-3	CAG-A--T-GC-G---G---T---..AGT-----C-----A-----GC	466
CI-330-4	CAG-A--T-GC-G---G---T---..AGT-----C-----A-----GC	466

CONSENSUS	ATATGTGGCAGAGAGTAGGACAAGCAATGTATGCCCCCTCCCATCCAAGGAGTAATAAGGTGTGAATCAAA	515
CI-3-1	-A-----A-----A-----C-----T-----	533
CI-3-2	-----A-----A-----C-----T-----	536
CI-3-3	-A-----A-----A-----C-----T-----	533
CI-45-3	-----A-----A-----A-----A-----	518
CI-13-3	-C-----A-----TG-----A-----C-----A-----TC-----	518
CI-13-4	-C-----A-----TG-----A-----C-----A-----TC-----	518
CI-14-5	GA-----G-----A-----A-----T-----	530
CI-14-13	-----G-----A-----A-----T-----	530
CI-14-21	GA-----G-----A-----A-----T-----	530
CI-20-1	-C-----C-----A-----AA-----A-----G-----	530
CI-20-2	-C-----C-----A-----AA-----A-----A-----	530
CI-20-3	-C-----C-----A-----AA-----A-----G-----	530
CI-20-5	-C-----C-----A-----AA-----A-----G-----	530
CI-21-1	-----C-----A-----A-----A-----A-----	524
CI-21-2	-----C-----A-----A-----A-----A-----	524
CI-21-4	-----C-----A-----A-----A-----A-----	524
CI-22-3	-C-----GA-----A-----AG-----CA-----T-----A-----CTC-----	524
CI-22-6	-C-----GA-----A-----AG-----CA-----T-----A-----CTC-----	524
CI-22-10	-C-----GA-----A-----AG-----CA-----T-----A-----CTC-----	524
CI-42-1	-G-----	533
CI-42-2	-G-----	533
CI-42-3	-G-----	533
CI-42-4	-G-----	533
CI-45-1	-----A-----A-----A-----A-----	518
CI-45-2	-----A-----A-----A-----A-----A-----	518
CI-47-1	-----A-----G-----A-----A-----T-----	533
CI-47-3	-----A-----G-----A-----A-----T-----	533
CI-47-8	-----A-----G-----A-----A-----T-----	536
CI-326-2	-----T-----C-----C-----	545
CI-326-3	-----T-----C-----C-----	545
CI-326-4	-----T-----C-----C-----	545
CI-327-1	-----A-----A-----C-----A-----	533
CI-327-2	-----A-----A-----C-----A-----	533
CI-327-3	-----A-----A-----C-----A-----	533
CI-327-4	-----A-----A-----C-----A-----	533
CI-329-1	GA-----A-----C-----A-----C-----	530
CI-329-3	GA-----C-----A-----C-----	530
CI-329-5	GA-----A-----C-----A-----	530
CI-330-1	-----T-----A-----AC-----TG-----	536
CI-330-2	-----T-----A-----AC-----TG-----	536
CI-330-3	-----T-----A-----AC-----TG-----	536
CI-330-4	-----T-----A-----AC-----TG-----	536

CONSENSUS	CATTACAGGACTACTATTAACAAGAGATGGAGGGGAATAATAATAGT.....ACAAATGAGACCTTCAGG	579
CI-3-1	-----G-----G-----AT-----	597
CI-3-2	-----G-----G-----AT-----	600
CI-3-3	-----G-----G-----AT-----	597
CI-45-3	T-----G-----T-----	582
CI-13-3	T-----G-----GCT-G-----T-GTA-----CA-----GT-----A	582
CI-13-4	T-----G-----GCT-G-----T-GTA-----CA-----GT-----A	582
CI-14-5	-----A-----A-----AGT.....	597
CI-14-13	-----A-----A-----AGT.....	597
CI-14-21	-----A-----A-----AGT.....	597
CI-20-1	-----A-G-G-----AAT.....G-----A	597
CI-20-2	-----A-G-G-----AAT.....G-----A	597
CI-20-3	-----A-G-G-----AAT.....G-----A	597
CI-20-5	-----A-G-G-----AAT.....G-----A	597
CI-21-1	T-----A-----T-GG-----A-----A	588
CI-21-2	T-----A-----T-GG-----A-----A	588
CI-21-4	T-----A-----T-GG-----A-----A	588
CI-22-3	T-----G-G-----T-TG-----C-CG-ACAGC...-C-C-----A	591
CI-22-6	T-----G-G-----T-TG-----C-CG-ACAGC...-C-C-----A	591
CI-22-10	T-----G-G-----T-TG-----C-CG-ACAGC...-C-C-----A	591
CI-42-1	T-----T-C-----G-----GG-----A-----T-CC-----T-----	597
CI-42-2	T-----T-C-----G-----GG-----A-----T-CC-----T-----	597
CI-42-3	T-----T-C-----G-----GG-----A-----T-CC-----T-----	597
CI-42-4	T-----T-C-----G-----GG-----A-----T-CC-----T-----	597
CI-45-1	T-----G-----T-----	582
CI-45-2	T-----A-----G-----T-----	582
CI-47-1	---C-----A-AGT...-TG-----T-----	600
CI-47-3	---C-----A-AGT...-T-----T-----	600
CI-47-8	---C-----A-AGT...-TG-----T-----	603
CI-326-2	-----G-----T-----	609
CI-326-3	-----G-----T-----	609
CI-326-4	-----G-----T-----	609
CI-327-1	T---T-----G-A-----T-C-----	597
CI-327-2	T---T-----G-A-----T-C-----	597
CI-327-3	T---T-----G-A-----T-C-----	597
CI-327-4	T---T-----C---G-A-----T-C-----	597
CI-329-1	-----G-----A-AATAGTG-----T-----	600
CI-329-3	-----G-----A-AATAGTG-----T-----	600
CI-329-5	-----G-----A-AATAGTG-----T-----	600
CI-330-1	-----G-----GC-----T---GT-G-G-----GC-----G-----A	600
CI-330-2	-----G-----GC-----T---GT-G-G-----GC-----G-----A	600
CI-330-3	-----G-----GC-----T---GT-G-G-----GC-----G-----A	600
CI-330-4	-----G-----GC-----T---GT-G-G-----GC-----G-----A	600

CONSENSUS	CCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAGTATAAAGTAGTAAAAATTGAAC	649
CI-3-1	-----G-----	667
CI-3-2	-----G-----	670
CI-3-3	-----G-----	667
CI-45-3	-----A-----G-----	652
CI-13-3	-----A-----A-----G-C-----	652
CI-13-4	-----A-----A-----G-C-----	652
CI-14-5	-----C-----	667
CI-14-13	-----C-----	667
CI-14-21	-----C-----	667
CI-20-1	-----C-----	667
CI-20-2	-----C-----	667
CI-20-3	-----C-----	667
CI-20-5	-----C-----	667
CI-21-1	-----G-----G-----	658
CI-21-2	-----G-----G-----	658
CI-21-4	-----G-----G-----	658
CI-22-3	-----A-----A-----	661
CI-22-6	-----A-----A-----	661
CI-22-10	-----A-----A-----	661
CI-42-1	-----A-----	667
CI-42-2	-----A-----	667
CI-42-3	-----A-----	667
CI-42-4	-----A-----	667
CI-45-1	-----A-----G-----	652
CI-45-2	-----A-----G-----	652
CI-47-1	-----	670
CI-47-3	-----	670
CI-47-8	-----	673
CI-326-2	-----C-----G-----	679
CI-326-3	-----G-----C-----G-----	679
CI-326-4	-----C-G-----G-----	679
CI-327-1	-----	667
CI-327-2	-----	667
CI-327-3	-----	667
CI-327-4	-----	667
CI-329-1	-----C-----	670
CI-329-3	-----T-----G-----	670
CI-329-5	-----A-----T-----G-----	670
CI-330-1	-----A-----A-----A-----	670
CI-330-2	-----A-----A-----A-----	670
CI-330-3	-----A-----A-----A-----	670
CI-330-4	-----A-----A-----A-----	670

HIVCI47

CONSENSUS	CACTAGGTGTAGCACCCACCCATOCAAAAAGAAGAGTGGTGGAGAGAGAAAAAGAGCAGTTGGACTGGG	719
CI-3-1	--A-----A-G-----	737
CI-3-2	--A-----A-G-----	740
CI-3-3	--A-----A-G-----	737
CI-45-3	-----A-----	722
CI-13-3	-----A-----AGG-----G-----A-----A-A--T-A--	722
CI-13-4	-----A-----AGG-----G-----A-----A-A--T-A--	722
CI-14-5	-----T-----	737
CI-14-13	-----T-----	737
CI-14-21	-----T-----	737
CI-20-1	-----T-A-G-----G-T--	737
CI-20-2	-----T-A-G-----G-T--	737
CI-20-3	-----T-A-G-----G-T--	737
CI-20-5	-----T-A-G-----G-T--	737
CI-21-1	--T--A-----ATG-----G-----A-A--	728
CI-21-2	--T--A-----ATG-----G-----A-A--	728
CI-21-4	--T--A-----ATG-----G-----A-A--	728
CI-22-3	--T--A-----A-G-----G-----C-----G--T-A--	731
CI-22-6	--T--A-----A-G-----G-----C-----G--T-A--	731
CI-22-10	--T--A-----A-G-----G-----C-----G--T-A--	731
CI-42-1	-----C-----	737
CI-42-2	-----C-----	737
CI-42-3	-----C-----	737
CI-42-4	-----C-----	737
CI-45-1	-----A-----	722
CI-45-2	-----A-----	722
CI-47-1	-----	740
CI-47-3	-----	740
CI-47-8	-----	743
CI-326-2	-----C-----	749
CI-326-3	-----C-----	749
CI-326-4	-----C-----	749
CI-327-1	-----C-----C-----G--A-A--	737
CI-327-2	-----G--A-A--	737
CI-327-3	-----G--A-A--	737
CI-327-4	-----C-----G--A-A--	737
CI-329-1	-----A--	740
CI-329-3	-----A--	740
CI-329-5	-----A--	740
CI-330-1	-----A-----A-G-----G-----	740
CI-330-2	-----A-----A-G-----G-----	740
CI-330-3	-----A-----A-G-----G-----	740
CI-330-4	-----A-----A-G-----G-----	740

CONSENSUS	A.....GCTGTCTTCCTTGGGTTCTTAGGAGCAGCAGGAAGCACTATGGGCGCGGCGTCAATAACGCTG	783
CI-3-1	-----T-----	801
CI-3-2	-----T-----	804
CI-3-3	-----T-----	801
CI-45-3	-----C-----	786
CI-13-3	-----A-G-----G-----G-----A-----G-----	786
CI-13-4	-----A-G-----G-----G-----A-----G-----	786
CI-14-5	-----A-----	801
CI-14-13	-----A-----A-----T-----	801
CI-14-21	-----A-----	801
CI-20-1	-CTGGGA-A-----C-----	807
CI-20-2	-CTGGGA-A-----C-----	807
CI-20-3	-CTGGGA-A-----C-----	807
CI-20-5	-CTGGGA-A-----C-----	807
CI-21-1	-----C-----A-----	792
CI-21-2	-----C-----A-----	792
CI-21-4	-----C-----A-----	792
CI-22-3	-----T-G--A-----A-T--G-G-----	795
CI-22-6	-----T-G--A-----A-T--G-G-----	795
CI-22-10	-----T-G--A-----A-T--G-G-----	795
CI-42-1	-----	801
CI-42-2	-----	801
CI-42-3	-----	801
CI-42-4	-----	801
CI-45-1	-----C-----	786
CI-45-2	-----C-----	786
CI-47-1	-----GT-----	804
CI-47-3	-----TGT-----	804
CI-47-8	-----GT-----	807
CI-326-2	-----G-C-----	813
CI-326-3	-----T-C-----	813
CI-326-4	-----T-C-----	813
CI-327-1	-----A-----	801
CI-327-2	-----A-----	801
CI-327-3	-----	801
CI-327-4	-----A-----	801
CI-329-1	-----A-----A-----A-----	804
CI-329-3	-----A-----	804
CI-329-5	-----A-----	804
CI-330-1	-----T-----	804
CI-330-2	-----T-----	804
CI-330-3	-----T-----	804
CI-330-4	-----T-----	804

HIVCI47

CONSENSUS	ACGGTACAGGCCAGACAATTATTGCTCTGGTATAGTGCAACAGCAGAGCAATTTGCTGAGGGCTATAGAGG	853
CI-3-1	-----C-----A-----	871
CI-3-2	-----C-----A-----	874
CI-3-3	-----C-----A-----	871
CI-45-3	-----C-----	856
CI-13-3	-----TG--C-----A-A-----	856
CI-13-4	-----TG--C-----A-A-----	856
CI-14-5	-----A-----C--C-----A-----	871
CI-14-13	-----A-----C--C-----A-----	871
CI-14-21	-----A-----C--C-----A-----	871
CI-20-1	-----C-----	877
CI-20-2	-----C-----	877
CI-20-3	-----C-----	877
CI-20-5	-----C-----	877
CI-21-1	-----C-----A-----T-----	862
CI-21-2	-----C-----A-----T-----	862
CI-21-4	-----C-----A-----T-----	862
CI-22-3	-----T-----A-A-----A--T-----	865
CI-22-6	-----T-----A-A-----A--T-----	865
CI-22-10	-----T-----A-A-----A--T-----	865
CI-42-1	-----A-----	871
CI-42-2	-----A-----	871
CI-42-3	-----A-----	871
CI-42-4	-----A-----	871
CI-45-1	-----C-----	856
CI-45-2	-----C-----	856
CI-47-1	-----C-----A-----	874
CI-47-3	-----C-----A-----	874
CI-47-8	-----C-----A-----	877
CI-326-2	-----G--A-----C-----	883
CI-326-3	-----G--A-----C-----	883
CI-326-4	-----G--A-----C-----	883
CI-327-1	-----C-----	871
CI-327-2	-----C-----	871
CI-327-3	-----C-----	871
CI-327-4	-----C-----	871
CI-329-1	-----C-----A-----	874
CI-329-3	-----C-----G-----	874
CI-329-5	-----C-----G-----	874
CI-330-1	-----C-----C--A-----C-----	874
CI-330-2	-----C-----C--A-----C-----	874
CI-330-3	-----C-----C--A-----C-----	874
CI-330-4	-----C-----C--A-----C-----	874

CONSENSUS	CTCAACAACATCTGTTGAAACTCAG	879
CI-3-1	-----	897
CI-3-2	-----	900
CI-3-3	-----	897
CI-45-3	-----	882
CI-13-3	-G-----C-----	882
CI-13-4	-G-----C-----	882
CI-14-5	-----	897
CI-14-13	-----	897
CI-14-21	-----	897
CI-20-1	-----	903
CI-20-2	-----	903
CI-20-3	-----	903
CI-20-5	-----	903
CI-21-1	-----G-----	888
CI-21-2	-----G-----	888
CI-21-4	-----G-----	888
CI-22-3	-G-----C-----A	891
CI-22-6	-G-----C-----A	891
CI-22-10	-G-----C-----A	891
CI-42-1	-----A-----	897
CI-42-2	-----A-----	897
CI-42-3	-----A-----	897
CI-42-4	-----A-----	897
CI-45-1	-----	882
CI-45-2	-----	882
CI-47-1	-----	900
CI-47-3	-----	900
CI-47-8	-----	903
CI-326-2	-----A-----G-----	909
CI-326-3	A-----A-----G-----	909
CI-326-4	-----A-----G-----	909
CI-327-1	-----A-----	897
CI-327-2	-----A-----	897
CI-327-3	-----A-----	897
CI-327-4	-----A-----	897
CI-329-1	-----AT-----G-----	900
CI-329-3	-----	900
CI-329-5	-----	900
CI-330-1	-----G--A-----G-----	900
CI-330-2	-----G--A-----G-----	900
CI-330-3	-----G--A-----G-----	900
CI-330-4	-----G--A-----G-----	900

HIVD808

LOCUS HIVD808 1236 bp DNA VRL 15-APR-1994
 DEFINITION Human immunodeficiency virus type 1 sample D808, env gene, partial cds (gpi20).
 ACCESSION U07102
 SOURCE Human immunodeficiency virus type 1, sample D808, PCR amplification of uncultured PBMC proviral DNA.
 REFERENCE 1 (bases 1 to 1236)
 AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Mahambre,G., Delwart,E.L., Mullins,J.I. and Ruebsamen-Waigmann,H.
 TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses
 JOURNAL J. Virol. 68, 2161-2168 (1994)
 COMMENT The variants in this study were isolated from HIV-1/HIV-2 dually infected individuals. Phylogenetic analysis of the HIV-1 and HIV-2 sequences indicates a significant degree of conservation within the two HIV types. The HIV-1 variants cluster with HIV-1 subtype C strains and are highly similar to sequences isolated in East and South Africa. Grez. et al. propose that the low degree of diversity between isolates (of a given type) indicates a relatively brief period of HIV-1 infection in India. Other sequences in this set are included in the alignment below (accession numbers U06716, U07098-U07101, and U07103).
 NCBI gi: 47174

BASE COUNT 489 a 211 c 267 g 305 t

ORIGIN 310 bp downstream from the beginning of env cds

1 atcagtttat gggatccaag cctaaagcca tgtgtaaagt tgaccccact ctgtgtcact
 env ->

61 ttagaatgta gaaatggttag cagtaatggt aatgggtacct acaatgagac ctacaatgag
 121 agcgtgaagg aaataaaaaa ttgctctttc aatgcaacca cagaagtaag agataggaag
 181 cagacagtgt atgcactttt ttatagactt gatatagtac catttaataa gaagaactct
 241 agtgagtatt atagattaat aaattgtaat acctcagcca taacacaagc ctgtccaaag
 301 gtcacttttg atccaattcc tatacactat tgcaactccag ctggttatgc gattctaaag
 361 tgtaatgata agacattcaa tgggacagga ccattgccaga atgtcaggac agtacagttt
 421 acacatggca ttaagccagt ggtatcaact caactactgt taatggtag ctacgcagaa
 481 agagagataa taattagatc tgaaaatctg acaaacatg tcaaaaceat aatagtacat
 541 ctttatcaat ctgtagaat tgtgtgtaca agacccaaca ataatacaag aaaaagtaca

V3 loop ->

601 aggataggac caggacaaac attctatgca acgggagaca taataggaga cataagacaa
 661 gcacattgta acattagtga agataaatgg aatgaaactt tacaagggt aagtaaaaaa

<- V3 loop

721 ttagcagaac acttcocctaa taaaacaata aaatttgcac cactctcagg aggggaccta
 781 gaaattacaa cacatagctt taattgtaga ggagaatttt totattgtaa tacatcaggc
 841 ctgtttaatg gtacatacat gcttaattgt acaaaaagta attcaagctc aatcatcaca
 901 atccccatgca gaataaagca aattataaac atgtggcagg aggtaggacg agcaatgtat
 961 gccctcccca ttgaaggaaa cataacatgt aaatcaata tcacaggact actattggta
 1021 cgtgatggag gaacagaatc aaataataca gagacattca gacctggagg aggagatatg
 1081 aggaacantt ggagaagtga attatataaa tataaagtgg tagaaattaa gccattggga
 1141 gtgcaccca ctacagcaaa aaggagagtg gtggagagag aaaaaagagc agtgggaata
 1201 ggagctgtgt tccctggggt cttgggagca gcaggaagca ctatgggcgc ggcgtcaata
 1261 acgctgacgg ta

env ->

CONSENSUS ATGAGAGTGAGGGGGATACTGAGGAATTATCAACAATGGGGGATATGGGGCATCTTAGGCTTTTGGATGT 70
 D1024 ----- 70
 CONSENSUS TAATGATTTGTAATGTGGTAGGAAACTTGTGGGTACAGTCTATTATGGGGTACCTGTGTGGAAAGAAGC 140
 D1024 ----- 140

CONSENSUS	AAAACTACTCTATTCTGTGCATCAGATGCTAAAGCATATGAGGAAGAAGTGCATAATGTCTGGGCTACA	210
D1024	-----	210
CONSENSUS	CATGCCTGTGTACCCACAGACCCCAACCCACAAGAAATGGTTTTGGAAATGTAACAGAAAATTTTAACA	280
D1024	-----	280
CONSENSUS	TGGGATC?AA?CCT?AAGCC	17
D1024	TGTGGAAAAATGACATGGTGAATCAGATGCATGAGGATCTAATCAGTCTA-----A--G--A----	350
D744	CGTGTA-----A--C--A----	26
D808	ATCAGTTTA-----C--G--A----	29
D868	-----C--G--G----	20
CONSENSUS	ATGTGTAAAGTTGACCCCACTCTGTGTCTACTTTAGAATGTAGA?A?GTTA.....??G?T	67
D1024	-----C-----G--A--G-----	400
D1044	-----T-----C--G--C-----	37
D744	-----A-----	66
D766	-----A--T-----CCAGTAAT--G--	39
D808	-----A--T-----GCAGTAAT--G--	90
D868	-----G--G-----ATGATACTAGAAATGGG--T--	90
NOF	-----A--T-----C--G--T--A-----AA--T--	48
CONSENSUS	AATG?TACCAATAAT.....?C?A?CAAT?A?AGGG??...GAAATAAAAAATTCCTCTTTCAATG	117
D1024	---C-----T-----C---A--TG---AA...-----T-----	454
D1044	---TT-----GC---ACCAATG--T--A---GCT--C--AA...-----G-----A-----	100
D744	G---C-----G-----G--T--C--G--AGTG---AA...---G--G-----	124
D766	---G---T--C---...GAGA--CTA---G--G--C--TGAAG--G-----	103
D808	---G---T--C---...GAGA--CTA---G--G--C--TGAAG-----	154
D868	---T-----GGGGAA-----	136
NOF	---C---G---C-----GT---A--C--T--ACTGGA-----T--A	106
CONSENSUS	CAACCACAGAAATAAGAGATAGGAAGCAGACAGTGTATGCACCTTTTTATAGACTTGATATAGTACCCT	187
D1024	-----G-----	524
D1044	-----A---C-----A-----	170
D744	-----N-----	194
D766	-----T-----G-----C-----CC-----T-----	173
D808	-----G-----	224
D868	-----	206
NOF	-----G---A---A---A-----A--C---G-----	176
CONSENSUS	TGATAATAAG.....AACTCTAGT?AG...TATTATAGATTAAATAAATTGTAATACCTCAGCCATA	244
D1024	A---G---C-----A--A--AGCA--G-----	579
D1044	---G---AGTAAGAACG-----AG---A-----	237
D744	---G---G--AATGAGAGA-----A---	261
D766	-A---G---A-G-TAGT-----	234
D808	-A---G-----G-----	282
D868	---AAGAAGAAC---G---	273
NOF	-A--G--G--G-----A--C---T-----C---C-----A-----	222
CONSENSUS	ACACAAGCCTGTCCAAAGGTCACCTTTGATCCAATTCCCTATACACTATTGTACTCCAGCTGGTTATGCOA	314
D1024	-----G-----	649
D1044	-----C--C-----C--C-----	307
D744	-----	331
D766	-----G-----	304
D808	-----C-----	352
D868	-----C-----G-----T-----	343
NOF	-----C-----T-----C-----T-----G-----	292

HIVD808

CONSENSUS	TTCTAAAGTGTAATAATAAGACATTCAATGGGACAGGACCATGCCATAATGTCAGCACAGTACAATGTAC	384
D1024	-----G-----	719
D1044	-----C-----	377
D744	-----	401
D766	-----G-----G-----T-----	374
D808	-----G-----G-----G-----G-T-----	422
D868	--A-----AG-----	413
NOF	-----A-----	362
CONSENSUS	ACATGGGAATTAAGCCAGTGGTATCAACTCAACTACTGTAAATGGTAGCCTAGCAGAAGGAGAGATAATA	454
D1024	-----	789
D1044	-----GT-----	447
D744	-----G-----T-----G-----	471
D766	-----A-----A-----	444
D808	-----C-----A-----	492
D868	-----G-----A-G-----	483
NOF	-----T-----G-T-----	432
CONSENSUS	ATTAGATCTGAAAATCTGACAAACAATGTCAAAACAATAATAGTACATCTTAATCAATCTGTAGAAATTG	524
D1024	-----T-----T-----C-----	859
D1044	-----C-----C-----G-A-----	517
D744	-----A-----	541
D766	-----C-----	514
D808	-----T-----	562
D868	-----C-----T-----G-----C-----	553
NOF	-----A-----G-----C-C-----A-----T-----	502
CONSENSUS	TGTGTACAAGACCCAAACAATAACAAAGAAAAAGTATAAGGATAGGACCAGGACAAACATTCTATGCAAC	594
D1024	-----G-----	929
D1044	---C-----	587
D744	A-----G-----	611
D766	-----C-----	584
D808	-----C-----	632
D868	--C-----A-----	623
NOF	-----G-----G-----G-----	572
CONSENSUS	AGGAGACATAATAGGAGACATAAGACAAGCACATTGTAACATTAGTAAAGATAAATGGAATGAAACTTTA	664
D1024	-----G-----T-----G-----C-----G-----	999
D1044	-----T-----	657
D744	---A-----G-----G-----C-----C-----G-----	681
D766	--G-----G-----	654
D808	-----G-----	702
D868	-----G-----T-----G-----A-C-----	693
NOF	-AAT-C-----A-GCT-----	642
CONSENSUS	CAAAGGGTAAGTAAAAAATTAGCAGAACACTTCCCTAATAAAACAATAAAATTTGCATCATCTTCAGGAG	734
D1024	-----G-----G-----C-----	1069
D1044	---T---G-A-----G-C-----C-----	727
D744	-----G-----	751
D766	-----AA-----G-----G-----	724
D808	-----	772
D868	-----G-----T-----C-----	763
NOF	G-CA---G-AG-----A-----GC---AAC-----	712

CONSENSUS	GGGACCTAGAAATTACAACACATAGCTTTAATTGTAGAGGAGAAATTTTCTATTGCAATACATCAAACCT	804
D1024	-----C-----	1139
D1044	-----T-----	797
D744	-----	821
D766	-----T-----GG--	794
D808	-----T-----GG--	842
D868	-----	833
NOF	---TT---G-----A--G---	782
CONSENSUS	GTTTAATGTACATAC.....ATGCTTAATGGTACA.....	833
D1024	-----A-----C-----	1170
D1044	-----A-----AATAGTACATAC---T-----	840
D744	-----A--G-----C-----TATATGTTTAATGGTACATACATGTTT	879
D766	-----G-----A-----	825
D808	-----G-----C-----	873
D868	-----G-----	861
NOF	-----TCT-C-A-----CGCAACAAATTTAAT	810
CONSENSUS	AAAAGTAATTCAA?CTCAACCATCACAAATCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGGAGG	902
D1024	-----A--C-----T--T--T-----G-----	1240
D1044	-C-----C--T--A-----	910
D744	--T---CAAGT-ATG-----T-----	949
D766	-----G--T-----G-----	895
D808	-----G--T-----G-----	943
D868	-----G-----T-----G--G-----	931
NOF	G---G--C--A-G-----C-----G-----	880
CONSENSUS	TAGGACGAGCA...ATGTATGCCCCCTCCCATTTGAAGGAAACATAACATGTAAATCAAATATCACAGGACT	969
D1024	-----	1307
D1044	-----C-----	977
D744	-----	1016
D766	-----C-----	962
D808	-----	1010
D868	-----	998
NOF	-----GA-GAGC-----C-----	950
CONSENSUS	ACTATTGGTACGTGATGGAGGAACAGA?.....TCAAATAATACA...GAGATA.....TTCAGACCT	1023
D1024	-----G-----G-----C-----C-----	1362
D1044	-----A-----C-----C-----	1029
D744	-----CG--GACAAGAG-----C-----	1077
D766	-----G-----C-GAGACA-----	1023
D808	-----A-----C-----	1065
D868	-----C-----G-----C-----	1053
NOF	-----AC-----AC-CAACACAA-----CACA-----	1014
CONSENSUS	GGAGGAGGAGATATGAGGAACAATTGCAGAACTGAATTATATAAATATAAAGTGGCTAGAAATTAAGCCAT	1093
D1024	-----	1432
D1044	-----G-----	1099
D744	-----C-----C-----	1147
D766	-----G-----	1093
D808	-----	1135
D868	-----C-----	1123
NOF	-----A-----A-G-----A-----	1084

HIVD808

CONSENSUS	TGGGAGTAGCACCCACTGCAGCAAAAAGGAGAGTGGTGGAGAGAGAAAAAGAGCAGTGGGAATAGGAGC	1163
D1024	-----	1502
D1044	-----A-C-----	1169
D744	-----G-----C-----	1217
D766	-----A-----T-----	1163
D808	-----A-----	1205
D868	-----T-----A-----G-----	1193
NOF	-----A-----G-T-----	1154
CONSENSUS	TGTGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCGGCGTCAATAACGCTGACGGTACAG	1233
D1024	---A-----	1569
D1044	-----	1236
D744	-----	1254
D766	-----1200-----	
D808	-----	1272
D868	-----C-----	1230
NOF	---C-----C-----	1224
CONSENSUS	GCCAGACAACCTGTTGTCTGGTATAGTGCAACAGCAAAGCAATTTGCTGAGAGCTATAGAGGCGCAACAGC	1303
NOF	-----	1294
CONSENSUS	ATATGTTGCAACTCACGGTCTGGGGCATTAAAGCAGCTCCAGACAAGAGTCTTGGCTATAGAAAGATACCT	1373
NOF	-----	1364
CONSENSUS	AAAGGATCAACAGCT	1388
NOF	-----	1379

LOCUS HIVDJ2581 1459 bp ss-DNA VRL 27-MAY-1993
 DEFINITION Human immunodeficiency virus type 1 (HIV-1) partial gag sequence, isolate DJ258.
 ACCESSION L11763
 SOURCE Human immunodeficiency virus type 1 cocultivated PCR-amplified PBMC DNA derived from a patient from Djibouti.
 REFERENCE 1 (bases 1 to 1459)
 AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Buell, E., Eddy, G., van der Groen, G., Franssen, K., Gershy-Damet, M., DeLays, R. and Burke, D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT 55 viral sequences were obtained from patients in fifteen countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zambia, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. These sequences were compared to another 15 previously sequenced isolates through phylogenetic analysis. With the inclusion of new gag sequences, Louwagie et al. report the emergence of three new HIV-1 genotypes, E-G, and the existence of an unclassified Belgian isolate, VI325. The env region of DJ258 has also been sequenced (L22939). This gag sequence (previously published on page I-A-194 AUG 93) groups with HIV1 A subtype gag sequences.

BASE COUNT 525 a 281 c 373 g 280 t

ORIGIN gag cds start

1 ATGggtgcga gagcgctcagt attaagtggg ggaaatttag attcatggga gaagattcgg
 gag start ->

```

61 ttaagggccag ggggaagaa aatatataga ctaaacatc tagtatggc gagcagggag
121 ctggaagat tgcacttaa cctggcctt ttagaacag cagaaggatg tcaacagtta
181 atggaacagt tacaatcagc tctcgggaca gaataaaaag aacttaaatc attatataat
241 accatagcag tcttttggtg cgtacatcaa aggatagaca taaaagacac caaggaagcc
301 ttagataaaa tagaggcagt acaaaataag agcaagcaaa agacacagca ggcagcggct
361 gccacaggaa gcagcagcca aaattaccct atagtgcaaa atgcacaggg gcaaatggta
421 catcagccca tgtcacctag gactttgaat gcattgggtga aggtaataga ggaagaggct
481 ttcagcccaa aagtaatacc catgtttaca gcattatcag agggagccac cccacaagat
541 ctaaatatga tgctaaacat agtgggggga caccaggcag caatgcagat gttaaaagat
601 accatcaatg aggaagctgc agaattgggac agggtagatc cagtacatgc agggcctatt
661 ccaccaggcc aatgagggga accaagggga agtgacatag caggaactac tagtaccctt
721 caagaacaaa taggatggat gacaagcaat ccacctatcc cagtgggaga catctataaa
781 agatggataa tcttgggatt aaataaaata gtaagaatgt atagccctgt cagcattttg
841 gacataagac aagggccaaa agaacccttt agagattatg tagataggtt ctttaaaact
901 ttgagagctg aacaagctac gcaggaggta aaaaactgga tgacagaaac cttgctggtc
961 caaatgcga atccagactg taagtccatc ttaagagcat taggaccagg ggctacatta
1021 gaggaatga tgacatcatg tcaggggagt ggaggacctg gccataaagc aagggttttg
1081 gctgagggca tgagtcaggt acaacaatcc aacataatga tgcagagagg cgattttagg
1141 ggccagagaa caataaagtg tttcaactgt ggcaagaag gacacctagc cagaatttgc
1201 aaggctccta gaaaaagggt ctgttggaaa tgtggaaagg agggacacca aatgaaagac
1261 tgtactgaga gacaggctaa ttttttaggg aaattttggc cttccaccaa ggggaggcca
1321 ggaattttc ctcaaagcag accagagcca acagcccccac tagcagagag ctttgggatg
1381 ggggaagaga taacctcccc tccgaagcag gaaccgaggg aacagggaact ataccctccc
1441 ttaacttccc tcaaatcac

```

gag ->

HIVDJ2582

LOCUS HIVDJ2582 3215 bp ds-DNA VRL 31-DEC-1993
DEFINITION Human immunodeficiency virus type 1 (HIV-1) env, vpU, and nef sequences, isolate DJ258
ACCESSION L22939
SOURCE Human immunodeficiency virus type 1, sample DJ258, PCR-amplified PBMC
REFERENCE 1 (bases 1 to 3215)
AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P. A., van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.
TITLE Genetic diversity of the HIV-1 envelope glycoprotein
JOURNAL Unpublished (1994)
COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 cds from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences that were published by Louwagie et. al. in AIDS 7, 769-780 (1993) and are represented in this compendium. This sequence and a corresponding gag sequence (L11763) cluster with A subtype sequences. See also accession numbers L22940-L22957 and L23064-L23065.
NCBI gi: 437342
BASE COUNT 1129 a 561 c 757 g 768 t
ORIGIN 85 bp upstream from the beginning of rev cds start
1 aactgcttgt agcaagtgtt actgtaaaaa atgctgctgg cattgccaat tgtgctttct
61 gaacaagggc ttaggcattt cctATGgcag gaagaagcgg agacgccgac gaggaactcc
rev cds start ->
121 tcagagtcgt caggaccatc aaaatcctgt accaaagcaG Tgagtagtaa taattaatat
5' sj /\n
181 atgtgATGaa ctcttttagaa atagctgcaa tagtaggact agtagtagca ttcacagcag
vpU cds start ->
241 ccatagttgt gtggaccata gtattaattg aatataggaa aataaggaaa cagaagaaaa
301 tagacaagtt gcttgataga ataagagaaa gagcagaaga tagtggcaAT Gagagtgatg
env cds start ->
361 gggatacaga agaattatcc actcttatgg agatggggta cgataatatt ttgaaTAATg
<- vpU
421 ataatttcta atgctaaaga gatgtgggtc acggtctact atggggatcc tgtgtggaga
481 gagcacaaga ccacctatt ttgtgcatca gatgctaaag catatgatac agaagtcacat
541 aatgtcttgg ctacacatgc ctgtgtaccc acagacccta accacaaga aatatatttg
601 aaaaatgtaa cagaaaagt ttacatgttg aaaaataaca tggtagaaca gatgcatgaa
661 gatataatta gtctatggga ccaaagccta aagccatgtg taaagttaac cctctcttgc
721 gttacttttag agtgtcataa tgtcagcaat agcaatacct ctagctacag gaacatatct
781 agtgaatgc gagaggaaat aaaaaactgc tctttcaata tgaccacaga actaagagat
841 aagcaacaga aaatgtctgc acttttttat agacttgatg tagtacaat tagtgaaagt
901 aacaatagtc aatataggtt aatcaactgt aatadctcag ccattacaca ggcattgtcc
961 aaggttaacct ttgagccaat tcccatatat tattgtgccc cagctgggtt tgcgattcta
1021 aagtgtaatg ataagaagt caatggaaac gggccatgca agaattgttag cacagtacaa
1081 tgcacacatg gaatcaagcc agcagtatca actcaactac tgttaaatgg tagtctagca
1141 gaaggagagg tagtgattag atctgaaaat atcacaaaa atgccccaaa cataatagta
1201 cagttgggta cgctgttaag aatcaattgt agcagacctg gcaacaatac aagaaaaagt
v3 loop ->
1261 gtacgtatag gaccagggca aacattctat gcaacaggtg acataatagg gcatataaga
1321 caagcacact gtaattgtcag tagatcaaaa tggaaatgaaa ctttacaaaa ggtggctaca
<- v3 loop
1381 caattaagga agcactttta cacaacaata atctttgcta actcctcagg aggggatata
1441 gaaattacaa cacatagttt taattgtaga ggagaatttt tctattgcaa tacatcaagt
1501 ctgtttaata gcacttgga tagcagctca attcacacaa attacagctc aaatgacaca

```

1561 gacaatataa ctctccaatg cagaataaag caaattgtaa atatgtggca gaaagttagga
1621 aaagcaatat atgccccccc catctcagga gaaataagggt gtgaatcana cattacagga
1681 ctaatatata caagagatgg agggagtaat aatagtacaa atgagatctt caggcctgga
1741 ggaggagata tgaggagataa ttggagaagt gaattatata agtataaagt agtaaaaaatt
1801 gaaccactag gtgtagcacc caccogtgc aaaaagaagag tggtagagag agaaaaaaga
1861 gcagttggac tgggagcttt ctctctggg ttcttaggag ctgcaggag cactatgggc
1921 gcggcgctca taacgctgac ggtacaggcc agacaattat tgtccggtat agtgcaacag
1981 cagagcaatt tgcagagggc tatagaggct cacaacatc tgttgaacct cagggtctgg
2041 ggcattaaac agctccaggc aagagtcctg gctctggaaa gttacctaa ggatcaacag
2101 ctcttaggaa ttctggggctg ctctggaaaa ctcatctgca ccaactactgt accctgggac
2161 tctagttgga gtaataaaac ttttaatgac atatgggata acatgacctg gctgcaatgg
2221 gataaagaaa ttagcaatta cacagacata atatatactc taattgaaga atcgcaaac
2281 cagcaggaaa agaataaaca agacttattg gcattggaca aatgggcagg tctgtggagt
2341 tggtttggca taacaaactg gctgtggtat ataagaatat ttataataa agtaggaggt
2401 ttaatagggt taagaatagt ttttgctgtg ctactataa taaatagagt taggcaggga
2461 tactcacctt tatcgttccA Gaccttaac caccactgga gggaaccacc cgacagaccc
      3' sj (tat, rev) /\
2521 gaaagaatcg aagaaggagg tggcgagcaa gacaaagaca gatcaatacg ctctgatgagc
2581 ggattcttag cacttgccctg ggacgatctg cggagcctgt gctctctcag ctaccacca
2641 ttgagagact ttgtcttgat tgcagttagg gctgtggaac ttctgggaca cagcagcctc
2701 aaggggactga gactgggggtg ggaagccctc cagtatctgg ggaatcttct actatactgg
2761 ggtcagggaac taagaaTAG tgcgtttaat ttgcttgata caatagcaat agtagtagct
      <- rev cds end
2821 aactggacag acagagttat agaaatagga caagaagtgg gtagagctat tctcaacata
2881 cctagaagaa tcagacaggg gcttgaaagg gctttgttaT AAcATGggtg gcaagtggtc
      <- env cds end
      nef cds start ->
2941 aaaaagcagc atagtgggat ggctaagggt tagggaaaga ataagacaaa cccctccagc
3001 agcaacagga gtaggagcag catctcaaga tttagataga catggagcaa ttacaagtag
3061 taatacagca gctactaatg ctgattgtgc ctggctagaa gcacaagaag aagaggaggt
3121 aggcctttcca gtcaggccgc aggtaccttt gagaccaatg acttataagg cagctttcga
3181 tctcggtctc tttttaaaag aaaagggggg actgg
      3' LTR ->

```

HIVDJ2591

LOCUS HIVDJ2591 1465 bp ss-DNA VRL 27-MAY-1993
 DEFINITION Human immunodeficiency virus type 1, isolate DJ259 from Djibouti, gag region.
 ACCESSION L11764
 SOURCE Human immunodeficiency virus type 1 (HIV-1) blood DNA, Djibouti isolate DJ259 (cocultivation on PBMCs).
 REFERENCE 1 (bases 1 to 1465)
 AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Buell, E., Eddy, G., van der Groen, G., Franssen, K., Gershy-Damet, M., DeLays, R. and Burke, D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Viral isolates were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate has also been sequenced (L22940). This gag sequence (previously published on page I-A-195 AUG 93) groups with HIV1 C subtype gag sequences.
 BASE COUNT 542 a 293 c 352 g 278 t
 ORIGIN gag cds start

1 ATGggtgcga gagcgtcaat attaagagggc gaaaaattag atgcctggga aaaaattaag
 gag start ->
 61 ctaaggccag gaggaaggaa acgtatatg ctaaaacacc tagtctgggc aagcagggag
 121 ctggaaaaat ttgcacttaa ccttagcctt ttagagacat cagaaggctg taaacaata
 181 ataagacagc tacaaccagc tcttcagaca ggaacagaag aacttagatc attatataat
 241 acagtagcaa ctctctattg tgtacatgca aggatagaga tacaagacac caaagaagca
 301 ctgacaaga tagaggaaga acaaaacaaa tgtcagcaaa aaacacagaa ggcagaaaca
 361 gctgacaagg gaaggtcag tcaaaatttt cctatagtag agaattctca agggcaaatg
 421 gtacaccagc cctatcacc tagaactttg aatgcatggg taaaagtaat agaggagaag
 481 gctttcagcc cagaagtaat acccatgttt acagcattat cagaaggagc caccacaaa
 541 gatitaaaca ccatgttaaa tacagtgggg ggacatcaag cagccatgca aatgttaaaa
 601 gataccatca atgaggaggc tgcagaatgg gataggttac atccagtgca tgcagggcct
 661 attgcaccag gccaaatgag agaaccaagg ggaagtgaca tagcaggaac tactagtaac
 721 cttcaggaac aaatagcatg gatgacaagt aaccccccta ttccagtggg agacatctat
 781 aaaaagatgga taattctggg gttaaataaa atagtaagaa tgtatagccc tgtcagcatt
 841 ttggacataa gacaagggcc aaaggaaact tttagagact atgtagatcg gttcttttaa
 901 accttaagag ctgaacaagc atctcaagat gtaaaaaatt ggatgacaga caccttggtg
 961 gtccaaaatg cgaacccaga ttgtaagacc attttaagag cattagggcc aggggcttca
 1021 ctagaagaaa tgatgacagc atgtcagggg gtgggaggac ctagccacaa agcaagagtg
 1081 ttggctgagg caatgagcca agcaaccaac acaaacataa tgatgcagaa aagcaatttt
 1141 aagggcccta gaaaaactat taaatgtttc aactgtggca aggaagggca cctagccaga
 1201 aattgcaggg cccttaggaa aaagggtgtg tggaaatgtg gaagggaagg acaccaaatg
 1261 aaagactgta ctgagagaca ggctaatttt ttagggaaaa tttggccttc caacaagggg
 1321 aggccaggga atttccttca gacagacact gagccaacag cccaccagc agagagcttc
 1381 ggggttcgagg agccgacccc cctcagaag caggagctaa aagacaagga accctacaag
 1441 gagcccttaa ctgcctcag atcac
 gag ->

LOCUS HIVDJ2592 3234 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, DJ259 from Djibouti, proviral DNA encoding env, tat, vpU, rev, and nef

ACCESSION L22940

SOURCE Human immunodeficiency virus type 1 (HIV-1), DJ259 from Djibouti, PCR amplification of proviral DNA from cocultivated PBMC's

REFERENCE 1 (bases 1 to 3234)

AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1994)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences that were published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 Human Retroviruses and AIDS compendium. This sequence and the corresponding gag sequence cluster with C subtype sequences. See also accession numbers L22939, L22941-L22957 and L23064-L23065.

NCBI gi: 437343

BASE COUNT 1134 a 558 c 786 g 756 t

ORIGIN 144 bp upstream from the beginning of rev cds start

1 ggcttcogtc tagatagatc ctagactaga gccctgggac catccaggaa gtcagcctaa
61 aactgcttgt aataagtgtt attgtaaaaa atgtagctat cattgtctag ttgtctttca
121 gacaaaaggt ttaggcattt cctATGgcag gaagaagcgg agacagcgac gaagcgctcc
rev cds start ->
181 tccaagcagt gaggatcatc agactcctat atcaaagcaG Taagtacaat agtagatgta
5' sj /\

241 ATGatagatt taccagcaaa agtagactat agattagcag taggagcatt aatagtagca
vpU start ->
301 ctaatcatag caatagtgtt gtggaccata gtatatatag aatataggaa attgttaaga
361 caaagacaaa tagactgggt aattgaaaga attagggaaa gagcagaaga cagtggcaAT
env cds start ->
421 Gaagtgatg gggacacaga ggaactggcc acggtggtgg gtatggggaa tcttaggctt
481 ttggatgttg ttgatttgTA Atgggaactt gtgggtcaca gtctattatg gggtacctgt
<- vpU end

541 gtggaaagat gcaaaccttc cctatttttg tgcacagat gctaaagcat ataaagtaga
601 agtacataat gtctgggcta cacatgcctg tgtcccaca gaccccagcc cacaggaaact
661 agctttggaa aatgtaacag aaaattttta catgtggaaa aatggcatgg ttgatcagat
721 gcatcaggat ataactcagtc tatgggatca aagcctaaag ccatgtgtaa agttgaccoc
781 actttgtgtt actttaaact gtagtaattg tcagagtaat gctacctaag agaatagcac
841 aaaagaagaa atgaaaattt gttctttcaa tgaaccaca gaactaagag ataagagaca
901 gaaagcatat gcacttttct atagacttga tatagtacca cttaatgaga ataatagtga
961 atatagatta ataaactgta atacctcagc cataacacaa gcctgtccaa aggtctcttt
1021 ggacccaatt cctatacatt attgtgctcc agctgggtat gogattctaa agtgtaggga
1081 taagaacttc aatggaaacag gccctgcaa taatgtcagc acagtacast gtacacatgg
1141 aattaaacca gtggtatcaa ctcaactatt gttaaatggg agtctagcag aagatgatat
1201 aataattaga tctgaaatc tgacaaccaa tgccaaata ataatagtac agcttaatga
1261 atctgtagca attaaactgta cgagacccaa caataatata agagaaagta taaggatagg
v3 loop ->
1321 accaggacaa acattctatg caacaggaga cataatagga gatataagac aagcacattg
1381 taacattagt agaaaagaat ggaatgaac ttttaagaag gtaagagaaa aattaanaac
<- v3 loop

1441 gctcttcaga actgcaaata taacatttca accatcctca ggaggggacc tagaaattac
1501 aacacatagc ttttaattgtg gaggagaatt tttctactgc aatacatcag acctgtttta


```

1561 taatcagagt atcacaaatg aaaacatcac acttcaatgc agaataaaac aaattataaa
1621 catgtggcag gaggtaggac gagcaatgta tggccctccc attaaaggaa taataacatg
1681 taaatcaaat gtccacaggac tgctactgac acgtgatgga gggacaaata ccacaacaga
1741 aggagagata ttcagacctg gaggaggaga tatgaggggac aattggagaa gtgaattata
1801 taaatacaaa gtggtagaaa ttaagccagt aggaatagca cccactaagg caaaaaggag
1861 agtggtaggg agagaaaaaa gagcagcact agggagctatg ttcccttgggt tcttggggagc
1921 agccggaagc actatgggag cagcatcaat aacgctgacg gtacaggcca gacaattatt
1981 gtctggtata gtgcaacagc aaagcaatct gctgagggct atagaggcgc agcagcatat
2041 gttgcaactc acggtctggg gcattaaagca gctccagaca agagtccctg ctatagaaag
2101 ataccttaag gatcaacagc ttctagggat ttggggctgc tctggaaaac tcctctgcac
2161 cactaatgtg ccttgggaacg ctagttaggag taataggtct caaaaggaga tttgggataa
2221 catgacctgg atacagtggg agagagaaat taataattat acaggactaa tatacgggtt
2281 gcttgaagac tcgcaaaaac agcaggagaa aaatgaacaa gatttactag cattggacaa
2341 atggcagaat ctgtggaatt ggttttagcat aacaaagtgg ctgtggtata taaaaatatt
2401 cataatgata gtaggaggct tgataggttt aagaataatt ttgtctgtgc tctctatagt
2461 gaatagagtt aggcaggggt actcacggct atcgtttcAG acccttacc caaacccgag
      3' sj (tat, rev) /\
2521 gggacccgac aggcccgag gaatcgaaga agaagggtga gagcaagaca gaggcagatc
2581 gattcgatta gtgagcggat tcttagcgct tgccctgggac gacctgogga gcctgtgcct
2641 cttcagctac cgcgattga gagacttaat cttgattgca gcgagagcgg tggaaactgct
2701 gggacagagg gggtagggaaa cctttaagta tctgggaagc cttgtgcagt attggggctct
2761 ggagctaaaa aagagtgcga ttagtctgct taataccaca gcaataacag tagctgaagg
2821 aacagatagg attataaaat tcatacaagg gctttggaga gctatctgca acatacctag
2881 aagaattaga cagggctttg aagcagcttt gcaatAaaAT Ggggagcacg tgggtcaaaa
      <- env
      nef cds start ->
2941 gtagtccagt gggatggcct gctgtaagag aaagaataag gcaaaactgat ccagcagcaa
3001 ggcaaaactga ttcagcaaca aagggagtag gagcagcgtc tcaagactta gacaaatag
3061 gggcacttac aagcagcaac acaccagcca ataatgctga ttgtgcttgg ctggcagcgc
3121 aagaggagga aggagaagta ggccttccag taagacctca ggtaccttta agaccaatga
3181 cttataaggg agcattcgat ctgggcttct ttttaaaaga aaagggggga ctgg
      3' LTR ->

```

LOCUS HIVDJ263A 3292 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, DJ263 from Djibouti, proviral DNA encoding env, tat, vpU, rev, and nef genes.

ACCESSION L22941

SOURCE Human immunodeficiency virus type 1 (HIV-1), DJ263 from Djibouti, PCR amplification of proviral DNA from cocultivated PBMC's

REFERENCE 1 (bases 1 to 3292)

AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,F.A., van der Groen,G., McCutchan,P.F.E., Eddy,G. and Burke,D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences that were published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype env sequences. See also accession numbers L22939-L22940, L22942-L22957 and L23064-L23065.

NCBI gi: 437344

BASE COUNT 1149 a 581 c 787 g 775 t

ORIGIN 149 bp upstream from the beginning of rev cds start

1 aggcaggaAT Ggagccggta gatcctagcc tagagccctg gaaccaccg ggaagtcagc
tat cds start ->

61 ctacaactgc ttgtagcaag tgttactgta aaatatgctg ctggcattgt caattgtgct
121 ttctgaacaa gggcttaggc atctctATG gcaggaagaa gggagacgc cgacgaggaa
rev cds start ->

181 ctcttcagag tcgtcaggat catcaaaatc ctgtaccaa gcaGTgagta gtaataatta
5' sj /\

241 gtaagtgtgA TGcaatcttt agaaataact gcaatagtag gactagtagt agcattcata
vpU cds start ->

301 gcagccatag ttgtgtgggc catagtattc attgaatata ggaaaataag gaaacagaag
361 aaaatagaca agttacttga tagaataaga gaaagagcag aagatagtgg caATGagagt
env cds start ->

421 gatggggatg cagaggaatt atccactctt atggaaatgg ggtacgatca tattttggac
481 aatgataatt tgTAatgctg aaaagttgtg ggtcacgggc tactatgggg taactgtgtg
← vpU cds end

541 gagaaacaca gagaccacc ctttttgtgc atcagatgct aaagcatatg atgcagaagt
601 acataatgtc tgggtacac atgcctgtgt accacagac cctaaccac aagaaataca
661 tttggaaaat gtaacagaaa agtttaacat gtggaaaaat aacatggtag aacagatgca
721 tgaagatata attagtctat gggaccaag cctaaagcca tgtgtaaagt taaccctct
781 ctgtgttact ttagattgtc ataattgtca cagcagcaat agcagcacct ctatcaacag
841 gaccatagat agtaaaatgc aagagggaat aaaaaactgc tctttcaata tgaccacaga
901 actaagagat aagaaacaga aagtgtctgc acttttttat agacttgatg tagtacaat
961 taatgaaagt agcaatagtc aatataggtt aatcaattgt aatacctcag ccattacaaa
1021 agcatgtcca aaggtaacct ttgagccaat tccatatac tattgtgccc cagctgggtt
1081 tgcaattcta aatgtaatg ataagaagt caatggaaca gggctatgca agaattgttag
1141 cacagtacaa tgcacacatg gaattaagcc agtagtatca acacaactac tgttaaatgg
1201 cagtctagca gaagaggagg tagtgattag atctgaaaat atcacaaaca atgccccaaa
1261 cataatagta cagttgggtg cgctgtgaag aatcaattgt accagaccta acaacaatac
v3 loop ->

1321 aagaaaaagt gtacgtatag gaccagggca aacattctat gcaacaggtg acataatagg
1381 ggatataaga caagcacatt gtaatgttag tagatcagaa tggaaatagaa ctttacaaca
← v3 loop

1441 ggtggctaca caattaagga agcaatttaa caaaacaata atctttgcta actcctcagg

```

1501 aggggatata gaaattacaa cacatagttt taattgtgga ggagaatttt tctattgcaa
1561 tacatcggaa ctgttttaata gcacttggaa tagcacttgg gataacagct caaatcacat
1621 agagtcaaat cacacagaag gcaatataac tctccaatgc aggataaage aaatttgtaa
1681 tatgtggcag aaagttaggac tagcaatgta tgcctctccc atctcaggag aaataagggtg
1741 taaatcaaac attacaggac tactactaac aagagatgga gggagttaata atagtacaaa
1801 tgagaccttc agccttgagg gaggagacat gagggacaat tggagaagtg aattatataa
1861 gtataaagta gtaaaaattg aaccactagg agtagcaccc acccgtgcaa aaagaagagt
1921 ggtggagaga gaaaaaagag cagttggact gggagctgtc ttccctgggt tcttaggagc
1981 tgcaggagac actatgggag cggcgtcaat aacgctgagc gtacaggcca gacaattatt
2041 gtctgggtata gtgcaacagc agagcaattt gctgagggct atagagggctc aacaacatct
2101 gttgaaactc acggtctggg gcattaaaca gctccaggca agagtccctgg ctctggaaag
2161 ttacctaaag gatcaacagc tcttaggaat tggggctgtc tctggaaaac tcatctgcac
2221 cactaatgta ccctgggaact ctagtggag taataaaaact tataatgaca tatgggataa
2281 catgacctgg ctgcaatggg ataagggaagt tagcaattac acagacataa tatataatct
2341 aattgaagaa tcgcagaacc agcaggaaaa gaatgaacaa gacttatttg cattggacca
2401 atgggcaaat ctgtggaatt ggtttggcat aacaaaatgg ctatggtata taaaaatatt
2461 cataatgata gtaggaggtt taataggttt aaggatagtt ttgctgtgac ttactataat
2521 aaatagagtt aggcagggat actcaccttt atcgttccAG acccttacc accaccagag
      3' ej (tat,rev) /\
2581 gggaccggac agaccgaaa gaatcgaaga aggaggtggc gagcaagaca gagacagatc
2641 agggcgatta gtgagcggtt tcttagcact tgcctgggac gatctgcgga gctctgcct
2701 cttcagctac caccgattga gagactttgt cttgattgca gtgagggctg tggaaacttct
2761 gggacacagc agcctcaagg gactgagact ggggtgggaa gccctccaat atctggggaa
2821 tcttctacta tactggggtc aggaactaaa gaaTAGtgct attaatgtgc ttgatacaat
      <- rev cds end
2881 agcattagca gtagctaatt ggacagacag agttatagaa ataggacaaa gagttggtag
2941 agctattctc aacataccta gaagaatcag acaggggtctt gaaagggctt tactaTAACA
      <- env
      nef cds start ->
3001 TGggtggcaa gtggtcaaaa agcagcatag tgggatggtc taagggttagg gaaagaataa
3061 gacagacccc tccagcagca acaggagtag gagcagcatc tcaagattta gatagacatg
3121 gagcaatcac aagtagtaat acagcagata ctaatgctga ttgtgcctgg ctagaagcac
3181 aagagggaaga ggaggtaggc tttccagtea ggccacaggt acctttgaga ccaatgactt
3241 ataaggcagc attcgtcttc ggcttctttt taaaagaaaa ggggggactg ga
      3' LTR ->

```

LOCUS HIVDJ264A 3328 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, DJ264 from Djibouti, proviral DNA encoding env, tat, vpU, rev, and nef genes.

ACCESSION L23064

SOURCE Human immunodeficiency virus type 1 (HIV-1), DJ264 from Djibouti, PCR amplification of proviral DNA from cocultivated PBMC's

REFERENCE 1 (bases 1 to 3328)

AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 proteins from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype env sequences. See also accession numbers L22939-L22957 and L23065.

NCBI gi: 437345

BASE COUNT 1160 a 583 c 797 g 788 t

ORIGIN 209 bp upstream from the beginning of rev cds start

1 tgtttattca ttccagaatt ggggtgtcaac atagcagaat aggcattatt cgagggagaa
61 gaggcaggaa TGgagccagt agatcctagc ctagagccct ggaaccaccc gggagtgac
tat cds start ->
121 cctacaactg cttgtagcaa ttgttactgt aaaatatgct gctggcattg ccaattgtgc
181 tttctgaaca agggcttagg catctcctAT Ggcaggaaga agcggagacy ccgacgagga
rev cds start ->
241 actcctcaga gtcgtcagga tcgtcaaat cctgtatcaa agcaCTgagt agtaataatt
5' 3' /\

301 agtatatgtg ATGaaatctt tagaaatagc tgcaatagta ggactagtag tagcattcat
vpU start ->
361 agcagccata gttgtgtgga ccataatagt cattgaatat aggaataata ggaacacagaa
421 gaaatagac aaattacttg atagaataag agaaagagca gaagatagtg gcaATGagag
env cds start ->
481 tgatggggat acagaggaat tatccactct tctggaaatg gggatatgata atattttgga
541 taatgataat ttgTAAtgct gaaaagttgt gggtcacggt ctactatggg gtacctgtgt
<- vpU cds end
601 ggagagatgc agaaaccacc ctattttgtg catcagatgc taaagcatac gatgtagaag
661 tacataatgt ctgggctaca catgcctgtg taccacaga ccctaacca caagaataac
721 atttgaaaaa tgtaacagaa aagttaaca tgtggaaaaa taacatggtg gaacagatgc
781 atgaagatat aattagtcta tgggaccaa gctaaagcc atgtgtagag ttaacccttc
841 tctgcgttac tttagattgt cataacaacg tcacctctaa caataggacc atatctaggg
901 acatgcaagg ggaagtaaaa aattgctctt tcaatatgac cacagaacta agagataaga
961 aacagcaagt gtctgcactt ttttatagc ttgatgtagt acaaatat gatagtagca
1021 atagtagtca atataggta atcaattgta atacctcagc cattaaacaa gcatgtccaa
1081 aggttaacct tgagccatt cccatacatt attgtgcccc agctgggttt gcgattctaa
1141 agtgtaatga taagaagttc aatgggaacag ggccatgcaa gaatgttagc acagtacaat
1201 gcacacatgg aatcaagcca gtagtatcaa ctcaactact gttaaatggc agtctagcag
1261 aagaagaggt agtgattaga tctgaaaata tcacaaacaa tgccaaaac ataattagtc
1321 agttgggtac gcctgtagga atcaattgta ccagacctaa caacaatata agaaagagtg
v3 loop ->
1381 tgcgtatagg accagggcaa acattctatg caacaggtga cataataggg gatataagaa
1441 aagcacattg taatgtcagt agatcaaat ggaataacac tttaaacag gtggctatac
<- v3 loop
1501 aattaaggaa gcactctaac acaacaataa tctttgctaa cccctcagga ggggatatag

```

1561 aaattacaac acatagtttt aattgtggag gagaattttt ctattgcaat acatcggaac
1621 tgtttaatag cacttgggat aacagcttaa atcacacaga gtcaaatcac acagaagaca
1681 atataactct ccaatgcaga ataaagcaaa ttgtaaatat gtggcagaaa gtaggacgag
1741 caatgtatgc cctcccatc ccaggagaaa taagggtgtga atcaaacatt acaggactac
1801 tattaacaag agatggaggg agtaataata gtacaaatga gactttcagg cctggaggag
1861 gagatatgag ggacaattgg agaagtgaat tatataagta taaagttagta aaaattgaac
1921 cactaggtat agcaccacc catgcaaaaag gaagagtggg ggagagagag aaaagagcag
1981 ttggactggg agctgtcttc ttgggtttct taggagcagc aggaagcact atggggcag
2041 cgtcaataac gctgacggtg caggccagac aattattgtc tggatatagt caacagcaga
2101 gcaatttgcg gagggtctata gaggtctcac aacatctgtt gaaactcacg gtctggggca
2161 ttaaacagct ccaggcaaga gtctctggctc tggaaagata cctaagggat caacagctcc
2221 taggaatttg gggctgctct ggaaaactca tctgcaccac tactgtacct tggaaactca
2281 gttggagtaa taaaacttat aatgacatat gggagaacat gacctggctg caatgggata
2341 aagaatttag caattacaca gacataatat ataattcaat tgaagaatcg cagaaccagc
2401 aggaaaagaa tgaaaaagac ttattggcat tggacaaatg ggaaagtctg tgggaattggt
2461 ttaacataac aaaatggcta tggatatata aatatctcat aatgatagta ggagggttaa
2521 taggggttaag aatagttctt gctgtgctta ctgtaataaa gagagttagg cagggatact
2581 cacttttacc attccAGacc ctaaccaacc accagagggg acccgacagg ccggaagaa
      3' sj (tat.rev) /\
2641 tcgaagaagg aggtggagag caagacaaag acagatccgt gcgattagtc agcggattct
2701 tagcacttgc ctgggacgat ctggggagcc tgtgcctctt cagctaccac cacttgagag
2761 actttgtctt gattgcagtg agggctgttg aacttctggg acacagcagt ctcaagggac
2821 tgagactggg ttggggagcc ctcaaatatc tggggaatct tctgacatac tgggggcagg
2881 aactaaagaa TAGtgctatt aatttgcttg atacaatagc aatagcagta gctaactgga
      <- rev cds end
2941 cagacagagt tatagaaata ggacagagag ttggttagagc tattcgtaac atacctgtaa
3001 gaatcagaca aggtctcgaa agggcttttg taTAACATGg gtggcaagtg gtcaaaaagc
      <- env cds end
      nef cds start ->
3061 agcttagtgg gatggcctaa ggttagggaa agaataatac agaccctcc agcagcaact
3121 ggagtaggag cagcatctca agatttagat agacatggag caatcacaag tagtaataca
3181 gcagaaacta atgctgattg tgcttggtta gaagcacaag aggcagagga ggtaggcttt
3241 ccagtcagac cacaggtaac ttgagacca atgacttata agggagcttt cgtctctggc
3301 ttctttttaa aagaaaaggg gggactgg
      3' LTR ->

```

LOCUS HIVDJ373A 3340 bp ds-DNA VRL 31-DEC-1993
DEFINITION Human immunodeficiency virus type 1, DJ373 from Djibouti, proviral DNA encoding env, tat, vpU, rev, and nef genes.
ACCESSION L23065
SOURCE Human immunodeficiency virus type 1 (HIV-1), DJ273 from Djibouti, PCR amplification of proviral DNA from cocultivated PBMC's
REFERENCE 1 (bases 1 to 3340)
AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
TITLE Genetic diversity of the HIV-1 envelope glycoprotein
JOURNAL Unpublished (1993)
COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with C subtype env sequences. See also accession numbers L22939-L22957 and L23064.
 MCBi gi: 437346
BASE COUNT 1191 a 568 c 791 g 790 t
ORIGIN 226 bp upstream from the beginning of rev cds start
 1 ggcttcgcgc tagatgctgt ctattcattt cagaattggg tgccctgcata gcagaatagg
 61 catcatgcga cagagaagag caagaaATGg agccagtaga tcccaaccta gagccctgga
 tat cds start ->
 121 accatccagg aagtcagcct aaaactgctt gtactaaatg ttattgtaaa aaatgtagct
 181 atcattgtct agtttgcttt cagacaaaag gcttaggcct tccctATGgc aggaagaagc
 rev cds start ->
 241 ggagacagcg acgaagcgct ccttcaagca gtgaggatca tcaaaatctt atatcaaagc
 301 aGTgagtacc aatagtaga tgtaATGata gacttactag caaaagtaga ttatagatta
 5' s' j / \ vpU cds start ->
 361 gcagtagcag catttataat agcatttata atagcaatag ttgtgtggac catagcatat
 421 atagaatata ggaaattggt aagacaaaaga aaaatagaca ggtaatttga aagaattagg
 481 gaaagagcag aagacagtgg caATGagagt gatggggata cagatgaact gtcaacaatg
 env cds start ->
 541 gtggataggg ggaatccttag gcttttggat gctgttgatg tgTAAtggga tggggaactt
 <- vpU cds end
 601 gtgggtcaca gtctattatg ggttacctgt gtggcaagaa gcaaacccca ctctattttg
 661 tgcatacagat gctaaagcat atgagacaga agtgcataat gtctgggcta cacatgcttg
 721 tgtacccaca gacccagacc cacaggaact agttatggaa aatgtaacag aaaattttaa
 781 catgtggaaa aatggcatgg tggatcagat gcatacaggat ataatacatt tatgggatga
 841 aggcctaaag ccatgtgtaa agttgacccc actctgtggt actttaaact gtagtaataa
 901 ggttactgct acgggtaatg gtagtggtac ctacaatagc acagaaaaag acatgagaaa
 961 ttgtctcttc aacataacca cagaactaag agataagaaa cagaaaaaat atgcactttt
 1021 ttataaactt gatatagtac cacttgggaa tgggaatgag gaatatcgat taataaattg
 1081 taataactca tccataaacac aagcctgtcc aaaggctctt ttggaaccaa ttcctatata
 1141 ttattgtgct ccagctgggt atgcgattct aaagtgtagg gatacaaaat tcaatggaac
 1201 aggacatgc aagaatgtca gtacagtaca atgtacacat ggaattaaac cagtggatgc
 1261 aactcaacta ttgttaaatg gtagtacagc agaggaggat ataataatta gatctgaaaa
 1321 tctgacaaac aatgccaaa taataatagt acagcttaac caatctgtag aaattaattg
 v3 loop ->
 1381 tacaagaccc aacaataata caaggcaag cataaggata ggaccaggac aaacatttcta
 1441 tgcaacagga gatataatag gagatataag acaagcacat tgtaacatta gtagacaaaa
 <- v3 loop
 1501 gtggaaagaa actttacaac aggtaaaagg aaaattaaaa gaacacttca ataaaacaat
 1561 aaaatttgca ccatcctcag gaggagacat agaaattaca acacatagct ttaattgtgg

```

1621 aggagaattt ttctattgta atacatcagc cttgttttaat gattcaatta caaatgcacc
1681 catttaataat tcgagtagag atgcaaacat cacacttcaa tgcagaataa aacaaattat
1741 aaacatgtgg caggaggttag gaogagcaat gtatgcccc cccattcaag gaataatagc
1801 atgtaaatca aatatacacag gactgctact gacacgtgat ggagggacaa atagcacaga
1861 cgagacattc agacctggag gaggagatat gagggacaat tggagaagtg aattatataa
1921 atacaaagtg gtagaaatta agccattagg actagcacc ctaaggcaa aaaggagagt
1981 ggtggagaga gaaaaaagag cagcactagc agctgtgttc cttgggttct tgggagcagc
2041 cggaagcact atggggcgag catcaataac gctgacggta caggccagac aattattgtc
2101 tggatatagt caacagcaaa gcaatttgc tgaaggctata gaggtgcac agcatatgtt
2161 gcaactcaca gtctggggca ttaagcagct ccagacaaga gtcttggcta tagaaagata
2221 cctgaaggat caacagctcc tagggatttg gggtgtgtct ggaaaactca tctgcaccac
2281 tgctgtgctc tggaaactca gttggagtaa taaatctcaa gaggagattt gggacaacet
2341 gacctggatg cagtgggata gagaatttaa taattataca gaaacaatat acaggttgcct
2401 tgaagtctcg caaaccacgc agggacaaaa tgaanaagat ttactagcat tggacaaatg
2461 gcaaaatctg tggagttggg ttgacataac aaagtggctg tggatatataa aaatattcat
2521 aatgatagta gggggcttga taggttttaag aataattttt gctgtgtctt ctatagttaa
2581 tagagttagg cagggatact caccattatc atttcAGacc cttatcccg acccaagggg
      3' sj (tat,rev) /\
2641 acccgacagg ccagaggaa tcgaagaaga aggtggagag caagacaaag gcagatccac
2701 tcgatttagt agcggattct tagcatttgc ctgggacgac ctggggagcc tgttcctttt
2761 cagctaccac cgattgagag acttaattct gattgtagcg agagtgggtg aacttctggg
2821 acagaggggg tgggaacccc ttaagtatct gggaagcctt gtgcaatatt ggggtctgga
2881 gctaaaaaag agtgctgtTA Gtctgcttga taccacagca atagcagtag ctgaaggaa
      <- rev cds end
2941 agataggatt atagaaataa tacaaagaat ttggagagct ttctgcaaca tacctagaag
3001 aataagacag ggcttttgaag cagcttttgc ATAAaATGgg gaacaagtgg tcaaagggat
      <- env cds end
      nef cds start ->
3061 ggctgtgtgt aagagaaaga ataagaaaaa ctaagccagc agcagaaaga gtagcagcag
3121 cagcagaagg agtaggggca gogtctcaag acttagataa atatggggca cttacaacca
3181 gcaacacacc cgccaataat gctgattgtg cctggctgga aacacaagag gaggaagaag
3241 aagtaggctt tccagtaaga cctcaggtac cattaagacc aatgacttat aaggagagcat
3301 ttgatctcag cttcttttta aaagaaaagg ggggactgga
      3' LTR ->

```

LOCUS HIVGER 384 bp ss-DNA VRL 27-JUN-1994
 DEFINITION Human immunodeficiency virus type 1 (HIV-1) gag sequence, isolate GER, partial cds.
 ACCESSION U11549
 SOURCE Human immunodeficiency virus type 1 PCR-amplified DNA from cocultivated PBMCs.
 REFERENCE 1 (bases 1 to 384)
 AUTHORS Emiliani, S., Blanc, M., and Devaux, C.
 TITLE p17gag sequence of HIV1GER, a new, highly replicative and highly cytopathic HIV Type 1 strain of D genotype
 JOURNAL AIDS Res. Hum. Retroviruses 10, in press.
 COMMENT Isolate HIVGER has an SI phenotype with elevated replication in both PBMCs and human hematopoietic cells, i.e. rapid/high HIV-1 strain. The virus was isolated from a 37 year old French patient who contracted the virus through heterosexual contact. HIVGER has the greatest nucleotide similarity to HIVNDK, a subtype D virus. This subtype designation obtained through phylogenetic analysis is consistent with the presence of an isoleucine in the amino acid sequence at position 68 in p17gag, a characteristic of subtype D sequences. Another characteristic common to both HIVNDK and HIVGER is the absence of three amino acids in the C-terminal region of p17gag. Emiliani et al. suggest that this sequence similarity may result in a high fusogenic potential.
 BASE COUNT 160 a 59 c 92 g 73 t
 ORIGIN Beginning of gag cds
 1 ATGggtgcga gagcgtcagt attaagcggg ggaaaattag atacatggga aagaattcgg
 gag cds ->
 61 ttacggccag gaggaagaa aaaatatgca ctaaaacatt tgatatgggc aagcagggag
 121 ctagaacgat ttaaacttaa tcttgacctt ttagagacat cagaaggctg taaacaaata
 181 ataggacagc tacaaccatc tattcaacaa ggatcagaag aaattagatc attgtataat
 241 acagtagcaa ccctctattg tgtacatgaa aggatagagg taaaagacac caaagaagct
 301 gtagaaaaga tggaggaaga acaaaacaaa agtaagaaaa aggcacagca agcagctgat
 361 agcagccagg tcagccaaaa ttat
 gag ->

HIVHYDA

LOCUS HIVHYDA 267 bp ss-DNA VRL 09-MAR-1994
 DEFINITION Human immunodeficiency virus type 1, sample HYDA (isolate 30008) from Hyderabad, India, C2-V3 of env region.
 ACCESSION L29091
 SOURCE Human immunodeficiency virus type 1 (HIV-1), sample HYDA from isolate 30008 from Hyderabad, India, PCR amplification of proviral DNA extracted from peripheral blood leukocytes (PBL's).
 REFERENCE 1 (bases 1 to 267)
 AUTHORS Baskar,P.V., Ray, S.C., Rao,R., Quinn,T.C., Hildreth,J.E.K. and Bollinger,R.C.
 TITLE Presence in India of HIV type 1 similar to North American strains
 JOURNAL AIDS Res. Hum. Retroviruses 10, 1039-1041 (1994)
 COMMENT These sequences, taken from four seropositive individuals, represent the first known subtype B infections in India. Phylogenetic analysis of the C2-V3 regions of the four Hyderabad sequences indicates a closer similarity to North American and European HIV-1 subtype B sequences than to subtype B sequences in Thailand and subtype C sequences from India and South Africa. Other sequences in this set are included in the alignment below (accession numbers L29091-L29094).
 NCBI gi: 459768

BASE COUNT 120 a 38 c 52 g 57 t
 ORIGIN 790 bp downstream from the beginning of env cds
 1 tagcagaaga agaggtagta attagatctg acaacttcac agacaatgct aaaaccataa
 env cds ->
 61 tagtacagct gaaagaatct gtagaatta attgtacaag acccggcaac aatacaagaa
 V3 loop ->
 121 gaagtataca tataggacca gggagagcat tttttgcaac aggagacata acaggagata
 181 taagacaagc acattgtaac cttagtagaa caaaatggaa tgaaacttta aaacagatag
 <- V3 loop
 241 ttataaagtt aagagaacaa tttcaga
 env cds ->

```

-----
CONSENSUS .....???CAGAAGAAGAGGTAGTAATTAGATCTGACAATTTACAGACAATGCTAAAC 52
HIVHYDA .....TAG-----C----- 55
HIVHYDB ..... 52
HIVHYDC ..... 49
HIVHYDD TGTAAATGGCAGTCTAG-----G----- 70

CONSENSUS CATAATAGTACAGCTGAACAAATCTGTAGAAATTAATTGTACAAGACCC?ACAAC?A?ATAA?AAGAAG? 117
HIVHYDA -----AG-----GG---A-T-C---G-----T 125
HIVHYDB -----G---G-G---G-----T 122
HIVHYDC -----G---A---AGT---A-----A 119
HIVHYDD -----A-A---G-G---A-----A 140

CONSENSUS AGAATACA?ATAGGACCAGGGAGAGCATTT?TTGCAACA...G?CACAG?AGGAGATACAAGACAAGCAC 180
HIVHYDA -T...-T-----T-----GGA-A--T-AC-----T----- 192
HIVHYDB --C-T--GG-----G---G---G---C----- 189
HIVHYDC -----T-----T-----AA---T----- 186
HIVHYDD -----A-G-----G-----G-----T----- 207

CONSENSUS AGTGTAACCTTAGTAGAGCAAAATGGAATGAAACTTTAAACAGATAGTTATAAAGTTAAGAGAACAATT 250
HIVHYDA -T-----A----- 262
HIVHYDB -----T-----T----- 256
HIVHYDC -----T----- 250
HIVHYDD ----- 277

CONSENSUS T?AGA 254
HIVHYDA -C--- 267
HIVHYDD -A 279

```

LOCUS HIVK1241 1477 bp ss-RNA VRL 20-APR-1993
 DEFINITION Human immunodeficiency virus type 1, isolate K124 from Nairobi, Kenya, gag region.
 ACCESSION L11769
 SOURCE Human immunodeficiency virus type 1 (HIV-1), Nairobi, Kenya isolate K124.
 REFERENCE 1 (bases 1 to 1477)
 AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Buell, E., Eddy, G., van der Groen, G., Fransen, K., Gershy-Damet, M., DeLeys, R. and Burke, D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, Brazil, Philippines, Djibouti, Uganda, Senegal, and Zaire) on four continents. The authors identified 7 distinct genotypes with an average gag inter-genotypic variance of 7% and an average intra-genotypic variance of 14%. The env region of this isolate (L22942) has also been sequenced; it groups with subtype D sequences. This gag sequence, previously published on page I-A-200 of AUG 93, tends to cluster with HIV1 A subtype gag sequences.

BASE COUNT 553 a 275 c 363 g 286 t

ORIGIN gag cds start

```

1 ATGgggtgcga gagcgctcagt attaagcggg ggaaaattag atgcatggga gaaaattogg
gag start ->
61 ttaaggccag ggggaaagaa aaagtataga ctgaaacatt tagtatgggc aagcagggag
121 ctggaaagat ttgcacttaa cctagcctt ttgaaacaa cagaaggatg tcaacaata
181 atggaacagt tacaatcatc tctcaagaca ggaacagaag aaattagatc attatttaac
241 gcagtagcag tccctatttg tgtacatcaa cggatagagg taaaagacac caaggaagct
301 ctagacaaaa tagaggaaat gcaasataag agcaaacaaa agacacaaca ggcagcagct
361 gacacaggaa acagcagcaa ggtcagccaa aattacccta tagtgcaaaa tgcacaaggg
421 caaatgatac accagagttt gtcacctagg actctgaatg catgggttaa agtaatagaa
481 gaaaaggctt tcagcccaga agtaataccc atgttctcag cattatcaga aggagccacc
541 ccacaagatt taaatatgat gctgaacata gtggggggac accaggcagc tatgcaaatg
601 ttaaaaaata ccatcaatga ggaagctgca gaatggggac ggttacatcc agtacatgca
661 gggcctatto caccaggcca gatgagagaa ccaaggggaa gcgacatagc aggaactact
721 agtacccttc aagaacaaat aggatggatg acaagtaacc caoctatcc agtgggagac
781 atctataaaa gatggataat cctgggatta aataaaatag taagaatgta tagccctgtt
841 agcatttttg acataaaaca agggccaaaa gaaccttca gagattatgt agatagggtc
901 tttaaaactc tcagagctga gcaagctaca caggaggtaa aaggttggat gactgaaaca
961 ttactgggtc aaaaagcaaa tccagattgt aagtccattt taagagcatt gggagcaggg
1021 gctacattag aagaaatgat gacagcatgc caggagtggt gaggaccagc tcataaagca
1081 agagtcttgg ctgaggcaat gagccaagca acaataata taaatgtgtc tgtaatgatg
1141 cagagaggca attttaaggg ccaaaagaaa atcattaaat gtttcaactg tggcaaagaa
1201 gggcacatag caaaaaattg cagggtctct aggaaaaggg gctgttggaa atgtggaagg
1261 gaaggacacc aaatgaaaga ttgcactgaa agacaggcta attttttagg gaaaatttgg
1321 ccttccaca agggaggcc agggaaactt cttcagagca gaccagaacc aacagcccca
1381 ccagcagaga gcttcgggtt tggggaggag ataacacct ctcagaaaca ggagccgaaa
1441 gacaaggaac tgtatccttt agcttccgtc aaatcac
gag ->

```

HIVK124A2

LOCUS HIVK124A2 2581 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, K124 from Kenya,
 proviral envelope glycoprotein (env) gene.
 ACCESSION L22942
 SOURCE Human immunodeficiency virus type 1 (HIV-1), K124 from Kenya,
 PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 2581)
 AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A.,
 van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT

These sequences were kindly provided prior to publication by Dr.
 Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory
 Rockville, Maryland. Twenty-one full length gp160 proteins from eight
 African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda,
 Zaire, and Zambia) were sequenced and analyzed with thirty-two
 previously published full-length gp160 env sequences. Sequences
 DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and
 VI525 have corresponding gag sequences published by Louwagie et. al.
 in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence
 tends to cluster with D subtype env sequences, although the corres-
 ponding gag sequence (L11769) clusters with A subtype gag sequences.
 See also accession numbers L22939-L22941, L22943-L22957 and L23064-
 L23065.

NCBI gi: 437347

BASE COUNT 905 a 451 c 611 g 614 t
 ORIGIN env cds start

```

1 ATGagagtga tggggataca gaggaattgt cagctcttat tgggaggggg aattatgac
env cds ->
61 ttggggatga taataatctg taatgctgca aacttgtggg ttactgttta ctatggggta
121 cctgtgtgga gagatgcaga gaccacttta tttgtgcat cggatgctaa agcatataag
181 acagaaatgc ataattgtctg ggtacacat gccctgtgtac ccacagaccc caaccacaa
241 gaaatacatt tggaaaatgt gacagaaaaat ttaacatgt ggaaaaataa catggtagaa
301 cagatgcata cagatataat cagtctatgg gaccaaagcc taaagccatg tgtaaagtta
361 acccctctct gcgttacttt aaattgtagt aatgtcaatg gcacgggtgg caacagcact
421 acgggtaaca acatcactag ggaggacata ggaatgaaaa actgctcttt caatataacc
481 acagaagtaa gagataggaa gaagcaagta tatgcacttt ttataaaact tgatatagta
541 caaatgaatg aaaatcagga taatagcagt agcgggtgaat atagatatag attaataaat
601 tgaataacct cagccattac acaggcgtgc ccaaaggat cctttgagcc aattcccata
661 cattattgtg ctccagctgg ttttgcgac ctaaagtgtg aggataagga gttcaatgga
721 acagggccat gcaagaatgt cagcacagtc caatgcacac atggaatcaa gccagtagta
781 tcaactcaac tgctgttaaa tggcagtcct gcagaaggag aggtgaaaat tagatctcaa
841 aatctcacaa acaatgccaa aactataata gtacaacttg tcaagcctgt gagaattaa
901 tgtaccagac ctaacaacaa tataagaaaa agtatacgtg taggaccagg acaagcattc
v3 loop ->
961 ttgcaacag gtgacataat aggaatatata agacaagcac agtgtaattgt taatggaaca
<- v3 loop
1021 gaatggaatg aaactttaca acagataagt acacaattaa aaaaacactt tatgaacaag
1081 acaataatct ttaaaccatc ctcgggaggg gacccagaaa ttacaacaca cagctttta
1141 tgtagagggg aatttttcta ctgcaataca tcaaaactgt ttaatagtgt atggggaaat
1201 tcaagtaatg tgacaagggt aaatgggaca aaggtaaaag agacaatcac actcccatgt
1261 aaaataaaac aaattataaa catgtggcag gaagtaggac gagcaatgta tgccccctcc
1321 attgaaggac tcatcaactg ttcatcaaat attacaggac tattgttgac aagagatggc
1381 ggtgaaaata atagtagtca gaatgagacc ttacagcctg gaggaggaga tatgagagac
1441 aattggagaa gtgaattata caaatataaa gtagtataaa ttgaaccact aggtctagca
1501 cccaccagcg caaagagaag agtgggtgaa agagaaaaaa gagcaatagg actgggagct
1561 atgttctctg ggttcttggg agcagcagga agcacgatgg gcgcagcgtc aatgacgctg
1621 acggtacagg ccagacagtt aatgtctggt atagtgaac agcaaaacaa ttgctgagg
1681 gctatagatg cgcaacagca tctgttgcaa ctacagctct ggggcattaa acagctccag

```

```

1741 gcaagagtc tggctgtgga aagctaacta aaggatcaac agctcctagg aatttggggt
1801 tgcctctggaa aacacatttg caccactaat gtgccttgga actctagctg gagtaataaa
1861 tctctacagg agacttggga taacatgacc tgggaaggagt gggaaagaga aattgacaat
1921 tatacagggt taatatacag ctttaattgaa aaatcgcaaa ccagacaaga aaagaatgaa
1981 caagagctat tacaattgga caaatgggca agtttgtgga attgggttag cataacaaaa
2041 tggctgtggt atataaaaat attcgtaatg atagtaggag gcttgatagg ttaaggata
2101 gtttttgeta tgccttcttt agtaaataga gttaggcagg gatattcacc tctgtcgttt
2161 cagaccctcc tcccagcccc gaggggacca gacAGcccc aaggaataga agaagaaggt
                               3' sj (tat,rev) /\
2221 ggagagcaag gcagagacag gtccattcga ttggtgaacg gattctccgc acttatctgg
2281 gacgacctga gggacctgtg cctcttcagc taaccaccgt tgagagactt aatcttaatt
2341 gcaacgagga ttgtggaact tctgggacgc agggggtggg aagccctcaa atatctgggg
2401 aatctctctg agtattggac tggggaacta aagaATAGtg ctattagctt agttgatacc
                               <- rev cds end
2461 atagcagtag tagtagctgg ctggacagac agtattataa aaataggaca aagacttggg
2521 agagcttttc tccacatacc tagsagaatc agacagggt tcgagagggc ttgtctaTAA
                               <- env
2581 c

```

HIVK986

LOCUS HIVK986 252 bp ss-RNA VRL 04-AUG-1994
 DEFINITION Human immunodeficiency virus type 1 (HIV-1) partial envelope sequence, V3 region, isolate K986.
 ACCESSION U12984
 SOURCE Human immunodeficiency virus type 1, PBMC DNA from a Kenyan patient, K986.
 REFERENCE 1 (bases 1 to 252)
 AUTHORS Janssens,W., Heyndrickx,L., Fransen,K., Temmerman,M., Leonaers,A., Ivens,T., Motte,J., Piot,P. and van der Groen,G.
 TITLE Genetic variability of HIV type 1 in Kenya
 JOURNAL AIDS Res. Hum. Retroviruses 10, (1994) In press, Nov issue.
 COMMENT This sequence set was derived from a 1990-1992 cohort study designed to determine risk factors for mother to child HIV transmission. In this study, samples were collected from twenty-two seropositive pregnant women and one infected infant(K88). Of the set, nineteen cluster with subtype A sequences and are shown in alignment following their consensus sequence. Classification of sequence K976 is problematic, although it seems to cluster most often with subtype A sequences. The three remaining sequences cluster with subtype D sequences; they are also shown in alignment below their consensus sequence. Five members of the set (K29, K31, K88, K98, and K112) were previously sequenced over the gag coding region (Louwagie et al. AIDS 7: 769-780). Four of these (K29, K88, K98 and K112) strains cluster with sequences of the same subtype for both the env and gag coding regions. However, K31 clusters with subtype D in gag and subtype A in env. Accession numbers for the sequence set are U12984-U13006.
 BASE COUNT 111 a 43 c 41 g 57 t
 ORIGIN 972 bp downstream from beginning of env cds
 1 gaagagataa taattagatc tgaaaatctc acaataatg ctaaaatcat aatagtacag
 env cds ->
 61 cttaatgagt ctgtaccaat tacttgcaca aggccttaca acaatacaag acaagggtaca
 V3 loop ->
 121 catataggac cagggcaaac actctttaca acaaaaataa taggagatat aagacaagca
 181 cattgtaata ttagtgaagc agaattggaac aaaactttac aactagtagc aagaaaatta
 <- V3 loop
 241 ggaggccttc tt
 env cds ->

CONSENSUS .A	GAAGAGGTAA?AATTAGATCTGAAAATATCACAACAATGCCAAAACCTATAATAGTACAA	59
HIVK 29	-----T-----	60
HIVK 31	A-----A-----T-----	60
HIVK 88	-----C-----AT-----	60
HIVK 98	-G-----A-----T-----AC-----T-----	60
HIVK 112	A-C-T-----A-C-----	60
HIVK 965	A-A-----A-----T-----T-----C-----	60
HIVK 967	-----G-----T-G-----T-----C-----C-----	60
HIVK 968	A-G--A--A-----T-----	60
HIVK 970	-G--A--G-----T-----AC-----	60
HIVK 973	-----T--T-----C-G-----T--A-----G	60
HIVK 976	-GG-AA--TG-----T--G-----AC-----G	60
HIVK 977	A--A--G-----T--G-----T-----C-----	60
HIVK 978	-G-T--GT-----C-----AC-----G	60
HIVK 979	A-----A-----T-----T-----T-----	60
HIVK 980	A-G-----A-----	60
HIVK 981	A--A-A--T-----T-----	60
HIVK 982	-----A--TG--G-----T-----AC-----	60
HIVK 983	-----A-----A--C-----	60
HIVK 984	-----A-----T-----A-----	60
HIVK 985	-----GTC-----T-----	60
CONSENSUS .D	GAAGAGATAATAATTAGATCTGAAAATCTCACAATAATGCTAAAATCATAATAGTACAG	60
HIVK 966	-----	60
HIVK 971	-----C-----	60
HIVK 986	-----	60
CONSENSUS .A	CTTG?CAAGCCTGTAAAAATTAATTGTACCAGACCTAACAACAATACAAGAAAAAGTATA	118
HIVK 29	-----TAG-----C-----C-G-----	120
HIVK 31	-----A-G-----T-----N-T-----	120
HIVK 88	-----C-C-----G-C-----T-----C-----	120
HIVK 98	T--AA--C-----G-----	120
HIVK 112	T--T--T-----G-C-----T-----	120
HIVK 965	-----TT-----GG-----C-----	120
HIVK 967	-----AC--C-----GC-----G-----	120
HIVK 968	-----TT--T-----G-C-----G-----T-----G-----	120
HIVK 970	T--A--C-----G-----G-----G-----	120
HIVK 973	-----CAAG-CC-----G-----C-----	120
HIVK 976	T--ACTG-----A-C-----T--C-----	120
HIVK 977	T--AAA-C-----C-----T-----C-----	120
HIVK 978	T--ACAG-----G-----	120
HIVK 979	-----C-----G-----C-----	120
HIVK 980	-----T-----G-----S-----	120
HIVK 981	-----T-G-----G-----G-----C-----	120
HIVK 982	-----AATG--T-----C-----GC-----	120
HIVK 983	T--AA-G--A--G-----T-----T-----TT--G-----	120
HIVK 984	-----T--C-----G-C--G-----T--G--T-----	120
HIVK 985	-----T-G-----G-T-----	120
CONSENSUS .D	CTTAATGAGTCTGTACCAATTAATTGCACAAGGCCCTACAACAATACAAGACAAGGTACG	120
HIVK 966	-----A-----A-----	120
HIVK 971	-----	120
HIVK 986	-----C-----A-----	120

HIVK986

CONSENSUS . A	CATATAGGACCAGGACAAGCATTCTATGCAACAGGTGACATAATAGGGGATATAAGACAA	178
HIVK 29	-----G--G-C-----A--	180
HIVK 31	-----T-----A--	180
HIVK 88	-G-----A-----A-----	180
HIVK 98	-----G-----T-----A-----	180
HIVK 112	-----A-----	180
HIVK 965	-G-----G-G--A-----A--	180
HIVK 967	-G-----T--C-----C-----	180
HIVK 968	-----G-A-----	177
HIVK 970	-----	180
HIVK 973	-G-----A-----	180
HIVK 976	-GCT-T-----C-----A-----	177
HIVK 977	-G-----G-----T-----A-----	180
HIVK 978	-C-----G-A-A-----	180
HIVK 979	TG-----A-----G--GA-----A-----	177
HIVK 980	-----	180
HIVK 981	-G-----	180
HIVK 982	A-----T-----A-----	180
HIVK 983	-----T-----AG--G--A-----	180
HIVK 984	-----T--G-A-CG-----T--A-----	177
HIVK 985	-G-----A-----G--A-----A-----	177
CONSENSUS . D	CATATAGGACCAGGGCAAACACTCTTT...ACAACAAAATAATAGGAGATATAAGACAA	177
HIVK 966	-G-----	177
HIVK 971	-----G-G-TA-A-----C-----	177
HIVK 986	-----	177
CONSENSUS . A	GCACATTGTAATGTCAGTAGATCAGAATGGAATAAACTTTACAAAAGGTAGCTA?ACAA	237
HIVK 29	---T-----C-----GG-----TG---	240
HIVK 31	-----T-----C---	240
HIVK 88	-----C-----GC-----A---	240
HIVK 98	-----T--A--C-----C--G-----TC-GC---	240
HIVK 112	-----A--C-----G-----G-----C---	240
HIVK 965	-----T-----C-----G-----GA---	240
HIVK 967	-----C-----G-----G-G-----T--T---	240
HIVK 968	-----GT--A--C-A-----C--C---	237
HIVK 970	-----T--A--C-----G--C-G-----C-A---	240
HIVK 973	-----A-----C-----C-----CA--G---	240
HIVK 976	---T-----A--C-A-A--A-----GCC-----A---	237
HIVK 977	-----G-----C-----T-A---	240
HIVK 978	-----GT-----C-----G-----C-A---	240
HIVK 979	-----GA-A--G--G-T-A--A--C--A--C---	237
HIVK 980	-----C-----G--C---	240
HIVK 981	-----GAC---	240
HIVK 982	-----G-----G--G-G-G-----C-G---	240
HIVK 983	-----G-----T-A-----T--C---	240
HIVK 984	-----C-----A-C-----CG--GG-C-----GAT---	237
HIVK 985	-----A-----T-----C---	237
CONSENSUS . D	GCACATTGTAACATTAGTGGAGCAGAATGGAACAAAACCTTTACAAC?GGTAGCTAAAAAA	236
HIVK 966	-----G-----T-----AG-----	237
HIVK 971	-----A-----A-----G---	237
HIVK 986	---T-----A-----TA--A-G---	237

CONSENSUS.A	TTAAGAAAATACTTT	252
HIVK 29	-----	255
HIVK 31	-----	255
HIVK 88	---G--C---T---	255
HIVK 98	-----C-----	255
HIVK 112	-----G---C----	255
HIVK 965	-----A-G--C-----	255
HIVK 967	-----G---C---GG	255
HIVK 968	-----C-----	252
HIVK 970	-----A--C-----	255
HIVK 973	-----Y-----	255
HIVK 976	---GAG---C---C	252
HIVK 977	-----C---GG	255
HIVK 978	-----C-----	255
HIVK 979	---AG---C-----	252
HIVK 980	-----C-----	255
HIVK 981	-----C-----	255
HIVK 982	-----C-GT----	255
HIVK 983	-----C-----	255
HIVK 984	-----C-----	252
HIVK 985	-----G--C-----	252
CONSENSUS.D	TTAGGAGACCTTCTT	251
HIVK 966	-----	252
HIVK 971	-----	252
HIVK 986	-----G-----	252

HIVKENYA

LOCUS HIVKENYA 2569 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, KENY proviral envelope glycoprotein (env) gene.
 ACCESSION L22943
 SOURCE Human immunodeficiency virus type 1 (HIV-1), isolate KENY from Kenya, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 2569)
 AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A., van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype env sequences. See also accession numbers L22939-L22942, L22944-L22957 and L23064-L23065. NCBI gi: 437348

BASE COUNT 896 a 462 c 607 g 604 t
 ORIGIN env cds start

```

1 ATGagagtga tggggacaca gaggaattgt cagcacttat tgagatgggg aactatgac
env cds ->
61 ttggggatga taataatctg tagtgcgtga gaaaacttgt gggttactgt ctactatggg
121 gtacctgtgt ggaaaaacgc agagaccacc ttatcttctg catcagatgc taaagcatat
181 gaaacagaaa aacataatgt ctgggctaca catgcctgtg taccacaga cccaaccca
241 caagaaatac atttggaaaa tgtgacagaa gactttaaca tgtggaaaaa taacatggta
301 gagcagatgc atacagatat aatcagttca tgggatcaaa gcttaaagcc atgtgttaag
361 ttaacccctc tctgcgttac cttaaatgtt agcaatgtca atgtcactgc cagcagtggt
421 acaacacacg ggggaacggg gagaatgaaa aactgctctt acaatataac cacagagcta
481 agggatagga gacagaaggt atattcactt ttttatagac ttgatgtagt accaattaat
541 gaaaataata gcagtaacaa tgagtataga ttaataaatt gtaatacctc agcattaca
601 caagcttgcc caaaggtaac ctttgagcca attcccatat attattgtgc cccagctggt
661 tttgcgatcc taaagtgtga agataagaag ttcaatggaa caggggccatg cacgaatgtc
721 agcacagtac aatgcacaca tgggaatcaag ccagtagtat caactcaact gctgttaaat
781 ggcagcttag cagaaggaga ggtaatgatt agatctgaaa ataccacaa caatgccaaa
841 aatataatag tacaatttgc cgagcctgta aaatataatt gtaccagacc taacaacaat
v3 loop ->
901 acaagaatga gtatcgtat aggaccagga caagcattct atgcaacagg tgacataata
961 gggaacataa gacaagcaca ttgtaattgt agtagagcag aatggaatac aactttgcaa
<- v3 loop
1021 aaggtagtca caaaatttag ggaatacttt gggaaacaaca aaacaataaa atttgctaac
1081 tcctcaggag gggatctaga aatcacaaca catagtttta attgtggagg agaatttttc
1141 tattgtaat catcaggcct gttaataatg acttggacca acatgcagga gtcaaatagc
1201 acggagtcaa atgacactat aactctccca tgcagaataa aaattataaa catgtggcag
1261 agagcaggac aagcaatgta tgccctctcc atccaaggaa taataaagtg tgtatcaaac
1321 attacaggac taatattaac aagagatggt ggggataata acagtgaag tgaaaccttc
1381 agacctggag gaggagatat gagggacaat tggagaagtg aattatataa gtataaagta
1441 gtaaaaattg aaccactagg agtagcacc accaaggcaa agagaagagt ggtggagaga
1501 gaaaaaagag cagttggaat aggagctgtc ttcttgggt tcttaggagc agcaggaagc
1561 actatggggc cggcgctaat aacgctgac gtacaggcca gacaattatt gtctggcata
1621 gtgcaacagc aaagcaattt gctgagggct atagaggctc aacaacatct gttgaaactc
1681 acggctcggg gcattaaaca gctccaggca agagtcctgg ctgtggaaag atacctaaag
1741 gatcaacagc tcctaggaat ttggggctgc tctggaaaaa tcatctgcac cactaatgtg

```

```

1801 ccoctggaact ctagttaggag caataaatct cagagtgaga tatggggagaa catgacctgg
1861 ctgcaatggg ataaagaaat tagcaattac acagacataa tatataatct acttgaagaa
1921 tcgcagaacc agcaggaaaa gaatgaacaa gacttattgg cattggacaa gtgggcaaat
1981 ctatggaatt ggtttgacat atcaaatctgg ctgtggtata taaaaatatt tataatgata
2041 gtaggaggct taataggatt aagaatagtt tttactgtgc tttctataat aatatagtt
2101 aggcagggat actcaccctt gtcttccAG accataccc cagccccaag gggctctgac
      3' sj (tat,rev) /\
2161 agggccgaaa gaatcgaaga agaagggtgga gagcaagaca gagacagatc gattcgatta
2221 gtgacgggat tottagcact tgcttgggac gatctacgga gctgtgctt cttcagctac
2281 caccgcttga gagacttcat ctgattgoc actgagactg tggaaacttct gggacacagc
2341 agtctcaggg ggttgagact ggggtgggaa ggctcaagt atctgtggaa tctctgttg
2401 tactggggtc aggaactaaa aattagtgt agtaatttgg ttgataccat agcaatagta
2461 gtagctggct ggacagatag ggttatagaa atagcacaag gaataggtag agctattctc
2521 cacataccta gaagaatcag acagggtttt gaaagagctt tgctataAac
      <- env cds end

```

HIVN759

LOCUS HIVN759 342 bp
 DEFINITION Human immunodeficiency virus type 1 (HIV-1) C2-V3 envelope sequence, sample N759 (patient THP01).
 ACCESSION U15576
 SOURCE Human immunodeficiency virus type 1 PCR-amplified PBMC DNA derived from Thai patient THP01
 REFERENCE 1 (bases 1 to 342)
 AUTHORS Kalish,M.L., Luo,C., Weniger,B.G., Limpakarnjanarat,K., Young,K., Ou,C. and Schochetman,G.
 TITLE Early HIV type 1 strains in Thailand were not responsible for the current epidemic
 JOURNAL AIDS Res. Hum. Retroviruses 10, (1994). In press, Nov issue
 COMMENT In this study, samples THP01-THP13 were collected from thirteen consenting Thai inmates, each classified as an IDU. Of these, ten (THP01-THP10) were diagnosed with infectious HIV in 1986 or 1987. The ten viral sequences were found to be similar to B subtype viruses found in Europe and the Americas. These samples were also found to be phylogenetically similar to the variants collected from IDUs infected in 1988 and 1989, however, they are distinguishable from the latter, which the authors designate B'. THP11 and THP12, from patients who were diagnosed with HIV in 1988 and 1992, were found to be subtype B' viruses, not subtype B. The remaining patient, THP13, was infected in 1989 with an E subtype virus. This subtype is predominantly found in Thai cases of sexual transmission. Kalish et al. suggest that the recent prevalence of subtypes B' and E over subtype B in Thailand may be explained by B' and E subtype infections of dynamic high-risk populations. They also propose that conditions in the Thai population may be more conducive to the transmission of subtypes B' and E than subtype B. Hence, phylogenetic comparison of Thai HIV-1 strains isolated in the mid-1980s and Thai HIV-1 strains isolated in the late 1980s indicates that the earlier virus was not responsible for the Thai epidemic beginning in 1988. Other sequences in this set are included in the alignment below (accession numbers U15577-U15588). Sequences kindly provided prior to publication by Dr. Marcia Kalish of the CDC Atlanta, Georgia.

BASE COUNT 145 a 49 c 66 g 82 t
 ORIGIN

```

1  acggacaata ctaaaaccat aatagtacag ctgaacaaat ctatagaaat taattgtaca
61  agacccaaca acaatacaag aaaggggtata catataggac cagggagagc attttataca
121 acaggagaaa tagtaggaga tataaggcaa gcacattgta accttagtag aacagaatgg
181 gataatactt taaaacagat agttggaaaa ttaagagaac aattagggaa caataaaaca
241 atagtcttta atcaatcttc aggagggggac ccagaaattg taatgcacag ttttaattgt
301 ggaggggaat tttctactg taattcaaca cagctgttta at

```

CONSENSUS	ACGGACRATACTAAAACCATAATAGTACAGCTGAATGAATCTATAGAAATTAATTGTACAAGACCCCAACA	70
HIVTHP01	-----CA-----	70
HIVTHP02	--A-----G-----T-----A-----T-----C-----	70
HIVTHP03	-----	70
HIVTHP04	--A-----CC-----	70
HIVTHP05	--A-----T-----	70
HIVTHP06	-----C-C-----	70
HIVTHP07	--A-----	70
HIVTHP08	-----C-C-----	70
HIVTHP09	-----A-NC-G-----	70
HIVTHP10	--A-----GT-----A-----A-----G-----	70
HIVTHP11	-----G-----G-GT-----G-----	70
HIVTHP12	-----G-----G-GT-----C-----G-----	70
HIVTHP13	--AA-----G-C-----G-C-T-----A-----G-----C-G-----C-----TC-----	70
CONSENSUS	ACAATACAAGAAAAGGTATACATATAGGACCAGGGAGAGCATTTTATACAACAGGAGAAATAATAGGAGA	140
HIVTHP01	-----G-----G-----	140
HIVTHP02	-----T-----CC-----G-----NG-----	140
HIVTHP03	-----C-----C-----	140
HIVTHP04	-----A-----C-----G-----C-----	140
HIVTHP05	-----	140
HIVTHP06	-----	140
HIVTHP07	-----	140
HIVTHP08	-----	140
HIVTHP09	-----G-G-----	140
HIVTHP10	-----A-----G-----AG-----G-----GC-----	140
HIVTHP11	-----C-----CA-----GG-----C-----	140
HIVTHP12	-----C-C-----CA-----GG-----C-----	140
HIVTHP13	-----C-A-----AC-----ACA-T-----C-----G-----C-----A-----	140
CONSENSUS	TATAAGACAAGCACATTGTAACCTTAGTAGAACAAAATGG?ATAATGCTTTAAAACAGATAGCT?AAAAA	208
HIVTHP01	-----G-----G-----G-----A-----T-GG-----	210
HIVTHP02	-----A-----G-C-----A-G-G-----T-G-----	210
HIVTHP03	-----G-----C-----G-----A-----	210
HIVTHP04	-----G-C-----G-GG-----AC-----	210
HIVTHP05	-----A-----G-----T-G-C-----	210
HIVTHP06	-----C-----G-----T-A-----	210
HIVTHP07	-----A-----G-----G-C-----	210
HIVTHP08	-----C-----G-----A-----	210
HIVTHP09	-----G-C-----A-----A-----G-C-----T-A-----	210
HIVTHP10	C-----T-----C-A-----G-GA-----A-----CA-C-C-----T-AC-----	210
HIVTHP11	-----C-----A-----CA-----G-----	210
HIVTHP12	-----G-----C-----AC-----CA-----G-----A-G-----	210
HIVTHP13	-----A-----T-----G-GA-----A-G-----A-G-AA-----G-----A-----	210
CONSENSUS	TTAAGAGAACAATTAGGGGAATAAAACA...ATAGTCTTTAATCAATCCTCAGGAGGGGACCCAGAAATTG	275
HIVTHP01	-----C-T-A-ACA-----	280
HIVTHP02	--G-----ATA-----T-----C-----T-----G-----	277
HIVTHP03	--G-----CC-----ACA-----	280
HIVTHP04	-----C-----	277
HIVTHP05	-----G-----T-----	277
HIVTHP06	C-----	277
HIVTHP07	-----T-----	277
HIVTHP08	C-----T-----	277
HIVTHP09	-----C-T-----TC-----G-----	277
HIVTHP10	-----G-TAA-----G-----C-----A-----	277
HIVTHP11	-----T-----C-C-----G-----	277
HIVTHP12	-----T-----C-----C-----G-----	277
HIVTHP13	---AC-G-C-TAA-G-C-T-----A-----C-C-C-C-----A-T-T-T-G-A-----	274

HIVN759

CONSENSUS	TAATGCACAGTTTAAATTGTGGAGGGGAATTTTCTACTGTAATACAAACAACTGTTTAAT	337
HIVTHP01	-----T-----G-----	342
HIVTHP02	-----	339
HIVTHP03	--GC-----T--G-----	342
HIVTHP04	--C-----C-----	339
HIVTHP05	--C-----	339
HIVTHP06	--C-----T-----	339
HIVTHP07	--C-----G-----	339
HIVTHP08	--C-----T-----	339
HIVTHP09	--C-----C-----	339
HIVTHP10	-----T-----T--G--GG-	339
HIVTHP11	-----T-----C-	339
HIVTHP12	-----T-----	339
HIVTHP13	C-----TCA-----A-----T-----AC-T-----	336

LOCUS HIVRJ484 417 bp RNA VRL 25-MAY-1994
 DEFINITION Human immunodeficiency virus type 1, sample RJ484 from Brazil, partial env cds, C2-V3 region.
 ACCESSION U00409
 SOURCE Human immunodeficiency virus type 1 (HIV-1), sample HIVRJ484 from Brazil, PCR amplified proviral DNA extracted from PBMCs.
 REFERENCE 1 (bases 1 to 417)
 AUTHORS Morgado,M.G., Sabino,E., Shpaer,E., Bongertz,V., Brigido,L., Guimaraes,M., Castilho,E., Galvao-Castro,E., Mullins, J.I., Hendry,R. and Mayer,A.
 TITLE V3 region polymorphisms in HIV-1 from Brazil: Prevalence of subtype B strains divergent from North American/European prototype and detection of subtype F
 JOURNAL AIDS Res. and Hum. Retroviruses 10, 569-576 (1994)
 COMMENT Submitted (04-AUG-1993) by Mariza G. Morgado, Instituto Oswaldo Cruz, Immunology, Av Brasil 4365, Rio de Janeiro, RJ 20000, Brazil. In an effort to assess the degree of HIV-1 genomic variation in Brazil, Morgado et al. collected HIV-1 samples from 28 infected individuals in Rio de Janeiro and four infected patients in Sao Paulo. The env V3 regions of these samples were PCR amplified and sequenced. The amino acid alignment of the sequences indicates that 26 of the 28 samples cluster with HIV-1 subtype B. RJ103 clusters with HIV-1 subtype F. Morgado et al. suggest that the remaining sequence, RJ101, may be a recombinant of HIV-1 subtypes B and F. RJ101 clusters with HIV-1 subtype B sequences except for the N-terminal region which is more similar to HIV-1 subtype F sequences (see <HIVRJSP3>). Phylogenetic analysis of the tetrapeptide tip of the V3 loop reveals a high degree of variation between North American/European and Brazilian sequences. Other sequences in this set are included in the alignment below (accession numbers U00400-U00427).
 MCB gi: 392389
 BASE COUNT 176 a 54 c 81 g 106 t
 ORIGIN 798 bp downstream from the beginning of env cds
 1 gaagaagtag tacttagatc tgaaaatttc aaggataatg taaaaccat aatagtagag
 env cds ->
 61 ctgaatgaat ctgtagaat taattgtaca agaccagca acaatacaag aaaaggtata
 V3 loop ->
 121 catataggat gggggagagc attttatgca acaggagaaa taacaggaga tataagaaaa
 181 gcatattgta acattagtag aacaaaatgg gaaaacactt taagacagat agctataaaa
 <- V3 loop
 241 ttagaagaac agtttaagaa caaaacaata atctttaatc aatcctcagg aggagaccca
 301 gaaattgtaa cacacagttt taattgtgga ggggaatttt tctactgtga tacaacaaag
 361 ctgtttaata gtacttgga tacaacacag ctgtctaata gttttaatag tactggg
 env ->
 CONSENSUSG?AGAGGTAGTAATTAGATCTGAAAATTTCACGAACAATGCTAAAACCATAATAG 54
 HIVRJ478AGAA-A-----G-----AG-----G----- 59
 HIVSPB4AGCA-A-A-----G-----AG-----G----- 59
 HIVSPB3AA-----A-----T-G-----A----- 50
 HIVSPB2-A-----A-----G-----A-----T----- 53
 HIVSPB1-A-----A-----A-----A-----A----- 53
 HIVRJ104A-----A-----A-----CTA---T-A----- 40
 HIVRJ103 AGCAGTCTAGCAGAA-A---TA---A---C---C-T-C---A---T-AG-T---A----- 70
 HIVRJ102-A---G-C-----A-----T----- 53
 HIVRJ101-A-A---C---C---A-T-G-T---A---T----- 53
 HIVRJ70-G-----A-----A----- 0
 HIVRJ636-G-----A-----T----- 47
 HIVRJ626AGAA-G-----A-----A----- 59
 HIVRJ625A-A-----T---G-----T----- 56

HIVRJ484

HIVRJ623	AGAA-G-----G-C-----G-----T-----	59
HIVRJ62	AGAA-G-----G-C-----T-----	59
HIVRJ59	AGAA-G-----C-T-----	59
HIVRJ54	A-AG-----A-G-----	56
HIVRJ49	AGAG-A-----A-----	59
HIVRJ484	A-A-----C-----G-T-----T-----	55
HIVRJ483	A-A-----C-----GT-----	55
HIVRJ477	A-----T-----	47
HIVRJ379	G-----C-C-----G-----T-----	49
HIVRJ27		0
HIVRJ19	A-G-----G-----A-----T-----	50
HIVRJ17	A-----	50
HIVRJ14	AGAA-A-----G-C-----G-----G-T-----	59
HIVRJ12	AGAA-A-----A-----C-----	59
HSURJ64	A-----C-T-A-----	55
CONSENSUS	TACAGCTGAACGAATCTGTAGAAATTAATTGTACAAAGACCCAACAACAATACAAGAAAAAGTATACA.TA	123
HIVRJ478	-----A-----C-G-----	128
HIVSPB4	-----AC-C-----T-----A-AGAG--	128
HIVSPB3	-----A-A-----G-----T-----A-----G-----	119
HIVSPB2	-----T-A-----G-----	122
HIVSPB1	-----T-----T-----G-----	122
HIVRJI04	-----T-----T-----G-----	109
HIVRJI03	-----CT-T-T-----C-G-G-----T-----C-T	139
HIVRJI02	-----G-----G-----	122
HIVRJI01	-----C-T-T-----A-----A-----	122
HIVRJ70	AT-CGTGAAG-----G-----G-----T	59
HIVRJ636	-----A-TC-----T-----C-----	116
HIVRJ626	-----T-----A-C-----G-----	128
HIVRJ625	-----T-A-----	125
HIVRJ623	-----A-----G-----G-----T	128
HIVRJ62	-----CT-----T	128
HIVRJ59	-----G-GG-----T	128
HIVRJ54	-----T-----T-----	125
HIVRJ49	-----T-----T-C-----T-G-----T-----	128
HIVRJ484	-----T-----G-----G-----	124
HIVRJ483	-----T-----G-----	124
HIVRJ477	-----T-----G-----	116
HIVRJ379	-----T-----T-----GG-----	118
HIVRJ27	A-----G-----	41
HIVRJ19	-----T-----G-----	120
HIVRJ17	-----T-----	119
HIVRJ14	-----C-----A-----G-----G-----	128
HIVRJ12	-----A-----T-----G-----G-----C-----	128
HSURJ64	-----A-----G-----T-----	124
CONSENSUS	TAGGAT?GGGG..AGAGCATTTTATGCAACAGGAGAAATAATAGGAGATATAAGACAAGCACATTGTAAAC	190
HIVRJ478	-G--G-----C-A--AT-G-----A-----	196
HIVSPB4	---CCA---A-----G-----	196
HIVSPB3	-G--G-----A-----AG-----T	187
HIVSPB2	-G--G-----A-----T-----	190
HIVSPB1	-G--G-----G-----	190
HIVRJI04	---TT---A-C--T-----C-----	177
HIVRJI03	---CCA-A..C-----C-C--A-G-C-----A	207
HIVRJI02	---GGT---A-----	190
HIVRJI01	---CCA-T-----C-----GT-----	190
HIVRJ70	---G---A-----	124
HIVRJ636	---C-CCA---C---GG-T-----	184
HIVRJ626	---CCA---C-----C-----	196

HIVRJ625	-----CCT-----GG-T-----G-----GC-CC-C-----	193
HIVRJ623	-----CCA-----G-----A-----A-----A-----TT-----C-----	196
HIVRJ62	-----G-----G-----G-----G-----G-----C-----	196
HIVRJ59	-----G-----CT-----C-----	196
HIVRJ54	-----CCA-----A-----	193
HIVRJ49	-G-----G-----A-----T-----T-----	196
HIVRJ484	-----G-----C-----A-----T-----	192
HIVRJ483	-G-----CTA-C-----G-----G-----	192
HIVRJ477	-G-----T-----C-----T-----	177
HIVRJ379	-----CCA-C-----A-----C-----	186
HIVRJ27	-----CCA-----G-GGG-CAT-AT-AG-AGA-----	106
HIVRJ19	-G-----CG-----GAG-----G-----G-----	190
HIVRJ17	-G-----CC-----G-----A-----A-----TC-----GC-----	184
HIVRJ14	-----G-----G-----A-----G-----TA-----C-----	196
HIVRJ12	-----G-----A-----C-----C-----	196
HSURJ64	-----G-----A-----A-----C-G-----A-----	192
CONSENSUS	CTTAGTAGAACAAAATGGAA?AACACTTTAAAACAGATAGCTA?AAAATTAAAGAGAACAAATTAA?AATA	257
HIVRJ478	-----G-----T-G-----C-----C-----GA-----GGG-----	266
HIVSPB4	-----G-A-G-----TG-T-----T-----T-C-G-----C-----G-----	266
HIVSPB3	A-----C-A-GT-GC-----C-GA-----GG-AG-----T-----GA-C-----T-CA-----	257
HIVSPB2	A-----	196
HIVSPB1	-----C-----G-CGC-----T-G-----G-----CGC-----	232
HIVRJ104	A-----G-----G-----C-----G-----C-A-A-----TG-----A-----T-GC-----	247
HIVRJ103	G-CA-G-----C-----T-A-TG-----G-----G-AAAGC-CG-CAGTCT-T-CTCT-G-----	277
HIVRJ102	A-----GA-----G-----A-----G-GA-----A-----A-----G-GTG-----	260
HIVRJ101	A-----C-----G-----C-----T-----T-----G-----A-----TC-----G-----G-----G-----	260
HIVRJ636	-----AGGT-C-C-----G-A-T-----A-----A-----G-----G-A-----	254
HIVRJ626	-----G-AC-----C-----G-----G-----CA-----T-----	266
HIVRJ625	-----C-----T-----CC-----T-----T-----CA-----GGG-----	263
HIVRJ623	-----A-G-----G-A-T-----A-----GA-----T-A-----	266
HIVRJ62	-----T-----T-----A-----GGG-----	266
HIVRJ59	-----AC-A-----G-A-A-----G-----T-----T-----CTC-----	266
HIVRJ54	-----C-----G-----C-----T-----T-----T-----C-----GGG-----	263
HIVRJ49	G-----GA-G-T-----A-----G-----G-----T-----GG-G-----	266
HIVRJ484	A-----G-----G-A-----G-----T-----GA-----G-----G-C-----	262
HIVRJ483	A-----G-----G-----G-----A-----A-----T-----	262
HIVRJ379	-----G-----TG-----G-----T-----G-----TGGA-----	256
HIVRJ19	A-----A-GA-----C-----C-----A-----T-A-----	260
HIVRJ14	A-----G-----GGA-----TG-----GC-----G-A-----G-----T-A-----	266
HIVRJ12	-----G-A-----G-A-----G-----A-----T-A-----	266
HSURJ64	--C-----G-----G-A-A-----T-----GOC-----	262
CONSENSUS	AA...ACAATA?TCCTTAATCAATCCTCAGGAGGGGACCCAGAAATTGTAAT?CACAGTTTAAATTGTGG	322
HIVRJ478	-----A-----C-----C-----	301
HIVSPB4	-----G-----T-----T-----	333
HIVSPB3	C-----TC-----AAA-----CTC-----A-----	290
HIVRJ104	C-----G-----A-----CA-GA-----GA-----	281
HIVRJ103	C-----AAA-----CCTC-----	297
HIVRJ102	-----A-----TC-----C-G-----G-----	301
HIVRJ101	-----A-----GC-----CC-----	299
HIVRJ636	-----CAA-----	299
HIVRJ626	-----A-----C-----G-----	333
HIVRJ625	-----G-----A-----G-----A-----GA-CAG-A-TGT-----	310
HIVRJ623	C-----GC-----G-----AA-----	301
HIVRJ62	-----GAA-----G-----AA-----	301
HIVRJ59	-----AC-----T-----	301
HIVRJ54	-----GC-----TC-----	301
HIVRJ49	-----C-----G-----A-----A-----	301

HIVRJ484

HIVRJ484	--...--A-----A-----CA-----	329
HIVRJ483	--...--G-----A-----G-----G-----	329
HIVRJ379	C-GAA-----A-----A--CAG	299
HIVRJ19	C-.....G-----	290
HIVRJ14	C-.....AAA---C-G.CCG-----A-----	300
HIVRJ12	C-.....G-----	311
HSURJ64	--...--G--A-----GA--CAG--A-TGT-AT	312
CONSENSUS	AGGGGAATTTTCTACTGTAATACATCAA??CTGTTTAATAGTACTTGGAAATACAACACAGCTGTCTAAT	390
HIVSPB4	-C-A--	339
HIVRJ626	-----	361
HIVRJ484	-----G---A--AG-----	399
HIVRJ483	-----CA-----	378
CONSENSUS	AGTTTTAATAGTACTGGG	408
HIVRJ484	-----	417

LOCUS HIVRJSP3 1270 bp DNA VRL 26-MAY-1994
 DEFINITION Human immunodeficiency virus type 1, sample RJ101, clone SP3 from Brazil, envelope glycoprotein, partial cds (V1-V5).
 ACCESSION U08975
 SOURCE Human immunodeficiency virus type 1 (HIV-1), sample RJ101, clone SP3, from Brazil, PCR amplified from proviral DNA.
 REFERENCE 1 (bases 1 to 1270)
 AUTHORS Sabino,E.C., Shpaer,E.G., Morgado,M.M., Korber, B.T.M., Diaz,R.S., Bongertz,V., Cavalcante,S., Galvao-Castro,B., Hendry,M., Mullins,J.I. and Mayer,A.
 TITLE Identification of HIV-1 proviral genomes recombinant between subtypes B and F in two epidemiologically-linked individuals from Brazil
 JOURNAL J. Virol. 68, 6340-6346 (1994)
 COMMENT Submitted (15-APR-1994) Allen Mayer, Irwin Memorial Blood Centers, Scientific Service, 270 Masonic Ave., San Francisco, CA 94118, USA. Sequence analysis of the gp120 and C2-V3 regions of sample RJ101 indicates a double recombinational event. A region of the C2-V3 gene of sample RJ101 clusters with HIV-1 subtype F while a large percentage of the gp120 region clusters with the HIV-1 B subtype. PBMC DNA extracted from the subject's sexual partner also coded the recombinant env genes. According to the authors, sample RJ101 reveals the plausibility of dual infection and recombination of distinct subtypes. Other sequences in this set are included in the alignment below (accession numbers U10019-U10026, U08953-U08960 and U08962-U08975). RJ103 is a Brazilian F subtype sample.
 NCBI gi: 487388

BASE COUNT 484 a 203 c 269 g 314 t

ORIGIN 345 bp downstream from the beginning of env cds

1 aaagccatgt gtgcaattaa cccactctg tgttacttta cattgtactg attgtaattgc
 env cds ->

51 taccaatatc aataatagca ccagtagtag tagttgcaat accaataata ccagcaccac
 121 taatagtagt ttgggggaca tggagggaga aataaaaacc tgctctttta atatcacctc
 181 agatataaaa gataaggtgc gaaaagaaca tgcacttttt aatagacttg atatagtacc
 241 aatagataag aataatagta aaagctatat gttgataaat tgtaacacct cagtccattac
 301 tcaggcctgt ccaaggtat cctttgaacc aatcccata cattattgta ccccggtctg
 361 ttttgcgatt ctaaagtgt atgateagaa gttcaatgga acaggattct gtaaaaatgt
 421 cagtcagtg caatgtacac atggaattga gccagtggtg tcaactcaac tgctgttgaa
 481 tggcagtgta gcagaagaag atatagtaat taaatccgaa aatttctogg acaatgctaa
 541 aacaataata gtacagctga acgaaactgt aaaaattgat tgtataagac caaacaacaa

V3 loop ->

601 tacaagaaa ggtatacata tgggatgggg gagaacattt tatgcaacag gaagaataat
 661 aggagatata agacaagcac attgtaattc tagcaaagta gcatggaaca gaactttaga

<- V3 loop

721 aaggatagct ataanaattaa ggaaccaatt taattacaac aatgataaaa attttaatca
 781 atcctcagga ggggacccag aaattgtaac gcacagtttt aattgtggag ggggaatttt
 841 ctactgtgat acaactcacc tgtttaatag tgcttggact aatgagagta ggactggcag
 901 agtcaaatgc actgaaaata ccacactctc atgcagaata aaacaatta taacaagtg
 961 gcagggaagta ggaaagcaaa tgtatgccc tccatcagt ggacaaatta gctgttcacc
 1021 aaatattaca gggctgctat tgacaagaga tgggtgtaag aataacgggc ccaacgagtc
 1081 cgaatcttc agacctggag gaggagatat gaagggcaat tggagaagtg aattatataa
 1141 atataaagta gtaaaaattg aaccattagg aatagcacc accaaggcaa ggagaagagt
 1201 ggtgcaaga caaaaagag cagtgggaac gctaggagac atgttccttg ggttcttggg
 1261 agcagcagga

env ->

CONSENSUS	CGATCAAAGCCTAAAGCCATGTGT?GAATTAACCCC?CTCTGTGTTACTTTACATTG?ACTGATGCTAAC	67
HIVSP3	-GC-----A-----T-----TG---T	58
HIVRJ103	-GA-G-----A-----A-----TT-----C-C-	60
HIVRJ101-10	-A-----G-----C-----	69

HIVRJSP3

HIVRJ101-05	-----A-----G-----C-----	70
CONSENSUS	?CCACTAGAAG?.....GGTACTAATAACACCAGTACTAATAGT.....ACTGTTAACA	114
HIVSP3	G-T--C-AT-TC.....-A-----G-----G-----G-----TGCAATACCAAT-A-ACC-G--	116
HIVRJ103	GT---CC---C.....A-A--CCTG.....	81
HIVRJ101-10	A-----T.....	118
HIVRJ101-05	A-----T.....	119
CONSENSUS	CCGCTGTTAGTAGTGGGGGAAAGATGGAGGGAGAAATAAAAAA?TGC?CTTTCAATATCACCACAAGCAT	182
HIVSP3	--AA-AA-----TT--G--C-----CC--T--T-----T--GAT--	186
HIVRJ103	G-AGA-CCA--G-C---C---C--TT-----G--GG-----A-G	131
HIVRJ101-10	-----G-----T--A-----	188
HIVRJ101-05	-----T--A-----	189
CONSENSUS	AA?AGATAAGGTGCAG?AAGAA?ATGCACCTTTTATATA?ACTTGAT?TAGTACCAATAG?CAATAATAGT	246
HIVSP3	--A-----GAA--C-----A--G-----A-----AT--G--A--	256
HIVRJ103	T-A-----CA--A--T-C-----GC-----A-----TAG--G--C---	201
HIVRJ101-10	--G-----G-----T-----A-----G-----A-----	258
HIVRJ101-05	--G-----G-----T-----A-----G-----G-----	259
CONSENSUS	A?T...AGTTATAGGTTGATAAATTGTAACACCTCAGTCATTA?ACAGGCCCTGTCCAAAGGTATCCTTTG	311
HIVSP3	-G-AAA--C---T-----CT-----	326
HIVRJ103	GGA...GAA--C--C-A--C-----T-----AC---C-----T-----T-GG-	268
HIVRJ101-10	-C-...-----A-----	325
HIVRJ101-05	-C-...-----A-----	326
CONSENSUS	AGCCAATTCCCATACATTATTGTT?CCCGGCTGGTTTTCGGATTCTAAAGTGTAATGATAAGAAGTTCAA	380
HIVSP3	-A---C--A-----AC-----	396
HIVRJ103	-T-----GCT--A-----A-----A-----	338
HIVRJ101-10	-----G-----	395
HIVRJ101-05	-----G-----A-----	396
CONSENSUS	TGGAACAGGACCATGTAAAAATGTCAGCACAGTACAAATGTACACATGGAATTAAGCCAGTGGTATCAACT	450
HIVSP3	-----TTC-----T---G-----G-----G-----	466
HIVRJ103	--G---G---C-CG-----A-----	408
HIVRJ101-10	-----	465
HIVRJ101-05	-----	466
CONSENSUS	TGGAACAGGACCATGTAAAAATGTCAGCACAGTACAAATGTACACATGGAATTAAGCCAGTGGTATCAACT	519
HIVRJ548-54	CTAC-TG-----	42
HIVRJ548-35	CTAC-TGTG-----C---A-----	42
HIVRJ548-82	GCAGTCTAGCAGAAGAAGA-----A-----	54
HIVRJ548-2	CTAC-TG-----A-----	42
HIVRJ548-169	CTAC-TG-----A-----	42
HIVRJ548-160	CTAC-TG-----A-----	42
HIVRJ548-155	CTAC-TGA---G-----	42
HIVSP3	CAACTGCTGTTGAATGGCAGTCTAGCAGAAGAAGAT--G--T-A--CG---T---C---	536
HIVRJ103	CAATTGTTGTTAAATGGCAGCCTAGCAGAAGA-GAT-----C-T-----A-----	478
HIVRJ101-10	CAATTGTTGTTAAATGGCAGCCTAGCAGAAGAAGA-----	535
HIVRJ101-05	CAATTGTTGTTAAATGGCAGCCTAGCAGAAGAAGA-----	536
HIVRJ549-113		35
HIVRJ549-111	CTAC-TG-----A-----	42
HIVRJ549-94	CTAC-TG-----	42
HIVRJ549-83	CTAC-TG-----A-----	42
HIVRJ549-78	CTAC-TG-----A-----	42
HIVRJ549-38	CTAC-TG-----	42
HIVRJ549-2		35
HIVRJ548-153	CTAC-TGT-----	42
HIVRJ548-147	CTAC-TGA-----A-----	42
HIVRJ548-130	A-----A-----	35
HIVRJ548-123	A---A---A-----	35
HIVRJ548-116		35

HIVRJ548-104	A-----	35
HIVRJ548-101	TG-----	35
HIVRJ548-100	TG-----	35
HIVRJ548-1	CTAC-TG-----	42
HIV1U08954	A-----A-----	35
HIVRJ548-95	-C-----A-----	33
HIV1U08953	A-----A-----A-----	35
CONSENSUS	CAAAAAGCATAATAGTACACCTTAATGAATCTGTAGAAATTAATTGTACAAGACCCAACAACAATACAAG	586
HIVRJ548-84	-----A-----	112
HIVRJ548-35	-----T-----G-----T-----	112
HIVRJ548-82	-----A-----A-----	124
HIVRJ548-2	-----	112
HIVRJ548-169	-----A-----G-----T-----	112
HIVRJ548-160	-----C-----	112
HIVRJ548-155	-----G-----	112
HIVSP3	-T---CA-----G-G-C-A---A---G---T---A---	606
HIVRJIO3	-----C-----T-----C-G-G-----T---	548
HIVRJIO1-10	-----T-----	605
HIVRJIO1-05	-----T-----A-----	606
HIVRJ549-113	-----T-G-----	105
HIVRJ549-111	-----A-----G-----	112
HIVRJ549-94	-----T-----	112
HIVRJ549-83	-----A-----G-----	112
HIVRJ549-78	-----A-----G-----	112
HIVRJ549-38	-----T-----	112
HIVRJ549-2	-----T-----	105
HIVRJ548-153	-----T-----A-----	112
HIVRJ548-147	-----A-----	112
HIVRJ548-130	-----C-----	105
HIVRJ548-123	-----T-----A-----GT-----	105
HIVRJ548-116	-----T-----G-----	105
HIVRJ548-104	-----G-----	105
HIVRJ548-101	-----T-----	105
HIVRJ548-100	-----T-----C-----	105
HIVRJ548-1	-----C-----G-----	112
HIV1U08954	-----A-----	105
HIVRJ548-95	-----	103
HIV1U08953	-----T-----A-----GT-----	105
CONSENSUS	AAAAAGTATACATATAGGACCAGCTAGAGCATTTTATGCAACAGGAGACATAATAGGAGATATAAGACAA	654
HIVRJ548-54	-----A-----C-A-----C-----C-----	182
HIVRJ548-35	-----	182
HIVRJ548-82	-----C-----T-----	194
HIVRJ548-2	-----	182
HIVRJ548-169	-----	182
HIVRJ548-160	-----	182
HIVRJ548-155	-----A-----	182
HIVSP3	---G---G---TGG-G-A---AGA---	676
HIVRJIO3	-----C-T-----AC-----A-----C-C-A-G	618
HIVRJIO1-10	-----A-----	675
HIVRJIO1-05	-----A-----	676
HIVRJ549-113	-----A-----	175
HIVRJ549-111	-----	182
HIVRJ549-94	-----A-----	182
HIVRJ549-83	-----	182
HIVRJ549-78	-----	182
HIVRJ549-38	-----A-----	182
HIVRJ549-2	-----A-----	175

HIVRJSP3

HIVRJ548-153	-----A-----A-----A-----A-----	182
HIVRJ548-147	-----A-----A-----A-----A-----	182
HIVRJ548-130	-----A-----A-----A-----A-----	175
HIVRJ548-123	-----A-----A-----A-----A-----	175
HIVRJ548-116	-----A-----A-----A-----A-----	175
HIVRJ548-104	-----A-----A-----A-----A-----	175
HIVRJ548-101	-----A-----A-----A-----A-----	175
HIVRJ548-100	-----A-----A-----A-----A-----	175
HIVRJ548-1	-----A-----A-----A-----A-----	182
HIV1U08954	-----A-----A-----A-----A-----	175
HIVRJ548-95	-----A-----A-----A-----A-----	173
HIV1U08953	-----A-----A-----A-----A-----	175
CONSENSUS	GCGTATTGTAACATTAGCAGAGCACAAATGGAATAATACTTTAGAACAAATAGCTATCAAATTA?AAGAAC	719
HIVRJ548-54	-----A-----C-----A-----A-----GG-----	252
HIVRJ548-35	-----A-----C-----A-----A-----GG-----	252
HIVRJ548-82	---C---T---A-A-T---AG---C---G---	264
HIVRJ548-2	-----A-----C-----A-----G-----	252
HIVRJ548-169	-----G-----ATA-----GG-----	252
HIVRJ548-160	---C---C---A---T---A-A---C---	252
HIVRJ548-155	-----A-----AA-----C-----	252
HIVSP3	--AC-----TC---A-T-GC---C-GA---AGG---A---AGGA-C-	746
HIVRJIO3	--CC-----AG-CATG-A-T---A-TG---GG-AAGGCA-CG---C-GTCT-	688
HIVRJIO1-10	-----A-----A-----A-----GG---G-	745
HIVRJIO1-05	-----A-----A-----A-----GG---G-	746
HIVRJ549-113	-----G-----T-T-G---A---	245
HIVRJ549-111	-----G-----T-G---A---	252
HIVRJ549-94	-----G-----T-G---A---	252
HIVRJ549-83	-----G-----T-G---A---	252
HIVRJ549-78	-----G-----T-G---A---	252
HIVRJ549-38	-----G-----T-G---A---	252
HIVRJ549-2	-----G-----T-G---A---	245
HIVRJ548-153	-----A-----AG-----G-----	252
HIVRJ548-147	-----A-----A-----G-----A-----A-A-----C-----	252
HIVRJ548-130	-----C-----A-----A-----A-----A-A-----C-----	245
HIVRJ548-123	-----AA-----AG-----A-----	245
HIVRJ548-116	-----A-----T-G---A---	245
HIVRJ548-104	-----A-----A-----GG-----	245
HIVRJ548-101	-----A-----A-----C-----	245
HIVRJ548-100	-----A-----A-----C-----	245
HIVRJ548-1	-----A-----A-----A-A-----C-----	252
HIV1U08954	-----A-----G-----A-----A-A-----C-----	245
HIVRJ548-95	---C-----A-----A-----GG-----	243
HIV1U08953	-----AG-----A-----	245
CONSENSUS	AATTTAAGAATAAA...AAAATAGTCTTTAACCAATCCTCAGGAGGGGACCCAGAAATTGTAATGCACAG	781
HIVRJ548-54	-----289-----	
HIVRJ548-35	-----289-----	
HIVRJ548-82	---C-A-C---A---TAGTAGTA-T-GC-	331
HIVRJ548-2	---G-T---A---289-----	
HIVRJ548-169	---A---289-----	
HIVRJ548-160	---C-A-C---A---289-----	
HIVRJ548-155	---C---C---A---289-----	
HIVSP3	---TT-C---CAATG-T-A-AAT---T-----	816
HIVRJIO3	-T-CTCT-GC-...-C---AAA---C-TC-----T-----AC-----T-	755
HIVRJIO1-10	---C---C-----C	812
HIVRJIO1-05	---C-----C	813
HIVRJ549-113	---C-----C	282
HIVRJ549-111	---C-----C	289

HIVRJ549-94	-----	289
HIVRJ549-83	-----C-----C-----	289
HIVRJ549-78	-----C-----C-----	289
HIVRJ549-38	-----	289
HIVRJ549-2	-----	282
HIVRJ548-153	-----C-A-C-----	289
HIVRJ548-147	-----C-A-C-----C-----A-----	289
HIVRJ548-130	-----C-A-C-----A-----	282
HIVRJ548-123	-----A-----A-----	282
HIVRJ548-116	-----C-----A-C-----	282
HIVRJ548-104	-----	282
HIVRJ548-101	-----C-A-C-----A-----	282
HIVRJ548-100	-----C-A-C-----A-----	282
HIVRJ548-1	-----C-A-C-----A-----	289
HIV1U08954	-----C-A-C-----C-----A-----	282
HIVRJ548-95	-----G-----C-----	280
HIV1U08953	-----A-----A-----	282
CONSENSUS	TTTAAATGTGGAGGGGAATTTTCTACTGTAAT?CAACAGAACTGTFTAATGGTACTTGGACTAATG?T	847
HIVRJ548-82	CCG-CTAGAG-T-CCAAGC--ACGTCGT--C-TGCACGT-TGTT-TAAGG--CTA-T-CATT--GC--T-	401
HIVSP3	-----G-A-----TC-T-----A-G-----AG	886
HIVRJ103	-----A-A-----C-A-T-----AC-AGTAT-C---AC	925
HIVRJ101-10	-----T-----G-----A---G-	882
HIVRJ101-05	-----T-----G-----A---G-	883
CONSENSUS	ACTAGGA?TGGTACT...?AAGGAAATAACACAATCACACTTCCATGCAGAAATAAACAAATTATAAACA	911
HIVRJ548-82	CAC----TG-AAC--GCCTCC	422
HIVSP3	-G---C---C-GAGTCA--T-C-C-G-A-AT-C-----CT-----	956
HIVRJ103	---TC-C-CT-C-A.....-TC-----G-----	865
HIVRJ101-10	---T---A---G-----	949
HIVRJ101-05	---T---A---G-----	950
CONSENSUS	TGTGGCAG?AAGTAGGAAGAGCAATGTATGCCCTCCCATCAGAGGACAAATTAG?TGTTTCATCAAATAT	979
HIVSP3	A-----G-----A-----T-----C-----	1026
HIVRJ103	-----G-----C-----G-----TGC---A-C---CC---AGC-----	935
HIVRJ101-10	-----A-----A-----	1019
HIVRJ101-05	-----A-----A-----	1020
CONSENSUS	TACAGGCTGT?TATT?A?AAGAGATGGTGAATAA?AAT...?...?AACCAGA?CGAGATCTTCAGA	1035
HIVSP3	-----C-----G-C-----G-T-C...GGGCCC-----TC---A-----	1093
HIVRJ103	---T---AC---G-C-----T---TG---AATGCTACTC-GAC---CT---C-----	1005
HIVRJ101-10	-----A---A-T-----C-----G-----	1080
HIVRJ101-05	-----A---A-T-----C-----G-----	1081
CONSENSUS	CCTGGAGGAGG?GATATGA?GGACAATTGGAGAAGTGAATTATATAAATATAA?GTAGTAAAAATTGAAC	1102
HIVSP3	-----A-----A-G-----A-----	1163
HIVRJ103	---G---AA-----AA-----G---A-G---G-----	1075
HIVRJ101-10	-----G-----G-----G-----	1150
HIVRJ101-05	-----G-----G-----G---C-----	1151
CONSENSUS	CATTAGGA?TAGCACCCACCA?GGCAAGGAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGA...AT	1167
HIVSP3	-----A-----A-----A---C-----ACGC-	1233
HIVRJ103	--C---A-----A---AA---CA---A-----C-----	1142
HIVRJ101-10	-----G-----G-----	1217
HIVRJ101-05	-----G-----G-----	1218
CONSENSUS	AGGAGCTATGTTCTTGGGTTCTTGGGAGCAGCAGGAAGCACTTAGT	1214
HIVSP3	-----C-----	1270
HIVRJ103	---G---T-----	1183
HIVRJ101-10	-----	1264
HIVRJ101-05	-----	1264

HIVSE364A

LOCUS HIVSE364A 3154 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, SE364 from Senegal,
 proviral DNA encoding env, vpU, rev, and nef genes.
 ACCESSION L22944
 SOURCE Human immunodeficiency virus type 1 (HIV-1), SE364 from Senegal,
 PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 3154)
 AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A.,
 van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr.
 Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory
 Rockville, Maryland. Twenty-one full length gp160 coding sequences
 from eight African countries (Djibouti, Gabon, Kenya, Senegal,
 Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with
 thirty-two previously published full-length gp160 env sequences.
 Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274,
 VI191, and VI525 have corresponding gag sequences published by
 Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993
 compendium. This sequence clusters with C subtype env sequences. See
 also accession numbers L22939-L22943, L22945-L22957 and L23064-L23065.
 Nucleotides 35 through 104 of this sequence have a perfect match to
 nucleotides 1 through 70 of a sequence, <BLVNADAA>, (Mansky and Temin,
 J. Virol. 68:494-499, 1994) that appears to be derived from a cloning
 vector related to pSG930 (U14118). This makes the first exons of tat
 and rev for this sequence problematic.
 NCBI gi: 437349
 BASE COUNT 1124 a 527 c 747 g 756 t
 ORIGIN 38 bp upstream from the beginning of rev cds start
 1 ccaagctatt taggtgacac tatagaatac tcaagctATG catcaagctt ggtaccgagc
 rev cds start ->
 insertion ->
 61 tcggatccac tagtaacggc cgccagtgtg ctggaattcg gcttaggata atcaaatcc
 <- insertion
 121 tatatcaaag caGtaagtac tacatagtag ATGtaattggt agatttacta gccaaaagta
 5' sj /\ vpU cds start ->
 181 gattatagat taggagtagg agcattgata gtagcactaa tcatagcaat aatttggtg
 241 gatcatagca tatatagaat ataggaaatt gttaagacaa aagaaattag attggttagt
 301 taagagaatt agggaaagag cagaagacag tggcaATGag agtgacgggg atgttgagg
 env cds start ->
 361 attgtcaacc atggtggata tggggcatct taggcttttg gatgcatcg atttaTAATg
 <- vpU
 421 tgggggggaa cttgtgggtc acagtatatt atgggggtacc tgtgtggaaa gaagcaaaa
 481 ctactctatt ttgtgcatcg gatgctaaag catatgagag agaagggcat aatgtctggg
 541 ctacacatgc ctgtgtaccc acagaccoca acccacaaga aatggtttgg gaaaatgtaa
 601 cagaatattt taacatgttg aaaaatgaca tggtagatca gatgcatgag gatataatca
 661 gtttatggga tcaaagccta aagccatgtg taaagttgac cccactctgt gtcactttta
 721 actgtaggaa tgttaccacc agcaataatg ccaccagtaa tgataacct aatggagaaa
 781 taaaaaattg ctctttcaat ataaccacag aactaagaga taagagaagg aatgaatatg
 841 cactttttta tagacttgat atagtaccac ttagtggggag taaaaacagt tccaattcca
 901 gtgagtatag atteataaat tgtaatacct cagccataac acaagcctgt ccaaaggtct
 961 ctttggaccc cattcctata cattattgtg ctccagctgg ttatgcgatt ttaaagtgt
 1021 ataataagac attcaatgga acaggacctt gcaataatgt cagcacagta caatgtacac
 1081 atggaattaa gccagtggtg tcaactcaac tactgtttaa tggtagtcta gcagaaggag
 1141 aaataataat tagatctgaa aatctgacaa ataatgccaa aacaataata gtacatctta
 1201 atgaatctat agaaattgtg tgtgcaagac ccaataataa tacaagaaaa agtatgagg
 v3 loop ->
 1261 taggaccagg acaaaccctt tatgcaacag gagacataat aggatata agacaagcac

```

1321 attgtaacat tagtggtaac tggaaatgcaa ctttagaaaa ggtaaaagga aaattacaag
      <- v3 loop
1381 aacacttccc tggtaaaaat ataagetttg aaccatcttc aggagggggac ctagaattta
1441 caacacacag ctttaattgt agaggagaat tttctatttg cgatacatca aagctgttta
1501 atgggtacaac acataactgca aattcaazaca tcacaatcca atgcagaata aaacagatta
1561 taaacatgtg gcagggagta ggacgagcaa tatatgcccc tccatttgcg ggaacataa
1621 catgtaaatc aaatatcaca ggactactac tgacacgtga tggagggaaca actaataata
1681 atacagagat attcagacct ggaggaggag atatgaggga caattggaga agtgaattat
1741 ataaatataa agtagtagaa attaagccat taggaatagc acccataag gcaaaaagga
1801 gagtgggtgga gagagaaaaa agagcagtggt gaataggagc tgtgttcttt ggggttcttg
1861 gagcagcagg aagcactatg ggcgcggtat caataacgct gacggtagac gccagacaat
1921 tgttgtcttg tatagtgcga cagcaaaagca atttgtctgag agctatagag ggcgaacagc
1981 atatgttgca actcacagtc tggggcatta agcagctcca gacaagagtc ctggctatag
2041 aaagatacct aaaggatcaa cagctctag ggatttgggg ctgctctgga aaactcatct
2101 gcaccactgc tgtgccttgg aactctagtt ggagtaataa atctctagag gatatttggg
2161 ataacatgac ttggatgcag tgggatagag aaattaataa ttacacaagt ataatataca
2221 gcttgcttga agaatcgcaa aaccagcagg agaaaaatga aaacgattta ctagcatttg
2281 acaattggaa caatctgttg aattgtttta acataacaaa gtggctatgg tatataaaaa
2341 tattcataat gatagtagga ggcttgatag gtttaagaat aatttttgcg gtgctctcta
2401 tagtgaatag agtcaggcag ggatactcac ctttgtcgtt ccAGaccctt atcccaaac
      3' sj (tat, rev) /\
2461 cgagggggacc cgacaggctc ggaagaatcg aagaagaagg tggagagcaa gacagagaca
2521 gatccattcg attagttagc ggattcttgt ccattgcctg ggaagacctg cggagcctgt
2581 gctctttcag ctaccaccga ttgagagact tcattattgat tgtagcgaga gcagtgggaa
2641 tcttgataca ggggggttgg gaaatcctta agtacctggg aagtcttgca cagtattggg
2701 gtttggaaat aaaaaagagt gctatTAGtc tgcttgatac catagcaata acagttagctg
      <- rev cds end
2761 aaggaacaga tagaattata gaattagtac aaagaatctg tagggctatc agcaacatac
2821 ctagaagaat aagacaggga tttgaagcag cgttgcaatA AaATGggggg caagtggta
      <- env cds end
      nef cds start ->
2881 aaatgcagta tagtgggatg gccagaggta agagaaagaa taagaagaac taatccagca
2941 gcaacgggag taggagcagc gtctcaagac ttagataaat atggggcact tacaagcagc
3001 aacacagccg ccactaatga ggcttgtgoc cagctggaag cgcaagagca ggaagaagag
3061 gtaggctttc cagtcagacc tcaggtaacct ttaagaccaaa tgacttataa gggagcattt
3121 gatctcagcc atttttttaa agaaaagggg ggac

```


HIVSE3651

LOCUS HIVSE3651 1474 bp ss-DNA VRL 27-MAY-1993
 DEFINITION Human immunodeficiency virus type 1, Senegalese isolate SE365, gag region.
 ACCESSION L11797
 SOURCE Human immunodeficiency virus type 1 (HIV-1) blood DNA, Senegalese isolate SE365 (cocultivation on PBMCs).
 REFERENCE 1 (bases 1 to 1474)
 AUTHORS Louwagie,J.J., McCutchan,F., Brennan,T., Peeters,M., Brennan,T., Sanders-Buell,E., Eddy,G., van der Groen,G., Fransen,K., Gershy-Damet,M., DeLeys,R. and Burke,D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate has also been sequenced (L22945).
 This gag sequence, previously published on page I-A-213 of the 1993 compendium, groups with HIV1 D subtype gag sequences.
 BASE COUNT 549 a 280 c 365 g 280 t
 ORIGIN gag cds start

```

1 ATGggtgcga gagcgtcagt attaagcggg ggacaattag atcgatggga aaaaattcgg
gag start ->
61 ttacggccag gaggaaagaa aaaatataag ctaaaacata tagtatgggc aagcagggaa
121 ctagaacgat ttgcacttaa tcctggcctt ttagagacat cagaaggttg tagacaaata
181 ataggacagc tacaaccagc tattcagaca ggatcagaag aacttagatc attatataat
241 acaatagcaa ccctctattg tgtacatgaa aagatagagg taagaagacac caaggaagct
301 gtagaaaaga tggagggaaga acgaaacaaa agtaagaaaa aggtacagca agcaacagct
361 gactcaggaa acaacagcca ggtcagccaa aattacocct tagtgcaaaa cctacagggg
421 caaatggtac accaagccct atcacctagg actttgaacg catgggtcaa agtaatagaa
481 gagaaggatt tcagcccaga agtgataccc atgtttacag cactatcaga aggagccacc
541 ccacaggatt taaacaccat gttaaacaca gtggggggac atcaagcagc catgcaaatg
601 ttaaaagaaa ccatcaatga ggaagctgca gaatgggata gattacatcc agtgcattga
661 gggcctattg caccaggcca gatgagagaa ccaaggggaa gtgatatagc aggaactact
721 agtacccttc aggaacaaat aggatggatg acaataatc caccatccc agtaggagaa
781 atctataaaa gatggataat cctgggatta aataagatag taagaatgta tagtcctgtc
841 agcatttttg acataagaca aggacccaaa gaacccttta gagattatgt agatcgggtc
901 tataaaaactc taagagccga gcaagctaca caggaggtaa aaaattggat gactgaatcc
961 ttgttgggtc aaaatgcgaa ccagattgtt aaaaatatct taaaagcatt aggaccaggg
1021 gctacattag aagaaatgat gacagcctgt caggggggtg gggaaaccag tcataaagca
1081 agagtcttgg ctgaggcaat gagccaagca acaagttcag ctgctgccat aatgatgcag
1141 agaggcaatt ttaagggta aaggaaaatc gtttaagtgt tcaactgttg caaagaaggg
1201 cacatagcaa gaatttgcaa ggccoctaga aaaaagggtt gttggaaatg tggacaggaa
1261 ggacacccaa tgaaggattg cactggaaga caggctaatt ttttagggaa aatttggcct
1321 tccagcaagg ggaggccagg gaactttcct cagagcagac cagagccaac agccccacca
1381 gcagagagct tcgggttttg ggaggagacg acccctctc agaaacagga gcagaaagac
1441 aaggaaactgt atccttttagc tcccccaaa tcac
gag ->

```

LOCUS HIVSE365A2 3315 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, SE365 from Senegal,
 proviral DNA encoding env, tat, vpU, rev, and nef genes.
 ACCESSION L22945
 SOURCE Human immunodeficiency virus type 1 (HIV-1), SE365 from Senegal,
 PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 3315)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A.,
 van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr.
 Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory
 Rockville, Maryland. Twenty-one full length gp160 coding sequences
 from eight African countries (Djibouti, Gabon, Kenya, Senegal,
 Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with
 thirty-two previously published full-length gp160 env sequences.
 Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274,
 VI191, and VI525 have corresponding gag sequences published by
 Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993
 compendium. This sequence and the corresponding gag sequence (L11797)
 cluster with D subtype sequences. See also accession numbers L22939-
 L22944, L22946-L22957 and L23064-L23065.
 NCBI gi: 437350

BASE COUNT 1175 a 575 c 783 g 782 t
 ORIGIN 209 bp upstream from the beginning of rev cds start
 1 tgtttattca ttccagaatt ggggtgtcaac atagcagaat aggcattact ccacagagaa
 61 gaataagaaa TGgatccagt agatccctaac cttagagccct ggaaccatcc aggaagtcag
 tat cds start ->
 121 cctaagaactc cttgtaacaa ctgttattgt aaaaagtgtt gttatcattg ccaaaattgc
 181 ttcatatacaa agggcttagg catctcctAT Ggcaggaaga agcggagaca gcgacgaaga
 rev cds start ->
 241 actcctcaag gcaatcagac tcatacaagt tctgtaccaa aacaGtaagt aatatatgta
 5' sj /\n
 301 ATGcagtott tagagatatt agcaatagca gcaattgtag tagcgctaact actagccata
 vpU start ->
 361 gttgtgtgga ccatagtatt catagaagtt aggaagataa gaaggcaag gaaatagac
 421 tggtaattg acagaataaa agaaagagca gaagatagtg gcaATGagag cgagggagat
 env cds start ->
 481 gaaaaggat tatcagcact tgtggagatg gggcaccatg ctcttggaa tgttgatgac
 541 atgTAGtggt gcagaaaaat tgtgggtcac agtttattat ggggtaccgg tgtggaaaaga
 <- vpU end
 601 agcaaccacc actctatatt gtgcacaga tgctaaatca tatgaaacag agaaacataa
 661 catctgggct acacatgcct gtgtacccac ggaccccaac ccacaagaaa tagaactggc
 721 aatgtgcaca gaaaacttta acatgtggaa aaataacatg gtggagcaga tgcattgaaga
 781 tataatcagt ttatgggac ccaagcctaaa accatgtgta aaattaaccc cactctgtgt
 841 cactttaaac tgcagggata tttctagtga cgcactagt aacaccatta gtaacgtcac
 901 tggaaatccg atgatgggga aaggagaaat gcaaaaactgt tctttcaata taaccacaga
 961 aataagagat aaaagacaga atgtatatc acttttttat agacttgatt tagtaccaat
 1021 aggcaatagt aataagaata gcaccaatta taccagttat agattaataa attgtaatac
 1081 ctcatgcatt aaacaggctt gtccaaaggt aaactttgag ccaattccca tacattattg
 1141 tgcaccagct ggttttgcaa ttctaaagtg taagataaag aagttcaatg gaacaggccc
 1201 atgcaaaaat gtcagtcacag tacaatgtac acatggcatt aagccagtgg tgtcaactca
 1261 actgctgtta aatggcagtc tagcagaaga agaataata attagatctg aaatctcac
 1321 aaataacgct aaaaccataa tagtacagtt taatgaatct gtaatgatta attgtacaag
 v3 loop ->
 1381 gccttacaac aataaaagac aaaggacacc tataggacta gggcaggtac tccatataac
 1441 aagggtaaaa ggagatataa gacaagcaca ttgtaacatt agtaagaaa gatggaataa
 <- v3 loop

```

1501 aactttacaa caggtagtta gaaaattaaa agacctcttt aacaaaacaa taatttttga
1561 accatccetca ggagggggacc cagaatttgc aacacacagt ttaatttga gaggggaatt
1621 tttctactgc aatcacatcag gactgtttta tagtacatgg aaacaaaaat cagggtgataa
1681 tattacactc ccatgcagaa taaaacagat tataaacctg tggcagagag taggaaaagc
1741 aatgtatgcc cctcccattg agggagtgat cagctgttca tcaaatatta cagggtactt
1801 attgctaaga gatggtggga tacataatac tagcaatgag accgagacct tcagacctgg
1861 aggaggagat atgaggggaca attggagaag tgaattatat aaatataaag tagtaagact
1921 cgaaccacta ggagttagcac ccaccacagc aaagagaaga gtggtgaaaa gagaaaaaag
1981 agcaatagga ctgggagtta tgttcccttg gttcttggga gcagcaggaa gcacgatggg
2041 cgcagcgtca atggcgtga cgttacaggg cagacagcta ttgtctggtg tagtgcaaca
2101 gcaaaacaat ttgctgagag ctatagaggg gcaacagcat ctgttgcaac tcacggctctg
2161 gggcattaaa cagctccagg caagaatcct ggctgtggaa agataacctaa gggatcaaca
2221 gctcctagga atttgggggt gctctggaag acacatttgc accactaatg tgccctggaa
2281 ctctagttag agtaataaat ctctaactga gatttggcaa aacatgacct ggatggagtg
2341 ggaagagaaa attgacaatt acacaggcct aatatatagc ttaattgagg aatcgcaaac
2401 ccagcaagaa aagaatgagc aagaattatt ggcattggac aagtgggcaa gtttgtggaa
2461 ttggtttgac ataacaaatt ggctgtggtg tataaaaaata ttcataatgg tagtagaagg
2521 cttaataggg ttaagaataa tttttgctgt gctttcttta gtaaatagag ttaggcaggg
2581 atactcactt ctgtcgtttc AGacctctt cccagcaccg aggggacccg acaggcccca
      3' sj (tat,rev) /\
2641 aggaatagaa gaagaaggtg gagagcaagg cagagacaga tcaattcgat tgggtgaaccg
2701 attctcagca ctatctctgg acggtcttgc gaacctgtgc ctcttcagct accaccggtt
2761 gagagactta atcttaattg cagcgaggat tgtggaactt ctgggaacga gaggggtggga
2821 agtccccaag tatctgtgga gtctcctgca gtattggata cagaaactga aaaaTAGtgc
                                     <- rev
2881 tattagcttg cttgatacca tagcaatagc agtagctgag gggacagata ggattataga
2941 cgtagtacaa agagctttag gagctattct tcacataccc acacgaataa gacaggggtt
3001 ggaaggggct ttactaTAA ATGgggtggca aatgggtcaa aagtagtata gtaggatggc
      <- env cds end
      nef cds start ->
3061 ctgctataag ggaagaata agaagaactg agccagcagc agatgggggtg ggagcagtat
3121 cgcgagacct ggaaaaacat ggggcagtca caagtagcaa taaaaaagat actaatgcgg
3181 actgtgcttg gctagaagca caagaagaga atgaaaaagt gggttttcca gtcaggcctc
3241 aggtaccttt aagaccaatg acttacaaag cagctgttga tcttagccac tttttaaag
3301 aaaagggggg actgg
      3' LTR ->

```

LOCUS HIVSM1451 1486 bp ss-RNA VRL 20-APR-1993
 DEFINITION Human immunodeficiency virus type 1, Somali isolate SM145, gag region.
 ACCESSION L11803
 SOURCE Human immunodeficiency virus type 1 (HIV-1), Somali isolate SM145. (cocultivation on PBMCs)
 REFERENCE 1 (bases 1 to 1486)
 AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Euell, E., Eddy, G., van der Groen, G., Fransen, K., Gershby-Damet, M., DeLays, R. and Burke, D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env cds of this isolate (L22946) has also been sequenced. This gag sequence, previously published on page I-A-214 of AUG 93, groups with HIV1 C subtype gag sequences.
 BASE COUNT 544 a 295 c 365 g 282 t
 ORIGIN gag cds start
 1 ATGGgtgcga gaggctcaat attaagagggc ggaaaattag atgcatggga aagaattagg
 gag start ->
 61 ttaagggccag ggggaaagaa acactatatg ataaaacacc tagtatgggc aagcagggag
 121 ctggaaagat ttgcacttaa ccttggcctt ttagagacat cagaaggctg taaacaaata
 181 atgaaacagc tgcaccagc tcttcagaca ggaacagagg aacttagatc attattcaac
 241 acagtagcaa ctctctattg tgtacatgaa aagatagagg tacgagacac caaggaagcc
 301 ttagacaaga tagaggaaga acaaaacaaa agtcagcaaa aaacacagca ggcaaaagcg
 361 gctgacggaa aggtcagtc aaattatcct atagtgcaga atctccaagg gcaaatggta
 421 caccaggcca tatcacctag aacottgaat gcatgggtaa aagtaataga agagaaggcc
 481 tttagcccg agataatacc catgtttaca gcattatcag aaggagccac ccgcgaagat
 541 ttaaacacca tgtaaatac agtgggggga catcaagcag ccatgcaaat gttaaaagat
 601 accatcaatg aggaggctgc agaattggat agattacatc cagtacaggc agggcctgtt
 661 gcaccaggcc aatgagaga cccaagggga agtgacatag caggaaactac tagtaccctt
 721 caggaacaaa tagcatgat gacaagtaac ccacctattc cagtgggaga catctataaa
 781 agatggataa ttctggggtt aaataaataa gtaagaatgt atagccctgt cagcatttta
 841 gacataaaac aagggccaaa ggaacccttt agagactatg tagaccggtt ctttaaaact
 901 ttaagagctg aacaagctac acaagatgta aaaaattgga tgacagacac cttgtttggtc
 961 caaatgcga acccagattg taagaccatt ttaagagcat taggaccagg ggcttcatta
 1021 gaagaaatga tgacagcatg tcaggggagt ggaggacctg ccacaaagc aagagtgttg
 1081 gctgaggcaa tgagccaagc aaacaatata aacatactga tgcaaagagg caattttaaa
 1141 ggctctaaaa gaattgttaa atgtttcaac tgtggcaaa gagggcacat agctagaat
 1201 tgcagggtc ctaggaaaaa gggtgtttg aatgttgaa aggaaggaca ccaaatgaaa
 1261 gactgtactg agaggcagc taatttttta gggaaaattt ggccttccca caaggggagg
 1321 ccaggggaatt tcttcagaa caggccagc ccaacagccc caccagcaga gagcttcagg
 1381 gtcaggccag agccaacagc ccaccagca gagagcttca ggttcagga gacaaccccc
 1441 gctccgaagc aggagctgaa agaaccctta acttccctca aatcac
 gag ->

HIVSM145A2

LOCUS HIVSM145A2 3228 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, SM145 from Somalia, proviral DNA encoding env, tat, vpU, rev, and nef genes.

ACCESSION L22946

SOURCE Human immunodeficiency virus type 1 (HIV-1), SM145 from Somalia, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 3228)

AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A., van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VII91, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence and the corresponding gag sequence (L21803) cluster with C subtype env sequences. See also accession numbers L22939-L22945, L22947-L22957 and L23064-L23065.

NCBI gi: 437351

BASE COUNT 1147 a 545 c 773 g 763 t

ORIGIN 105 bp upstream from the beginning of rev cds start

1 ccattccagga tgcacacctt aaactgcttg taataattgt tattgttaa c tctgttagcta
61 ccattgcta gtttgcttcc agacaaaagg ctttaggcatt tcctATGgca ggaagaagcg
rev cds start ->

121 gagacagcga cgaagcgctc ctccagagcag tgaggatcat caaaatccta tatcaaagca
181 GTgagtatct gtaATGttaa atttatttagc aggagtagat tatagaatag caataggagc
5' sj /\ vpU start ->

241 atttagtgta gcactaatca tagcaatagt tctgtggacc atagtatata tagaatatag
301 gaaattgtta agacaaagaa aaatagactg gttagttaaa agaattaggg aaagagcaga
361 agacagtggc aATGagagtg aggggggatac tgagggaattg gcaacaatgg tggatatggg
env cds start ->

421 gcattcttagg cttttggatg ttaattgattt gTAGtggggg aggggaacttg tgggtcacag
<- vpU end

481 tctattatgg ggtacgggtg tggagagaag caaaaactac tctgttttgt gcattcagatt
541 ctaaagcata tgagagagaa gtgcataatg tctggggcgac acatgcctgt gtacccacag
601 accccagccc acaagaaatg cttttgggaa atgtaacaga aaattttaac atgtggaaaa
661 atgatatggg ggtacagatg catgaggata taatcagttt atgggatcaa agtctaaagc
721 catgtgtaaa gttgacccca ctctgtgtca ctttaaatg tacagaagct aatattaata
781 ttaccagtgt tggcaatggt accagtaatg acaataagga agaaataaaa aattgctctt
841 tcaatgcaac cacagaatta agtgataaga aacagaaagt gtatgcactt ttttatagac
901 ttgatatagt accacttaat gaggggtgaa acaactctag gaagtataga ttaataaatt
961 gcaatacctc agccataaca caagcctgtc caaagggtct ttttgaccca attcctatac
1021 attattgtgc tccagctggg tatgcgattc taaagtgtaa taatgagaca ttcaatggga
1081 caggaccatg ccgttatgtc agcacagtac aatgtacaca cgggattaag ccagtggtat
1141 caactcaact actgttaaat ggtagcctag cagaaggaga gataatgatc agatctgaaa
1201 atctgacaaa caatgccaaa acaataatag tacacttaa ccaatctgta gaaattaggt
v3 loop ->

1261 gtacaagata cgccaataat acaagaaaaa gtgtaagaat aggaccagga caaacattct
1321 atacaaatga cataatagga gacataagac aagcacattg taacattagt ggagataaat
<- v3 loop

1381 ggaacagaa cttacacag gtaggaaaaa aattagcaga acacttcctt aataaaacaa
1441 taaagtttga accatcctca ggaggggata tagaaattac aacacatagc ttttaattga
1501 gagggagaatt tttctattgc aatacatcag gctgttttaa tagtacatac aatcctaata
1561 acacagggaag tgagtcaaac aacaacatca cactccaatg cagaataaaa caaattataa

```

1621 acatgtggca ggaggttaga caagcaatgt atgccccccc cattaagga aacataacat
1681 gtacatcaaa tatcacagga ctactattgc tgcgtgatgg aggaatgat ccaaaaagaa
1741 ctgaagagca aaataacaca gagagattca gacctggagg aggagatat agagacaatt
1801 ggagaagtga attatataaa tataagggtgg tagaaattaa accattagga atagcaccca
1861 ctggggcaaa aaggagagtg gtggagagag aaaaaagagc agtgggaata ggagctgtgt
1921 tccctggggtt cttgggagca gcaggaagca ctatggggcg ggctcaata acgctgacgg
1981 tacaggccag acaactgttg tctggtatag tgcaacagca aagtaatttg ctgagagcta
2041 tagaggcgca acagcatatg ttgcaactca cggctctgggg cattaagcag ctccgggcaa
2101 gagtcttggc tatagaaaga tacctagagg atcaacagct cctagggatt tggggtgct
2161 ctggaaaact catctgcacc actgctgtac cttgggaactc cagttggagt aatagaactc
2221 aagaagaaat ttgggaaaac ctgacctgga tgcaatggga taaagaaatt agtaattaca
2281 cagacacaat atacaggttg cttgaagtct cgcaaaacca gcaggaaaaa aatgaaaag
2341 atctattagc attggacagt tggggaagtc tgtggaattg gtttgacata acaaaatggc
2401 tgtggtatat aaaaatattc ataataatag taggaggctt gataggttta agaataattt
2461 ttgttgtgct ctctatagtg aatagagtta ggcagggata ctacctttg tcgtttcAGa
                                     3' sj (tat,rev) /\
2521 cccctacccc aaacccgagg ggacccgaca ggctcgggaag aatcgaagaa gaaggctggag
2581 agcaagacaa agacagatcc attcgattag tgaacggatt cttagcactt goctgggacg
2641 acctacggag cctgtgtctc ttcagctacc accgattgag agacttcata ttggtggcag
2701 cgagagcggt ggaacttctg ggacgcagca gtctcagggg actacagaga ggggtgggaag
2761 ccccttaagta tctgggaaat cttggccagt attggggtct ggagctaaaa aagagtgcta
2821 ttagtgtgct tgataccata gcaataacag tagctgaagg aacagatagg attatagaat
2881 tagcacaaag aatttgtaga ggtgtccgca acatacctac aagaataaga cagggttttg
2941 aagcagcttt gctATAAAT Ggggggcaag tggtaaaaga gcagaatagt tggatgggct
      <- env cds end
      nef cds start ->
3001 aatataagaa aaagaataag aagaactgag ccagcagcag agggagtagg agcagcatct
3061 aaagacttag ataatatgg agcacttaca aacagcaaca caccaccaa taatgctgat
3121 tgtgcttggc tgcaagcgca agaggaggaa gaggaagtag gctttccagt caaacctcag
3181 gtgectttaa ggccaatgac tcataaggca gcatttgatc tcagcttc

```

HIVUG2661

LOCUS HIVUG2661 1465 bp ss-DNA VRL 27-MAY-1993
 DEFINITION Human immunodeficiency virus type 1, Ugandan isolate UG266,
 gag region.
 ACCESSION L11798
 SOURCE Human immunodeficiency virus type 1 (HIV-1) blood DNA, Ugandan
 isolate UG266.
 REFERENCE 1 (bases 1 to 1465)
 AUTHORS Louwagie,J.J., McCutchan,F., Brennan,T., Peeters,M., Brennan,T.,
 Sanders-Buell,E., Eddy,G., van der Groen,G., Fransen,K.,
 Gershy-Damet,M., DeLeys,R. and Burke,D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1
 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation
 Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850.
 Samples were obtained from 15 different countries (Belgium, Burundi,
 Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the
 Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four
 continents. From this study the authors identified 7 distinct
 genotypes with an average gag intra-genotypic variance of 7% and an
 average inter-genotypic variance of 14%. The env region of this
 isolate has also been sequenced (L22947); it tends to cluster with
 D subtype sequences. This gag sequence, previously published on page
 I-A-218 of the 1993 compendium, groups with A subtype gag sequences.
 BASE COUNT 591 a 281 c 318 g 275 t
 ORIGIN gag cds start

```

1 ATGggtgcga gagtgtcagt attaagcgag gggaaattag atgcatagga gaaaattcgg
gag start ->
61 ttaaggccag agggaaagaa aaaatataaa ctgaaacatc tagtgtaggc aagcagggag
121 ctgaaaaaat tcgcacttaa ccctagcctt ttagaaacaa cagaaggatg tcaaagaata
181 ctagaacaat tacaaccagc tctccagaca ggagcagaag aacttagatc cttatataat
241 acagtagcag tctctatttg cgtacatcaa aggatagaga taaaagacac caaggaagct
301 ttgatataaa tagaggaat gcaaaataag agcaagcaaa agacacagca ggctgcagct
361 aacacaggaa acagcaacaa ggtcagtcga aattacccca tagtgcaaaa tacacsaaga
421 caaatgacat accaggcctt gtcacctagg accttaaatg cataggtaaa agtagtagag
481 gaaaaggctt tcagcccaga agtaataccc atgttttcag cattatcaga aggagccacc
541 tcacaagatt taaatatgat gctgaatgta gtagaaggac accaggcagc tatgcaaatg
601 ctaaaggata ccataaatga ageagctgca gaataggaca ggttacatcc agtgcaagca
661 aggcctattc cacctggcca gatgagagaa ccaagaggaa gtgacatagc aggaactact
721 agtaccgttc aagaacaagt gagatagatg acaagcaacc cacctatccc agtaggagaa
781 atctatagaa gatagataat cctaggacta aataaaatag taagaatgta tagccctgtt
841 agcatttttag acataagaca agagccaaaa gaacccttta aagaatatgt agataagttc
901 tttaaaactc tcagagctga gcaagctaca caggatgtaa aaaactagat gacagaaacc
961 ttgctagtcc aaaatgcgaa tcagatttgt aagtcatttt taagagcctt aagaccaaaag
1021 gctacattag aagaatgat gacagcatgc caaagagtaa gaggaccag ccataaagca
1081 caggttttag ctgaggcaat gagtcaagta caacagacaa gcataatgat gcagagagggc
1141 aatttttaag gccagaaaag aattaaagtg ttcaactgtg gcaaaagaag acacctagcc
1201 agaaattgca aggcacctaa gaaaaagggc tgttagaaat gtaagaaaga aggacaccaa
1261 ataaaagact gcactgaaag acaggctaatt tttttaaaga aaatttggcc ttccagcaag
1321 gggaggccaa ggaactttcc tcagagcaga ccagagccaa cagccccacc agcagagatc
1381 ttgaggatga aagaagagat gacctccctt ccgaacaga agcaaaaaga caaggaacag
1441 actcctttaa tttcctcaa atcac
gag ->

```

LOCUS HIVUG266A2 3026 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, UG266 from Uganda, proviral DNA encoding env, vpU, and nef genes.
 ACCESSION L22947
 SOURCE Human immunodeficiency virus type 1 (HIV-1), UG266 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 3026)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with D subtype sequences whereas the corresponding gag sequence (L11798) clusters with A subtype sequences. See also accession numbers L22939-L22946, L22948-L22957 and L23064-L23065.
 NCBI gi: 450599

BASE COUNT 1070 a 520 c 731 g 705 t
 ORIGIN 169 bp upstream from the beginning of env cds start
 1 atgtaataaaa tccttttagaa atctggggcaa tagtagcact gatagtggcg ctaatccttag
 61 caatagtgtgt gtgggctata gtatccatag agtgtagaaa attaaaaagg caaaggaaaa
 121 tagactggtt gattgataga ataagagaaa gagaagaaga cagtggcaAT Gagagtgaag
 env cds start ->
 181 gggataccga ggaattatca aaacttattg agatggggaa ctataatcct gggaatgata
 241 ataatttgtg gtactgcaga caacttctgtg gttactgtct actatggggg acctgtgtgg
 301 aaagatgcag agaccacett atttggggca tcagatgcta aagcatatga tacagaagtg
 361 cataatgtct gggctacaca tgctgtgtga ccaacagacc ccaacccacg agaaataaaa
 421 ctgaaaaatg tcacagaaaa cttaaacatg tggaaaaata acatggttga gcggatgcat
 481 gaggatataa tcagtttatg ggagcaaaag ctaaaaccat gtgtaaaatt aacccactc
 541 tgtgtcacct taaactgcag tgaaggagg gttacagata caaatggcac tattggaaat
 601 ggcaactattg gggacatgag aaactgtctt tccaatataa ccacagaat aagagataag
 661 aagaagcaag tatatgcact tttttataaa cttgatgtgg taaccaataga tgataataat
 721 agtactaata ccagctatag gttataaat tgtaataact cagccattac acagacgtgt
 781 ccaaggtaa cttttgagcc aattcccatt cattattgtg cccagctgg atttgcaatt
 841 ctaaaatgta atgataagaa gttcaatggg acggggccct gcaaaaacgt cagcacagtg
 901 cagtgtacac atggaattag gccagtagta tcaactcaac tgttgttgaa tggcagctca
 961 gcagaagaag aaatagtaat tagatctgaa aatctcaca ataatgctaa aatcataata
 1021 gtacagctta atgagtctgt aataattaat tgcacaaggc cctacaacaa tataaaaaata
 v3 loop ->
 1081 caaaggacac ctataggacg agggcaagca ctctttacaa caagaagaaa aggaataaaa
 1141 ggcacaagcac attgtaatat tagtagagaa ggatggaata aaactttaca acaggtagca
 <- v3 loop
 1201 aaaaaattaa gagacottat taacaagaca acaataatct ttaaacatc ctctgggaggg
 1261 gaccagaaa ttacaacaca cagctttaat tgtggagggg aatttttcta ctgcaataca
 1321 tcaggactgt ttaataatag tgtatggagg aatgattatc catcaaatag tacaggagta
 1381 aatgagagcg taatcacact cccatgcaaa ataaaacaaa ttataaacct gtggcaggga
 1441 gtaggacag caatgtatgc cctcccatt gaaggactaa tcaaatgttc atcaaatatt
 1501 acaggactat tattgacaag agatgggtgt gtcaataata gtcagaatga gaccttcaga
 1561 cctggaggag gagatatgag ggacaatttg agaagtgaat tatacaaata taagtaata
 1621 agacttgaac cactaggtct agcacctacc agggcaaga gaagagtgtt ggaaagagaa
 1681 aaaaagcaa taggactagg agctgtgttc cttgggttct tgggagcagc aggaagcacg


```

1741 atggggcgag cgtcactgac gctgacggta caggccagac tgttatgtc tggatatagt
1801 caacagcaaa acaatttgct gatggctata gagggcgaac agcatctatt gcaactcaca
1861 gtctggggca ttaaacagct ccaggcaaga gtcttggtg tggaaagcta cctaaaggat
1921 caacagctcc taggaatttg gggttgctct ggaaaacaca ttgcaaccac tagtgtgcc
1981 tggaaactcta gctggagtaa taaatctcta gaagagattt ggaataacat gacctggatg
2041 gagtgggaaa gagaattga caattacaca gggttaatat acagcttaat tgaagaatcg
2101 caagtccagc aagaaaagaa tgaacaagaa ttattgcaat tggacaaatg ggcaagttt
2161 tggaaattggt ttaacataac acaatggctg tggatatata aatattcat aatgatagta
2221 ggaggttga taggtttaag gatagctctt gctgtgctt ctttagtaaa tagagttagg
2281 cagggatatt cacctctgtc gcttAGacc ctctcccag ccccgagggg acccgacagg
      3' sj (tat, rev) /\
2341 cccgaaggaa cagaagaaga aggtggagag caaggcagag gtagatccat tggattagt
2401 agcggattct tagcattgc ctgggacgac ctgcggagcc tgtgctctt cagctaccac
2461 cgcttgagag acttactctt aattgcaacg aggatgttg aactctggg acgcaggggg
2521 tgggaagccc tcaaatactg gtggaatctc ctacagtatt ggattcagga actaagaat
2581 agtgcattta acttgtttaa taccacagca atagcagtag ctgaggggac agatagggct
2641 atagaaatag tacaagagac tgttagagct attcttaaca taccacaaag aataagacag
2701 ggcttcgaaa gggctttact aTAAaATGgg tggcaaatgg tcaaaaagta gtatagttg
      <- env cds end
      nef cds start ->
2761 atggactgct ataagggaaa gaataagaaa aactgatcca gcagcagaag gggtagggagc
2821 agtatctcgg gacctggaaa aacatggggc aatcacaagt agcaacacag cadagactaa
2881 ccttgactgt gcttgggtag aagcacaaga agaggacggg gaagtgggtt ttccagtcag
2941 acctcaggtt ccgttacgac caatgactta caaggagact ctagatctga gccactttt
3001 aaaagaaaag gggggactgg aaagcc
      3' LTR ->

```

LOCUS HIVUG2681 1465 bp ss-DNA VRL 27-MAY-1993
DEFINITION Human immunodeficiency virus type 1, Ugandan isolate UG268, gag region.
ACCESSION U11799
SOURCE Human immunodeficiency virus type 1 (HIV-1) blood DNA, Ugandan isolate UG268.
REFERENCE 1 (bases 1 to 1465)
AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Buell, E., Eddy, G., van der Groen, G., Fransen, K., Gershby-Damet, M., DeLeys, R. and Burke, D.
TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
JOURNAL AIDS 7, 769-780 (1993)
COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate has also been sequenced (U22948). This gag sequence, previously published on page I-A-219 of the 1993 compendium, groups with HIV1 C subtype gag sequences.
BASE COUNT 539 a 288 c 357 g 281 t
ORIGIN gag cds start
1 ATGggtgcga gagcgctcagt attaagagggc ggaaaattag atacctggga aaaaattaag
gag start ->
61 ttaaggccag ggggaaagaa atgctatatg atgaaacacc tagtctgggc aagcagggag
121 ctgggaagat ttgcacttaa ctctggcctt ttagaaacat cagaaggctg taacacaata
181 atgaaacagc tacaaccagc tcttcagaca gggacagatg aacttagatc attatataac
241 acagtagcaa ctctctattg tgtacataaa gggatagaag tacgagacac caagggaagcc
301 ttgacaaga tagaggaaga acaaaacaag attcagcaaa aaacagaaac agctgacaaa
361 ggaaaggcca gtcaaaatta tcttatagtg cagaatctcc aagggcaaat ggtacaccag
421 gccctatcac ctagaacttt gaatgcattg gtaaaagtaa tagaggaaaa ggctttcagc
481 ccagaggtaa taccatggtt tacagcatta tcagaaggag ccaccccaca agattttaac
541 accatgctaa atacagtggg gggacatcaa gcagccatgc aaatgttaaa agacaccatc
601 aatgaggagg ctgcagaatg ggtatagtta catccagtgc atgcagggcc tgttgacca
661 ggccaaatga gageaccaag ggggaagtgc atagcaggaa ctactagtac ccttcaagaa
721 caaataaatt ggtgacaag taaccaccc attccagtgg gagaaatcta taaaagatgg
781 ataattctgg ggttaaataa aatagtaaga atgtatagcc ctgtcagcat tttggacata
841 aaacaagggc caaaggaacc ctttagagac tatgtatagc ggttctttaa aaccttaaga
901 gctgaacaag ctacacaaga tgtaaaaaat tggatgacag acaccttggt ggtccaaaat
961 gcgaaccagc attgtaagac cattttaaga gcattagggc caggagcttc attagaagaa
1021 atgatgacag catgtcaggc agtgggagga cctggccaca aagcaagagt gttggctgag
1081 gcaatgagcc aagcaaacaa tataaacata atgatgcaga gaagcaattt taagggccct
1141 aagagaatta ttaaattgtt caatgtgtgc aaggaaggcc acctagccag aaattgcagg
1201 gccctagga aaaaaggctg ttggaaatgt ggaagggaag gacaccaaat gaaagactgt
1261 actgagagac aggccaatct tttagggaag atttggcctc ccaacaaggg gaggccaggg
1321 aatttccttc agagcagacc agaaccaaca gcccccaccg cagagccaac agccccaccg
1381 gcagagagct tcaggttcga ggggacaact cccctccga agcaggagcc gaaagacagg
1441 gaacccttaa ttccctcaa atcac
gag ->

HIVUG268A2

LOCUS HIVUG268A2 3228 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, UG268 from Uganda, proviral DNA encoding env, vpu, rev, and nef genes.

ACCESSION L22948

SOURCE Human immunodeficiency virus type 1 (HIV-1), UG268 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 3228)

AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence and the corresponding gag sequence (L11799) cluster with C subtype sequences. See also accession numbers L22939-L22947, L22949-L22957 and L23064-L23065.

NCBI gi: 437352

BASE COUNT 1153 a 546 c 760 g 769 t

ORIGIN 100 bp upstream from the beginning of rev cds start

1 caggaagtca gctaaaaact ccttgacta agtggtttgg taaaaaatgt agttatcatt
61 gtcaagtttg ctttctgaac aaaggcttag gcattctcctA TGgcaggaag aagcggagac
rev cds start ->

121 agogacgaaa cgctcctcca agcagtggag atcatcaaaa tcttatatca aagcaGTAag
5' sj /\

181 tattacatgg tagatgtaAT Gttaaattta ctacagagg tagactatag aataggaata
vpu start ->

241 ggagcattgc taatagcact aatcatagca atagttgtgt ggatcatagt atatatagaa
301 tataggaaac tgttaaaaca aagaagaatg gacaggttaa ttaaaagaat tagggaaaga
361 gcagaagata gtggcaATGa gattgatggg gatacagagg aattgtcaac aatgggtggat
env cds start ->

421 atgggggaatc ctaggctttt ggatatTAAt gatttgtaat gtgatgggga acttgtgggt
-< vpu cds end

481 cacagtctat tatgggggtac ctgtgtggaa agaagcaaca actactctat tttgtgcac
541 agatgctaaa gcatatgaga cagaagtga taatgtctgg gctacacatg cctgtgtccc
601 cacagacccc aaccacaaag aaatagtctt ggaatgtga acagaaggtt ttaacatgtg
661 gaaaaaatgac atgggtggatc agatgcatac ggatgtaac agtctatggg atcaaaagcct
721 aaagccatgt gtaaagtga cccactctg tgttacttta aattgtacaa atgttaatgt
781 taatattacc aataatgcta atgctaccaa tagcccttac gaaaatggga aactcatgga
841 acagggagaa ataaaaaatt gctctttcaa tgaaccaca gaaataagag ataagaaca
901 gacagcgcac gcactttttt ataaacttga ttagtatca cttgaagggg atagtaaac
961 ttacagatta ataaattgta atacctcagc cataacgcaa gcctgtccaa aggtcacttt
1021 ggacccaatt cctatacatt attgtgtctc agctgtttat gcgattctaa agtgtaataa
1081 taaaacattt aatggaacag gacctgttaa taatgtcagc acagtacaat gtacacatgg
1141 aattaaacca gtgatatcaa ctcaactatt gttaaatgg agtatagcag aagaagaat
1201 aataattaga tctgaaaatc tgacaaataa tgccaaata ataatgtac agcttaacaa
1261 atctgtagaa attaatgtg cgagacccaa caataatata agagaaagta taaggatagg
v3 loop ->

1321 accaggacaa acattctatg caacaggaga cataatagga gatataagac aagcatactg
1381 taacattagt agaatgaat ggaatataac tttaaatgg gtaagagaaa aattaaaaag
-< v3 loop

1441 gcacttccct aataaaacaa taaactttac tcaacctca ggaggggact tagaaattac
1501 aacacatagc ttttaattga gaggagaatt tttctattgc aatacatcaa gcctgtttaa

```

1561 tagtagtgat aataataata gtaccatcat cacactccca tgcagaataa aacaaattat
1621 aaacatgtgg cagggggtag gaagagcaat gtatgccctt cccattaaag gaaaaataac
1681 atgtagatca aatatcacag gactactact gacacgtgat ggaggggaga caagtggagc
1741 aaatagcaca gagacattca gaactggagg aggagatatg agggacaatt ggagaagtga
1801 attatataaa tataaagtgg tagaagttaa gccactagga gtacgaacca ctanggcasa
1861 aaggagagtg gtggagagag aaaaaagagc agtgggaata ggagctgtgt tccttgggtt
1921 cttgggagcc gcaggaagca ctatgggccc agcatcaata acgctgacgg tacaggccag
1981 acaattattg totggtatag tgcaacagca aaacaatttg ctgagggcta tagaggcgca
2041 acagcatatg ttgcaactca cggctctgggg cattaagcag ctccagacaa gactcctggc
2101 tatagaaga tacctacagg atcaacagct cctagggatt tggggctgct ctggaaaact
2161 catctgcacc actgctgtgc cttggnaactc tagttggagt aataaatctc taggggatat
2221 ttgggataac atgacctgga tgcagtggga tagagaaatt agtaattata caggaacaat
2281 atacaggttg cttgaagact cacaaaacca gcaggagaaa aatgaaaaag atttactagc
2341 attggacaaa tggcaaatc tgtggagttg gtttgacata acaaaatggc tgtggtatat
2401 aaaaatattc ataattgat taggaggctt aataggttta agaataattt ttgctgtgct
2461 ctctatagtg aatagagtta ggaagggata ctaccatta tegtctcAGa ccttaccctt
                                     3' sj (tat,rev) /\
2521 aaccccgagg ggaccgcaca ggcccggaga aatcgaagaa gaagggtggag agcaagacag
2581 agacagatcc gtgagattaa tcagcggatt cttagcactt gcctgggacg atctgcggag
2641 cctgtgctc ttcagctacc gccgcttgag agacttacta ttgattgcag cgagagcagt
2701 ggaacttctg ggacgcagca gtctcagggg actacagagg ggtgggaag cctTAAgta
                                     <- rev
2761 totgggaagc cttgtgcagt attgggggtca agaactaaaa aagagtgtca ttagtttgcct
2821 tgataccata gcaatagcag tatctgaagg aacagataga attatagaag tgggacangg
2881 aattgggaga gctatctctc acatacctag aagaataaga caaggctttg aagcagcttt
2941 gcaatTAAaAT Ggggagcaag tggatcaacat gtagtccagt gggatggcct gatgtaagag
                                     <- env cds end
nef cds start ->
3001 aaagaataaa accgactgct ccagcagcag agggagtagg agcagcgtct caggacttag
3061 acaaacatgg ggcactgaca accagcaaca cccccccaa taatgctgat tgtgcttggc
3121 tggaggcgca agaggatgag gaggtaggct ttccagtcac acctcaggtc cctttaagac
3181 caataactta taaggcagca ttgatctca gcttcttttt aaaagaaa

```

HIVUG269A

LOCUS HIVUG269A 3132 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, UG269 from Uganda, proviral DNA encoding env, vpu, and nef genes.

ACCESSION L22949

SOURCE Human immunodeficiency virus type 1 (HIV-1), UG269 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 3132)

AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A., van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with D subtype env sequences. See also accession numbers L22939-L22948, L22950-L22957 and L23064-L23065.

COMMENT NCBI gi: 437353

BASE COUNT 1090 a 547 c 764 g 731 t

ORIGIN 16 bp upstream from the beginning of rev cds start

1 gcttaggcatt ttctATGgc aggaagaagc ggagacagcg acgaagatct cctaaaggcg
rev cds start ->

61 gtcaggctca tcaagttcct ataccaaagc aGtaagtagt acATGtatgc aacctttagt
S' sj /\ vpu ->

121 gatattagca atagtagcat tagtagtagt gctaatacta gcaatagttg tgtggactat

181 agtagtcata gagggtataa gattacaaaa gcaagaaaaa atagactggt taattgatag

241 aataagagaa agagcagaag atagtggcca TGagagttag ggagatagag aggaattatc
env cds start ->

301 agcacttggt gagatggggc acgatgctcc ttgggatggt gatgatatgt aggggtacag

361 gagagtgtgt ggtcacagtt tattatgggg tactgtgtgt gaaagaagca accactactc

421 tattctgtgt atcagatgct aaatcatata aaaaaggggc acataatatc tgggctacac

481 atgectgtgt accaacagac cccaacccac gagaaTAaa ttagaaaatg tcaccgaaaa
<- env premature stop

541 ctttaacatg tggaaaaata atatggtgga gcagatgcac gaggatataa tcagtttatg

601 ggatgaaagc ctaaaaccat gtgtaaaatt aaccccactc tgtgtcactt taagtgcac

661 tgaagccaaa ttcaatgtca ctagagccac taatggaaag aacactgtta gaaatgccac

721 taaggggggc ctagaatga aagactgtta ttcaatgta accacagaaa taagagataa

781 gaagaagcaa gtacgtgcac ttttttataa acttgatgta gtacaaatga ataattgtga

841 gagtattaac accagctata gattaataaa ttgtaattcc tcagccatta cacaggcgtg

901 tcctaagata acctttgagc caattcccat acattattgt gccccagctg gatttgcaat

961 tctaaaatgt accgataaga agttcaatgg gaagggtcca tgcaaaaatg tcagcacagt

1021 acagtgtaca catgggatta agccagtagt gtcaactcaa ctgttgttga atggcagtc

1081 agcagaagaa gagataataa ttagatctga aaatctcaca aataatgcta aaatcataat

1141 agtacagctt aatgcgtctg taaceattac ttgcacaagg cctacacca agacaagaca

1201 cagggcacaa gggcgagcat ggtggacaac aggtataaca ggagatataa gacaagata

1261 ttgttaacatt agtaagcaa aatggcagga aactttacaa caggtagcta aaaaattagg

1321 agaccttctt aacaagacac aaataatttt tgaaccatcc tcgggagggg acccagaagt

1381 taaaaaacac agctttaatt gtggagggga atttttctac tgcgatacat caagactggt

1441 taatgatagt tataatacat cagcattaaa tagtacacag ggaaatgaga gtgaaaacat

1501 cacactccca tgcagaataa aacaaattat aaacatgtgg cagggaagtag gaaaagcaat

1561 gtatgccctt cccattgaag ggataatcag atgttcatca aatattacag gactattggt

1621 gacaagagat ggtggtataa ctaacaatag taatgagacc ttcagacctg gaggaggaaa

1681 tatgagagac aattggagaa gtgaattata caaatataaa gttagtaaga ttgaaccact

1741 aggtctagca cccaccacgg caaagagaag agtgggtgga agagaaaaaa gagcaatagg

```

1801 actgggagct 'ctgttccttg ggttcttggg agcagcagga agcacaatgg ggcagcgtc
1861 actgaogctg acgggtacagg ccagacaagt attgtccggc atagtgcac accaaaacaa
1921 ttctgctgagg gctatagagg cgcaacaaca tctgttgcaa ctacagtcct ggggcattaa
1981 acagctccag gcaagagtc cggctgtgga aagctaccta ggggatcagc agtccttggg
2041 aatttggggg tctctaggaa gacacatttg caccactaat gtgccctgga actctagctg
2101 gagtaataga tctctagagg agatttggca taacatgacc tggatggagt gggaaagaga
2161 aattgacaat tacacaggtt taatatagag ctgtattgaa gaatgcacaa ccagacaaga
2221 aaaaaatgaa caagaactat tgcaattgga caaatgggca agtttgtgga attggtttag
2281 cataacaaac tggctgtggt atataagaat atctataatg atagtaggag gcttgatagg
2341 ttttaaggata gtttttggctg tgttttcttt agtaaataga gttaggcagg gatattcacc
2401 tctctcggtt cAGacccctc tcccagcccc gaggggaccc gacaggcccg aaggaataga
3' sj (tat.rev) /\
2461 agaagaaggt ggagagcaag gcagaggcag atccattcga ttagtgagcg gattctcagc
2521 acttatctgg gacgacctga ggaacctgtg cctcttcagc taccaccgct tgagagactt
2581 aatcttaatt gcaacgagaa ttgtggaact tctgggacgc agggggtggg aagccatcaa
2641 atatctgtgg aatctcttaa aatattggat tcaggaacta aagaaTAGtg ctattagctt
<- rev cds end
2701 gtttgatacc atagcaatag tagtagctga ggggacagat agggttatag aagtaataca
2761 aagagctgtt agagctgttc ttcacatacc cagacgaata agacagggtt ttgaaagggc
2821 ttactaTAA aATGggtggc aaatggtcca aaaagtagta tagttggatg gctgctata
<- env cds end
nef cds start ->
2881 agagacagaa taagaagaac tgatccagca gcagaggggt gggagctgca tctagggact
2941 tggaaaaaca tggggcaatc acaagtagca atacaccaca cacaattct gactgtgct
3001 ggctagaagc acaagaagag gacggggacg tgggttttcc agtcagacct caggtaccgt
3061 taagaccaat gacttacaag gcggctgtag atctgagcca ctttttaaaa gaaaaggggg
3121 gactggaaag cc
3' LTR ->

```

HIVUG273A

LOCUS HIVUG273A 3337 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, UG273 from Uganda, proviral DNA encoding env, tat, vpU, rev, and nef genes.

ACCESSION L22957

SOURCE Human immunodeficiency virus type 1 (HIV-1), UG273 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 3337)

AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype env sequences. See also accession numbers L22939-L22956 and L23064-L23065. NCBI gi: 437354

BASE COUNT 1171 a 598 c 775 g 793 t

ORIGIN 209 bp upstream from the beginning of rev cds start

1 tgtttattca tttcagaatt ggggtccaac acagcagaat aggcattatt cgagggagaa
61 gagtcagggA TGgatccggT agatcctaac ctagagccct ggaatcatcc ggaagtaag
tat cds start ->
121 cctacaactc cttgtagcaa gtgttactgt aaagtgtgtt gctatcattg ccagttttgc
181 tttctgaaaa aaggcttagg catctcctAT Ggcaggaaga agcggagtcA gcgacgagga
rev cds start ->
241 actcctcaaa gcaataagga tcatcaaaat cctataccaa agcaGtaagt attagtaatt
5' sj /\

301 aatataatga ATGcttccct taaatatctg tgcAatagta ggactgatag tagcaactaa
vpU cds start ->
361 cttagcaata gttgtgtgga ctatagtagg tatagactat aagaaattgt taaagcaaa
421 aaaaatagac aggttaattg agagaataag agaaagagca gaagacagtg ccaATGagag
env cds start ->
481 tgaaggggat acagaggaat tggcagcact tgttgagatg gggaaactata atcttggggt
541 tgataacaat ctgTAGtact gcaaaaaact tgtgggttac tgtatactat ggggtacctg
<- vpU cds end

601 tgtggagaga cgcagagacc accctatctt ggcacacaga tgctaaagca tatgagacag
661 aaatgcataa tgtctgggct acacatgccc gtgtacccac agaccctaac ccacaagaaa
721 tacgttttga aaatgtgaca gaagagttta acatatggaa aaataacatg gtagagcaga
781 tgcatacaga tataatcagt ctatgggacc aaagcctaaa gccatgtgta aagttaaccc
841 ctctctgagt tactttaaat tgtaccaatg ccaacctcac caatgccaat aaaaatgcca
901 aggtcaacag cggcacgaat ggcaccatcc agagcatgga aagagaaata aaaaactgct
961 ctttcaatat gaccacagaa ctaagggata agacacggag agtatattca cttttttata
1021 aacttgatgt agtacaaatt aaaaatagta gtgagtatag attaataaat tghtaatcct
1081 cagccattac acaggcttgt ccaagggtat cctttgagcc aattcccata cattattgtg
1141 cccagctgg ttttgogac ctaaaatgta atgataagga gttcaatgga acagggctat
1201 gcaagaatgt cagtcagta caatgcacac atggaatcag gccagtagta tcaactcaac
1261 tgctgttaaa tggcagtcta gcagaaacag aggtaaaaat taaatctgaa aatatctcag
1321 acaatgctaa aaccttaata gtacaactta ccacgcctgt aaaaattaat tgtactagac
v3 loop ->
1381 ctggcaacaa tacaagaaca agtgtacgta taggaccagg acaagcattc tatgcaacag
1441 gtgatataat aggagatata agacaagcac attgtaattg cagtegatca gaatggaag
<- v3 loop

1501 aaacttttga aaaggtagtc aaacaattaa gaacacactg gaacaaaaca ataacttta

```

1561 caaactcctc aggaggggat ttagaaatta caacacatag ttttaattgt ggaggagaat
1621 ttttctattg taatacagca ggccctgttta atagtacttg gaataaaaat gacagcataa
1681 agtcaaatga cacaagtc aatgacatta taactcttca atgcagaata aagcaaatga
1741 taaatatgtg gcagagagta ggacgagcaa tgtatgccc tcccatcaaa ggagtaataa
1801 ggtgtgaatc aaacattaca ggactactat taacaagaga tgggtggaat gggagcaata
1861 caaatgaaac cttcagacca ggaggaggag atatgagggg caattggaga aatgaattat
1921 ataagtataa agtagtaca attgaaccat taggagtagc acctaccca gcaagagaa
1981 gagtgggtgca gagagaaaaa agagcagttg gactaggagc tgtattcatt gggttcttag
2041 gagcagcagg aagcactatg ggcgcggtt caataacgt gacggtacag gccagacaat
2101 tattgtctgg catagtgc aaagcaagca atttgcctg agctatagag gctcaacagc
2161 atctgttgaa actcacagtc tggggcatta aacaactcca ggcaagactc ctggtctgtg
2221 aaagatacct aagggatcaa cagctcttag gaatttgggg ctgctctgga aaactcatct
2281 gcaccactac tgtgcccctg aactctagtt ggagtaataa atctcaagt gagatatggg
2341 ataacatgac ctggctgcaa tgggataaag aaattagtaa ttacacaaac ataatatg
2401 atctaattga agaagcgag aacagcagg aaaagaatga acaagactta ttggcattgg
2461 acaagtgggc aagtctgtg agttggtttg acatatcaaa ttggtctgtg tatataagac
2521 tatttataat gatagtagga ggttaatat gattaagaat agttcttctg gtgttttctg
2581 taataaatag agttaggaag ggatactcac ctgtgtcatt tcAGaccat accccaacc
3' sj (tat.rev) /\

2641 caagggatct cgacaggccc gaaagaatcg cagaagaagg tggagagcaa gacagagaca
2701 gatcgattcg cttagtcagc ggattcttag cacttgctg ggacgacctg cggagcctgt
2761 gctcttccag ctaccacgc ttgagagact tcatcttgat tgtcgagagg actgtgaaac
2821 ttctgggaca cagggggtgg gaaggaatca agtatctgtg gaaectcttg ttgttttggg
2881 gtcgggaact caaaatTAGt gctattaatt tgcttgatac catagcaata gcagtagctg
<- rev cds end

2941 gctggacaga tagagttata gaaataggac aaagaattgg tagggctttt cttaacatac
3001 ctagaagaat caggcagggc ttcgaaagag ctttactata AcATGgtag caagtggta
<- env cds end

nef cds start ->

3061 aaaagttggc ctcaggttag ggaaagaatg agacgagctc ctccagcagg acaagctcct
3121 ccagcagcaa aaggagtagg accagtatct caagatttag aaaaacatgg agcagtcaca
3181 agcagtaata taaatcaccg gagttgcacc tggctgcaag cgcaagaaga agatgatgag
3241 gtaggctttc cagtcaggcc acaagtacct ctaagaccaa tgactttcaa agcagctttt
3301 gatctcagcc actttttaaa agaaaagggg ggactgg
3' LTR ->

```


HIVUG2741

LOCUS HIVUG2741 1474 bp ss-DNA VRL 27-MAY-1993
 DEFINITION Human immunodeficiency virus type 1, Ugandan isolate UG274, gag region.
 ACCESSION L11801
 SOURCE Human immunodeficiency virus type 1 (HIV-1) blood DNA, Ugandan isolate UG274.
 REFERENCE 1 (bases 1 to 1474)
 AUTHORS Louwagie, J.J., McCutchan, F., Brennan, T., Peeters, M., Brennan, T., Sanders-Buell, E., Eddy, G., van der Groen, G., Fransen, K., Gershy-Damet, M., DeLeys, R. and Burke, D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. The authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate has also been sequenced (L22950). This gag sequence, previously published on page I-A-221 of the 1993 compendium, groups with HIV1 D subtype gag sequences.
 BASE COUNT 555 a 276 c 361 g 282 t
 ORIGIN gag cds start
 1 ATGgggtgcga gagcgctcagt attaagcggg ggaaaattag atgaatggga aaaaattcgg
 gag start ->
 61 ttacgggccag ggggaaacaa aaaatataaa ctaaaacata tagtatgggc aagcagggag
 121 ctggaacgat ttgcaattaa tcttggcctt ttagaacat cagaaggctg taaacaata
 181 atgggacagc tacaaccagc tcttaagaca ggatcagagg aaatcaaatc attatttaat
 241 acagtagcaa cctctattg tgtacatgaa aggataaagg taacagacac caaggaaagt
 301 ttagacaaga tagaggaaga acaaaccaaa agtaagaaa aggcacagca agcaacagct
 361 gacacaagaa acagcagcca ggtcagccaa aattatccta tagtgcaaaa cctacagggg
 421 caaatggtag accagggccat atcacctaga actttgaatg catgggtaaa agtaatagag
 481 gagaaggctt ttagcccgaa agtaataccc atgtttacag cattatcaga aggagccacc
 541 ccacaggatt taaacacccat gctaaacaca gtggggggac atcaagcagc catgcaaatg
 601 ctaaaagaaa ccatcaatga ggaagctgca gaattgggata gggatcatcc agtgcagca
 661 ggacctattg caccaggaca aatgagagaa ccaaggggaa gtgatatagc cggaactact
 721 agtacccttc aggaacaaat aggatggatg acaagcaatc cgcctatccc agtaggagaa
 781 atctataaaa gatggataat cctaggatta aataaaatag taagaatgta tagccctgtc
 841 agtatatttg acataagaca aggaccaaag gaacccttta gagactatgt agatcgggtc
 901 tataaaactc taaggggcga gcaagcttca caggatgtaa aaaattggat gactgaaacc
 961 ttgttgggtc aaaatgogaa cccagattgt agaactatct taaaagcatt gggaccaggg
 1021 gctacattag aagaaatgat gacagcatgt cagggagtgg gaggaccag tcataaagca
 1081 agagtcttgg ctgaggcaat gagccaagcc acaaatgcaa atactgctat aatgatgcag
 1141 agaggttaatt ttaagggccc aaggaaaatc attaatgtt tcaactgtgg caaagaaggg
 1201 cacatagcaa aaaattgtag ggctcctagg aaaaagggtc gttggaatg tggaaaggaa
 1261 ggacacaaaa tgaaagagtg cactgaaaga caggctaatt ttttagggaa aatttggcct
 1321 tcccacaagg ggaggccagg gaacttcctt cagagcagac tggagccaac agccccacca
 1381 gcagagagct tggggtttgg ggagaagata acacctctc agaaacagga gcagaaagac
 1441 aaggagccat gtcccttagc tccctcaaa tcac
 gag ->

LOCUS HIVUG274A2 3288 bp ds-DNA VRL 31-DEC-1993
DEFINITION Human immunodeficiency virus type 1, UG274 from Uganda, proviral DNA encoding env, tat, vpu, rev, and nef genes.
ACCESSION L22950
SOURCE Human immunodeficiency virus type 1 (HIV-1), UG274 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.
REFERENCE 1 (bases 1 to 3288)
AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
TITLE Genetic diversity of the HIV-1 envelope glycoprotein
JOURNAL Unpublished (1993)
COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence and the corresponding gag sequence (L11801) cluster with D subtype sequences. See also accession numbers L22939-L22949, L22951-L22957 and L23064-L23065.
 NCBI gi: 437355
BASE COUNT 1164 a 566 c 794 g 764 t
ORIGIN 144 bp upstream from the beginning of rev cds start
 1 ggcttcgcgc tagatagatc ctagactaga gccctggaac catccaggaa gtcagcctag
 61 cactccttgt aacaaatgct attgtaaaaa gtgctgctat cattgccaat tgtgtctcat
 121 aacgaaaggc ttaggcatct cctATGgcag gaagaagcgg agacagcgac gcagacctcc
 rev cds start ->
 181 tcaaaagcagt caagctcctc aagatcctat accaaagcag Taagtagtac atgtaATGaa
 5' sj /\ vpu start ->
 241 accttttagtg atattagcaa tagtagcatt agtagtagcg ctaataatag caatagttgt
 301 gtggattata gtattcatag agtgtagaag attaaaaagg caaagaaaaa tagactgggt
 361 aattgataga ataagtgaag gagcagaaga tagtggcaAT Gagagtgagg gagatagaga
 env cds start ->
 421 ggaattatca gcacttgtgg atatggggca ccatgctcct tggaatgttg atgacataTA
 481 Gtgttgacga acagttgtgg gtcacagtgt attatggggg acctgtgttg aaagaagcaa
 <- vpu end
 541 ccactactct attttgtgca tcagacgcta aatcatatat agcagaggca cataatatct
 601 gggctacaca tgctgtgtga ccacacagac ccgacccacg agaaataaga ctagaaaatg
 661 tcacagaaaa ctttgacatg tggaaaaata acatggttga gcaaatgcat gaggatataa
 721 tcagttttatg ggatcaaagc ctaaaaccat gtgtaaaatt aacccactc ttgtgtcactt
 781 taactgcac tgaatggaag aataccactg aacaggagaa gaataccgct gtaccgggga
 841 gtacaacatc tgacacagac atgagaaaact gctctttcaa tataaccaca gagttaagag
 901 ataagaagga gcaagtatat gcactttttt ataaacttga tgtggtacaa atagatgata
 961 ataatagtag taatagtact aactatagat tagtaaatgt tactacctca gccattacac
 1021 aagcgtgtcc aaaggtaacc tttgagccaa ttcccataca ttattgtgac ccagctggat
 1081 atgcgattct aaaatgtaat gataagaagt tcaatgggac gggctccact caaaacgtca
 1141 gcacagtaca gtgtacacat gggatcaggc cagttagtgc aactcaactg ttgttgaatg
 1201 gcagtctagc agaagaagag ataataatta gatctgaaaa tctaacaast aatgctaaaa
 1261 tcataatagt acagcttaat gagtctgtaa cgattgattg cgcaaggccc tacgaaaata
 v3 loop ->
 1321 taagacaaag gacacccata ggaacagggc aagcactcta tacaacacaa ggaagaaaaa
 1381 agataggaaa gataggacaa gcacattgta atgttagtaa agcaggatgg gaaaaaactt
 <- v3 loop
 1441 tacaacaggt agctaaaaat taaaagacct ttttaaggca gataatataa cttttaaacc
 1501 atcctcagga ggggacccag aaattacaac acacagcttt aattgtggag gggaattttt
 1561 ctactgcaat acatcaagac tatttaataa tagtggatgg gaaaattata atatattcaa

```

1621 tagtacaggg ggaatgaga caatcagact ccaatgcaga ataaaaacaaa ttataaacat
1681 gtggcaggaa gtaggaaaag caatgtatgc cctcccatt ataggacaaa tcaaatgttc
1741 atcaaatatc acaggactat tggttaacaag agatgggtgt gtataataata gtagtaatga
1801 gactagcagt aaaaatgaga tcttcagacc tggaggagga gatatgagag acaattggag
1861 aagtgaatta tacaatatata aagtagtaag aattgaacca ctaggctggg caccaccag
1921 ggcaaaaagg agagtgggtg aaagagaaaa aagagcaata ggactgggag ctttgttctt
1981 tgggttcttg ggaacagcag gaagcacgat gggcgagcg tcaatgacgc tgacggtaca
2041 ggccagacag gtattgtctg gtatagtgca acagcaaac aatttgcctg gggctataaa
2101 ggcgcaacag catctgttgc aactcacagt ctggggcatt aaacagctcc aggcagaat
2161 cctggctgtg gaaagatacc taaggatca acagctccta ggaatatggg gttgctctgg
2221 aaaacacatt tgcaccacta ctgtgccctg gaactctagc tggagtaata gatctctaga
2281 ggagatttgg aataacatga cctggatgga gtgggaaaga gaaattgaca attacacagg
2341 gttaatatac accttaattg aagaatcgca aaacagcaa gaaaagaatg aacaagaact
2401 attgcaattg gataaatggg caagtgtgtg gaattgggtt agcataacaa aatggctgtg
2461 gtatataaga atattcataa tgatagtagg aggcctgata ggtttaagaa tagtttttgc
2521 tgtgctttct ttagtaaata gagttaggca gggatattca cctctgtcat ttcaGacctt
                                     3' ej (tat,rev) /\
2581 cctcccagcc ccgaggggaa ccgacaggcc cgaagaata gaaagaaga gtggagagca
2641 aggcagaggc agatccattc gattgggtgaa cggattctca gcacttatct gggacgatct
2701 gaggaacctg tgcctcttca gctaccgctg cttgagagac ttactcttaa ttgcagcgag
2761 gattgtggaa atcctgggac gcagggggtg gaaagccatc aatatctgtt ggaatctcct
2821 gcagtattgg attcaggaac TAAagaatag tgctattagc ttgtttgata ccacagcaat
                                     <- rev cds end
2881 agcagtagct gaagggacag ataggattat agagatagta caaagagctg ttagagctgt
2941 tcttaacata ccacacgaa taagacaggg cttggaaagg gctttgctaT AAAATGggtg
                                     <- env cds
                                     nef cds start ->
3001 gcaaatggtc aaaaagtagt atagttggat ggcttgctat aagggaaaga atgagaagaa
3061 ctgctccagc agcagaaggg gtgggatcag tatctcggga tctggaaaaa catggggcaa
3121 tcacaagtag taatacagcc catactaata ctgattgtgc ctgggtagaa gcacaagaag
3181 aggaagagga agtgggtttt tcagtcagac ccaggtacc gttaagacca atgacttaca
3241 agggggctgt agatctgagc cactttttaa aagaaaaggg gggactgg
                                     3' LTR ->

```

LOCUS HIVUG275A 3311 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, UG275 from Uganda, proviral DNA encoding env, tat, vpU, rev, and nef genes.
 ACCESSION L22951
 SOURCE Human immunodeficiency virus type 1 (HIV-1), UG275 from Uganda, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 3311)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SB365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype env sequences. See also accession numbers L22939-L22950, L22952-L22957 and L23064-L23065. NCBI gi: 437356

BASE COUNT 1152 a 579 c 789 g 791 t
 ORIGIN 209 bp upstream from the beginning of rev cds start
 1 tgtttattca tttcagaatt ggggtgcacac atagcagaat aggcattatt cgaaggagaa
 61 gagtcaggga TGgatccggt agatcctaac cttagagccct ggaaccatcc ggggaagtcag
 tat cds start ->
 121 cctacaactc cttgtaacaa gtgttactgt aaaaagtgtt gctatcattg tacagtttgc
 181 tttctgaaca aaggcttagg catctctat ATGgcaggaaga agcgggaaga gcgacgaaga
 rev cds start ->
 241 actctctaat gcaataagga tcatcaaaat cctataccaa agcaGtaagt attagtaatt
 5' sj /\

301 agtatATGta atgtctctct tgcctctctt gcatactctg gcaatagtag gactggtagt
 vpU start ->
 361 ggcactaatc cttagcaATAG ttgtgtggac tatagtaggt atagaatata agagattgct
 <- vpU end

421 aaagcaagg aaaatagaca gggttaattaa gagaataagt gaaagagcag aagacagtgg
 481 caATGaaagc gatggggaca cagatgaatt gtcaaaactg gtggagatgg ggaactatga
 env cds start ->
 541 tcttggggat gctaataatt tgtagtactg cagaagactt gtgggttact gtctactatg
 601 gggtaacctg atggagagac gcagagacca ccttattttg tgcacagat gcgaagcat
 661 atgatacaga aatgcataat gtctgggcta cacatgctg tgtacctaca gaccccaacc
 721 cacaagaaat aaatttggga aatgtgacag aaaaatttaa catgtggaaa aataacatgg
 781 tagagcagat gcatgaagat ataactcagtc tatgggatca aagcctaag ccatgtgtaa
 841 agctaaccoc tctctgcgtt acttttaatt gtagtgtatg caataatgtc actgtcacca
 901 atcatagtat tagcagtgc atgaaaggag aaataaaaaa ctgctcttcc aatatgacca
 961 cagaattaag ggataagcaa cagaaaatac atccactttt ttatagactg gatgtagtac
 1021 caattaatgg cagcaacaac aggcagtata gattaataaa ttgtaatact tcagccatta
 1081 cacaggcttg tccaaaagta tcccttgagc caattcccat acattatttg gccccagctg
 1141 gtttttgcgat tctaaagtgt aatgatcagg agttccatgg aacagggata tgttaaatg
 1201 tcagcacagt acaatgcaca catggaatca agccagtagt atcaactcag ctgctgttaa
 1261 atggcagttt agcagaagga aaggtaagga tttagtctga aaatatcaca aacaatgcc
 1321 aaatatatct ggtataactt accacgcctg taacaattaa ttgtaccaga cetaacaaca
 v3 loop ->
 1381 atacaagaac aagtgtacgt ataggaccag gacaatcatt ctatgcaaca ggtgacataa
 1441 taggggatat aaacaagca cattgtataa tcagtagagc agaattggaat gaaactttac
 <- v3 loop

1501 aaaaggtagt cagccaatta agaacacact ttggcaacaa aacaataatc tttggtaact

```

1561 cctcaggagg ggatatagaa ataacaacac atagttttta ttgtggaaga gaatttttct
1621 attgtgacac atcaggcctg tttaatagca cttggaatgg caataccagc aagcaggagt
1681 caagtagcac agagtcaa atggaatata ctctccaatg cagaataaag caaattataa
1741 atatgtggca gagagcagga aaagcaatat atgccccctc tatcaaagga gtaataagggt
1801 gtgtatcaaa cattacagga ctaatatata caagagatgg tgggggtaat aacagtgaaa
1861 atgaaacctt cagacctgga ggaagagata tgaggggaca ttggagaagt gaattatata
1921 agtataaagt agtaaaaatt gaaccactag gagtagcacc caccagggca aggaggagag
1981 tgggtggagag agaaaaaaga gcagttggag taggagctgt ctctcttggg ttcttaggag
2041 cagcaggaag cactatgggc gcggcgtcaa taacgctgac ggtacaggcc agacaattat
2101 tgtctggcat agtgcaacag caaagcaatt tgctgaaggc tatagaggct caacagcatc
2161 tgttgaact caoggtctgg ggcattaaac agctccaggc aagagtcctg gctgtggaaa
2221 gatacctaag ggatcaacag ctcttaggaa tttggggctg ctctggaaaa ctcatctgca
2281 ccaactaatgt gccctggaac tctagttgga gtaataaatc ccagagtgat atatgggaca
2341 acatgacctg gctgcaatgg gataaggaaa taacaatta cacagagcta atatataggc
2401 tacttgaaga atgcgagaac cagcaggaaa agaataaaca agacttattg gcattggaca
2461 aatggggcaaa tctgtggaat tgggtttggca tatcaaaact gctgtggtat ataagactat
2521 ttataatgat agtaggaggc ttaataggat taagaatagt ttttgctgtg ctttctgtaa
2581 taaatagagt taggcaggga tactcacctt tgcgtttccA Gaccatacc ccaaacccaa
      3' sj (tat, rev) /\
2641 ggggtctcga caggccccga agaatacgaag aagaagggtg aaagcaagac agagacagat
2701 cgattcactt agtgagcgga ttcttagcac ttgcttggga cgacctggcg agcctgtgcc
2761 tcttcagcta ccaccgattg agagacttca tcttgattgc agcgaggact gtggaacttc
2821 tgggacaggg cctcaagtat ctgtggaatc tactagcata ttgggggtcg gaactaaaaa
2881 tTAgTgctat tagtttgttt gacaccatag cagtagtaat agctgggtgg acagatagag
      <- rev cds end
2941 ttatagaatt aggacaaaga ctttgttagag ctattttaca catacctaga agaatacagc
3001 aaggcttcga aagggtcttg ctaTAACATG gggggcaagt ggtcaaaaaa aaaatagggg
      <- env cds end
      nef cds start ->
3061 atgggtctgag gttagggaaa aaataagaca aactcctcca gcagcaaaag gagtaggagc
3121 agtatctcaa gatthagata aacatggagc agtcacaagc agcaatacaa atcatcctag
3181 ttgtgacctg ctggaagcac aagaggaaga ggaaggtagg ttccagtc gccacaggt
3241 acctctaagg ccaatgactt tcaagggagc ttttgatctc agccacttt taaaagaaaa
3301 ggggggactg g
      3' LTR ->

```

LOCUS HIVVAU 2631 bp RNA VRL 06-JUL-1994

DEFINITION Human immunodeficiency virus type 1, isolate VAU from France, env cds.

ACCESSION X80020

SOURCE Human immunodeficiency Virus type 1 (HIV-1), isolate VAU from France, proviral DNA extracted from PBMCs, PCR amplified.

REFERENCE 1 (bases 1 to 2631)

AUTHORS Charneau,P., Borman,A.M., Quillent,C., Guetard,D., Chamaret,S., Cohen,J., Remy,G., Montagnier,L. and Clavel,F.

TITLE Isolation and envelope sequence of a highly divergent HIV-1 isolate: definition of a new HIV-1 group

JOURNAL Virology 205:247-253 (1994).

COMMENT Submitted (04-JUL-1994) to the EMBL/GenBank/DBSJ databases. F. Clavel, Institut Pasteur, Unite d'Oncologie Virale, Dept Sida et Retrovirus, 25 rue du Dr. Roux, 75015 Paris, FRANCE

The VAU, ANT-70 and MVP-5180 samples define the recently identified O clade of outlier HIV-1 genomes. The VAU virus was isolated from a French woman who claimed to have never left Europe. Hence, the VAU sample indicates the spread of the group O HIV-1 virus outside of Africa. Genomic comparison of VAU with the prototypical HIV-1 isolate, HIVLAI, reveals a high degree of variability (50%) with the most extreme divergence occurring within the V3 region. Twenty-nine potential N-linked glycosylation sites were identified in the VAU env protein sequence; thirteen of these sites are conserved between VAU and M group HIV-1 variants. In addition, the presence of 19 conserved cysteines suggests the basic structure is also conserved. Phylogenetic analysis of the three O group env sequences indicates that they are separated by distances comparable to the distances that exist between M group subtypes.

BASE COUNT 940 a 485 c 576 g 630 t

ORIGIN env cds start

1 ATGacagcga ttatgaaagc aatggggaag aggaacagga agttagggat ctggtgcttg

env start ->

61 attttggett tgataatccc atgtttgagc tgtaaccaac tatatgccac agtctattct

121 ggggtacctg tatgggaaga tgcaaaacca acattgttct gtgcttcaga tgctaaactg

181 acaagcactg aacagcataa tatttgggca acacaagcct gtgttccac agaccccagt

241 ccaaatgaat atgagctaaa aaatgtgaca ggtaaattca atatatggaa aaattatata

301 gttagccaaa tgcacgaaga cattatagat ttgtgggacc agagttaaa acctgtgtt

361 caaatgactt tcttgtgtgt acaaatgaat tgtacagata tcaaaaatag tattaatacc

421 acaaacagtc ccttaaacct aaacaataca aaagagggtga aacagtgtga ctttaattga

481 actacagtgc tcaaaagaaa acaggagaaa aaacaggctc taktctatgt gacagatttg

541 gttaagatta acgccacatc aatgaaaca atgtatagat taattaatg taactccaca

601 accatcagga aggcctgtcc aaaggatatc tttagcccca tcccatata ctattgtgct.

661 ccagcgggat gtgccatctt taagtgtaat gaaacaggat ttaatggaac aggtctctgt

721 aaaaacgtta cagttagttac ttgtacacat ggcatacaac caacagtaag taccacaata

781 atactaaatg ggacactctc taagggaat ataacaatca tgggaaagaa tatttcagac

841 agtggggaga acatcctaata aaccctaata actaatataa caatagcatg tgagagacca

V3 loop ->

901 ggaaatcaga caatacaaaa gataatggca ggtccaatgg ctltggtacag catggccctt

961 agtaatacaa agggggatag aagggcagct tattgtaatt atagtccac tgactggaac

<- V3 loop

1021 aaagccttaa aaacataaac tgaagatat ttagaacttg tagaatataa tcaaaactgat

1081 gttaccatga aattcggtaa tcacagtggg gaagatgcag aagtaacaaa tttctttttt

1141 aactgtcatg gagaattctt ttattgtaac acaaatcggc tgtttaatca taccttttcc

1201 tgcaagaaga atatgaccaa taacaagatc aattgtacta atattagcaa taatagcaat

1261 ggcactcagg caataccttg caggttgaga caagtagtaa gggactggat gaggggagga

1321 tggggacttt atgcacctcc catcccagga aacctagtat gcaggtcaaa cataactgga

1381 atgattctac aattggacac gccatggaat aaacacatc ctaacagcac cacccttaga

1441 ccaggagggg gagatatgaa agatatatgg agaactcaat tgttcaata taaagtagta

HIVVAU

```

1501 agagtaaaac cttttagtgt agcaccaaca aaaattgcaa ggccaactat aggaactaga
1561 tctcatagag agaaaagagc agcagggtttg gcaatgctat tcttggggat tctaagtgc
1621 gcaggaagca ctatgggcgc agcggcaaca gcgctgacgg taoggaccca gcatctgata
1681 aagggtatag tgcaacagca ggataacctg ctgagagcaa tacaggccca gcaacacttg
1741 ctgaggccat ctgtatgggg tattagacaa ctccgagctc gcctgctagc cttagaaacc
1801 tttatacaga atcagcaact ccttaacctg tggggctgca agaatagact aatctgctac
1861 acatcagtaa agtggataa aacatgggga ggagataatg aatcaatttg ggatgagtta
1921 acatggcagc agtgggatca acagataaac aacgtaagct ccttcataa tgaaaaata
1981 caagaggcac aagaacaaca ggagaaaaat gagaaagaat tgctggagtt agatgaatgg
2041 gcctctatct ggaattggct tgacataact aaatgggtgt ggtatataaa aatagctata
2101 atcatagtag gagcactaat aggtgtaaga gtagttatga tagtacttaa tctagtaag
2161 aacattaggc agggatatca accctctctg ttacagatcc ccaccaaca acaagcggaa
2221 gtaggaacgc caggaggaac aggagaagga ggtggagacg aagacaggcg cagggtggact
2281 ccattgccgc aaggggtctt gcctctgttg tacacggacc tcaggacaat aatcttgtgg
2341 atttaccacc tcttgagcaa cttagcctca gagatccaga agttgatcag acacctggga
2401 cttggactat ggatcatagg gcagaggaca attgaagctt gcagactctt taaagctata
2461 atacaatact ggctacaaga attgcaact agtgcataca atctactaga tactgttgca
2521 gtggcagttg ctaattggac tgacagcaca atcttaggca tacaagcat agggagaggg
2581 attcttaaca taccagaag gattagacag ggccttgaac gactcctgtt a

```

env ->

LOCUS HIVV11911 1453 bp ss-RNA VRL 20-APR-1993
 DEFINITION Human immunodeficiency virus type 1, Zairean isolate VI191, gag region.
 ACCESSION L11783
 SOURCE Human immunodeficiency virus type 1 (HIV-1), Zairean isolate VI191.
 REFERENCE 1 (bases 1 to 1453)
 AUTHORS Louwagie,J.J., McCutchan,F., Brennan,T., Peeters,M., Brennan,T., Sanders-Buell,E., Eddy,G., van der Groen,G., Franssen,K., Gershy-Damet,M., DeLeys,R. and Burke,D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate has also been sequenced (L22952). This gag sequence, previously published on page I-A-195 of the 1993 compendium, groups with LBV21-7 as subtype G, also found in Gabon and Taiwan.

BASE COUNT 535 a 282 c 363 g 273 t

ORIGIN gag cds start

1 ATGggtgcga gagcgtcagt attaagcggg ggaactag atgcatggga aaaaattcgg
 gag start ->

61 ttgaggccag ggggaagaa acaatataga ataaagcacc tagtatgggc aagcagggag
 121 ctggaaaaat ttgcacttaa ccttggcctt ttagaataca cagaaggctg tcaacaata
 181 atgagacagt tgcaaccagc tctccagaca ggaacagagg agattaatc actatataat
 241 acaatagcaa ctctctattg tgtacatcaa aggtatagaga taaaagacac ccaagaagct
 301 ttacagggaag tggagaaggc acaaaagaaa agtcagcaaa aaacacaaac ggagggaagg
 361 aacagcagtc aagtcagtc aaattatcct atagtgcaga atgcacaagg gcaaatggt
 421 caccagccct taacacctag aactttgaat gcattggtaa aagtaataga agaaaagac
 481 ttcatgccag aagtaatacc catgttttca gcattatcag aaggagccac cccacaagat
 541 ttaaatacca tgctaaacac agtggggggg catcaagcag ctatgcaaat gctaaaggat
 601 actatcaatg aggaggctgc agagtgggac aggtatcac caccacaggc aggacctatt
 661 ccaccaggcc agataaggga accgagggga agtgatatag caggaaactac tagtacccta
 721 caggaaacaa tcagatggat gaccagtaat ccacctatcc cagtgggaga catttataaa
 781 agatggataa tcttgggatt aaacaaaata gtaagaatgt atagccctgt cagcattttg
 841 gatataagac aaggggccaa agaacccttt agagattatg tagataggtt ctttaaaaca
 901 ctgagagctg agcaagctac acagggaagta aaaggctgga tgacagacac cttgttgggc
 961 caaantgcga cccagattg taagaccatc ttaagagcat taggaccagg agctacacta
 1021 gaagaaatga tgacagcatg ccaggagggtg ggaggaccgc gccataaagc aagagtttta
 1081 gctgaagcaa tgagccaagc atcaggtgta gcagcagcca tgatgatgca aaaaagcggg
 1141 ttttaaggcc cgaaaagaac tattaagtgt ttcaactgtg gcaagggaagg acatctagcc
 1201 agaaattgca gggcccttag gaaaagggc tgttggaaat gtggaaggga aggcacatcaa
 1261 atgaaagact gcacagagag acaggctaatt ttttagggga gaatttggcc ttcccacaag
 1321 gggaggccgg ggaattttct tcatatcagg acagagccaa cagccccacc tgcagaagag
 1381 gagataaact cctctccgaa gcaggaaacc agggacaagg aactatatcc tcttttagtt
 1441 tccctcaaat cac

gag ->

HIVVII191A2

LOCUS HIVVII191A2 2566 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, VII191 from Zaire, proviral DNA, envelope glycoprotein (env) gene.
 ACCESSION L22952
 SOURCE Human immunodeficiency virus type 1 (HIV-1), VII191 from Zaire, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 2566)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VII191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with A subtype sequences, whereas the corresponding gag sequence (L11793) clusters with G subtype sequences. See also accession numbers L22939-L22951, L22953-L22957 and L23064-L23065.
 NCBI gi: 437357

BASE COUNT 878 a 458 c 610 g 620 t

ORIGIN env cds start

```

1 ATGagagtga aggggatata gatgaattgg caacacttgt ggaagtggg aactttgatc
env start ->
61 cttgggttgg ggataatttg taatgctgaa aattttgtgg tcacagtatt ctatggggta
121 cctgtgtgga aagatgccga gaccacctg ttttgtgcat cagatgctaa agcatatagt
181 gaagagaaac acaatgtctg ggctacacat gcctgtgtac ccacagaccc cgacccaaaa
241 gaaataaatt tggtaaatgt gacagaagac ttcaacatgt ggaataataa aatggtagag
301 cagatgcagg aagatataat cagtctatgg gaccagagcc taaagccatg tgtgaagtta
361 accctctctc gcgttacttt agaattgcat aacagcacca cccagataa caataccacc
421 cagccacta tcatagatga aacaacaaac tgcacttaca atatgaccac agaattaaga
481 gataaggtac aaaaagtaca ttcaactttt tatagacttg atgtagtaca aattaatgga
541 attaatggaa ctatgaataa taatagtgtc cagtatagac taataaattg taatacctca
601 gccattacac aggcctgttc aaaggtatcc tttagagcaa ttcccatata ttattgtgcc
661 ccggtctggt ttgcatctc aaaaatgtaga gataaggagt tcaatggaac agggccatgc
721 aggaatgtca gctcagtaca atgcacgcat ggaatcaagc cagtagtate aactcagota
781 ctgttaaatg gcagctctagc agaagaaaag ataattgatta gatctgaaaa cttcacagat
841 aatgggaaaa ccataatagt acagcttgct gagcctgtaa aaattaattg taccagacct
v3 loop ->
901 agcaacaata caagaaaagg tatacatata ggaccaggac gagcattcta tgcaacaggt
961 cagataacag gggatataag acaagcacat tgtaacatca gtagaacaga atggaataaa
<- v3 loop
1021 actttacacc aggtagctat acaattaaga gaacacttta taactaagaa cacaacaata
1081 atctttaata gctcctcagg aggggatttg gaaattacaa cacacagttt taattgtgga
1141 ggagaatttt tctattgcag tacatcagcc ctgtttaata gcacttggga aaatggcact
1201 caggtggcaa atgacacggg atcaaatgac actataactc tcccatgcag aataaagcaa
1261 attgtaagaa tgtggcagag agtaggacaa gcaatgtatg cccctcccat ccaaggaaatg
1321 ataagatgtg aatcaaacat tacaggatta atattaacaa gagatgggtg tgggaataac
1381 agtacaaatg aaaccttcag acctggagga ggagatatga gggacaattg gagaagtga
1441 ttatataagt ataaagtatg aaaaattgaa ccattaggag tagcaccacg cagggcgaag
1501 agaagagtgg tggcagagaa aaaaagagca gttggactgg gagctgtttt ccttgggttc
1561 ttaggagcag cagggaagcag tatgggcgca gcgtcactaa cgctgacggg acaggccaga
1621 caattattgt ctggcatagt gcagcagcag agcaatctgc taagggtctat agaggctcaa
1681 catcatctgt tgcagctcac tgtctggggc attaaacagc tccaggcaag agtctctggt

```

```

1741 ctggaaagat accttaggga tcaacagctc ctaggaattt ggggatgctc tggcaaactc
1801 atctgcccc aataatgtgcc ctggaactat agctggagta ataatctta tgatgatata
1861 tggaaacaaca tgacctgggt gcaatgggat aaagaattg acaattacac acaattaata
1921 tatggtctaa ttgaagaatc gcagaaccag caggaaaaga atgaacaaga cttattggca
1981 ttggacaagt gggcaagtct gtggagttgg ttgacatat caaattgggt gtggtatata
2041 agaatattta taatgatagt aggaggtttg ataggtttaa gaatagtttt tgotgtgctt
2101 tctataataa atagagttag gcagggatac tcacctttgt cattccAGac cctccccac
                               3' sj (tat,rev) /\
2161 caccagaggg aaccogacag gcccgaggga atogaagaag gaggaggaga gcaagacaga
2221 ggcagatcca ttgattagt gagcggattc ttagctcttg cctgggacga cctacggagc
2281 ctgtgccttt tcagctacca ccgattgaga gacttcatat tgattgtagc gaggattgtg
2341 gagactctgg gacgcagggg gtgggagggt ctcaaatac tgtggaatct cctgttatat
2401 tggggtcggg aactgaaaat tagtgctatt actttgcttg atgctacagc aatagcagta
2461 gctaggtgga cagataggat tatagaaata ggacaaagaa ttgttagagc tattctcaac
2521 ataccacaga gaatcagaca gggagcggaa aaagctttgc aaTAAa
                                <- env cds end

```

HIVVI5251

LOCUS HIVVI5251 1471 bp ss-RNA VRL 20-APR-1993
 DEFINITION Human immunodeficiency virus type 1, Gabonese isolate VI525, gag region.
 ACCESSION L11792
 SOURCE Human immunodeficiency virus type 1 (HIV-1), Gabonese isolate VI525.
 REFERENCE 1 (bases 1 to 1471)
 AUTHORS Louwagie,J.J., McCutchan,F., Brennan,T., Peeters,M., Brennan,T., Sanders-Buell,E., Eddy,G., van der Groen,G., Franssen,K., Gershy-Damet,M., DeLays,R. and Burke,D.
 TITLE Phylogenetic analysis of gag genes from seventy international HIV-1 isolates provides evidence for multiple genotypes
 JOURNAL AIDS 7, 769-780 (1993)
 COMMENT Kindly provided prior to publication by Henry M. Jackson Foundation Research Laboratory, 1500 East Gude Drive, Rockville, MD 20850. Samples were obtained from 15 different countries (Belgium, Burundi, Gabon, The Ivory Coast, Kenya, Rwanda, Zaire, Thailand, Somalia, the Philippines, Djibouti, Brazil, Uganda, Senegal, and Zaire) on four continents. From this study the authors identified 7 distinct genotypes with an average gag intra-genotypic variance of 7% and an average inter-genotypic variance of 14%. The env region of this isolate(L22953) has also been sequenced. This gag sequence, previously published on page I-A-195 of the 1993 compendium, groups as subtype H in gag but G in env.

BASE COUNT 543 a 279 c 361 g 288 t

ORIGIN gag cds start

1 ATGggtgcga gaggctcagt attgagcggg ggaagattag atgcttggga gaaattcgg
 gag start ->

61 ttaaggccag ggggaaagaa acaatatagg ctanaacatc tagtatgggc aagcagagag
 121 ctggaaagat ttgcacttaa ccctggcctt ttagaaacat cagaaggctg tctacagata
 181 atagaacagc tacagccagc tcttaagaca ggaacagaag aacttcaatc attatttaat
 241 ctattagcag tctctatttg tgtacatcaa agaatagata taaaagacac caaggaagct
 301 ttaagtaaga tagaggaagt acaaaacaac agacagcaaa aaacacagca agcaacagct
 361 gataaggaaa aagacaaaaa ggtagtcaa aattatcccc tagtacagaa tgcccaaggg
 421 cagatggtac accaggccat atcacctagg accttaaatg catgggtaaa agtagtagaa
 481 gaaaaggctt ttagcccaga agtaatacct atgttttcag cattatcaga aggagccacc
 541 ccacaagact taaatgccat gctaaataca gtgggggggc atcaagcagc catgcaata
 601 ctaaaagata caatcaatga ggaagctgca gaatgggata ggctacatcc agtgcattga
 661 gggcctattc caccaggcca aatgagagaa ccaaggggaa gcgatatagc aggaactact
 721 agtacccttc aggaacaaat agcatggatg acaggaatc cagctatccc agtgggagac
 781 atctataaaa gatggataat cctgggatta aataagatag taagaatgta tagtccgttc
 841 agcatttttg acataaaaca aggaccacaa gaaccttta gagactatgt agacagggtc
 901 tttaaaactt taagagctga gcaagccaca caggatgtaa agaattggat gacagacacc
 961 ttgttggtcc aaaatgcgaa tccagattgc aagaatattt taaaagcatt aggaacaggg
 1021 gcttcaatag aagaatgat gacagcatgt caggaggtgg gaggacctag ccacaaagca
 1081 agagttttgg ctgaggcaat gagccaagta acaataacac atgcagccat aatgatgcag
 1141 aaaggcaact ttaagggtca aagaagaatt gttaaatgtt tcaactgttg caaagaagga
 1201 cacatagcta gaaattgcag ggcccctagg aaaaagggtt gttggaaatg tggaggggaa
 1261 ggacatcaga tgaaagattg cacagagaga caggctaatt ttttagggaa aatctggcct
 1321 tccagcaaaag ggaggccagg gaattttctt cagagcaggc cagagccaac agccccacca
 1381 gcagagagct ttgggttcgg agaggagatg accccctctc cgaagcagga gccccagggac
 1441 aaggaacctc ccttagcttc cctcaaatca c

gag ->

LOCUS HIVVI525A2 2604 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, VI525 from Gabon, proviral DNA, envelope glycoprotein (env) gene.
 ACCESSION L22953
 SOURCE Human immunodeficiency virus type 1 (HIV-1), VI525 from Gabon, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 2604)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence clusters with G subtype sequences, however the corresponding gag sequence (L11792) clusters with H subtype sequences. See also accession numbers L22939-L22952, L22954-L22957 and L23064-L23065.
 NCBI gi: 437358

BASE COUNT 904 a 458 c 601 g 641 t

ORIGIN env cds start

1 ATGagagtga gggggataca gaagaaattgg caaaacttgt ggaattgggg gactttgatc
 env start ->

61 cttgggttgg tgataattag tagtgccctca aataacttgt ggggtcacagt ctactatggg
 121 gtacctgtgt ggaagatgc agaaacccct ctattttgtg catctgatgc taaatcatat
 181 agtactgaaa gccataatgt ctgggctaca catgctgtg taccacaga cccaaccca
 241 caagagatat ctctggaaaa tgtaacagaa aattttaaca tgtggaaaaa taaatgggta
 301 gaacagatgc atgaggatat aatcagttta tgggatgaaa gottaaagcc atgtgtagag
 361 ctaacccttc totgtgttac tttaaactgt actaaggtaa cctacaatag cactgcagcc
 421 aacaatagat catccactgt caccacccct actagtactg taaccaacaa tagcactgtg
 481 agtggcaaag aagaatgac aaactgctct ttcaatgtaa ccacagaact aagagataag
 541 cagaaggcag aatacgcgct tttctataaa cttgatgtag taccattga taataataga
 601 aatgatagtt ataggctaat acattgtaat gtctcaacca ttaaacaggc ttgtcccaag
 661 gtatcttttg aaccattcc catacattat tgtgtccag ctggttttgc aattttaaag
 721 tgtagggaga agaatttcac tggaaacagga ctatgtaaaa atgtcagtac agtacaatgt
 781 acacatggaa ttaagccagt ggtatcaact caactactgc tgaatggcag tttagcagaa
 841 aaagaaatta taattagatc tgaaaacttc acaacaatg ccaaatgcat aatagtgcag
 901 cttaatgaaa ctgtagaat taattgtacc agaccaaca acaatacaag aaaaagtata

v3 loop ->

961 aaattcggaa caggacgagt gttatatgca acaggtgcc aataaggaaa tataagacaa
 1021 gcacattgta atgttagtac agaaaaatgg ggtaatgtg tacagaatgt agctgcaaaa

<- v3 loop

1081 ctacaaaaag aatttaataa aaacataacc ttttaactcat cctcaggagg ggacctagaa
 1141 attacaacac atagttttta ttgtagagga gaattttttt attgtaagac atcagagctg
 1201 ttttaatagta cttggttgtt taatagtact gggctgggta atagtacttg gaatgatata
 1261 cgaaaagata ttacactgca atgcaaaata aaacaaattg tgagaatgtg gcagagagtg
 1321 ggacaagcaa tgtatgcccc tcccatcgca gcaggaaaca ttgaatgtag atcaagcatt
 1381 acaggactaa tattaacaa agatgggtggg aataataata ataccaatga gaccttcaga
 1441 cctggaggag gagacatgag gggcaattgg agaagtgaat tatataagta taagtagta
 1501 aaatcaaat cactaggagt agcaccaccc agggcaagga gaagagtggg ggagagagaa
 1561 aaagagacaa ttggaactgg agctgtcttc cttgggttct taggagcagc aggaagcact
 1621 atgggcgcgg cgtcaataac gctgacggta caggtcagac aattattgtc tggcatagtg
 1681 caacagcaaa gcaatttgcg gagggctata gaggcgcagc agcatctgtt gcaactcaca

HIVV1S25A2

```

1741 gtctggggca ttaaacagct ccaggcaaga gtcttggttg tagaaagatt cctaaaggat
1801 caacagcccc tagggatttg gggctgctct ggaaaactca tctgcaccac taatgtgccc
1861 tggaaatacta gttggagtaa taaatcctat aatgacattt gggataacat gacctggatg
1921 gaatgggaaa gggaaatcag caattacaca aaacaaatat acagcctaast tgaagaatcg
1981 cagatccagc agggaaaagaa tgaacaagac ttattggcat tggacaaatg ggcaagtctg
2041 tgaatttggg ttgacatata aagctggcta tggatatata aaatatattat aatggtagta
2101 ggagggtttaa taggtttaag aatagttttt gctgtgcttt ctatagtaaa tagagtcagg
2161 cagggatact cacctctgtc attccAGacc cttaccatc accaagggga acccgacagg
      3' sj (tat,rev) /\
2221 ctgaaaagaa tgaagaagg aggtggcgag cagcacagaa ccagatccac gogatcagcg
2281 aaaggattct taccacttgt ctgggacgac ctctggagcc tgtgcctttt cagctaccgc
2341 ctcttgagag acttactctt aattgtagcg aggactgtgg aacttctggg gagaaggggg
2401 tgggaagccc tcaaatatct gtggaatctt ctgcagtatt ggggacagga actaaagaat
2461 agtgccattc atttgcttaa caccacagca atagcagtag ctgaggggaa agatagggtt
2521 atagaaatag tgcaaatagc ttgcagagct gttcttcaca tacctagaag aataagacag
2581 ggctttgaaa gggctttgct aTAA
      <- env cds end

```

LOCUS HIVZAM181 1477 bp ss-RNA VRL 06-OCT-1992
 DEFINITION Human immunodeficiency virus type 1, ZAM18 core protein (gag) gene, 5' end.
 ACCESSION L03705
 SOURCE Human immunodeficiency virus type 1, individual_isolate ZAM18 from Zambia RNA.
 REFERENCE 1 (bases 1 to 1477)
 AUTHORS McCutchan,F., Ungar,B.L.P., Hegerich,P.A., Roberts,C.R., Fowler,A.K., Hira,S.K., Perine,P.L. and Burke,D.S.
 TITLE Genetic Analysis of HIV-1 Isolates from Zambia and an Expanded Phylogenetic Tree for HIV-1
 JOURNAL J. Acquir. Immune Defic. Syndr. 5, 441-449 (1992)
 COMMENT See related env sequence (L22954). Both the ZAM18 gag and env sequences cluster with C subtype sequences. This sequence was previously presented on page I-A-170 of SEP92.
 NCBI gi: 325636

BASE COUNT 554 a 293 c 353 g 277 t

ORIGIN

```

1 ATGggtgoga gaggctcaat attaaaaggg ggaaaattag atacatggga aaaaattagg
gag start ->
61 ttaaggccag ggggaaagaa acactatatg ataaaacacc tagtatgggc aagcagggag
121 ctggaaagat ttgcacttaa ccttggcctt ttagagacat cagaaggctg taaacaaata
181 ataaaacagc tacaaccatc tattcagaca ggaacagagg aacttaaatc attattcaac
241 acagtatgaa ctctatggtg tgtacatgaa gatataacag tacgagacac caaggaagcc
301 ttagacaggt tagaggaaga acaaaacaaa agtcagcaaa aaacaaaaac acagcaggca
361 aaagcggctg acggacaggt cagtcaaaat tacctatag tgcagaatct ccaagggcaa
421 atgatacacc aggcacatc acctagaact ttgaatgcat gggtaaaagt aatagagggg
481 aaagccttta gccagagat aatacccatg ttacagcat tatcagaagg agccacccca
541 caagatttaa acaccatgtt aaatcacgtg gggggacatc aagcagccat gcagatgtta
601 aaagatacca tcaatgagga ggctgcagaa tgggatagat tacatccagt acatgcaggg
661 cctaatgcac caggccaatt gagagaacca aggggaagtg acatagcagg aactactagt
721 acccttcagg aacagatagc atggatgaca aataacccac ctgttccagt gggagacatc
781 tataaaagat ggataattct ggggttaaat .aaaatagtea gaatgtatag cctgtgcage
841 attttggaca taaaacaagg gccaaaggaa ccttttagag actatgtaga ccggttcttt
901 aaacttttaa gagctgaaca agctacacaa gatgtaaaaa attggatgac agaaaccttg
961 ttggtccaaa atgcgaaccc agattgtaag accattttta gagcattagg accagcggct
1021 acattagaag aaatgatgac agcatgtcag ggagtgggag gacctagcca caagcaaga
1081 gtgttagctg aggcaatgag ccaagcaaac aatacaaaaca tactgatgca gagaagcaat
1141 tttaaaggac ctaaaagaag tgtaaatgt ttcaactgtg gcaaggaagg gcacatagcc
1201 aaaaattgca gggcccttag gaaaaaaggg tgttggaaat gtggaaagga aggacaccaa
1261 atgaaagact gtaccgagag gcaggctaatt tttttagggg aaatctggcc ttcccacaag
1321 gggagggccag ggaatttctt tcagagcagg ccagagccaa cagcccttag gccagagcca
1381 acagcccccac cagcagagag cttcaggttc gaggagacga ccccgctcc gaatcaggag
1441 ccgaaagaca gggaacccctt aacttccctc aaatcac

```

HIVZAM18A

LOCUS HIVZAM18A 2544 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1, ZAM18 from Zambia, proviral envelope glycoprotein (env) gene.

ACCESSION L22954

SOURCE Human immunodeficiency virus type 1 (HIV-1), ZAM18 from Zambia, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 2544)

AUTHORS Louwagie,J.J., Janssens,W., Mascola,J.J., Fischer,C.L., Hegerich,P.A., van der Groen,G., McCutchan,F.F.E., Eddy,G. and Burke,D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence and the corresponding gag sequence (L03705) cluster with C subtype sequences. See also accession numbers L22939-L22953, L22955-L22957 and L23064-L23065. NCBI gi: 437359

BASE COUNT 904 a 439 c 598 g 603 t

ORIGIN env cds start

1 ATGagagtga gggagatact gaggaattgg caacaatggt ggatatgggg catcttaggc
env start ->

61 ttttggatgg taatgaacta taatgtggta ggagacaact tgtgggtcac agtctattat
121 ggggtacctg tgtggaaga agcaaaaact actttattct gtgcacaga tgctaaagca
181 tatgagagag aagtgcataa tgtctgggct acacatgect gtgtaccac agacccaac
241 ccacaagaaa tagttttggg aaatgtaaca gaaatttta acatgtggaa aaatgacatg
301 gtggatcaga tgcattgagg tataatcagg ttatgggata aaagcttaaa gccatgtgta
361 aagttgaccc cactctgtgt cactttagaa tgtggaaatg ttaatgttac ccattgagaat
421 agcacgaagg gggaaatgaa aaattgctct ttcaatgcaa ccacagaact aaaagataaa
481 aaacagagag tgtatgcact tttttataaa ctgtatatag taccacttaa tgagaataac
541 aactctagtg aggaacttag tgaatataga ttaataaatt gtaatacctc agccataaca
601 caagcctgtc caaaggtcac tttggatcca attcctatac attattgtgc tccagctgga
661 tatgcgattc taaagtgtaa taataagaca ttcaatggga caggaccatg ccataatgtc
721 agcacagtac aatgtacaca cggaaatcaag ccagtggat caactcaact actgttaaat
781 ggtagcctag cagaagaaga gataataatt aggtctgaaa atctaacaaa caatgccaaa
841 acaataatag tacatcttaa tgaatctgta gaaattgtgt gtacaagacc cagcaataat
v3 loop ->

901 acaagaaaaa gtataaggat aggaccagga caagcattct atgcaacagg tggcataata
961 ggaacataa gacaagcaca ttgtaacatt agtaaagaga actggaataa aactttacaa
-< v3 loop

1021 aaggtaggaa aaaaattagc agagcacttc cctaataaaa caataaaatt tgaccaaacac
1081 tcaggagggg acctagaaat tacaacacat agctttaatt gttagaggaga atttttctat
1141 tgcaatacat caaacctgtt taatagtaca tataagccta atgatacaaa tagtacatat
1201 aatcctaatt atacaatcac actcccatgc agaataaaaac aaattataaa catgtggcag
1261 ggggtaggac aagcaatgta tgcacctccc attgcaggaa acataacatg taatcaaat
1321 atcacaggac tactattgac aogggatgga gggctcaaat ataccacaaa cacagagaca
1381 ttacagacct gaggaggaga tatgagggac aattggagaa gtgaactata taatatataa
1441 ctggttagaa ttaaacattt gggcatagca cccactgagg caaaaaggag agtgggtgag
1501 acgagagaaa aaagagcagt gggaaatagga gctgtgttcc ttgggttctt gggagcagca
1561 ggaagcacta tgggcgcggc gtcaataacg ctgacggcac aggcagaca agtgttgtct
1621 ggtatagtgc aacagcaaag caatttgctg agagccatag aggcgcaaca gcatctgttg
1681 caactcacgg tctggggcat taagcagctc cagacaagag tcttggttat ggaagatac

```

1741 ctaaaggatc aacagctcct agggatttgg ggctgctctg gaaaactcat ctgcaccact
1801 gccgtgtctt ggaactccag ttggagtaat aagtctcaa cagatatttg ggataacatg
1861 acctggatgc agtgggtag agaatattagc aattacacag acacaatata caggttgctt
1921 gaagaatcgc aaattcagca ggaagaatat gaaaaagact tattagcatt ggacagttgg
1981 aaaaatctat ggaattggtt tgacataaca aattggctgt ggtatataaa aatattcata
2041 atgatagtag gaggcttgat aggettaaga ataagttttg ctgtgctttc tatagtgaat
2101 agagttaggc agggatactc acctttgtcg ttccAGacc ttaccccgaa cccaagggga
      3' sj (tat, rev) /\
2161 cccgacaggc tcggaagaat cgaagaagaa ggtggagagc aagacagaga caaatccatt
2221 cgcttagtgg acggattctt agcacttttc tgggacgacc tgcggagcct gtgccttttc
2281 agctaccacc gattgagaga cttcatattg attgcagcga gacgagcgga acttctggga
2341 cgcagcagtc tcaggggact acagaggggg tgggaagccc ttaagtatct ggaagtctt
2401 gtgcagtatt ggggtctgga actaaaaaag agtgcatta gtctgggtga tgccatagca
2461 ataacagtag ctgaaggaac agataggatt atagaattca tacaagaat ttgtagagct
2521 atccgcaaca tacctagaag aata
      env ->

```


HIVZAM184

LOCUS HIVZAM184 3307 bp ds-DNA VRL 31-DEC-1993

DEFINITION Human immunodeficiency virus type 1. ZAM184 from Zambia, proviral DNA encoding env, tat, vpu, rev, and nef genes.

ACCESSION L22955

SOURCE Human immunodeficiency virus type 1 (HIV-1), ZAM184 from Zambia, PCR amplification of proviral DNA from cocultivated PBMC's.

REFERENCE 1 (bases 1 to 3307)

AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.F.E., Eddy, G. and Burke, D.

TITLE Genetic diversity of the HIV-1 envelope glycoprotein

JOURNAL Unpublished (1993)

COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, K124, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. The subtype classification of ZAM184 is uncertain at this time. See also accession numbers L22939-L22954, L22956-L22957 and L23064-L23065.

NCBI gi: 437360

BASE COUNT 1133 a 594 c 794 g 786 t

ORIGIN 226 bp upstream from the beginning of rev cds start

1 ggcttcogtc tagatgctgt ttattcattt cagaattggg tgccaacata gcagaatagg
61 cattactcga agaagaagag taagggATGg atccagtaga cctaaacta gagccctgga
tat cds start ->

121 accatccggg aagtcagcct aaaactgctt gtaccaagtg ttattgtaaa aagtgttgc
181 atcattgccc agtgtgcttt ctaaacaaag gcttaggcct ctctATGgc aggaagaagc
rev cds start ->

241 ggagaccccg acgaagatct cctcaaaaca gcgaggatca tcaaaatcct ataccaagc
301 aGTAagtagt agtaattaat atatgtaATG ttaccttag caacattgtc aatagtagga
5' sj /\ vpu cds start ->

361 ctgatagtag cgttaattct agcaatagtg gtatggacta tagtatttat agaataaaa
421 aaaattaaag agcaaaagaa aatagactgg ctaatcgaga gaataagtga aagagcagaa
481 gacagtgcc ATGagagtga gggggatacc gaagaattgg caacaatggt ggatatggg
env cds start ->

541 catcttaggc ttttggatgt taatgatttg TAGtgtggtg ggaacttgt gggtcacagt
<- vpu cds end

601 ctattatggg gtacctgtgt ggagagaagc aaaaactact ctattctgtg catcagatgc
661 taaagcatat gagagagaag tgcataatgt ctgggctaca catgcctgtg taccacaga
721 ccccaacca caagaatag ttttggaaa tgtaacagaa aattttaaca tctggaaaaa
781 tgacatgggt gatcagatgc atgaggatat aatcagtttg tgggatcaa gccataagcc
841 atgtgtaaaa ttaacccctc tctgcgtcac tttaaattgt agtgcgcca ctagtaccaa
901 tagcaacacc actgaagaaa taaaaaactg ctcttacagg actaccacag aactaaaaga
961 taagacgctg agagtctatt cactgtttta taaacttgat gtagtacaaa ttgatccaag
1021 taataatact atgtatagac taataagtrg taatacctca gccatcgcac aagcttgtcc
1081 aaaggtatcc tttgagccaa ttctatatac ttattgtacc ccagctgggt ttgcgacct
1141 aaagtgtaa gatacggat tcaatggaac agggtcagtc aagaatgta gctcagta
1201 atgtacacat ggaattagc cagtagtacc aactcaactg ctgttgatg gcagctcagc
1261 agaaggagag gtaatgatta gatctgagaa tattacagac aatgccaaa acataatagt
1321 acagtttact gaacctgtac aaattagttg taccagaccc agcaacaata caagaagaag
v3 loop ->

1381 tatccgcttt ggaccaggac aagccttcta tacaatgac ataatagggg atataagaca
1441 agcatattgt agtatcaacg aaacaaaatg gaatgccact ttacaaaagg tagctaaaca
<- v3 loop

1501 attacaaaac tacttcccta acaaaacaat aagctttgct aactcttcag gaggggacct

```

1561 agaaattaca acacatagtt ttaattgtgg aggagaatTT ttctattgca atacatcaaa
1621 cctgtttaat agtacataca atcctaatag tacagaaagt aattcagact taaacatcac
1681 actcccatgc agaataaaac aaattataaa catgtggcag agagttaggac gagcaatgta
1741 tgccccctcc attgcaggag taataaagtg tacatcaaac attacaggaa taatattgtc
1801 aagagatggg gggaacagtg ggaacagtac caatgggact gagaccttca gacctggagg
1861 aggagatatg agggacaatt ggagaagtga attatataag tataaagtag taaaactaga
1921 accactagga gtagcaccca ccaagccaag gagaagagtg gtggagagag aaaaaagagc
1981 agttggactg ggagctgtct tccttgggtt cttgggagca gcaggaaaga ctatggggcg
2041 ggcatacata acgtgacggg tacaggccag acaactgttg tctggtatag tgcaacagca
2101 aagcaatttg ctgaaggcta tagaggctca acagcatctg ttgaaactca cggctctggg
2161 cattaaacag ctccaggcaa gagccctggc tatggaaaga taactacagg atcaacagct
2221 cctgggaatt tggggctgct ctggaaaact catctgcacc actactgtgc cctggaaactc
2281 tagttggagt aataagactc aggatgagat ttgggacaaa atgacctggt tgcaatggga
2341 caaagaaatt agcaattace cagaataat ataccagcta cttgaggact cgcagaacca
2401 gcaggaaaag aatgagcaag acttattggc attagacaaa tgggcaaat tgtggaattg
2461 gtttgacata acaagatggc tatggtatat aaaaatattt ataataatag taggaggctt
2521 gataggatta agaatagtta ttgtatata ttctgtagta aatagagtta ggcagggata
2581 ctacaccttg tcatttcAGa tccctacccc acaccagag ggtctcgaca ggcccggag
      3' sj {tat,rev) /\
2641 aatgaagaa ggaggtggag agcaaggcag agacagatcg attcgattag tgagcggatt
2701 cttagcgctt gcctgggacg atctgaggag cctgtgcctc ttcagctacc accgcttgag
2761 agattgcac ttgattgcag cgaggactgt ggaacttctg ggacacagca gtctcaaggg
2821 actgagactg ggggtgggaag gcctcaagca tctatggaat cttctgttat attggggcta
2881 ggaattgaag aatagtgtca tTAGtttact tgataccata gcaatagcag tagctgagtg
      <- rev cds end
2941 gacagatagg attatagaaa taggacaaag agcttgacga gctattcgca atatacctac
3001 aagaataaga cagggtcttg aaagggtttt actgTAAaAT Gggaggcaag tggtcacaaa
      <- snv cds end
      nef cds start ->
3061 gtagcatagt gggatggcct gaagttaggg agagaatgaa aagaactcct ccagcagcag
3121 aaggagttag agcagtgtct caagatttag ctgcacgtgg agcagtcaca agcagtaata
3181 cagcacatac taatcctgat tgtgcctggc tggaaagcga agaagagggg gagggagtag
3241 gcttcccagt caagccacag gtacctttaa gaccaatgac ctacaagggg gctgtggatc
3301 tcaccac

```

HIVZAM201

LOCUS HIVZAM201 1488 bp ss-RNA VRL 06-OCT-1992
 DEFINITION Human immunodeficiency virus type 1, ZAM20 core protein (gag) gene, complete cds.
 ACCESSION L03707
 SOURCE Human immunodeficiency virus type 1, individual_isolate ZAM20 from Zambia RNA.
 REFERENCE 1 (bases 1 to 1488)
 AUTHORS McCutchan,F., Ungar,B.L.P., Hegerich,P.A., Roberts,C.R., Fowler,A.K., Hira,S.K., Perine,P.L. and Burke,D.S.
 TITLE Genetic Analysis of HIV-1 Isolates from Zambia and an Expanded Phylogenetic Tree for HIV-1
 JOURNAL J. Acquir. Immune Defic. Syndr. 5, 441-449 (1992)
 COMMENT See related env sequence (L22956). Both the ZAM20 gag and env sequences cluster with C subtype sequences. This sequence was previously published on page I-A-172 of SEP 92.
 NCBI gi: 325640

BASE COUNT 543 a 293 c 373 g 279 t
 ORIGIN

```

1 ATGgggtgga gagcgtcaat attaagaggg gaaaaattag ataatggga gaaaattagg
gag start ->
61 ttaaggccag ggggaagaa acattatatg ataaaacatt tagtatgggc aagcaggag
121 ctggaaagat ttgcacttaa ccctggcctt ttagagacag cagaoggctg taacaaaata
181 ataagacagc tacatccagc tcttcagaca ggaacagagg aacttagatc attatataac
241 acagtagcaa ctctctattg tgtacatgca gggatagagg tacgagacac caaggaagcc
301 ttagacgaga tagaggaaga acaaaaacaa tctcagcaaa aaacacagca ggcaaaaacg
361 gctgacggaa aagtcagtca aaatcttctt atagtacaga atcttcaagg gcaaatggta
421 catcaggcca tatcacttag aactttgaat gcattgggtta aagtaataga gggagaaggc
481 ttttaagccag aggtaatacc catgtttaca gcattatcag aaggagccac cccacaagat
541 ctaaacacca tgtaaatac agtgggggga caccaagcag ccatgcaaat gttaaaagac
601 accatcaatg aggaggctgc agaattggat agaatacatc cagtacatgc agggcctgtt
661 gcaccaggcc aaatgagaga accaagggga agtgacatag caggaactac tagtaccctt
721 caggaacaaa tagcatgat gacagctaac ccacctgttc cagtgggaga aatctataaa
781 agatggataa ttctggggtt aaataaaaata gtaagaatgt atagccctgt cagcattttg
841 gacataagac aagggccaaa ggaacctttt agagactatg tagaccggtt ctttaaaact
901 ttaagagctg aacaagctac acaagatgta aaaaactgga tgacagacac cttgttggtc
961 caaaatgga acccagattg taagaccatt ctaagagcat taggaaccag ggctacatta
1021 gaagaaatga tgacagcatg tcagggagtg ggaggacctt gccacaaagc aagagtgttg
1081 gctgaggcaa tgagccaaac aaacaggcaa atgtaatgat gcagagaggc aatttttaag
1141 gccctaaaag aactgtttaa tgtttcaact gtggcaagga agggcacata gccagaaatt
1201 gcagggtctc tcggaaaaag ggctgttggg aatgtgggag ggaaggacac caaatgaaag
1261 attgtactga gagacaggct aatttttttag ggaaaatttg gccttcccaa aaggggaggg
1321 cagggaaattt cctccagaac aggccagagc caacagcccc accagcagag aacttcaggc
1381 caacagcccc accagcagag agcttcaggt tcgaggagac aacccccgtt ccgaagcagg
1441 agccgagaga cagtccctaca ggggaacctt taactgcctt caaatcac

```

LOCUS HIVZAM20A 2625 bp ds-DNA VRL 31-DEC-1993
 DEFINITION Human immunodeficiency virus type 1, ZAM20 from Zambia, proviral envelope glycoprotein (env) gene.
 ACCESSION L22956
 SOURCE Human immunodeficiency virus type 1 (HIV-1), ZAM20 A from Zambia, PCR amplification of proviral DNA from cocultivated PBMC's.
 REFERENCE 1 (bases 1 to 2625)
 AUTHORS Louwagie, J.J., Janssens, W., Mascola, J.J., Fischer, C.L., Hegerich, P.A., van der Groen, G., McCutchan, F.P.E., Eddy, G. and Burke, D.
 TITLE Genetic diversity of the HIV-1 envelope glycoprotein
 JOURNAL Unpublished (1993)
 COMMENT These sequences were kindly provided prior to publication by Dr. Joost Louwagie of the Henry M. Jackson Foundation Research Laboratory Rockville, Maryland. Twenty-one full length gp160 coding sequences from eight African countries (Djibouti, Gabon, Kenya, Senegal, Somalia, Uganda, Zaire, and Zambia) were sequenced and analyzed with thirty-two previously published full-length gp160 env sequences. Sequences DJ258, DJ259, KI24, SE365, SM145, UG266, UG268, UG274, VI191, and VI525 have corresponding gag sequences published by Louwagie et. al. in AIDS 7, 769-780 (1993) and in the 1993 compendium. This sequence, and the corresponding gag sequence (L03707) cluster with C subtype env sequences. See also accession numbers L22939-L22955, L22957, and L23064-L23065.
 NCBI gi: 437361

BASE COUNT 910 a 455 c 620 g 640 t

ORIGIN env cds start

1 ATGagagtga gggggatact gaggaattgt caacaatggt ggatatgggg catcttaggc
 env start ->

61 ttttggatgt taatgatttg taatgggggtg ggggacttgt gggtcacagt ctattatggg
 121 gtacctgtgt ggaaggaagc aaaaactact ctattctgtg catcagatgc taaagcttat
 181 gagacagaag tgcataatgt ctggggtaca catgcctgtg tgcccacaga tcccagccca
 241 caagaaatgg ttttggcaaa tgaacagaa aactttaaca tgtggaaaaa tgacatgggtg
 301 gatcagatgc atgaggatat aatcagttta tgggatcaaa gcttaaaaggc atgtgtaaaa
 361 atgacaccac totgtgtaac tttagaatgt ggaatgtga ctgttcccaa ggataatagt
 421 acctacaatg gtgatataca tagtaatact acctgaaag aagaatgag aaattgctot
 481 ttcaatgcaa ccacagtaat aaatgataag acaagtaagg tgcattgact tttttatagg
 541 gttgatatag taccacttaa tgaatgagaat agtaacaaga ctactgcaa ctctactaag
 601 ctactaagt atataattaat aaattgtaat acctcaacca taacacaagc ctgtccaaag
 661 gtcagtttgg gcccaattcc tatacattat tgtgtccag ctggatttgg gattctaaag
 721 tgaaccaata agacattcaa tgggacaggg ccattgccata atgtcagcac agtacaatgt
 781 acacatggaa ttaagccagt ggtatcaact caactactgt taaatggtag cctatcagaa
 841 gaagggataa taattagatc tgaatatctg acaacaatg cccaacaat aatagtacag
 901 cttaataaat ctatagagat tgtgtgtgca agacctggca ataatacaag aaaaagtata

v3 loop ->

961 aggataggac caggccaac attctttgca acaggtgcca taataggaga cataagacaa
 1021 gcacattgta acattagtag aagtgaatgg aatgaacat taccaaatgt tagtaggaaa

<- v3 loop

1081 ttagcaggac acttccctaa taggaccata atatttgaac catcctcagg aggggacctta
 1141 gaaattacaa cacatagctt taattgtaga ggagaatttt tctattgcaa tacctcaatc
 1201 ctgtttaata gatcacactt aggtctgttt aatggtacag aaaataatcc accccaaaac
 1261 atcaactatcc catgcagaat aaaccaaatt ataaacatgt ggcaggagggt aggaaaagct
 1321 atgtatcccc ctcccattgc aggaacata acatgtatat caaatatcac aggactacta
 1381 ttgacacgtg atggaggaat agatacaaat agtaaaaaa tcacacatga aacaatggc
 1441 acaaggaacg agacattcag acctggagga ggagatatga gggacaattg gagaagtga
 1501 ttatataaat ataaagtagt agaaattaa ccactaggaa tagcaccac tgggcccacaa
 1561 aggagagtgg tggagaagca gatagagaag agagcagtgg gaataggagc tgtgttctct
 1621 ggggttcttg gagcagcagg agcaactatg ggcggggcgt cactgacgt gacggtacag
 1681 gccagacaag tgttctctgg tatagtcaa cagcaaaagca atttgctgag agctatagag
 1741 gcccaacagc atatatgtca actcaaggtc tggggcatta agcagctcca ggcaagagtc

HIVZAM20A

```

1801 ctggccatag aaagatacct aaaggatcaa cagctcctag ggatttgggg ctgctctgga
1861 aaactcatct gcaccactgc tgtgccttgg aactccagtt ggactaataa aacttacgaa
1921 tatatttggg ataacatgac ctggatgcag tgggagaaag aaattagtaa ttactcagac
1981 acaatatata agttgcttga agactcgcaa acccagcagg aacaaaatga aaaagagtta
2041 ttagcattgg acaagtggaa tattctgtgg aattggtttg acatatcaaa atggctgtgg
2101 tatataagaa tattcataat gatagtagga ggcttgatag gtttaagaat aatttttgct
2161 gtgctatcta tagtgaatag agttaggcag ggatactcac ctttgctggt tcAGaccctt
                               3' sj (tat,rev) /\
2221 accccgaaca tcagggggga acccgacagg ctcggaagaa tgaagaaga aggtggagag
2281 caagacagag acagatccat tcgattagtg aacggattct tagcacttgt ctgggacgac
2341 ctgcgagacc tgtgccgttt cagctaccac caattgagag acttoacatt gattgcagcg
2401 agagcggtgg aacttctggg acggagcagt ctcaggggac tacagagggg gtgggaagcc
2461 cttaaatata tgggaagtct tgtgcagtac tggggtctgg aactaaaaaa gagtgccatt
2521 agtttgcttg ataccatagc aatagcagta gctgaaggaa cagataggat tatagaatta
2581 gtacaaagaa tttgtagagc tatccgtaat atacctagaa gaata
                               env ->

```

CONSENSUS .A	ATGGGCGCGAGAACTCCGTCCTTGAGAGGGAAAAAGCAGA?GAATTAGAAAAaTtAGGTTACGGCCCGCG	72
HIV2ROD	-----T-----G-----C-----	73
HIV2NIH2	-----T-----	73
HIV2ISY	-A-----G-----T-----	73
HIV2ST	-----C-----G-----	73
HIV2BEN	-----C-----G-----	73
HIV2CAM2	-----C-----G-----	73
HIV2D194	-----C-----G-----AA--	73
HIV2GH1	-----T-----G-----A--	73
CONSENSUS .B	ATGGGCGCGAGA?GCTCCGTC?TGTCAGGGAAGAAAAACAGATGAATTAGAAAAaTTAGGTTACGGCCCGCG	71
HIV2D205	-G-----C-----	73
HIV2UC1	-A-----T-----	73
CONSENSUS .D	ATGGGcGcGAGAAaCTCCGTCCTTGTCAGGGAagaAAGCAGATGAATTAGAAAAaTTAGGcTACGaCCCggCG	73
SIVMM251	-----	73
SIVMM32H	-----	73
SIVMM1A11	-----AA--	73
SIVMM142	-----A--	73
SIVMM239	-T-----AA--	73
SIVMNE	-----	73
SIVSMH4	-----G-----G-----T-----G-----AA--	73
SIVSMH9	-T-T-----T-----G-----	73
SIVSMMPBJ	-----T-----G-----	73
SIVSTM	-G-----A-----G-----T-----G-----	73
SIVSTM22579	-G-----A-----G-----T-----G-----	73
CONSENSUS .A	GaAAGAAAAa?TAcaggtTAAACATaTTGTGTGGGCAGCGAAtgAAATTgGACAgATTcGGATTggCAGAGAG	144
HIV2ROD	-----G-----C-----A-----A-----	146
HIV2NIH2	-----G-----A-----A-----A-----	146
HIV2ISY	-G-----A-----A-----AA--	146
HIV2ST	-----A-----T-----C-----	146
HIV2BEN	-----G-----A-----T-----A-----	146
HIV2CAM2	-----A-----A-----	146
HIV2D194	-----GA-----G-----A-----	146
HIV2GH1	-----G-----C-----C-----A-----	146
HIV7924A	-----A-----	36
HIV60415K	-----A-----	36
HIVFA	-----C-----T-----	36
CONSENSUS .B	GAAAGAAAAaGTAT??CTAAACAT?TA?TATGGGCAGtGAACGAATTgGACAGATTtGGATTaGCaGAgag	138
HIV2D205	-----A-----ATG-----G-----G-----A-----	146
HIV2UC1	-----G-----TGT-----A-----A-----	146
HIV60667K	-----A-----T-----A-----A-----G-----A-----A	36
HIVJA	-----A-----G-----	36
HIV7312A	--CA--T-----A-----C-----G-----A-----A	36
HIVFT	-----T-----C-----	36
HIV7810A	--CA-----C-----	36
HIVON	-----G-----C-----	36
CONSENSUS .C	AGCGAACGAATTGGACAGATTTCGATTAGCAGAAAG	36
HIV22238	-----	36
CONSENSUS .D	GaAAGAAAAaGTAcAtGtTGAAGCATgTAGTaTGGGCAGCAAAcGAATT?GA?AGATTTCGATTaGCAGAAAG	144
SIVMM251	-----A-----T-----	146
SIVMM32H	-----A-----T-----	146
SIVMM1A11	-----A-----T-----	146
SIVMM142	-----A-----T-----	146
SIVMM239	-----A-----T-----	146
SIVMNE	-G-----A-----T-----	146
SIVSMH4	-----T-----G-----C-----	146
SIVSMH9	-----C-----A-----G-----C-----	146
SIVSMMPBJ	-G-----G-----CA-C-----A-----G-----C-----	146
SIVSTM	-G-----A-----G-----C-----G-----C-----	146
SIVSTM22579	-G-----A-----G-----C-----G-----C-----	146
HIVFO784	-----G-----C-----	36
CONSENSUS .E	AGCCAGAGAAATTAGACAGATTTCGATTGGCAGAAAG	36
HIVPA	-----	36

HIV2/SIV GAG

CONSENSUS .A	CCTGTTGGAGTCAAAAGAAAGGTTGCCAAAaaATTCTtaccAGPTTTAGATCCAcTaGTACCGACAGGgTCAGAA	217
HIV2ROD	-----G-----T-----A-G-----T-----	219
HIV2NIH2	-----A-----T-----A-----A-----T-----	219
HIV2ISY	-----G-----	219
HIV2ST	-----G-----C-G-----A-----	219
HIV2BEN	-----A-----G-----A-----T-----	219
HIV2CAM2	-----A-----G-----A-----A-----A-----T-----	219
HIV2D194	-----A-----G-----C-----A-----	219
HIV2GH1	-----T-----G-----A-----T-----C-----	109
HIV7924A	-----G-----	109
HIV60415K	-----G-----A-C-----A-----	109
HIVFA	-----G-----A-C-----A-----	109
CONSENSUS .B	CCTGTTGGGaaTCAAAAGAAAGGATGTCACAAAATTCtaccGtCTTAGctCCATTAgTaCCAACAGGgTCAGAA	211
HIV2D205	-----A-----AA-----G-----	219
HIV2UC1	-----G-----C-----T-----CG-A-----AG-----	219
HIV60667K	-----G-----C-----T-----CG-A-----AG-----	109
HIVJA	-----C-----G-----A-G-----G-----	109
HIV7312A	-----G-----	109
HIVFT	-----G-----C-----	109
HIV7810A	-----G-----C-----	109
HIVON	-----T-----	109
CONSENSUS .C	CCTGTTGGAGACCAAGAAAGGATGCCAAAAATTTTAACAGTGTAGAACCATTAGTTCCGACAGGTTTCAGAA	109
HIV22238	-----	109
CONSENSUS .D	CCTGTTGGAGaaacAAAGAAAGGATGTCAAAAATTCtaccGtCTTAGctCCATTAgTgCCaACAGG?TCAGAA	215
SIVMM251	-----A-----C-----	219
SIVMM32H	-----A-----C-----	219
SIVMM1A11	-----A-----C-----	219
SIVMM142	-----A-----C-----	219
SIVMM239	-----A-----C-----	219
SIVMNE	-----A-----C-----	219
SIVSMNH4	-----T-----T-----T-----A-----T-----T-----G-----T-----	219
SIVSMNH9	-----T-----T-----T-----A-----T-----T-----G-----T-----	219
SIVSMNPBJ	-----T-----T-----T-----A-----T-----T-----G-----T-----	219
SIVSTM	-----TCA-----G-----TA-AA-----AA-----T-----	219
SIVSTM22579	-----TCA-----G-----T-----T-----AA-A-----AA-----T-----	219
HIVFO784	-----A-----T-----G-----T-----C-----A-----G-----T-----	109
CONSENSUS .E	CCTGTTGGAAAACAAGGAAGGATGTCACAGAATACTGACAGTACTAGAACCATTAATGCCAACAGGCTCTGAA	109
HIVFA	-----	109
CONSENSUS .A	AATTTAAaAAAGcCTTtTtTAAtACtqtCTGCGTcaTtTGCGTcaTaCACGCAGAGAGAAAGtGAAAGAtACTG	290
HIV2ROD	-----T-----	292
HIV2NIH2	-----C-----	292
HIV2ISY	-----A-----AC-----T-G-----	292
HIV2ST	-----C-----T-G-----C-----	292
HIV2BEN	-----C-----T-G-----	292
HIV2CAM2	-----C-----A-----T-G-----	292
HIV2D194	-----C-----A-----T-G-----	292
HIV2GH1	-----C-----A-----T-G-----	292
HIV7924A	-----C-----AG-----C-----	182
HIV60415K	-----C-----GC-----	182
HIVFA	-----C-----T-----A-----C-----T-G-----	182
CONSENSUS .B	AATTTAAaAGcCTTTTAAATAcTGTCTGCGTcaTtTtTGCGTgCACGCAGAGAGAAAGTGAAGATACTG	284
HIV2D205	-----TC-----T-----C-----A-----	292
HIV2UC1	-----C-----C-----	292
HIV60667K	-----G-----C-----C-----G-----	182
HIVJA	-----T-----TG-----	182
HIV7312A	-----GG-----A-----A-----C-----	182
HIVFT	-----TC-----C-----	182
HIV7810A	-----G-----A-----	182
HIVON	-----C-----C-----	182
CONSENSUS .C	AATTTAAAGAGCCTTTATAATACCTGCGTCACTTGGTGTtTACACGCAGAACAGAAAGTGAACATACAG	182
HIV22238	-----	182
CONSENSUS .D	AATTTAAaAGcCTTTTAAATAcTGTCTGCGTcaTtTtTGCGTgCACGCAGAGAGAAAGTGAAGATACTG	287
SIVMM251	-----A-----	292
SIVMM32H	-----A-----	292
SIVMM1A11	-----A-----	292
SIVMM142	-----A-----	292
SIVMM239	-----A-----	292
SIVMNE	-----G-----	292
SIVSMNH4	-----C-----G-----T-----T-----	292
SIVSMNH9	-----G-----C-----C-----C-----T-----T-----	292
SIVSMNPBJ	-----G-----C-----C-----C-----T-----T-----	292
SIVSTM	-----G-----T-----T-----T-----T-----C-----T-----	282
SIVSTM22579	-----G-----T-----T-----T-----C-----T-----	292
HIVFO784	-----A-----C-----TT-A-----G-----	182
CONSENSUS .E	AATTTAAaAGcCTTTTAAAGcTGTCTGCGTcaTtTtTGCGTgCACGCAGAGAGAAAGTGAAGATACAG	182
HIVFA	-----	182

HTV2/SIV GAG

CONSENSUS.A	AaGaAGCAaaa..	assaataGtacaGAGACATCTAGtGGCAGAAACaggAACTgCAGAGAAAATGCCAAatAc	344
HIV2ROD	---G-----	---C-----G-G-----	362
HIV2NIH2	---G-----	---C-----	362
HIV2ISY	---G-----	---G-----GG-----A-----T	362
HIV2ST	---C-----	---AG-----AA-----A-----G-	362
HIV2BEN	-G-----G-	-C-----C-----T-----	362
HIV2CAM2	---G-----	---G-----C-----T-----G-----	362
HIV2DI194	---G-----	---C-----C-----	362
HIV2GHI	---C-----	---C-----G-----A-----G-	362
HIV7924A	---G-----	---G-----	252
HIV60415K	---G-----	---C-----	252
HIVFA	---G-----	---C-----	252
CONSENSUS.B	AGGAAGCAAAA..	AAGaTaGcACAGAGcCATCTAGtGGtggACaCagaa.....??TAAAtgCCagCTaC	344
HIV2D205	---A-----	---C-----C-----GAA	353
HIV2UC1	---A-----	---C-----C-----GAG	353
HIV60667K	---C-----	---A-----T-----GAA	243
HIVJA	---C-----	---C-----A-----G-----C-----	243
HIV7312A	---C-----	---T-----A-----T-----CA-TA-G-	243
HIVF7	---T-----	---T-----G-----G-----	243
HIV7810A	---C-----	---C-----A-----	243
HIVON	---G-----	---T-----G-----C-----	243
CONSENSUS.C	AGGAACGAAAC..	GAGGTAGTAGAGAGACATCTAGCGGCAGAAACTAAAAATGCAGAAAAATGCCAGCCAC	252
HIV22238			252
CONSENSUS.D	AgGaAGCAAAA..	CAGaTaGtGcAGAGACA?CTaGTGgTgGAAACaGGAAcagCAGa7A7?ATGCCab?AAC	352
SIVMM251	---C-----	---A-----CT-----A-----	362
SIVMM32H	---C-----	---A-----CT-----A-----	362
SIVMM1A11	---A-----	---C-----A-----TT-----A-----	362
SIVMM142	---C-----	---A-----A-----CT-----A-----	362
SIVMM239	---C-----	---A-----A-----CT-----A-----	362
SIVMNE	---C-----	---A-----CT-----A-----	362
SIVSMH4	---A-----	---T-----C-----GA-----GC-----	362
SIVSMH9	-A-----	-A-----T-----C-----AA-----GC-----	362
SIVSM4PBJ	-A-----	-A-----T-----C-----AA-----GC-----	362
SIVSTM	---G-----	---AG-G-AA-----T-T-----A-G-T-----T-A-C-AA-----TGCC	362
SIVSTM22579	---G-----	---AG-G-AA-----T-T-----A-T-----T-C-AA-----TGCC	362
HIVF0784	---G-----	---T-----T-----C-AA-----TC-----	252
CONSENSUS.E	AGGAAGCGAAAAACATGTACACAGCACCACCTAGTGGCAGGAGGTAAAACTACAGAGAACTGCCTCTCTCA		255
HIVFA			255
CONSENSUS.A	AAGTAGACCAACAGCACCACCTAGcGgGaaaggaggaAACTtCCCcGTgCAaca??TgAGgCGCAACTAt		429
HIV2ROD	---T-----	---A-----G-----T-A-----A-----T-----C	432
HIV2NIH2			426
HIV2ISY	---G-----	---G-----A-----CT-----	432
HIV2ST	---C-----	---A-----A-----A-----G-C-----	432
HIV2BEN	---T-----	---A-----A-----A-----CG-T-----	432
HIV2CAM2	---T-----	---G-----T-A-----AGC-----A-----	432
HIV2DI194	---T-----	---G-----G-----C-----	432
HIV2GHI	---G-----	---A-----T-A-----GACA-GC-----T-----	435
HIV7924A			322
HIV60415K	---G-----	---T-----T-----	322
HIVFA	---GT-----	---T-----GACA-GC-----T-----C	325
CONSENSUS.B	AagTA7AocAaCAGCacCacCTAGcGGC.....aGAAATATCCAGTCAGCA7..ggTaGC?GGCAATTA7		405
HIV2D205	---A-----	---G-----G-----A-----C-G-T-----C-C	417
HIV2UC1	---G-----	---G-----C-----G-----A-----T-----C-C	417
HIV60667K	---G-----	---T-----G-----AA-----T-A-----T	307
HIVJA	---G-----	---G-----C-----G-----C-----C	307
HIV7312A	-GAC-A-----G-----GA-T-----T-----C-----A-----A-----T-----T		307
HIVF7	---A-T-----	---T-----T-----C-----C-----T-----	307
HIV7810A	---A-----	---G-----C-----C-----	307
HIVON	---G-----	---C-----A-----C-----T-----	307
CONSENSUS.C	AAGTAGACCAACAGCTCCATCTAGTGGCGGCAGAGGAAATTACCCAGTCAGCAA..GTAGCAGGCAATTAT		322
HIV22238			322
CONSENSUS.D	aAGTAAGACCAACaGCACCA7cTAGtGpCagAGGAgGAAATTACCCAGTAcA7CAA..TAgGtGGTAA7TA7		418
SIVMM251	---T-----	---T-----C-----A-----A-----C-----	432
SIVMM32H	---T-----	---T-----C-----A-----A-----C-----	432
SIVMM1A11	---C-----	---T-----C-----A-----A-----C-----	432
SIVMM142	---TT-----	---C-----A-----A-----C-----	432
SIVMM239	---T-----	---C-----A-----A-----C-----	432
SIVMNE	---T-----	---A-----G-----C-----	432
SIVSMH4	---C-----	---C-----G-----G-----C-----T-----	432
SIVSMH9	---C-----	---C-----G-----G-----C-----T-----	432
SIVSM4PBJ	---C-----	---A-----G-----G-----C-----T-----	432
SIVSTM	---C-----	---C-----G-----C-----C-----T-----	432
SIVSTM22579	---C-----	---G-----G-----C-----C-----T-----	432
HIVF0784	G-----	---C-----A-----G-----G-----G-C-----A-----T-----	322
CONSENSUS.E	GAGTAGACCAACAGCGCCACCTAGTGGAGGTGGAGGGAATTATCCAGTCACACAG..ATAGGAAATAACTAT		325
HIVFA			325

HIV2/SIV GAG

CONSENSUS .A	accCATgTgCCgCTgAGtCCCCGAACtCTAAATGCCtTGGGTaAAAtTagTaGAgGaaAagAagTTCGGGGCAG	502
HIV2ROD	-----A-A-----C-----C-----A-----	505
HIV2N1H2	-----A-A-A-----G-----C-----G-----A-----	499
HIV2ISY	GT-----C-----C-----G-----G-----A-----	505
HIV2ST	-----A-A-----C-----C-----C-----	505
HIV2BEN	GT-----A-----C-----A-----G-----	505
HIV2CAM2	---T-----A-----C-----C-----C-----	505
HIV2D194	---T-----C-----C-----C-----C-----	505
HIV2GH1	---T---C-----A-----C-----C-----	508
HIV7924A	-----A-A-----C-----C-----G-----A-----G-----	395
HIV60415K	-----A-A-----C-----C-----C-----A-----	395
HIVFA	GT-----A-A-A-----C-----A-----T-----	398
CONSENSUS .B	gtCCACcTGCcGTAAgCCCCGaaACcTFAAATGCCtTGGGTaAAAtTagTaGA?GaaAAGAAGTTCGGGGCag	477
HIV2D205	-----A-----T-----G-----A-----	490
HIV2UC1	-----A-----T-----G-----A-----	490
HIV60667K	TC-----T---G-----G-----G-----AGA	380
HIVJA	-----C-----C-----G-----A-----	380
HIV7312A	-----G---AT-----C-----G-----A-----	380
HIVFT	-----G-----C-----A-----A-----	380
HIV7810A	---T-----C-----G---T---A-----	380
HIVON	-----G-----C-----A-----A-----	380
CONSENSUS .C	GTCCACCATCCTCTAAGTCTAGAACACTGAATGCATGGGTAAAGTTAGTAGAAGAAAAGAGTTTGGGGCAG	395
HIV22238	-----A-----T-----G-----A-----	395
CONSENSUS .D	gtCCACCTaCCatTaAGcCCgAGAACaTFAAATGCCtTGGGTaAAATTTgaTaGAgGaaAagAAATTTGGaGCAG	491
SIVMM251	-----G-----	505
SIVMM32H	-----G-----	505
SIVMM1A11	-----G-----	505
SIVMM142	AC-----A-----G-----	505
SIVMM239	-----G-----	505
SIVMNE	AC-----G-----	505
SIVSMH4	-----C---T---A-----T-----AG---A---G---A---G---	505
SIVSMH9	AC-----T---T---A-----T-----G---A---G---A---G---	505
SIVSMH9BJ	AC-----T---T---A-----T-----A---G---A---G---	505
SIVSTM	-----A---C---G---A---G---AG-G---T---	505
SIVSTM22579	-----A---C---G---A---G---AG-G---T---	505
HIVF0784	-----GC-G---C---A---T---G---AG-G---C---A---C---G---	395
CONSENSUS .E	GTGCACTCTCCACTGTCCCAAGAACATTGAATGCATGGGTCAAATTAGTAGAAGAAAAGAAATTTGGAGCAG	398
HIVFA	-----A-----C-----	398
CONSENSUS .A	AAGTAGTGCcAGGATTtCagGCaCTCTCAGAAGGCTGCACGCCCTATGATATtAAtCaaTGCTTAAATTTGTGT	575
HIV2ROD	-----C---C-----	578
HIV2N1H2	-----G-----	572
HIV2ISY	-----G-----	578
HIV2ST	-----	578
HIV2BEN	-----	578
HIV2CAM2	-----C-----	578
HIV2D194	-----C-----	578
HIV2GH1	-----A---C---C-----	581
HIV7924A	-----C---G-----C-----	468
HIV60415K	-----C-----C-----	468
HIVFA	-----C-----C-----	471
CONSENSUS .B	AaGTaGTaCCAGGATTtCagGCaCTCTCAGAAGGCTGCACGCCCTATGATATtAAtCaaTGCTTAAATTTGTGT	549
HIV2D205	-----T-----A-----	563
HIV2UC1	-----A---T-----A-----	563
HIV60667K	-C---G-----A---C-----T-----	453
HIVJA	-----G-----G-----A-----	453
HIV7312A	-----G---T-----G---A---T-----	453
HIVFT	-----G---A---G---G-----	453
HIV7810A	-----G---T-----A-----	453
HIVON	-----G---A---G---G-----C-----	453
HIV26H0A	-----T-----	36
CONSENSUS .C	AGGTAGTACCAGGATTTCAGGCACTGACAGAAGGCTGCACGCCCTATGATATtAAtCaaTGCTTAAATTTGTGT	468
HIV22238	-----A-----	468
CONSENSUS .D	AaGTaGTGCcAGGATTtCagGCaCT?TCaGAAAGGCTGcaccCCCTaTGAcATtAATCagaTGCTTAAATTTGTGT	563
SIVMM251	-----G-----	578
SIVMM32H	-----A---G-----	578
SIVMM1A11	-----G-----T-----	578
SIVMM142	-----T-----G---CT-----	578
SIVMM239	-----G---T-----	578
SIVMNE	-----G-----	578
SIVSMH4	-G-----A---G---A-----T-----T-----A---C-----	578
SIVSMH9	-G-----A---G---A-----T-----Y---T---C-----A---C-----	578
SIVSMH9BJ	-G-----A---G---A-----T-----C-----A---C-----	578
SIVSTM	-----G-----A-----T-----C-----AC-----	578
SIVSTM22579	-----G---T-----G---A-----T-----C-----A-----	578
HIVF0784	-G-----G-----A---G-----T---G-----C-----C-----	468
CONSENSUS .E	AGGTAGTGCcAGGCTTCCAGGCATTATCAGAAGGATGCACGCCCTATGACATCAACCAGATGCCTTAAATTTGCCT	471
HIVFA	-----A-----	471

CONSENSUS .A	GGGCGACCAACCAAGCAGCCTATGCCAATAATACAGGGAATATTAATGAAGAGCAGCAGAGTGGGATgctcaa	646
HIV2ROD	-----C-----G-----G-----C-----G-----A-----TG-----	651
HIV2NIH2	-----G-----A-----G-----C-----C-----TGC-----	645
HIV2ISY	-----G-----A-----C-----C-----TA-----	651
HIV2ST	-----T-----C-----A-----G-----C-----T-----G-----G-----	651
HIV2BEN	-----T-----C-----A-----G-----C-----T-----G-----G-----	651
HIV2CAM2	-----T-----C-----A-----G-----C-----T-----G-----G-----	651
HIV2D194	-----T-----C-----A-----G-----C-----T-----G-----G-----	651
HIV2GH1	-----T-----C-----A-----G-----C-----T-----G-----G-----	654
HIV7924A	-----T-----C-----A-----G-----C-----T-----G-----G-----	541
HIV60415K	-----T-----C-----A-----G-----C-----T-----G-----G-----	541
HIVFA	-----T-----C-----A-----G-----C-----T-----G-----G-----	544
CONSENSUS .B	AGGAGAACATCAGGCAGCCATgcaaaTtATTAGGGAATTAATCAATGAGGAAGCAGCAGACTGGGACCAgCag	622
HIV2D205	-----C-----G-----C-----G-----G-----A-----	636
HIV2UC1	-----C-----G-----C-----G-----G-----A-----	636
HIV60667K	-----C-----G-----C-----G-----G-----A-----	526
HIVJA	-----G-----A-----G-----C-----G-----G-----A-----	526
HIV7312A	-----C-----A-----G-----C-----G-----G-----A-----	526
HIVPT	-----C-----A-----G-----C-----G-----G-----A-----	526
HIV7810A	-----C-----A-----G-----C-----G-----G-----A-----	526
HIVON	-----C-----A-----G-----C-----G-----G-----A-----	526
HIV2EHOA	-----C-----A-----G-----C-----G-----G-----A-----	109
CONSENSUS .C	AGGAGATCATCAGGCTGCTATGCGATTATTAGAGAAATTATCAATGAGGAAGCAGCCGATTGGGATGAAACAC	541
HIV22238	-----C-----G-----C-----G-----G-----A-----	541
CONSENSUS .D	gGGAAGA?CATCAAGCAGC?ATGCAGATtATcAGAga?ATtATAAatGAGGAGGCTGCaGATtGGGActtgCag	633
SIVMM251	-----C-----G-----T-----T-----C-----C-----	651
SIVMM32H	-----C-----G-----T-----T-----C-----C-----	651
SIVMM1A11	-----C-----G-----T-----T-----C-----C-----	651
SIVMM142	-----C-----G-----T-----T-----C-----C-----	651
SIVMM239	-----C-----G-----T-----T-----C-----C-----	651
SIVMNE	-----A-----T-----A-----C-----T-----A-----A-----	651
SIVSMH4	A-----A-----G-----C-----A-----T-----G-----A-----A-----C-----T-----A-----A-----	651
SIVSMH9	A-----A-----G-----C-----A-----T-----A-----A-----A-----T-----T-----A-----A-----	651
SIVSMH9	A-----A-----G-----C-----A-----T-----A-----A-----A-----T-----T-----A-----A-----	651
SIVSMH9	A-----A-----G-----C-----A-----T-----A-----A-----A-----T-----T-----A-----A-----	651
SIVSTM	-----G-----G-----C-----A-----A-----G-----C-----A-----T-----A-----A-----	651
SIVSTM22579	-----G-----G-----C-----A-----A-----G-----C-----A-----T-----A-----A-----	651
HIVF0784	A-----A-----G-----C-----A-----A-----G-----C-----T-----A-----A-----TCA-----	541
CONSENSUS .E	GGGCGAACATCAGGCAGCCATGCCAAATTATCAGAGAGATAATCAATGAAGAAGCAGCAGATTGGGACCTACAG	544
HIVFA	-----C-----G-----C-----G-----G-----A-----	544
CONSENSUS .A	CATCCAAATACCA...GGCCCTTtCCAGCGGGCAGCTTAGAGAG?CCAGAGGGGCTCTGACATAGCAGGAACAA	715
HIV2ROD	-----C-----G-----G-----A-----G-----G-----	721
HIV2NIH2	-----C-----G-----G-----A-----G-----G-----	715
HIV2ISY	-----C-----G-----G-----A-----G-----G-----	721
HIV2ST	-----C-----G-----G-----A-----G-----G-----	721
HIV2BEN	-----C-----G-----G-----A-----G-----G-----	721
HIV2CAM2	-----C-----G-----G-----A-----G-----G-----	721
HIV2D194	-----C-----G-----G-----A-----G-----G-----	721
HIV2GH1	-----C-----G-----G-----A-----G-----G-----	724
HIV7924A	-----C-----G-----G-----A-----G-----G-----	611
HIV60415K	-----C-----G-----G-----A-----G-----G-----	611
HIVFA	-----C-----G-----G-----A-----G-----G-----	614
CONSENSUS .B	CACCCGpTcaCCa...GGCCCACTGCCAGCAGGACAGCTTAGGGATCCAGGGGCTCTGACATAGCAGGAACCA	691
HIV2D205	-----C-----G-----G-----A-----G-----G-----	706
HIV2DC1	-----C-----G-----G-----A-----G-----G-----	706
HIV60667K	-----C-----G-----G-----A-----G-----G-----	596
HIVJA	-----C-----G-----G-----A-----G-----G-----	596
HIV7312A	-----C-----G-----G-----A-----G-----G-----	596
HIVPT	-----C-----G-----G-----A-----G-----G-----	596
HIV7810A	-----C-----G-----G-----A-----G-----G-----	596
HIVON	-----C-----G-----G-----A-----G-----G-----	596
HIV2EHOA	-----C-----G-----G-----A-----G-----G-----	179
CONSENSUS .C	CACCCAGTACCA...GGCCCACTGCCAGCAGGACAGCTTAGGGATCCAGGGGCTCTGACATAGCAGGAACCA	611
HIV22238	-----C-----G-----G-----A-----G-----G-----	611
CONSENSUS .D	CAGCCACaCCa??TG?TCCac?ac?A?caGGAca?CTAG?GAGCC?t?agGaTCaGATATtGCAGGAACaA	695
SIVMM251	-----C-----A-----A-----G-----G-----G-----C-----	710
SIVMM32H	-----C-----A-----A-----G-----G-----G-----C-----	710
SIVMM1A11	-----C-----A-----A-----G-----G-----G-----C-----	710
SIVMM142	-----C-----A-----A-----G-----G-----G-----C-----	710
SIVMM239	-----C-----A-----A-----G-----G-----G-----C-----	710
SIVMNE	-----A-----C-----A-----A-----G-----G-----G-----C-----	710
SIVSMH4	-----G-----T-----C-----G-----G-----A-----A-----AAG-----C-----T-----	721
SIVSMH9	-----R-----G-----G-----AT-----C-----G-----A-----A-----AAG-----C-----T-----	721
SIVSMH9	-----G-----G-----G-----GAT-----C-----C-----A-----A-----AAG-----C-----T-----	721
SIVSTM	-----T-----G-----G-----TG-----C-----G-----G-----A-----A-----AAG-----C-----T-----	724
SIVSTM22579	-----T-----G-----G-----TG-----C-----G-----G-----A-----A-----AAG-----C-----T-----	721
HIVF0784	-----T-----G-----G-----TG-----C-----G-----G-----A-----A-----AAG-----C-----T-----	611
CONSENSUS .E	CATCCAAAGA...GGGCAACCGCCAGCAGGGGCTTAAGAGACCCATCAGGATCAGACATAGCAGGAACAA	611
HIVFA	-----C-----G-----G-----A-----G-----G-----	611

HIV2/SIV GAG

CONSENSUS .A	CAAGCACAGTAgAaGaaCAGATCagTGGATGTTTAGggccAcaAAATCCcgTaCCaGTaGG?AaCATCTATAG	787
HIV2ROD	-----T-----A-----	794
HIV2NIH2	-----G-----A-----A-----	788
HIV2ISY	-----G-----A-----A-----	794
HIV2ST	-----G-----T-----A-----	794
HIV2BEN	-----T-----A-----G-----C-----	794
HIV2CAM2	-----A-----G-----G-----	794
HIV2D194	-----T-----A-----A-----C-----G-----G-----C-----	794
HIV2GH1	-----A-----G-----G-----G-----C-----	797
HIV7924A	-----G-----AG-----C-----T-----G-----A-----	684
HIV60415K	-----T-----A-----A-----	684
HIVFA	-----T-----A-----G-----A-----G-----G-----C-----	687
CONSENSUS .B	CCAGCACAGTAgGAGGaaCagATACagTGGATGTaCAGggccCaAAATCCTgTcCCaGTaGaaAACATtTATAG	764
HIV2D205	-----G-----G-----	779
HIV2UC1	-----A-----A-----T-----G-----G-----C-----	779
HIV60667K	-----A-----A-----G-----	669
HIVJA	-----G-----GG-----	669
HIV7312A	-----G-----GT-----G-----A-----G-----	669
HIVFT	-----A-----CAA-C-----T-----G-----G-----C-----	669
HIV7810A	-----G-----G-----A-----A-----	669
HIVON	-----A-----A-----CAG-C-----T-----G-----G-----C-----	669
HIV2EHOA	-----A-----AC-----G-----G-----C-----	252
CONSENSUS .C	CCAGCACGGTAgAGAAACAGATACAGTGGATGTATAGGCACAGAACCCAGTACCGGTAGGCAACATCTATCG	684
HIV22238	-----	684
CONSENSUS .D	CTAGtaCagtaGAtGAACAAATcCagTGGATGTaCaG7CAaCAGaAaCCcATACcAGTAgGCaAaATtTACAG	767
SIVMM251	-----T-----A-----	791
SIVMM32H	-----T-----A-----	791
SIVMM1A11	-----T-----A-----	791
SIVMM142	-----A-----A-----	791
SIVMM239	-----T-----A-----	791
SIVNIE	-----A-----A-----	791
SIVSMH4	-----A-----G-----A-----T-----	794
SIVSMH9	-----C-----S-----C-----T-----A-----Y-----G-----G-----A-----T-----C-----	794
SIVSMMPBJ	-----C-----C-----C-----T-----A-----T-----G-----G-----A-----T-----T-----C-----	794
SIVSTM	-----CT-----CC-----G-----T-----A-----C-----G-----G-----A-----T-----T-----G-----A-----C-----	797
SIVSTM22579	-----G-----T-----A-----C-----C-----G-----G-----A-----T-----T-----G-----A-----C-----	794
HIVFO784	-----C-----G-----G-----A-----A-----G-----T-----T-----A-----T-----C-----T-----	684
CONSENSUS .E	CCAGTACCCCGCAGAAACAAATAGAGTGGATGTACAGGAATCCAAATCCATCCCTCTGGAGACATCTATAG	684
HIVFA	-----	684
CONSENSUS .A	aAGaTGGaTcCagATAGGgTgCagAAGTGTGTcAGgAtGTACAaCCaACcAACaTccTAGAcaTaaaaCag	860
HIV2ROD	-----AT-----G-----	867
HIV2NIH2	G-----T-----A-----A-----C-----T-----T-----A-----	861
HIV2ISY	G-----A-----A-----A-----T-----T-----A-----	867
HIV2ST	-----A-----A-----C-----	867
HIV2BEN	-----A-----A-----A-----A-----T-----T-----	867
HIV2CAM2	-----G-----	867
HIV2D194	G-----A-----A-----T-----T-----TG-----G-----G-----	867
HIV2GH1	-----A-----T-----T-----G-----G-----	870
HIV7924A	-----A-----C-----T-----TG-----	757
HIV60415K	-----A-----C-----G-----	757
HIVFA	-----G-----G-----A-----A-----T-----G-----G-----G-----G-----	760
CONSENSUS .B	aAGaTGGATcCAa?TAGGgTgCagAAATGTGTcCGgATGTaCAATcCTACTAATATAcTaGAcATAAaCaAa	836
HIV2D205	-----T-----AT-----A-----C-----C-----T-----G-----G-----	852
HIV2UC1	-----A-----A-----C-----C-----T-----G-----G-----	852
HIV60667K	-----GA-----A-----C-----C-----G-----G-----	742
HIVJA	G-----T-----G-----T-----C-----G-----G-----	742
HIV7312A	-----GC-----AT-----G-----TG-----	742
HIVFT	-----C-----A-----C-----T-----TG-----	742
HIV7810A	G-----C-----T-----A-----G-----T-----C-----G-----	742
HIVON	-----A-----T-----T-----C-----C-----T-----TG-----	742
HIV2EHOA	-----G-----GT-----A-----G-----	325
CONSENSUS .C	GAGATGGATCCAACTAGGCCCTGCAAAAGTGTGTcAGAAATGTATAATCCAACTAACATTCTAGACGTGAAACAG	757
HIV22238	-----	757
CONSENSUS .D	gAGaTGGATcCAaCT?GGGcTgCAaAAATGTGTcAGAAATGTATAAACCcAaCaAaCATTCTaGATgTaaaaCAA	839
SIVMM251	-----G-----	864
SIVMM32H	-----G-----	864
SIVMM1A11	-----G-----	864
SIVMM142	-----G-----	864
SIVMM239	-----G-----	864
SIVNIE	-----G-----G-----	864
SIVSMH4	A-----G-----T-----A-----C-----G-----A-----T-----G-----	867
SIVSMH9	R-----G-----T-----A-----C-----G-----C-----A-----T-----T-----G-----G-----	867
SIVSMMPBJ	A-----G-----A-----C-----G-----A-----T-----T-----G-----G-----	867
SIVSTM	-----T-----A-----C-----A-----C-----T-----GTT-----T-----C-----G-----	870
SIVSTM22579	-----T-----G-----A-----A-----C-----T-----GTT-----T-----G-----CA-----G-----	867
HIVFO784	-----G-----A-----A-----G-----C-----T-----T-----G-----A-----	757
CONSENSUS .E	AAGATGGATCCAGCTAGGGCTCCAGAAATGTGTcAGAAATGTATAATCCAAACAAACATTCTGACCTCAAACAG	757
HIVFA	-----	757

CONSENSUS . A	GGACCAAAAGGAacCgTTcCAaAGcTATGTaGATAGaTTCTACAAaAGCtTgAGGGCAGAACAAACAGATccAG	933
HIV2ROD	-----G-----	940
HIV2NIH2	-----A-----	934
HIV2ISY	-----A--GT-----G-----A-----G-----G-----	940
HIV2ST	-----G--A-----G--T-----G-----	940
HIV2BEN	-----A-----	940
HIV2CAM2	-----GT--A-----	940
HIV2D194	-----A--T-----G-----C-----C--A-----C--G-----	940
HIV2GH1	-----G-----G-----G--C--G-----T-----G-----	943
HIV7824A	-----T--A--T-----	781
HIV60415K	-----A--G-----	781
HIVFA	-----T-----G-----	784
CONSENSUS . B	GGgCCAAAaGAaCCcTTCCAAAgCTaTGATAGATGATTCTACAAaAGCTTACGGGCAGAgCAaACAGACCCAG	909
HIV2D205	--A--G--G-----A-----	925
HIV2UC1	--A-----T-----C-----	925
HIV60667K	-----G--G-----A--	766
HIVJA	--A-----	766
HIV7312A	-----	766
HIVPT	-----G-----	766
HIV7810A	--A-----	766
HIVON	-----G-----	766
HIV2EHOA	-----	398
CONSENSUS . C	GCTCCAAAAGAACCATTTCAGAGT	781
HIV2238	-----	781
CONSENSUS . D	GGgCCAAAAGAgcCaTTtCAGAGcTATGTAGAcAGgTTCTACAAaAGtTtAGAGGCAGAgCAaACAGATccAg	912
SIVMM251	-----C-----G-----	937
SIVMM32H	--A-----C-----G-----	937
SIVMM1A11	-----G-----G-----	937
SIVMM142	-----	937
SIVMM239	-----G-----G-----	937
SIVMNE	-----C-----	937
SIVSMH4	--A-----A--C--T--A-----C-----G-----C--	940
SIVSMH9	--A-----T-----A-----T--A-----G--C-----C--G--	940
SIVSMPEBJ	--A-----A-----T--A-----G--C-----C--G--	940
SIVSTM	-----AA--G--C--A--T-----G--G-----G-----T	943
SIVSTM22579	--A-----A--G--C--A--T-----G--G-----G-----	940
HIVFO784	-----A--G--C-----	781
CONSENSUS . E	GGGCCCAAGAATCTTTTCAGAGC	781
HIVPA	-----	781
CONSENSUS . A	CaGTAAaAAATGGATGACCCAAAC?CTOCTATTaCAgaatgCCaACCCAGAcTGCaAgTTAGTAcTAAaAGG	1004
HIV2ROD	-----G--G-----A-----G-----A-----G-----	1013
HIV2NIH2	-----G--G-----A-----A-----A-----T-----A-----	1007
HIV2ISY	-----G--G-----G--G--G--G--A...T--G-----G-----	1010
HIV2ST	-----A-----G--d-----	1023
HIV2BEN	-----G--A-----C-----	1013
HIV2CAM2	-----A-----G--G-----A-----G-----	1013
HIV2D194	-T-----G--A-----C-----T-----	1013
HIV2GH1	-----G--C-----G--A-----C-----	1016
CONSENSUS . B	CAGTgAAAAATTGGATGACACAAACACTGCTgATTTCAGAATGCTAACCCAGATTGCAAgTTAGTgCTPAAGGG	982
HIV2D205	-----A-----A-----A-----	998
HIV2UC1	-----A-----A-----A-----	998
HIV2EHOA	-----A-----A-----A-----	471
CONSENSUS . D	CAGTaaAGaATTGGATGACcCaAACACTGCTgATtCAAAATGCTAAcCCaGATtGcAAGcTaGTgCTgAAgGG	985
SIVMM251	-----T-----A-----	1010
SIVMM32H	-----	1010
SIVMM1A11	-----	1010
SIVMM142	-----	1010
SIVMM239	-----	1010
SIVMNE	-----	1010
SIVSMH4	-----C-----T--A-----C-----	1013
SIVSMH9	--S--G--A-----C-----G--C--T--AT--G-----C-----	1013
SIVSMPEBJ	--G--A-----C-----G--C--T--AT--G-----C-----	1013
SIVSTM	--G--A-----GG-----A-----T-----C-----AT--G--AT--A--	1016
SIVSTM22579	--G--A-----G-----C-----A--C-----T-----C-----AT--G--AT--A--	1013

HIV2/SIV GAG

CONSENSUS.A	AcTaGGGATGAACTCCCTACCTAGAGAGATGCT?ACCGCTGTCAAGGGgTAGGtGGaCCAGGCCAgAAaGCC	1076
HIV2ROD	-----C-----G-----G-----T	1086
HIV2NIH2	-T--A-----A--G--A-----G-----T	1080
HIV2ISY	-----A-----A-----A-----G-----	1083
HIV2ST	-----A-----A-----A-----	1086
HIV2BEN	--G-----C--C-----A-----C-----	1086
HIV2CAM2	-T-----A--T-----A--A-----T	1086
HIV2D194	-----C--C--G-----G--T-----C--A-----	1086
HIV2GH1	--G-----C--C-----G--T-----A-----	1089
CONSENSUS.B	CTTgGGAAATGAATCCCACcTTAGAGGAAATGCTaACgCCCTGCCAaGGGATAGAGGCCCAGGCCAGAAAGGCA	1055
HIV2D205	-----G-----	1071
HIV2UC1	-----A-----G-----	1071
HIV2EHOA	-----C-----A-----A-----G-----	544
CONSENSUS.D	?CTGGGtatgAATCCCACcctAGAAGAsATGCTgACgGC?TGTCAaGGaGTaGGgGGgCCaGGACAgAAggCt	1056
SIVMM251	G-----G-----T-----A-----	1083
SIVMM32H	A-----G-----T-----A-----	1083
SIVMM1A11	G-----G-----T-----	1083
SIVMM142	G-----C-----C-----	1083
SIVMM239	G-----G-----T-----G-----	1083
SIVMNE	G-----T-----A--A-----A--A-----	1083
SIVSMH4	T-----C-----TT-----A--C-----G-----A-----A-----	1086
SIVSMH9	T-----Y--A-----TT-----T--A--C-----G-----R-----A--AS-----	1086
SIVSMMPBJ	T-----A-----TT-----T--A--C-----G-----A-----A-----A-----	1086
SIVSTM	T-----C-----G-----T--A--A-----G--T-----A-----A-----	1089
SIVSTM22579	T-----C-----G-----T--A--A-----G--T-----A-----A-----	1086
CONSENSUS.A	AGaCTAATGGCCAGAAAGCCcTaAAGGAGGccctTgg?acCagcCCCTATCCCATTtGCAGCAGCCCAaCag...	1144
HIV2ROD	---T-----G---G---A---T---A---G---T---C-----G-----	1155
HIV2NIH2	---T-----T-----T-----C-----G-----	1149
HIV2ISY	---T-----T-----A---CG-----	1152
HIV2ST	---T-----T-----C-----C-----	1155
HIV2BEN	---G-----T-----A---T---G---AG-----A-----	1155
HIV2CAM2	---T-----T-----G---C-----C-----G-----	1155
HIV2D194	---T-----T-----TT---ACG-----A-----	1155
HIV2GH1	---C-----G-----TT---AC-----C-----A-----	1158
CONSENSUS.B	AGGCTAATGGCTGAGGCCcTTAAAGAGGCCcTaACACCTgCacCCATaCCGTTTGCTGCCGTCAACaAaag	1128
HIV2D205	---C---A-----C-----T-----	1144
HIV2UC1	-----T-----A-----G-----T---CA---AT-----C-----C---A	1144
HIV2EHOA	-----T-----A-----G-----T---CA---AT-----C-----C---A	614
CONSENSUS.D	AGatTaATGGCAGAAAGCCcTgAAAGAGGCCcTogcaCcAG?gCcAcTcCCT?TtGCAGCAGCCCAaCAGaagG	1128
SIVMM251	-----A-----GA-----G-----	1156
SIVMM32H	-----T-----A-----G-----	1156
SIVMM1A11	-----T-----A-----G-----	1156
SIVMM142	-----C-----A-----G-----	1156
SIVMM239	-----T-----A-----G-----	1156
SIVMNE	-----T-----G-----A-----	1156
SIVSMH4	---G-----AT-----A---GAGG---AT---A-----A-----T-----A-----	1159
SIVSMH9	---C-----G---AT-----T---A---GA---G---G-----A-----R-----T-----A-----	1159
SIVSMMPBJ	---C-----G---AT---A---T---A---GA---G---G-----A-----T-----A-----	1159
SIVSTM	---C---G-----T-----T---CA---AC-----G---C---C-----C-----	1162
SIVSTM22579	---C---G-----T-----T---CA---AC-----G---C---C-----C-----	1159
CONSENSUS.A	AgAA?GgcAaTtAagTg?TGGAATGtGGaAAGGA?GG?CACTCGBCaAgACAaTGccGaGCaCC	1205
HIV2ROD	-----A---T---A---C-----C-----A---G-----	1220
HIV2NIH2	-----A---A---T---C-----C-----A---A-----GA---C---G-----	1214
HIV2ISY	-----A---G-----T-----A---G-----	1217
HIV2ST	-----G---A-----C-----C-----A---G-----	1220
HIV2BEN	-----A---A-----G---AT---C-----C-----G---A-----G-----	1220
HIV2CAM2	-----G---A---G-----A---C-----G-----G---G-----	1220
HIV2D194	-----G-----G---T-----G---A-----G---A-----G-----	1220
HIV2GH1	-----A---T-----G---T-----C---C-----G---A-----G-----G-----	1223
CONSENSUS.B	cAGgaaAGagagGACAcGTGACATGCTGGAAGTGTGGCAAgg?ggGACAcACAGCCAgGCCaATGCAGGGCCCC	1200
HIV2D205	---G-----ACA-----	1217
HIV2UC1	-----G-----TA-----A-----	1217
HIV2EHOA	G---C---CG---AGA-----C-----C-----T-----G-----A-----	687
CONSENSUS.D	GACca...aGAAAGccAaTtAagTGTgGAATGTgGAAAGaaGGaCACTCTGCAAggCAaTgcAgagcCCC	1198
SIVMM251	-----G-----G-----	1226
SIVMM32H	-----G-----	1226
SIVMM1A11	-----G-----	1226
SIVMM142	-----G-----	1226
SIVMM239	-----G-----	1226
SIVMNE	-----G-----A-----	1226
SIVSMH4	---A---G---A---C-----A---G---G-----AA-----G-----	1229
SIVSMH9	---A---R---AT---W---R---C---R---A---G-----A---T-----	1229
SIVSMMPBJ	---A---AT-----A---G-----A-----	1229
SIVSTM	---C---G---A---G---A---C-----C---G---A---AA---G---T---A---G-----	1229
SIVSTM22579	---C---G---A---G---A---C-----G---A---AA---G---T---A---G-----	1226

		pol orf ->	
CONSENSUS .A	tAGAAGaCagGGCTGCTGGAAgTGTTGGcAAG?CAGGACacaTCATGgCAAAcTGCCCaGA?AGACAGGCagGT		1276
HIV2ROD	-----G-----T-----C-----A-----T-----		1293
HIV2NIH2	-----T-----T-----TG-----T-----T-----		1287
HIV2ISY	-----A-----T-----T-----T-----T-----		1290
HIV2ST	-----A-----A-----G-----A-----A-----G-----		1293
HIV2BEN	-----C-----C-----A-----G-----A-----		1293
HIV2CAM2	-----G-----C-----A-----T-----T-----G-----		1293
HIV2D194	C-----T-----T-----G-----A-----A-----		1293
HIV2GH1	-----A-----A-----G-----A-----A-----		1296
CONSENSUS .B	TAGAAGaCAGGGATGCTGGAAATGTGGAAaAacaAGGACACATCATGTCAAAATGCCAGAAAGACAGGCGGT		1273
HIV2D205	-----AC-----		1290
HIV2UC1	-----C-----G-----		1290
HIV2EHOA	-----G-----C-----G-----		760
CONSENSUS .D	aAGAAGACagGG?TGcTGGAAATGTGGAAaAaatgGgCCATgttATGgCCAAATGCCAgAcAGACagGcgGGT		1270
SIVMM251	-----A-----A-----A-----A-----A-----		1299
SIVMM32H	-----A-----A-----A-----A-----A-----		1299
SIVMM1A11	-----A-----A-----A-----A-----A-----		1299
SIVMM142	-----A-----A-----A-----A-----A-----		1299
SIVMM239	-----A-----A-----A-----A-----A-----		1299
SIVMNE	-----C-----C-----C-----C-----C-----		1299
SIVSMH4	T-----C-----G-----C-----T-----T-----A-----		1302
SIVSMH9	T-----R-----C-----G-----GCA-----T-----C-----A-----		1302
SIVSMH9BJ	T-----C-----G-----GCA-----T-----C-----A-----		1302
SIVSTM	G-----A-----T-----C-----CA-----CAG-----A-----TA-----		1302
SIVSTM22579	G-----A-----C-----T-----CC-----CAG-----A-----TA-----		1299
CONSENSUS .A	TTTTTAGGgaTgGGcCcatGGGGAAAGaGcCcCGCAAcTTCCcCGtGgcCCaAgttcCgCaGGGGCTg...		1345
HIV2ROD	-----AC-----T-----G-----C-----C-----A-----		1362
HIV2NIH2	-----T-----AC-----G-----C-----C-----A-----		1356
HIV2ISY	-----C-----T-----A-----T-----C-----AG-----T-----A-----		1359
HIV2ST	-----T-----C-----T-----T-----T-----A-----C-----		1362
HIV2BEN	-----T-----C-----T-----T-----A-----CC-----T-----		1362
HIV2CAM2	-----T-----C-----T-----T-----A-----A-----A-----		1362
HIV2D194	-----T-----C-----C-----C-----C-----T-----		1362
HIV2GH1	-----T-----C-----C-----C-----C-----T-----		1365
CONSENSUS .B	TTTTTAGGGTTaGgaCCCTGGGGAAAGaAGCCCTGGCAAcTTCCcCaTgacCCaAGtGcCtCAGGGagTgacaC		1346
HIV2D205	-----		1363
HIV2UC1	-----C-----C-----T-----G-----G-----CA-----C-----GA-----AGTG-----		1363
HIV2EHOA	-----C-----C-----T-----G-----G-----CA-----C-----GA-----AGTG-----		830
CONSENSUS .D	TTTTTAGGcctTTGGcCCaTGGGGAAAGaAGCCCGCAAcTTCCcCATGGCCcAaaTgCaTCAGGGGCTG????		1339
SIVMM251	-----T-----T-----T-----T-----T-----G-----		1368
SIVMM32H	-----T-----T-----T-----T-----T-----G-----		1368
SIVMM1A11	-----T-----T-----T-----T-----T-----G-----		1368
SIVMM142	-----T-----T-----T-----T-----T-----G-----		1368
SIVMM239	-----T-----T-----T-----T-----T-----G-----		1368
SIVMNE	-----T-----T-----T-----T-----T-----G-----		1368
SIVSMH4	-----G-----G-----G-----G-----G-----C-----		1371
SIVSMH9	-----G-----G-----G-----G-----G-----C-----		1371
SIVSMH9BJ	-----G-----G-----G-----G-----G-----C-----		1371
SIVSTM	-----T-----T-----A-----C-----A-----C-----		1371
SIVSTM22579	-----T-----T-----C-----C-----A-----C-----ACGC		1372
CONSENSUS .A	AcACCAACAGCACCCcCa7t7GA		1366
HIV2ROD	G-----G-----		1385
HIV2NIH2	T-----G-----		1379
HIV2ISY	A-----G-----		1382
HIV2ST	GA-----A-----		1385
HIV2BEN	T-----T-----GGCA-----		1385
HIV2CAM2	T-----T-----T-----G-----		1385
HIV2D194	T-----A-----A-----		1385
HIV2GH1	T-----A-----A-----GCA-----		1388
CONSENSUS .B	CATCTGCaCCCCGATGaaCCCCAGCagagGGCATGACACCTCgGGGgGGCaCaCCATCTGCGCCCCCTGCAGA		1419
HIV2D205	-----G-----G-----		1436
HIV2UC1	-----G-----G-----		1436
HIV2EHOA	-----G-----TTC-----A-----T-----TT-----A-----		903
CONSENSUS .D	????????.....????????.....AcgCCAACCTGCtCCCCCAGagGA		1362
SIVMM251			1391
SIVMM32H			1391
SIVMM1A11			1391
SIVMM142	A-----		1391
SIVMM239	T-----		1391
SIVMNE			1391
SIVSMH4	TA-----		1394
SIVSMH9	A-----Y-----		1394
SIVSMH9BJ	A-----		1394
SIVSTM	C-----T-----A-----		1394
SIVSTM22579	CAACTGCC.....CCTCCAGAG.....T-----C-----T-----TA-----		1412

HIV2/SIV GAG

CONSENSUS .A	LCCAGcAGcGGAcCTacTgGAgAaaTatATGCAGCAaGGGA?Aa?gCAGAG?GAG.....CAGAGA	1420
HIV2ROD	-----T-----A--GA--A-----	1446
HIV2NIH2	-----A--G--A--G-----	1440
HIV2ISY	-----G--C-----G--AA--A-----	1443
HIV2ST	C---T---A-----A---G---C-----A---G---A-----	1446
HIV2BEN	-----C---A---GT---G-----G---A---G-----G	1446
HIV2CAM2	-----C---T-----A---G---G-----	1446
HIV2D194	-----GT-----G---A---A-----G	1446
HIV2GH1	-----T---T---A---G-----G---G---A-----G	1449
CONSENSUS .B	TCCAGCAGcGGAGATGCTgAAaAgcTACATGcAgATgGGGAgacAgCAGAGaGAG.....AgCcGA	1480
HIV2D205	-----T-----A-----	1497
HIV2UC1	-----A-----	1497
HIV2EHOA	-----A---G---A---AC---A---AGA---AG---A---A---	964
CONSENSUS .D	CCCAGCTGcGGATCTgCTaAagAacTACATGcAg?TGGGCAagcAGCAGAGaGAA?????????AgCAGA	1422
SIVMM251	-----T-----	1452
SIVMM32H	-----C-----AAGCAGAGAGAA-----	1464
SIVMM1A11	-----T-----AAGCAGAGAGAA-----	1464
SIVMM142	A-----CT-----	1452
SIVMM239	-----T-----AAGCAGAGAGAA-----	1464
SIVMNE	-----A-----T-----A-----A-----	1452
SIVSMH4	T-----A---G---T-----A---A---GAA---G---G-----A---	1455
SIVSMH9	T-----A---G---T-----A---G---GAAG---G---G-----A---	1455
SIVSMMPBJ	T-----A---G---T-----A---A---GPAG---G---G-----A---	1455
SIVSTM	-----C-----G---GA---GT-----C-----A-----G-----G	1455
SIVSTM22579	-----C-----G---GA---GT-----C-----A-----G-----G	1473
CONSENSUS .A	gAgAGaCCaTACAAAGA?GTGACaGAGGACTTaCTGCaccTCGAgcAgggagAgCaCCa?ACAGAGaG????a	1492
HIV2ROD	-----G---A-----G-----T---G---CCAC	1519
HIV2NIH2	-----G---G---A-----GTT-----C-----TG---G-----	1510
HIV2ISY	C---A-----A-----G---T-----A-----C-----	1513
HIV2ST	-----G-----A-----C---G---A---ACA-----TG-----	1516
HIV2BEN	-----G-----G-----G-----A---A-----TC-----G	1516
HIV2CAM2	C-----G-----A-----A-----G---C---TG---G-----	1516
HIV2D194	-----G---G---G-----A-----G---CC---G---G-----	1516
HIV2GH1	-----G---G---A-----A---AG---TC-----G	1519
CONSENSUS .B	GAGAGACCCtTACAAGGAGGTG????????????????????	1501
HIV2D205	-----	1518
HIV2UC1	-----	1518
HIV2EHOA	-----ACAGAGGATTtGCTGCACCTC-----	1006
CONSENSUS .D	gaGAagCCtTACAAGGAGGTg????????????????????	1443
SIVMM251	-----	1473
SIVMM32H	-----A-----	1485
SIVMM1A11	-----A-----	1485
SIVMM142	-G-----	1473
SIVMM239	-----	1485
SIVMNE	AG-----	1473
SIVSMH4	-GA-----	1476
SIVSMH9	-GA-----A-----	1476
SIVSMMPBJ	-GA-----A-----	1476
SIVSTM	A---CA---C-----	1476
SIVSTM22579	A---CA---ACAGAGGATTtGGTGCACCTC-----	1515
CONSENSUS .A	cgACAGAGGACTTGCTgCACCTCAATTCTCTCTTTGGAAAAaGACCCAG	1539
HIV2ROD	-A-----	1566
HIV2NIH2	-----	1557
HIV2ISY	-----C-----	1560
HIV2ST	-----	1563
HIV2BEN	A-----	1563
HIV2CAM2	-A-----	1563
HIV2D194	-----A-----	1563
HIV2GH1	-----	1566
CONSENSUS .B	..acagaggatttgctgcacctCAATTCTCTCTTTGGAGAAAGACCAg?AG	1548
HIV2D205	-----	1563
HIV2UC1	-----	1563
HIV2EHOA	-----T---	1033
CONSENSUS .D	..acagaggatttgctgcacctCAATTCTCTCTTTGGAG?AGACCAg	1487
SIVMM251	-----G-----	1518
SIVMM32H	-----G-----	1530
SIVMM1A11	-----G-----	1530
SIVMM142	-----G-----	1518
SIVMM239	-----G-----	1530
SIVMNE	-----A-----	1518
SIVSMH4	-----A-----	1521
SIVSMH9	-G-----A-----	1521
SIVSMMPBJ	-G-----A-----A-----	1521
SIVSTM	-----A-----	1521
SIVSTM22579	-----	1530

CONSENSUS .A	GTGTTGGAATTGTGGAAAGGAGGGACACTCGGCGAAACAGTCCCGAGCAGCCAGAGACAGGGCTGCTGGAA?	72
HIV2ST	-----A-----A	19
HIV2BEN	-----G	19
HIV2D194	-----G	73
HIV2GHI	-----A	19
CONSENSUS .B	AAGACAGGGATGCTGGAAA	19
HIV2D205	-----	19
HIV2UC1	-----	19
HIV2EHOA	-----	10
CONSENSUS .D	GTGTTGGAATTGTGGGAAAGAGGGACACTCTGCAAGGCAATGCCAGAGcCCCaAGAAAGACAGGGATGCTGGAAA	73
SIVMM251	-----G-----	73
SIVMM142	-----G--A-----	73
SIVMM239	-----	73
SIVMM32H	-----A-----	73
SIVMNE	-----A-----C-----	73
SIVSTM	-----G--G-----A--T--T-----	28
CONSENSUS .A	TGTGGCAAGTCAGGACACaTCATGGCAAA?TGCCC?GAAAGACAGGCAGSTTTTTAGGGaTgGGcCcatGGG	142
HIV2ROD	-----AC-----T-----	32
HIV2NIM2	-----T-----T-----AC-----	32
HIV2ISY	-----T-----C--T--A-----	32
HIV2ST	-----G-----G-----T-----	92
HIV2BEN	-----C-----C--G-----T-----C-----	92
HIV2CAN2	-----G-----T-----C-----	32
HIV2D194	-----T-----C--G-----C-----	146
HIV2GHI	-----A--G-----A--A-----T-----	92
CONSENSUS .B	TGTGGAAaAaCaAGGACACATCATGTCAAAATGCCCAAGAGACAGGCgGGTTTTTTAGGGTTaGGaCCCTGGG	92
HIV2D205	-----AC-----	92
HIV2UC1	-----T-----	92
HIV2EHOA	-----C--G-----C--C-----	83
CONSENSUS .D	TGTGGAAaAAaTGGaCCATgtcATGGCCAAATGCCCAgAcAGACAGGcGGTTTTTTAGGGcTTGGcCCaTGGG	146
SIVMM251	-----T-----	146
SIVMM142	-----A-----T-----	146
SIVMM239	-----T-----	146
SIVMM32H	-----T-----	146
SIVMNE	-----C-----G-----A-----T-----	146
SIVSMNH4	-----A-----G-----	35
SIVSMNH9	-----C--A-----G-----	44
SIVSMNPBJ	-----C--A-----A-----	44
SIVSTM	-----C--CA--G-----CAG-----A-----T-----	101
CONSENSUS .A	GAAAGaAGcCcCGCAAcTTCCCCGtGgcCCaAgTtcGgCaGGGGCTgAcACCAACAGCACCcCCa?t?GATCC	213
HIV2ROD	-----G--G-----	105
HIV2NIM2	-----G-----C-----C-----A-----T--G-----	105
HIV2ISY	-----T--C--AG--T-----A-----A--G-----	105
HIV2ST	-----T-----T-----A--C-----GA--A--C-----	165
HIV2BEN	-----T-----A-----CC--T-----T-----T--GGCA-----	165
HIV2CAN2	-----T-----A-----A-----T--G-----	105
HIV2D194	-----C-----C-----C--T-----T-----A--A-----	219
HIV2GHI	-----C-----T--C-----T-----GCA-----	165
CONSENSUS .B	GAAAGaAGCCCTCGCAAcTTCCCCatgacCCaAGtgcCTCAGGGagTgacaCCATCTGCaCCCCGATGaACCC	165
HIV2D205	-----	165
HIV2UC1	-----G-----	165
HIV2EHOA	-----T-----GT--G--CA--C-----GA--AGTG-----G-----	153
CONSENSUS .D	GAAAGaAGCCCGCAAcTTCCCCATGGCCCaAaTgCaTCAGGGGCTGaagCCAACtGCTCCcCCAGagGaACCC	219
SIVMM251	-----T-----G-----	219
SIVMM142	-----T--G-----A--A-----	219
SIVMM239	-----T--G-----T-----	219
SIVMM32H	-----T--G-----	219
SIVMNE	-----G-----C-----TA-----T-----	108
SIVSMNH4	-----G--C-----A-----Y-----T-----	117
SIVSMNPBJ	-----G--C-----A-----T-----	117
SIVSTM	-----A--C-----A--C-----C--T-----A-----	174

HIV2/SFV POL

CONSENSUS. A	AGCAGTGGACCTacTgGAGaaaTatATGCAGCAaGgA?Aa?	gCAGAG?GAGCAGAGagAg	271
HIV2ROD	-----T-----A-----G-----A-----A-----	-----A-----A-----	166
HIV2NIH2	-----T-----A-----G-----A-----A-----G-----	-----A-----G-----	166
HIV2ISY	-----G-----C-----G-----A-----A-----A-----C-A	-----A-----A-----	166
HIV2ST	---T---A-----A-----G-----C-----A-----G-----A-----	-----A-----A-----	226
HIV2BEN	---C---A---GT---G-----G-----A-----A-----G-----G-----	-----G-----G-----	226
HIV2CAN2	---T-----T-----A-----G-----A-----G-----C-----	-----G-----C-----	166
HIV2DI194	---GT---G-----G-----A-----A-----A-----G-----	-----A-----G-----	280
HIV2GH1	---T---T---A-G---G-----G-----G-GA-----G-----G-----	-----G-----G-----	226
CONSENSUS. B	AGCAGagGGCATGACACCTCgGGGgGGGAcacCATCTGCGCC	CCCTGCAGATCCAGCAGtG	226
HIV2D205	-----TTC-----A-----T-----TT-----A-----A-----	-----A-----	226
HIV2UC1	-----TTC-----A-----T-----TT-----A-----A-----	-----A-----	214
CONSENSUS. D	AGCTGTGGATCTgCTaAagAacTACATGcAg?TGGGCAagca?????????GCAGAGaGaaAgCAGagaG		279
SIVMM251	-----T-----C-----G-----G-----G-----G-----	-----G-----	280
SIVMM142	-----T-----C-----G-----G-----G-----G-----	-----G-----	280
SIVMM239	-----T-----C-----G-----G-----G-----G-----	-----G-----	292
SIVMM32H	-----T-----C-----G-----G-----G-----G-----	-----G-----	292
SIVMM1E	-----A-----T-----A-----A-----A-----A-AG-	-----A-AG-	280
SIVSMH4	---A---G---T---A---A---GAA-----G---G-A-----	-----G---G-A-----	169
SIVSMH9	---A---G---T---A---G---GAAG-----G---G-A-----	-----G---G-A-----	178
SIVSMFPBJ	---A---G---T---A---A---GAAG-----G---G-A-----	-----G---G-A-----	178
SIVSTM	---C---G-GA-GT---C-----A-----G-----GA-----	-----G-----GA-----	235
CONSENSUS. A	AGaCCaTACAAaGA?GTGACaGAgGACTTtACTGCaccT...CGAgcAgggagAgacCaCCa?aCAGaGa??gga		336
HIV2ROD	-----G---A-----G-----T-----G-----GCCAC		236
HIV2NIH2	---G---G---A-----GTT-----C-----TG---G-----		233
HIV2ISY	-----A-----G-----T-----A-----C-----C-----		233
HIV2ST	---G---A-----C-----G-----A-ACA-----TG-----		293
HIV2BEN	---G---G---G---G---G-----A-----TC-----G-----		293
HIV2CAN2	---G---G---G---G---G-----A---A-----G-C---TG---G-----		233
HIV2DI194	---G---G---G---G---G-----A-----G---CC---G---G-----		347
HIV2GH1	---G---G---G---A-----A-AG-----TC-----G-----		293
CONSENSUS. B	GAGATGCTgAaaAgcTACATGcAgaTgGGGAgacAgCAGAGaGAGAgCcGAGAGAGACCCCTACAAGGA...GG		296
HIV2D205	-----T-----A-----A-----A-----A-----		296
HIV2UC1	-----A-----A-----A-----A-----A-----		296
HIV2EHQA	---A---G-A---AC-A---AGA---AG---A-A-----		284
CONSENSUS. D	AagCCtTACAAAGGAGGT		296
SIVMM251	-----A-----A-----A-----A-----A-----		297
SIVMM142	-----A-----A-----A-----A-----A-----		297
SIVMM239	-----A-----A-----A-----A-----A-----		309
SIVMM32H	-----A-----A-----A-----A-----A-----		309
SIVMM1E	-----A-----A-----A-----A-----A-----		297
SIVSMH4	---GA-----A-----A-----A-----A-----		186
SIVSMH9	---GA-----A-----A-----A-----A-----		195
SIVSMFPBJ	---GA-----A-----A-----A-----A-----		195
SIVSTM	---CA---C-----A-----A-----A-----A-----		252
CONSENSUS. A	cgACAGAGGACTTGTCTgCACCTCAATTCTCTCTTTGGAAAaGACCAGTAgTCACAGCatacaTtGAGGgtCAG		409
HIV2ROD	---A-----A-----A-----A-----A-----		309
HIV2NIH2	-----C-----C-----C-----C-----C-----		306
HIV2ISY	-----C-----C-----C-----C-----C-----		106
HIV2ST	-----C-TG-----C-----C-----C-----		366
HIV2BEN	A-----O-----C-----A-----C-----		366
HIV2CAN2	---A-----C-----C-----C-----C-----		366
HIV2DI194	-----A-----T-----C-----A-----C-----		406
HIV2GH1	-----C-----TA-----C-----C-----		366
CONSENSUS. B	TGACAGAGGATTTGTCTGCACCTCAATTCTCTCTTTGGAGAAGACCAGTAgTCACaAGCAtgTATcGAGGGTCag		369
HIV2D205	-----G-----AC---T-----A-----		369
HIV2UC1	-----G-----AC---T-----A-----		169
HIV2EHQA	-----G-----AC---T-----A-----		357
CONSENSUS. D	gACaGAGGATTTGTCTGCACCTCAATTCTCTCTTTGGAGAAGACCAGTAgTCACcGCTcACATTCGAaGgaCAG		369
SIVMM251	-----G-----G-----G-----G-----G-----		169
SIVMM142	-----G-----G-----G-----G-----G-----		369
SIVMM239	-----G-----G-----G-----G-----G-----		181
SIVMM32H	-----G-----G-----G-----G-----G-----		381
SIVMM1E	-----G-----G-----G-----G-----G-----		169
SIVSMH4	-----CT-C-----A-----A-----A-----		258
SIVSMH9	---A---G---CT-C-----A-----A-----A-----		267
SIVSMFPBJ	---A---G---CT-C-----A-----A-----A-----		267
SIVSTM	---A---C---G---C---G---T-----		322

CONSENSUS . A	CCaGTAGAAAGT?TT?TAGACACAGGCTGA?GACTCAATAGTACGAGGaaTAgAgTTAGGGa?CAATTAta	477
HIV2ROD	-----C-GT-----C-----A-----	382
HIV2NIH2	-----T-GT-----A-----A-C-----G-----	379
HIV2ISY	-----T-AC-----C-----G-----	379
HIV2ST	-----T-GC-----C-----CG-----G-----	439
HIV2BEN	--G-----C-AC-----T-----A-----GA-----C-----	439
HIV2CAM2	-----T-GT-----T-----A-----G-----	379
HIV2D194	--G-----C-AC-----A-----T-----GA-----C-----	493
HIV2GH1	--G-----C-AT-----T-----G-----C-----GA-----CG-----	439
CONSENSUS . B	tCAGTAgAAGTATTACTAGACACAGGAGcTGacGACTCAATAGTAgCAGGgATAGAAATTAGGtAGCAATTACA	442
HIV2D205	-----T-----	442
HIV2UC1	C-----G-----A-----	442
HIV2EHQA	-----T-----C-----	430
CONSENSUS . D	CCcgtAGAAAGTAtTatTaGAtACAGGGCTGAtGATTCtATtGTAgCaGGaATaGAgTtAGGtCcAcAtTAtA	441
SIVMM251	-----G-----A-----	442
SIVMM142	-----A-----	442
SIVMM239	-----C-G-----A-----	454
SIVMM32H	-----C-----A-----T-----	454
SIVMNE	-----C-----	442
SIVSMH4	--C-----C-----A-----G-----A-G-----A-----	331
SIVSMH9	--CA-----G-----A-----G-----A-----	340
SIVSMMPBJ	--C-----A-----G-T-A-G-----A-----	340
SIVSTM	-----C-----G-C-----C-C-A-A-G-G-G-----C-----G-T-----A-C-----	397
CONSENSUS . A	gtCCAAAaTAGTAgGGgGGAATAGGGGATTcATAAAtACCAAaGAAtatAAAaATgTAgAAATAAAaAGTactT	550
HIV2ROD	-C-----G-----G-----T-----	455
HIV2NIH2	-----A-----G-----G-----C-----	452
HIV2ISY	-----A-----G-----G-----G-----	452
HIV2ST	-----G-----T-----C-----C-----G-----T-----	512
HIV2BEN	C-----T-----C-----C-----	512
HIV2CAM2	-C-----G-----A-----G-----G-----T-----	452
HIV2D194	-C-----G-----A-----G-----G-----	566
HIV2GH1	T-----G-----C-----GAT-----A-----	512
CONSENSUS . B	CCCCAAAaTAGTAgGGgGGAATAGGGGTTcATAAAtACCAAaGAATACAAaATGTAGAAATAGAAAGTAGT	514
HIV2D205	-----A-----	515
HIV2UC1	-----G-----	515
HIV2EHQA	-----T-----A-T-----T-----A-----	503
CONSENSUS . D	CCCCaAAAaTAGTAgGGgGGAATAGGGGTTcATAAAtACCAAaGAATACAAaATGTAGAAATAGAAAGTtTT	514
SIVMM251	-----G-----	515
SIVMM142	-----G-----	515
SIVMM239	-----G-----	527
SIVMM32H	-----G-----	527
SIVMNE	-----	515
SIVSMH4	-----T-----C-C-----C-C-----T-G-----A-C-----	404
SIVSMH9	-----T-----C-C-----C-C-----T-G-----A-A-----	413
SIVSMMPBJ	-----T-----C-C-----C-C-----T-G-----A-A-----	413
SIVSTM	-----GG-----T-----A-----A-----G-TT-G-----C-----G-A-----	470
CONSENSUS . A	AaaTAAAaG?GTaagaGCCACcATAATGACAGG?GAeACCCCaATCAACATtTTTGCCAGAAAtAtCTGAcA	621
HIV2ROD	AG-C-G-----C-C-----	528
HIV2NIH2	-GG-----G-----A-----T-----G-----	525
HIV2ISY	AG-----T-----C-----	525
HIV2ST	-----A-----T-----C-----	585
HIV2BEN	-----A-----A-----C-----	585
HIV2CAM2	-GG-----G-----G-----T-----C-C-----	525
HIV2D194	-----A-----A-----G-----	639
HIV2GH1	-----A-----A-----C-----C-----	585
CONSENSUS . B	gGGAAAaAGAGTAAGgGCAActaTAATGACAGGgGAcACCCCAATAAACATTTTGGCAGAAATATTTAAAT	587
HIV2D205	-----T-----	588
HIV2UC1	-----	588
HIV2EHQA	A-----A-----AG-----G-----	576
CONSENSUS . D	AGGCAAAaggATtAAAaGGgACaATcATGACaGGgGAcACCCgATTAAcATTTTGG?AGaAAttTpcTaACA	586
SIVMM251	-----T-----G-----	588
SIVMM142	-----T-----A-----	588
SIVMM239	-----T-----	600
SIVMM32H	-----T-----A-----	600
SIVMNE	-----T-----	588
SIVSMH4	-----GGTA-----G-A-----T-----G-A-T-----A-----T-----C-----	477
SIVSMH9	-----GTA-----G-A-----T-----A-T-----A-----C-----C-----	486
SIVSMMPBJ	-----GTA-----G-A-----T-----A-T-----A-----T-----C-----C-----	486
SIVSTM	-----G-AA-----A-----T-----T-----A-----C-----C-G-CC-AT-G-----	543
HIVFORTC2	-----C-----A-----	17

HIV2/SIV POL

CONSENSUS .A	gCctTtaGGCATGTCATTAAAtcTaccAgTcGCCAAgaTAGAaCCAAaPAAAAaTAAtgcTaaAGCCAGGgAAaG	694
HIV2ROD	-----AG-----G-----	601
HIV2NIH2	--T-----C--G-----T-----	598
HIV2ISY	-----G-----T-----A-----G-----C---CAT-----A---	598
HIV2ST	-----G-----A-----G-----	658
HIV2BEN	-----T-----T-----G-----G---CAT-G-----	658
HIV2CAM2	-----A-----	598
HIV2D194	A-----C-----T---C-----G---CAT-G-----	712
HIV2GH1	-----T-----A-----G-----C-T-G-----	658
HIV2S52907	A-GG-----AG-----G-----A---	52
CONSENSUS .B	AccTTaGGCATGACTcTAAATTTCCCAAGTaGCAaaGaTAGAaCCAGTAAAGTc?AgTtTAAAgCCTGgaAAAG	659
HIV2D205	-----G-----T-----G-----G-----TG-----A-----	661
HIV2UC1	--A-----A-----A-AC-----G-----	661
HIV2EHOA	-G-----G-----C-----A-----	649
CONSENSUS .D	GctcTpGggATGTCctTAAAtcTtCCcaTaGCTAAgaTaGagCCTgTAAAggt?ac?ttTAAA?CCAGGaaAa?G	655
SIVMM251	-----C---C-----G-----G-----	661
SIVMM142	-----GTCGC-C-----G-----	661
SIVMM239	-----P-----A-----CG-C-----G-----G-----	673
SIVMM32H	-----P---C-----A-----CG-C-----G-----G-----	673
SIVMNE	-----A-----T-C-----C-----G-----A-----	661
SIVSMMH4	---A---C---T---C-----G-----A-----A-AC---A-----A---	550
SIVSMMH9	---A---C---C-----G---R---A-----A-AC---A-----A---	559
SIVSMPBJ	---A---T---C-----A-----A-AC---A-----A---	559
SIVSTM	---T---AT---T---AG---A---A---A---G---G---A---A---A---	616
HIVFORTC2	-AAT-A-A-----T---AG-G-C-AA---A-A---A-GA---A---G-TG-	90
CONSENSUS .A	ATGGaCCAAgactgAgACAATGGCCctTaACAAaAGhaAAAATAGaAGCAcTAAaAGAAATCTGTGcGAaPAAAT	767
HIV2ROD	-----A-----	674
HIV2NIH2	-----A-A-----T-----	671
HIV2ISY	---G---A---AA-----G-----G-----	671
HIV2ST	-----A-----T-----G-----G-----	731
HIV2BEN	-----G---A-----C-----G-----G-----	731
HIV2CAM2	-----T-----G-----G-----	671
HIV2D194	-----G---A-----C-----T-----	785
HIV2GH1	-----C-----G-----G---G---C-----	731
HIV2S52907	-----A-----T-----	125
CONSENSUS .B	ATGGGgCCAAAaATCAGACAATGGCCctTATCChaaGaaAagATAcTaGCCCTCAaaGAAATCTGTGAAAAaAT	732
HIV2D205	-----G-----GG-----	734
HIV2UC1	---A-----T-G-----G-----G---	734
HIV2EHOA	-----C-----G---A-----	722
CONSENSUS .D	atGGACCAAAAT?A?P?CagTGGCCatTATCaAAAaGAAAaGATagTtGCATTaAgAGAAATcTGTGAAAA?AT	724
SIVMM251	T-----G-AG-----	734
SIVMM142	-----G-AG-----	734
SIVMM239	-----G-AG-----	746
SIVMM32H	-----G-AG-----	746
SIVMNE	-----G-GG-----	734
SIVSMMH4	-A-----A-GA-----GC-----A-----A---	623
SIVSMMH9	-----A-GA-----GC-----G-----A-----A---	632
SIVSMPBJ	-----A-GA-----GC-----A-----G-----A---	632
SIVSTM	-----A-A-AA---A-----AG---C---A---G-----A---	689
HIVFORTC2	-----A-GA-----C-----G-----A---AA-----A-----A---	163

CONSENSUS.A	GGAAAAaGaaGGcCagCTAGaaGaaGCaCCTCCAACTAATCCTTATAATACCCcACATTTGCAATcAaGAAA	840
HIV2ROD	-----G-----	747
HIV2NIH2	-----	744
HIV2ISY	---G---T---G---C---T---T---	744
HIV2ST	---G---G---G---G---	804
HIV2BEN	---G---G---	804
HIV2CAM2	-----G---	744
HIV2D194	---GG---G---A---	858
HIV2GH1	---G---G---A---G---	804
HIV2S52907	-----G---	198
CONSENSUS.B	OGAAAAgGAaGGcCag?TAGAAGaaGC?cC?CCtACTAATCCATACAaCaCCCCACCTTTtGCCATAAA?Aag	801
HIV2D205	-----AT-----A-C-----G---	807
HIV2UC1	-----C-----G-C-----A---	807
HIV2EHOA	---A---T---G---G---TT-G---C---A---	795
HIV7312AR	-----G---C-----AT-T-C-----T---G-A	71
CONSENSUS.D	GGAAaAagGAtGGtCagT?TGagGaaGCTCC?CCgACCAATCCaTACAACACCCcACaTTTGCTATAAAAGaAa	795
SIVMM251	-----G-----C-----C---	807
SIVMM142	---A-----G-----C-----T---	807
SIVMM239	-----G-----C-----	819
SIVMM32H	-----G-----C-----	819
SIVMNE	---G-----C-----C---	807
SIVSMMH4	---A---C---A---C-T-A---G-T---T---G	696
SIVSMMH9	---G-A---C---A---R---T-A---T---T---R-G	705
SIVSMPBJ	---G-A---C---A---C-T-A---T---T---T---G	705
SIVSTM	---G---G---C-A-A---T-A-A---T---T---T---762	
HIVFORTC2	---G---A---C-A-A-A---T---C---A---A---G	236
CONSENSUS.E	AGAAGGAGGGACAAATTAGAACGAGCCCCGCCACTAATCCTTATAATACACCAACCTTTTGCTATTAGGAAA	71
HIVPAR	-----	71
CONSENSUS.A	AAGGACAAaAACAAATGGAGgATGCTaATAGATTTtAGaGaaCTAAAcAagGTaACTCAaGATTTcACAGaa	913
HIV2ROD	-----C-----	820
HIV2NIH2	---A---C---G-G---G---	817
HIV2ISY	-----G-----GG	817
HIV2ST	---A---C---	877
HIV2BEN	---G---G---T---G---	877
HIV2CAM2	---G---G---G---A---C---	817
HIV2D194	---G---A---G---G---G---	931
HIV2GH1	---G---C---G---G---	877
HIV2S52907	-----C-----	271
CONSENSUS.B	AaaGAtAAaAACAAATGGAGaATGCTAATAGA?TT?AGaGaatTAAAcAAGGTaACCCaAGa?TTtACAGaag	871
HIV2D205	---G---G---C---T---G---G---C---C---	879
HIV2UC1	---G---T---T---T---G---C---	880
HIV2EHOA	---C---G---T---C---C---A---G---	868
HIV7312AR	---G---C---G---T---A---	144
CONSENSUS.D	aAaGAcas?AacAAaTGGAG?ATGCTaATAGATTtAG?GAAC?AaAtAagGTcACTCAaGAcTTTACaGAAG	865
SIVMM251	---T-G---A---G---G---G---	880
SIVMM142	---G-T-A---A---G---G---G---	880
SIVMM239	---G-T-G---A---G---G---G---	892
SIVMM32H	---T-G---A---G---G---G---	892
SIVMNE	---G---G---A-A---G---G---	880
SIVSMMH4	---A-T---G---A---T-G---	769
SIVSMMH9	R---GGA-T---G---A---A---T---C---T---	778
SIVSMPBJ	---G---G-T---G---A---A---T---C---T---	778
SIVSTM	---G---A---G---G---C---A---A-G-A---T---A	835
HIVFORTC2	---A-G---G---G---C---A---G-C-A---A---	309
CONSENSUS.E	AAAGATAAGAGTAATGGAGGATGCTTATAGATTTCAGGGAGCTAAATAAGGCCAACACAGSACTTCACAGAAG	144
HIVPAR	-----	144

HIV2/SIV POL

CONSENSUS .A	TtCAGtTAGGAATTCCACACCCaGCAGGatTgGCCAAaAGAgaaGaATtLaCTgTtactTAGATGTAGGgGATGC	986
HIV2ROD	-----G--G-----G-----	893
HIV2NIH2	-----G-----G-----	890
HIV2ISY	-----G--A-----GT-----A-----	890
HIV2ST	-C-----C-----G--A-C-----C-----	950
HIV2BEN	---C-----G---C-----A--G--CT--A--T-----	950
HIV2CAM2	-C-----G-----A-----G-----	890
HIV2D194	---C-----G-----A--G-----	1004
HIV2GH1	---C-----G-----A--G--C-----	950
HIV2S52907	-----G-----G-----G-----	344
CONSENSUS .B	tCCAACTGGGTATTCCcCACCCgGCAGG7t?ggcAGaAAA?aggAGAATAACAGTa?TAGATGT?GGAGATGC	939
HIV2D205	GT-----TGGCAG-A-----GG-----A-----G-----	950
HIV2UC1	-----G-T-----A-----C-----G-----	953
HIV2BHOA	---G-----T---A---ACT---TC---G-AA-----C-----A-----	941
HIV7312AR	-----A-----C-G-C-----A-----GA-----A-----	217
CONSENSUS .D	TcCA?tTAGGaaTaCCACACCCtGcAGGACTAGCAAAaAGGa?aAGGaTcACAGTa?TgGATgTAGGtGATGC	935
SIVMM251	---A-----T-----C-----A-----	953
SIVMM142	---A-----A-----T-----C-----A-----C-----	953
SIVMM239	---A-----A--A--T-----C-----A-----	965
SIVMM32H	---A-----A-----T-----C-----A-----	965
SIVMNE	---A--G-----AAG-----C-----C-----	953
SIVSMMH4	-A--G-----G-----G--A--GG-----T-----	842
SIVSMMH9	-A--G-----G-----R-----G--A--G-----GT-----	851
SIVSMMPB3	-A--G-----G-----G--A--G-----GT-----	851
SIVSTM	-T--GC-----T-----G-----A-----T-----G-----	908
HIVFORTC2	-T--G--G--G--T-----A--G-----T--A-----G-----	382
CONSENSUS .E	TCCAACTGGGTATTCCACACCCAGCAGGTC?AAAAGAAATGGAGCAGATAACAGTGT?TAGATATAGGGGATGC	217
HIVPAR	-----	217
CONSENSUS .A	tTAcTTTCCATACCcTAcATGAgGatTTTAGaCagTatACTGCATTtACtCTaCCAtCagTaAACAAATgca	1059
HIV2ROD	-----C-----CA-----G-----	966
HIV2NIH2	-----C-----C-----	963
HIV2ISY	C-----C--T-----G-----	963
HIV2ST	-----A-----G-----C-----A-----T	1023
HIV2BEN	C--T-----A-----G-----C-----G-----ATG	1023
HIV2CAM2	-T-----T-----A--C-----C-----G-----	963
HIV2D194	C-----A-----G--A-----C-----C-----	1077
HIV2GH1	C-----G-----A-----G-----C-----C-----	1023
HIV2S52907	-----C-----CA-----G-----	417
CONSENSUS .B	CTAcTTcAGT?TCCACTAGatCCA?ACTTCAGaCa?TATACAGCATTcAC?TTGCCAtCa?TAAa?AATGCA	1006
HIV2D205	-----A-----A-----G-----C-----G-----C-----	1023
HIV2UC1	-----A-----A-----G-----C-----A-----C-----	1026
HIV2BHOA	-----G-----G-----A-----T-----G-----T-----	1014
HIV7312AR	-----G-----C--G-----G--A-----T-----A-----T-----	290
CONSENSUS .D	ATAcTTcTCcATACCtctAGATgAaGaaTtTAGGCAGTAcACTgCcTtTAcTtTACCATCaGTAAAtCAATgca	1008
SIVMM251	-----T-----	1026
SIVMM142	-----T-----	1026
SIVMM239	-----	1038
SIVMM32H	-----	1038
SIVMNE	-----T-----	1026
SIVSMMH4	-----G-----C-----	915
SIVSMMH9	-----T-----C-----A-----C-----G-----C-----	924
SIVSMMPB3	-----T-----C-----A-----C-----G-----C-----	924
SIVSTM	-----T-----C-----G--GC-----T-----T--C--C-----C-----	981
HIVFORTC2	---C--T--T---CT---A--G--C---A--T--CA--T---CC---G---C---CG--	455
CONSENSUS .E	CTATTTTCGGTACCCTAGACCCAGAATTTAGACAGTATACAGCATTcACCATACCCTCGGTAAATAATCAA	290
HIVPAR	-----	290

CONSENSUS .A	GAaCCAGgAAAAAGATAtaTATatAAaGTCTtaCCaCagGGaTGGAAgGGaTCACCAGCAATtTTTCAstaCa	1132
HIV2ROD	-----C-----G-----C-----	1039
HIV2NIH2	-----C-----C-----G-----G-----	1036
HIV2ISY	-----C-----C-----G-----G-----	1036
HIV2ST	-----C-----C-----G-----G-----	1096
HIV2BEN	-----C-----G-----A-----A-----	1096
HIV2CAM2	-----C-----G-----A-----A-----	1036
HIV2D194	--G--A-----G-----G-----A-----A-----C-----T-----	1150
HIV2GH1	-----G-----G-----A-----A-----C-----C-----	1096
HIV2S52907	-----C-----G-----C-----C-----	490
CONSENSUS .B	GAaCCAGGAAAgAGaTAcATTTatAAAGT?CTACCACA?GGaTGGAAAGGG?TCCCCagCAATTTtCaATACT	1076
HIV2D205	-----T-----A-----G-----G-----AGT-----G-----	1096
HIV2UC1	-----T-----A-----G-----G-----AGT-----G-----	1099
HIV2EHOA	--A-----TC-----C-----G-----A-----A-----C-----G-----A-----	1087
HIV7312AR	-----A-----G-----C-----G-----A-----A-----	363
CONSENSUS .D	GAaCCAGGAAAgGATACAT?TATaAGGTtCTgCCTCagGGaTGGAAgGGaTCACCAGCCAT?TTTCAatAcA	1077
SIVMM251	--G-----C-----T-----C-----C-----C-----	1099
SIVMM142	--G-----C-----T-----C-----C-----C-----	1099
SIVMM239	--G-----C-----T-----C-----C-----C-----	1111
SIVMM32H	--G-----GC-----T-----A-----C-----C-----C-----	1111
SIVMNE	-----C-----T-----G-----C-----C-----C-----	1099
SIVSMNH4	-----A-----C-----AT-A-----A-----G-----T-----T-----G-----T-----	988
SIVSMNH9	-----A-----C-----AT-A-----A-----G-----A-----A-----T-----T-----RC-T-----	997
SIVSMMPBJ	-----A-----C-----AT-A-----A-----G-----A-----A-----T-----T-----GC-T-----	997
SIVSTM	-----A-----C-----C-----A-----A-----A-----A-----T-----T-----T-----	1054
HIVFORTC2	-----GA-----C-----A-----G-----A-----A-----A-----T-----T-----G-----	528
CONSENSUS .E	GAACCAGGAAAGATACATCTACAGGTCTTACCACAAGGGTGGAAAGGGATCCCCAGCAGTCTTTTCAGGAGA	363
HIVPAR	-----	363
CONSENSUS .A	caATGAGgCaggTcTTAGAAcCaTTCAGAAAAGCAACccaGAtgTCATtCaTcCagTACATGGATGATAT	1205
HIV2ROD	-----A-----A-----AAG-----	1112
HIV2NIH2	-----A-----GAG-----	1109
HIV2ISY	-----A-----G-----G-----	1109
HIV2ST	-----G-----A-----C-----	1169
HIV2BEN	-----A-----T-----C-----C-----	1169
HIV2CAM2	-----G-----T-----	1109
HIV2D194	TG-----AA-----T-----C-----C-----C-----A-----	1223
HIV2GH1	-----A-----T-----C-----C-----A-----	1169
HIV2S52907	-----A-----A-----AAG-----C-----	563
CONSENSUS .B	CCATGagaaAGGTA?TAGA?CCTTTCAGAAagGCCAACA??GATGTCA?cATAAT?CAGTACATGGATGACAT	1143
HIV2D205	-----T-----T-----T-----GC-----T-----T-----	1169
HIV2UC1	-----G-----T-----T-----GC-----T-----T-----	1172
HIV2EHOA	--GC-----C-----C-----A-----AT-----C-----C-----	1160
HIV7312AR	-----C-----C-----C-----G-----AT-----OC-----C-----	436
CONSENSUS .D	CtaTGAGaaATgTgcPaGAaCCCTTCAGGaa?GCaAatcCaGATGTgACctPa?TCCA?TA?ATGGATGAcAT	1146
SIVMM251	-----C-----G-----G-----G-----T-----	1172
SIVMM142	-----C-----G-----G-----G-----T-----	1172
SIVMM239	-----C-----G-----G-----G-----T-----	1184
SIVMM32H	-----C-----G-----G-----G-----T-----	1184
SIVMNE	-----G-----G-----G-----G-----T-----	1172
SIVSMNH4	-----AT-----T-----A-----A-----C-----GA-----A-----C-----	1061
SIVSMNH9	--R-----AT-----T-----T-----A-----A-----C-----GA-----A-----C-----	1070
SIVSMMPBJ	-----CT-----T-----T-----A-----A-----C-----GA-----A-----C-----	1070
SIVSTM	--A-----A-----AT-----G-----A-----GA-----C-----C-----A-----TC-----GA-----A-----C-----T-----	1127
HIVFORTC2	--A-----G-----A-----A-----A-----T-----CT-----C-----A-----T-----A-----A-----C-----	601
CONSENSUS .E	CCATGAGACAGGTACTAGAGCCCTTCAGAAAAGCAATTCGGACGTTCTACTCATTCAGTACATGGACGACTT	436
HIVPAR	-----	436

HIV2/SIV POL

[illegible]

CONSENSUS .A	CAACTAAATGGAGgtTGCAGaaAaTACAatTgCCcCAaAAAGaaTaTGGACAGTcAAtGACATCCaAAAGCT	1424
HIV2ROD	-----G-----	1331
HIV2NIH2	-----A-----GC-----G-----	1328
HIV2ISY	---C---A---A---G---G---G---T---G---A---	1328
HIV2ST	-----C---A-G-----G---G-----A---	1388
HIV2BEN	-----C---C---C---G---C-----	1388
HIV2CAM2	-----C---C---G---GG-G-----	1328
HIV2D194	-----AC-----A-T-G-----A---	1442
HIV2GH1	-----C---T---C---G-----C---G---	1388
HIV2S52907	-----G-----	782
CONSENSUS .B	CAAA?AagTGGAAaCTGCAAAAAATACAaCT?CCAGAAAa?GA?GTTTGGACAGT?AATGacATTCA?AAaCT	1352
HIV2D205	---A---G---G---A---A---G---CA---A---	1388
HIV2UC1	---A-G---G---G---A---G---A---A---	1391
HIV2EHOA	---G-A---G---A---A---G---A---G---GT---	1379
HIV7312AR	---G---G---A-A---GG-G---A---G---	655
CONSENSUS .D	CaACaAAaTGGAA??TgCAaAA?ATAGAgTgCCCaCAaAgAGAgaccTGGACAGTgAATgAtATaCAgAA?LT	1357
SIVMH251	-G-----GT-----G-----G-----	1391
SIVMH142	-G-----GT-----G-----G-----	1391
SIVMH239	-----GT-----G-----G-----	1403
SIVMH32H	-----GT-----G-----G-----	1403
SIVMNE	-----GT-----G-----A-----G-----	1391
SIVSMH4	---C---AC---G-A---A---C---A-A---	1280
SIVSMH9	-C-C---AC---A---G---A---A---A---A---	1289
SIVSMMPBJ	-C-C---AC---A---A---A---A---A---A---	1289
SIVSTM	---G---AC---A---A---A-T-G---CGTT-----C-A-AC	1346
HIVFORTC2	---G---AC-A---A---C---G---GAT---A---C-T---AC-	820
CONSENSUS .E	CCAAGAAATGGAACTACAAAAAATAGAGCTTCCTGAAAAAGAAAGATGGACAGT?AATGACATTCAAGATT	655
HIVPAR	-----	655
CONSENSUS .A	aGTgGGGTGTccTaAAATGGGCaCCaCAaaTCTaccCAGGgATAAAgACCAaacActTaTGTAggtTaATrAgA	1497
HIV2ROD	-----C-----	1404
HIV2NIH2	-----C-----	1401
HIV2ISY	-----C---G-----A-----A-C-----	1401
HIV2ST	G-----G-----G-A-----C-----	1461
HIV2BEN	---A---T---G---TT---A---A---AC-----	1461
HIV2CAM2	---A---TT-G---C---G---G---T---A---G---T---AA---G---	1401
HIV2D194	---A---A---TT-G---C---G---G---T---A---G---T---AA---G---	1515
HIV2GH1	G---A---T---G---T---A---A---T---G---C---C-A---	1461
HIV2S52907	-----C-----	855
CONSENSUS .B	aGTgGGGTATTAAAcTGGGCAGCTCAACT?TT?CCtGGaATtAaGACaAGGCacATATG?AaACTaATLAgG	1422
HIV2D205	G-A-----C-T-----C-----	1461
HIV2UC1	-----C-T-----T-----	1464
HIV2EHOA	-----T---T---C---G---G---C---T---T---A---	1452
HIV7312AR	-----T---C---C---C---A---C---G---G---C-A-	728
CONSENSUS .D	aGTAGGgGTatTAAATGGGCAGC?CAAAATTATCCAGG?ATAAAaACcAAACATCTCTG?Ag?LTAAATLAgA	1426
SIVMH251	-----T-----T-----T---G-----	1464
SIVMH142	-----T-----T-----T---G-----	1464
SIVMH239	-----T-----T-----T---G-----	1476
SIVMH32H	-----T-----T-----T---G-----	1476
SIVMNE	-----T-----T-----T---G-----	1464
SIVSMH4	-----GC---A---A---G-T---T---C---AC---C---	1353
SIVSMH9	-----GC---A---A---G-T---T---C---A---C---	1362
SIVSMMPBJ	-----GC---A---A---G-T---T---C---A---C---	1362
SIVSTM	G---G---C---C---A---A---T---T---G---C-AA---C---	1419
HIVFORTC2	G---C---C---A---A---A---G---C---A---C---	893
CONSENSUS .E	AGTAGGAATACTAAATGGGCAGCACAAGTATATCCAGGAATCAAAACAAAAAATTTATGCAAGCTCATCAGA	728
HIVPAR	-----	728

HIV2/SIV POL

CONSENSUS . A	GGAAAAATGACACtCACaGaaGaaGTaCagTGGACAGAAtTAGCAGAAgCaGagCTaGaaGAAAAcAaaATtA	1570
HIV2ROD	-----G-----G-----G-----	1477
HIV2NIH2	-----G-----G-----	1474
HIV2ISY	---G---C---G-----G---G---	1474
HIV2ST	-----G-----G---A-----C-	1534
HIV2BEN	-----G-----C-----	1534
HIV2CAM2	-----G---A-----A-----G---	1474
HIV2D194	-----G-----G-----G---A-----	1588
HIV2GH1	-----G-----G-----	1534
HIV2S52907	-----G-----	928
CONSENSUS . B	GGAAAGATGACCCtTAACAGaaGA?GTaCA?TGGACaGAA?TaGCAGagGCAGAgT.TaCAGGA?AA?AAAA?TCA	1490
HIV2D205	-----A---G-----C-----A-----C-----G---T-----	1534
HIV2UC1	-----A---A-----T-----T---G-----A---C-----G---C-----	1537
HIV2EHOA	-----G-----A-----T---T---G-----A---C-----A---C-----	1525
HIV7312AR	---G---G---G---G---C-----C-----A---G-----A---T-----	801
CONSENSUS . D	GGaAAaAtGaCt?TaACaGagGaaGTtCAGTGGACTGagATGGCAGagGCAGaaTATGa?GaaAAcAA?ATaA	1496
SIVMM251	-----C-----G-----A-----	1537
SIVMM142	-----C-----A-----G-----A-----	1537
SIVMM239	-----C-----A-----G-----A-----	1549
SIVMM32H	-----T-----A-----G-----A-----	1549
SIVMNE	-----C-----G-----A-----	1537
SIVSMNH4	-----T-----A---G-----A---C---G-----	1426
SIVSMNH9	-----T-----A---G-----A---C---G-----	1435
SIVSMNFB3	-----T-----A---G-----A---C---G-----	1435
SIVSTM	---T---G---G---AC---G---T---G---A---A---CA---A---T---	1492
HIVFORTC2	---A---CT---G---A---C---G---A---G---C---G---T---	966
CONSENSUS . E	GGAAAAATGACCCTAACAGAGAGGTTTcAGTGGACAGAATTGTcAGAGGCAGAAATTGGCAGAAAATTAAGATTa	801
HIVPAR	-----	801
CONSENSUS . A	TctTaAGcCaggAAcAaGagGGAcacTatTAcCAaGAaGAAaAggAgcTAGAaGCAACagTccaaAAagatCA	1643
HIV2ROD	--C-----C-----G-----	1550
HIV2NIH2	-----A-----A-----G-----A-----G-----	1547
HIV2ISY	-----G-----G-----T-----G-----	1547
HIV2ST	-----A---TG-----G---G---C-----	1607
HIV2BEN	---G-----A---T---T-----A---G-----A-----AGC---	1607
HIV2CAM2	-----A-----C---G---G-----T-----G-----	1547
HIV2D194	-----T-----TC---C---T---G---G---AC-----ATC---AGC---	1661
HIV2GH1	-----A---T-----A---G-----A-----A-----	1607
HIV2S52907	---C-----G-----T-----A-----G-----	1001
CONSENSUS . B	Tc?TAGaACagGAaCA?GAAGGATCtTAcTACAA?GAAGGGTACctcTAGAAGCAACAGT?CAGAAAAacT	1559
HIV2D205	---T-----A-----G---A-----G-----A-----	1607
HIV2UC1	---TT-----A-----G-----A-----A-----	1610
HIV2EHOA	---C-----A---G---G-----T---A-----T-----G-----T---	1598
HIV7312AR	---C---G-----G-----T-----A-----G-----T---	874
CONSENSUS . D	TTCTcCAGTCagGAaCaaGAaGGatgtTA?TAcCaaGAaggcAAgCCatTAGA?GC?ac?GTaaTaAAagagTCA	1565
SIVMM251	-----T-----A---C---G-----	1610
SIVMM142	-----T-----A-----A---C---G---G-----	1610
SIVMM239	-----T-----A---C---G-----	1622
SIVMM32H	-----T-----G-----A---C---G-----	1622
SIVMNE	-----T-----G---C---G-----A-----	1610
SIVSMNH4	-----A-----C-----A---A---A---G---A---A-----	1499
SIVSMNH9	---C---A---C---G-----C---T---G---G---A---C---G---AC---A-----	1508
SIVSMNFB3	-----A-----C---T---G---G---A---C---G---A---A-----	1508
SIVSTM	-----G---G---GCAA---C---G---A---TC---G---A---A---G---G---AGA---	1565
HIVFORTC2	---A---A---G---G---C---C---T-----A---A---C-----A---A---A---G---AGA---	1039
CONSENSUS . E	TTCTAAGTCAGGGACAGGAGGGGAAGATACTATAGGGAAGAAGAAACTTAGAAGCAACAGTACTAAAGAACCA	874
HIVPAR	-----	874

CONSENSUS .A	AGacaATCagTGGaCATatAaaATACACCAGGaaGa?AaaaTtCTaAAAGTaGGaAAatagcaAAagataAAA	1715
HIV2ROD	---G-----A-----G-G---	1623
HIV2NIH2	-----GG-----GG-G-----C---GT---G---ATATGC---AGAT---	1620
HIV2ISY	-----G-----G-A-----	1620
HIV2ST	-----G-----G-GA-----G---	1680
HIV2BEN	--GAC---A---C-----G---C-----G-----	1680
HIV2CAM2	-----G-----G-----A-----A-----	1620
HIV2D194	-----G---C-----G---G-GG-----G---G-----	1734
HIV2GH1	-----C-----G---C-----G-----	1680
HIV2S52907	-----C-----G-----G-----	1074
CONSENSUS .B	AGCAAAatCAGTGGACATACAAaATTCATCAGGGaATAaa?TccTAAAAGTAGGAAAATATGCAAAGGTPAAA	1631
HIV2D205	-----G-----	1680
HIV2UCL	-----G-A-TT-----	1683
HIV2ENOA	-----G-----G-----	1671
HIV7312AR	---C-----G-----	947
CONSENSUS .D	GGACAATCA?TGGTCtTALAAaATtCACaAGaGACAAAaTACTgAAAAGTaGGaAAaTtTGCAAagaTaAAa	1637
SIVMM251	-----G-----G-----	1683
SIVMM142	-----G-----G-----	1683
SIVMM239	-----G-----G-----	1695
SIVMM32H	-----G-----G-----	1695
SIVMNE	-----G-----G-----A-----	1683
SIVSMH4	---T---A---A---A---G---T---G---T---	1572
SIVSMH9	---T---A---A---G---A---C---T---	1581
SIVSMH9	---T---A---A---C---C---T---	1581
SIVSTM	---A---C---G---A---T---C---G---A---A---	1638
HIVFORTC2	---T---A---C---C---T---G---T---A---G---T---G---A---C---	1112
CONSENSUS .E	AGATAATCAGTGGAGCTATAAGATTTCATCAGGGAGACAGAATTCTTAAAGTAGGAAAATTTCCCAAAATAAAA	947
HIVPAR	-----	947
CONSENSUS .A	aAtaccCatacCaAtgGggTCagatT?tTaGCACAGGTAGTtCAaAAAAATAGGAAAaGaaGCACTaGTCATtT	1787
HIV2ROD	--C-----AA-----G-----G-----	1696
HIV2NIH2	--ATA--CATA-C-AC-----G-----	1693
HIV2ISY	-----C-----G---G---G-----G-----	1693
HIV2ST	-----C---C---A---C-CC-----A-----	1753
HIV2BEN	-----AC-----G-----G-----	1753
HIV2CAM2	C-----A---A---A-----	1693
HIV2D194	-----T-----C-AC-----A---C-----G-----G-----	1807
HIV2GH1	-----AC-----G-----G-----C-----	1753
HIV2S52907	--C-----AA-----G-----G-----	1147
CONSENSUS .B	AACACccCACACCAA?GG?GTAAGAcTAcTgGCaCATGTAGT?CA?AAAATAGGAAAaGAAAGCccTaGTCATCT	1700
HIV2D205	-----G-----C---G-----T---G-----C-----	1753
HIV2DCL	-----C---G-----T---G-----	1756
HIV2ENOA	-----T-----T---A-----T---T-----C---A-----G-----AT-G-----	1744
HIV7312AR	-----T---A---A---A---C---A---	995
CONSENSUS .D	AAtACACATAC?AATGGAGTtAGa?TAtTaGCACATGTaaTaCagAAaATAGGAAA?GAaGCAaTagTgATcT	1707
SIVMM251	-----C-----C-----G-----	1756
SIVMM142	-----C-----C-----G-----	1756
SIVMM239	-----C-----G---C-----G-----	1768
SIVMM32H	-----C-----C-----G-----	1768
SIVMNE	-----C-----C-----G-----	1756
SIVSMH4	-----A-----C---T---C-----C---G-G-----A---C---A---T---	1645
SIVSMH9	-----A---C---CT-----G-G-----A---A---A---T---	1654
SIVSMH9	-----A-----C---T-----G-G-----A-----A---T---	1654
SIVSTM	--C-----A-----T---G-----G---T---G-----G---A-----	1711
HIVFORTC2	-----A-----T-----G---A-----A---G-----A---T---	1185
CONSENSUS .E	AATACCCATACAAATGGAGTAAGACTATTAGCTAGTGTAGTACAGAA	995
HIVPAR	-----	995

HIV2/SIV POL

CONSENSUS .A	GgGGACGaaTaCCaAAATtCACcTaCCaGTaGagAGaGA?acCTGGGAgCagTGGTGGGATAACTAcTGCCA	1859
HIV2ROD	-----A-T-----	1769
HIV2NIMZ	-----G-----T-----	1766
HIV2ISY	-----A-----G-----A-----T-----	1766
HIV2ST	-----A-----T-----A-----	1826
HIV2BEN	-----G-----G-----G-----	1826
HIV2CAM2	-A-----T-----GCT-----	1766
HIV2D194	-----G-G-----G-----C-----A-----	1880
HIV2GH1	-----G-----C-----T-----G-----G-----	1826
HIV2S52907	-----G-----C-----G-A-T-----A-----	1220
CONSENSUS .B	GGGGAGAGATACCAgTgTTCCATCTgCCAGTAGAAAGAGAGACA?GGGAcCAGTGGTGGACAGATTACTGGCA	1773
HIV2D205	-----	1826
HIV2UC1	-----A-----	1829
HIV2EHOA	-----A-----A-----T-----	1817
CONSENSUS .D	GGGGAcagGgTcCCaAaaTtcCACtTaCCagT?GagAgagATgTaTGGGAACA?TGGTGGaCAGAcTAATTGGCA	1778
SIVMM251	-----T-----G-----G-----	1829
SIVMM142	-----T-----AG-----G-----	1829
SIVMM239	-----T-----AG-----G-----	1841
SIVMM32H	-----T-----T-----AG-----G-----	1841
SIVMNE	-----T-----A-----G-----	1829
SIVSMH4	-----G-----G-----T-G-----A-A-----AA-T-----A-----T-----	1718
SIVSMH9	-----G-----G-----T-T-G-----A-----R-----AA-T-----A-----T-----	1727
SIVSMPEJ	-----G-G-G-----T-T-G-----A-----AA-T-----A-----T-----	1727
SIVSTM	-----A-----G-----A-----G-C-----A-----G-----T-----	1784
SIVSMLIB1	-----	1
HIVFORTC2	-----G-T-G-----C-----A-A-A-----T-----A-----T-----	1258
CONSENSUS .A	agTgACATGGaTCCAGAcTGGGACTTcgTaTCTACCCCaCCacTgGTCAGgcTagcaTTTAACcTgGTagga	1932
HIV2ROD	-----G-----G-----G-----	1842
HIV2NIMZ	-A-----T-----C-----	1839
HIV2ISY	-----C-----GT-----AA-----	1839
HIV2ST	-----A-----G-----A-----T-----GAA-----	1899
HIV2BEN	-----G-----T-----A-----A-----A-----	1899
HIV2CAM2	G-----AC-----T-A-----G-----	1839
HIV2D194	-----G-----G-----GA-----T-----	1953
HIV2GH1	-----G-----A-----A-----A-----	1899
HIV2S52907	-----G-----G-----G-----G-----	1293
CONSENSUS .B	AGTAACcTGGATCCCAGaTGGGAcTTTGTCTCaACCCCaCCATTaAaTAAGAcTAGCCTaAACCTaGTCAAA	1846
HIV2D205	-----G-----C-----G-----	1899
HIV2UC1	-----T-----G-----	1902
HIV2EHOA	-----GT-----T-----G-----	1890
CONSENSUS .D	gGTaACcTGGATACCAgAaTGGGAcTTtAT?TCAaACcC?CC?LTaGTaAGAcTAGTcTTcAAAcTaGTaAAg	1848
SIVMM251	-----G-G-----C-----G-A-A-----G-----	1902
SIVMM142	-----G-----C-C-----A-A-----G-----	1902
SIVMM239	-----G-----C-----A-GC-----G-----	1914
SIVMM32H	-----C-----A-AC-----G-----	1914
SIVMNE	-----GA-----C-----A-AC-----G-----	1902
SIVSMH4	A-----T-----C-G-G-----T-C-----C-----C-----A	1791
SIVSMH9	A-----C-G-G-----T-C-----C-----C-----C-----A	1800
SIVSMPEJ	A-----T-----C-G-G-----T-C-----C-----C-----T-----A	1800
SIVSTM	-----A-----G-A-C-----T-T-----T-----T-----A	1857
SIVSMLIB1	-----G-A-----A-C-----C-T-G-T-G-----C-G-----	73
HIVFORTC2	A-----G-A-T-----T-TC-----T-C-----T-T-CT-G-----A	1331

CONSENSUS .A	GATCCTaTACcAGGcgCAGAGACcTtTACACaGATGGATCCTGCAATAggCAaTCAaaaGaaGaaAAAGCAG	2005
HIV2ROD	-----T-----	1915
HIV2NIH2	--G--G-----A--T-----G-----G-----	1912
HIV2ISY	-----G-----T-----	1912
HIV2ST	---C---T---A---A---G---	1972
HIV2BEN	-----A-----A-----G-----G-----	1972
HIV2CAM2	-----A-----	1912
HIV2D194	-----A-----T-----A-----A-----G-----	2026
HIV2GH1	-----A-----A-----G-----G-----C	1972
HIV2S52907	-----T-----	1366
CONSENSUS .B	GACCCCTAGAAggggtAGAAACcTACTACACAGATGGaTCCTGcAAGAggCCTCAAAGGaaGaaAAAGCAG	1919
HIV2D205	-----AG-----G-----T-----A-----	1972
HIV2UC1	-----AA--G-----	1975
HIV2EHOA	-----A-----T-----T-----A-----A-----G-----	1963
CONSENSUS .D	GA?CCTaTagAgggag?AGAAACcTtTATgtAGATgGATCaTGTaatAgacAgTCaAaaGaaGGaAAAGCAG	1919
SIVMM251	--C-----A-----A-----AC-----A-----	1975
SIVMM142	--C-----A-----A-----G-----A-----	1975
SIVMM239	--C-----A-----A-----AC-----A-----	1987
SIVMM32H	--C-----A-----A-----AC-----A-----	1987
SIVMNE	--C-----A-----A-----A-----A-----	1975
SIVSMH4	--G-----C-----C-----A-----C-----G-----G-----A-----	1864
SIVSMH9	--G-----C-----C-----A-----A-----C-----G-----G-----A-----	1873
SIVSMFPBJ	--G-----C-----C-----A-----C-----G-----G-----A-----G-----	1873
SIVSTM	--A--CT--G--A--AC-----A-----G-----C-----C--G--A--T-----	1930
SIVSMLIB1	--A--CC-----T-----C-----C-----G--A--T--G--T--A--G-----	145
HIVFORTC2	--A--CC-----AAGAC-----C-----C--G--CC--C-----A--T-----	1404
CONSENSUS .A	GATATgTaACAGATAGAGGaaAgAGACAagGTaAaa?tATAGAGCaAACTaCCAATCAgCAaGCAGAAaTAGA	2077
HIV2ROD	-----G--A-----GAA--C-----C-----	1988
HIV2NIH2	-----A-----G-----G-----G-----	1985
HIV2ISY	-----A-----A-----A-----GGA-----C--A-----	1985
HIV2ST	---CA-----G--GGC-----C-----	2045
HIV2BEN	-----A-----A-----G-----A-----G-----	2045
HIV2CAM2	-----A-----A-----C-----A-----G-----	1985
HIV2D194	-----G-----G-----G--G-----AT-----C-----	2099
HIV2GH1	-----G-----G--G--C-----A--G-----	2045
HIV2S52907	-----G--A-----GAA--C-----	1439
CONSENSUS .B	GATATGTACaGACAGGGGAAaaGATAagGTTAAAggttTAGAaCagACAACAAATCAaCAAGCAGAAaCTTGA	1992
HIV2D205	-----T-----C-----	2045
HIV2UC1	-----C-----G-----	2048
HIV2EHOA	-----G-----A-----CCA-----A-----G-----G-----	2036
CONSENSUS .D	G?PAtcT?ACaGAtAGGgGcAaAgacAAagTAAAggtgtTagAACAgACTACTaAtCAacAACAGAAaTgGA	1990
SIVMM251	--A-----C-----A-----	2048
SIVMM142	--A-----C-----C-----G-----	2048
SIVMM239	--A-----C-----	2060
SIVMM32H	--A-----C-----	2060
SIVMNE	--A-----C-----C-----	2048
SIVSMH4	--C--G--G--G-----G--C--C--T-----C--C-----G-----	1937
SIVSMH9	--C--G--G-----A--G--G--GAC--C--C--G-----C-----G-----	1946
SIVSMFPBJ	--C--G--G-----A--G--G--G--C--C--G-----C-----G-----	1946
SIVSTM	--C--C--A--T-----A--A--A-----CT-----A-----A-----	2003
SIVSMLIB1	--G--C--A-----C--A--A--G-----G--CT-----A-----C-----GA-----GC--A-----	218
HIVFORTC2	--T--G--A-----A--G--A--AC-----CC--C-----A-----C-----C--A-----	1477

HIV2/SIV POL

CONSENSUS .A	AGCCTTTGcgATGgCacTaaCAGACTCAGGCCaAAaGtAAAtATTaTAGTAGAcTCACAgTATGTAAATGGGa	2150
HIV2ROD	-----G-----	2061
HIV2NIH2	-----A-----C-----A-----G	2058
HIV2ISY	-----A-----G-----C-----G-----	2058
HIV2ST	-----G-----G-CC-C-----A-----	2118
HIV2BEN	--T--CC--GG--C--G--C-----T-----G	2118
HIV2CAM2	-----C-----	2058
HIV2D194	-----OG-----C-G-----C-----G	2172
HIV2GH1	-----A--G-----G-----C-----	2118
HIVFA	-----	7
HIV25132P1	-----G-----	7
HIV2S52907	-----G-----	1512
CONSENSUS .B	AGCATTTCGAcTACacTACagGAcTCAGgACCACAAcTCAAcATCAGTAGAGAcTCACAATATGTGATGGGA	2065
HIV2D205	-----T-----T-ACA-----A-----T-----	2118
HIV2UC1	-----T-----C-----	2121
HIV2EHOA	-----G-----T-----	2109
HIV7312AI	-----	7
HIVJA	-----	7
HIVOM	-----	7
CONSENSUS .C	GTTAGGG	7
HIV2POLB7	-----	7
CONSENSUS .D	AGCaTTTctcaTgGCatTg?caGAcTCAGG?CCAAAggCAAAATATTaTaGTAGATTCaCAATAcGTcATgGGa	2061
SIVMM251	-----A-----G-----A-----	2121
SIVMM142	-----A-----G-----C-----	2121
SIVMM239	-----A-----G-----	2133
SIVMM32H	-----A-----G-----	2133
SIVMM6	-----G-----G-----	2121
SIVSMMH4	--C--CTATC-A--C--AG--T--G--A--A-----C-----C	2010
SIVSMMH9	--C--TATC-A--C--AG--T--A--A-----C-----C	2019
SIVSMMPB3	--C--CTA-C-A--C--AG--T--A--A-----C-----C	2019
SIVSTM	---CC---C--G-G---T---A---G-G---C---G---	2076
SIVSMMLIB1	-----T-----A-CA-----T--G-A-----C-----C-----G	291
HIVPOLC12	-----C-A---	7
HIVFORTC2	-----T-----ACA--T--T-----C-----C-----C-----	1550
CONSENSUS .A	AtagttagcAgGCCAaCCaCaGagTCAGaaAgTAgAATAGTaAaccAaATCATAGaAGaaATGATAAAGAAgG	2223
HIV2ROD	--CAGT--A-----G-A-----T-G-----A-----A-	2134
HIV2NIH2	-----G-A-----T-G-----A-----A-	2131
HIV2ISY	-----GG-T--A--G-----TA-----G-----A-----	2131
HIV2ST	-----G-----A-----T-----A-----	2191
HIV2BEN	-----G-----A-----G-----G-----	2191
HIV2CAM2	-----G-----A-----T-----A-----	2131
HIV2D194	-----A-----T-G-C-----A-----	2245
HIV2GH1	-----T-----A-----T-G-C-----A-----	2191
HIVFA	-----T-----A-----T-G-C-----A-----	80
HIV25132P1	-----A-----A-----A-----T-G-C-----A-----	80
HIV2S52907	-----A-----A-----G-----G-----A-----	1585
CONSENSUS .B	ATA?TAGCTGgACAgCCAACAGAAACaGAgTCACCAa?AGTaaataAaATAATTGAAGAAATGATCAAAAAGG	2137
HIV2D205	--A--C-----A-----GCA-----A-----	2191
HIV2UC1	--G-----C-----G-----C-----A-----	2194
HIV2EHOA	--G-----C-----A-----G-----GAG-----	2182
HIV7312AI	--G-----G-----	80
HIVJA	--A-----	80
HIVOM	--A-----A-----G-----	80
CONSENSUS .C	ATTATAGCAGGGCAGCCAACAGAAACAGAGTCCCCACTAGTAAACAGAAATCATAGAAAGAA?TAATAAGGAACG	80
HIV2POLB7	-----	80
CONSENSUS .D	AtaaTa?CAGGaca?CCLACaGAATCaGagAgcAgg?TAGTAAACCA?ATaATAGAgAaATGATLAAAAAgg	2130
SIVMM251	-----A-----TGC-----C-----T-----A-----T	2194
SIVMM142	-----A-----TGC-----C-----T-----A-----C-----A	2194
SIVMM239	-----A-----TGC-----C-----T-----A-----T	2206
SIVMM32H	-----A-----TGC-----C-----T-----A-----T	2206
SIVMM6	-----A-----TGC-----C-----T-----A-----A	2194
SIVSMMH4	-----G-----T-G--C--T-----A--T--AT-----G-----	2083
SIVSMMH9	-----G-----T-A--C--T-----A--T--AT-----G-----GA-----	2092
SIVSMMPB3	-----G-----T-A--C--T-----A--T--AT-----G-----	2092
SIVSTM	--C--A-----G--C-----G-----A--T-----T-----G-----	2149
SIVSMMLIB1	--TG--G-----G-----T-----A--AT--ACC-----T--A-----G-----C-----A	364
HIVPOLC12	--TG--GG-----A-----A--AT-----G--T--G-----	80
HIVFORTC2	--TG--GG-----A-----A--AT-----T--G-----A-----	1623

CONSENSUS.A	AAGCAATCTATGTCGCaTGGGTCCCAAGCCCATAAAGGCATAGGAGGAAACCAgGAagTaGACCaTCTAGTaAG	2296
HIV2ROD	-----C-----G-----T-----G-----	2207
HIV2NIH2	-----G-----C-----T-----A-----A-----T-----C-----	2204
HIV2ISY	-----G-----C-----T-----A-----A-----T-----C-----	2204
HIV2ST	-----T-----G-----T-----C-----	2264
HIV2BEN	-----G-----T-----G-----	2264
HIV2CAM2	-----G-----T-----G-----	2204
HIV2D194	-----G-----C-----G-----	2318
HIV2GH1	-----G-----C-----G-----	2264
HIVFA	-----C-----	153
HIV25132P1	-----G-----A-----G-----	153
HIV2852907	-----C-----G-----T-----G-----	1658
CONSENSUS.B	AagCaATATATGTAGGATGGGTGCCAGCTCACAAAGGACTAGGTGGTAATCAAGGAAGTAGACCACCTAGTaAG	2210
HIV2D205	-----G-----A-----G-----	2264
HIV2UC1	-----G-----	2267
HIV2EHOA	-----AA-----A-----G-----G-----G-----	2255
HIV7312AI	-----	153
HIVJA	-----	153
HIVON	-----	153
CONSENSUS.C	AAGCTGTTTACATCAGCTGGGTGCCAGCTCACAAAGGCTCTCGAGGAAATCAAGGAAGTAGACCACCTAGTaAG	153
HIV2POLB7	-----	153
CONSENSUS.D	?aGCaATtTATgTaGCATGGGTaCCaGCaCACAAAGgaATAGGAGGaAAcCAaGAagTAGaCCA7cTaGTcAG	2201
SIVMM251	C-A-----T-----A-----C-----	2267
SIVMM142	C-A-----G-----T-----A-----C-----	2267
SIVMM239	C-A-----T-----A-----C-----	2279
SIVMM32H	C-A-----T-----A-----C-----	2279
SIVMNE	C-A-----G-----T-----A-----C-----	2267
SIVSMMH4	A-----T-----T-----C-----T-----T-----C-----G-----	2156
SIVSMMH9	A-----T-----T-----RR-----A-----T-----T-----T-----	2165
SIVSMMPB7	A-----T-----T-----T-----T-----T-----T-----G-----	2165
SIVSTM	A-----T-----T-----T-----G-----G-----T-----TT-----G-----A-----	2222
SIVSMMLIB1	C-----T-----A-----T-----T-----T-----T-----G-----T-----TT-----G-----	437
HIVPOLC12	AG-----A-----CA-----T-----T-----T-----T-----C-----	153
HIVFORTC2	AG-----A-----CA-----T-----A-----G-----T-----C-----	1696
CONSENSUS.A	TCAGGGCATCAGACAAgTatTaTtccctgGAAaAAATAGAGCCCGCtCA7GAaGAACatGAAAAATATCATAgC	2368
HIV2ROD	-----T-----G-----G-----G-----	2280
HIV2NIH2	-----T-----A-----G-----A-----G-----G-----	2277
HIV2ISY	-----A-----G-----G-----G-----G-----G-----	2277
HIV2ST	-----A-----G-----A-----G-----G-----	2337
HIV2BEN	-----A-----G-----G-----A-----G-----	2337
HIV2CAM2	-----A-----C-----G-----A-----G-----P-----G-----G-----C-----	2277
HIV2D194	-----G-----A-----G-----A-----C-----	2391
HIV2GH1	-----G-----A-----G-----C-----A-----	2337
HIVFA	-----G-----G-----G-----A-----	226
HIV25132P1	-----T-----T-----C-----A-----G-----	226
HIV2852907	-----T-----G-----G-----AATA-----	1686
CONSENSUS.B	tCAAGGAATtAGgCAGgTtCtTaTtccTAgAAAAgATAGAACCAgCaCAaGAaGAaCATGAAAA?TAtCATggc	2282
HIV2D205	-----C-----A-----CT-----G-----A-----C-----G-----G-----A-----	2337
HIV2UC1	-----G-----G-----G-----G-----G-----G-----G-----G-----	2340
HIV2EHOA	C-----A-----AA-----C-----T-----A-----T-----A-----AAT	2328
HIV7312AI	-----C-----G-----G-----A-----	226
HIVJA	-----G-----G-----G-----	226
HIVON	-----T-----G-----C-----	226
CONSENSUS.C	TCAAGGAATtAGGCAAGTATTTCTTAgAAAAATAGAACCAgCACAGGAAGAGCATGAAAGATTTTCATAGC	226
HIV2POLB7	-----	226
CONSENSUS.D	tCAAGg?ATtAGACaAGT?CT?TTctTAgAaAAaATaGAaCCAGCgCaGAaGAaCATga?AAaTAcCATAG?	2269
SIVMM251	-----G-----G-----T-----C-----G-----G-----G-----T-----T-----T-----	2340
SIVMM142	-----G-----T-----C-----G-----G-----G-----AGT-----T-----	2340
SIVMM239	-----G-----T-----C-----G-----G-----G-----T-----T-----T-----	2352
SIVMM32H	-----G-----T-----C-----G-----G-----G-----G-----T-----G-----T-----	2352
SIVMNE	-----G-----T-----C-----G-----G-----G-----T-----T-----T-----	2340
SIVSMMH4	C-----G-----A-----C-----A-----A-----G-----A-----G-----C-----	2229
SIVSMMH9	-----AA-----C-----A-----G-----G-----G-----A-----C-----	2238
SIVSMMPB7	-----A-----C-----A-----G-----G-----G-----A-----C-----	2238
SIVSTM	-----G-----A-----G-----C-----A-----T-----A-----T-----C-----	2295
SIVSMMLIB1	-----A-----T-----A-----G-----C-----G-----G-----T-----C-----	510
HIVPOLC12	C-----G-----G-----C-----G-----TC-----T-----C-----G-----C-----	226
HIVFORTC2	C-----A-----G-----C-----G-----TC-----T-----C-----G-----T-----	1769

HIV2/SIV POL

CONSENSUS .A	AaTaTaaaAGaaCTaaccCATAAAATTTGGaaTaCCcCaactTaGtGCAAGaCA?ATAGTAAACaCATGtgCCc	2440
HIV2ROD	---G---GT-T---A-TT---G-A---T---	2353
HIV2NIH2	---GT---A-G---A---	2350
HIV2ISY	---G---G---T---A-C-G---	2350
HIV2ST	---G---T---C-G---A-T---A---CA---	2410
HIV2BEN	---T---TT---A---G---T---	2410
HIV2CAN2	---G---G---TG---AT---A---A---	2350
HIV2D194	---G---	2464
HIV2GH1	---G---A---G---	2410
HIVFA	---G---T---A---G---	299
HIV25132P1	---C---G---A-A---	299
CONSENSUS .B	AATGTAAAGAAaCTaGTeCATAAAATTTGGaaT?CcacaactTAGTgGCAAAACaATACTAAACCTCTGTGATA	2354
HIV2D205	---G---C---C---T---G---T---	2410
HIV2UC1	---T---C-C---G---	2413
HIV2EHOA	---C---G-T---G---T---	2401
HIV7312AI	---G---G---T---C---	299
HIVJA	---C-ACAGT---	299
HIVON	---C---G---A---	299
CONSENSUS .C	AATGCCAAAGAGCTGACACACAAAATTTGGTATACCTAAGCTTAGTGGCTAAACAGATAGTAAATACTTGTCTATG	299
HIV2POLB7	---	299
CONSENSUS .D	AAtgTaaaAgAaTgGTaTtCaaATtTgG?LTaCC?AgacTaGtaGCaAaaCagATagTaGAcACaTGTGAtA	2340
SIVMM251	---A---C---A---G---C---C---	2413
SIVMM142	---CA---C---C---G---C---	2413
SIVMM239	---A---C---A---G---C---C---	2425
SIVMM32H	---A---C---A---G---C---C---	2425
SIVMNE	---A---C---G---C---	2413
SIVSMH4	---T---T-G---C---	2302
SIVSMH9	---A---R-T---T-G---A---	2311
SIVSMMPBJ	---T---T-G---	2311
SIVSTM	---G---GC-A-C---G---CA---A-G---G-A---G-T---C---	2368
SIVSMHL1B1	---A---G-T---A-CA---A-T-G---G---C---C---	583
HIVPOLC12	---A---C-A---T---G---A-A-T---G---A---A---T---	299
HIVFORTC2	---A---C-A---TC-G---A-A-T---G---A---T---C---	1842
CONSENSUS .A	AatgcCAaCAGAAAGG?GA?GCTATACATGGGCAaGTAAATGCAGAA?TaGGc??TGGCAAAATGGACTgCAC	2507
HIV2ROD	---T---G-A---C---ACT---	2426
HIV2NIH2	---TGTG---G-G---C---ACT---	2423
HIV2ISY	---G---A---G-A---C---ACC---	2423
HIV2ST	---T-G---G-G---T---ACT---	2403
HIV2BEN	---A-A-C---A---GTT---	2483
HIV2CAN2	---AT---G-G---G-G---ACC---	2423
HIV2D194	---A-A-C---A---GTT---	2537
HIV2GH1	---A-G-C---G---A---TGTC---	2483
HIVFA	---A-G---A---TGTC---	372
HIV25132P1	---A-A-C---C---GTC---	372
CONSENSUS .B	AATGcCAACAAAAGGgGAAGCTgTTCATGGACAGGTAAATgCAGAACTAGG?ACaTGGCagATGGACTGTAC	2426
HIV2D205	---A---C---G---	2483
HIV2UC1	---A---	2486
HIV2EHOA	---A---T---G---A---	2474
HIV7312AI	---T---A---G---A---	372
HIVJA	---A-G---	372
HIVON	---A---	372
CONSENSUS .C	TGTGTCAACAGAAAGGAGAGGCCATACATGGGCAAGTAAATGCAGATTTAGGAACATGGCAAAATGGACTGTAC	372
HIV2POLB7	---	372
CONSENSUS .D	AaTGtCA?CagAAaGGAGaaGCLaTaCatGGgCagGtAAATgCaGaaCTaGGgActTGGCAaATGGACTGTAC	2412
SIVMM251	---T---T---T---	2486
SIVMM142	---T-A---T---C---	2496
SIVMM239	---T---G---C---T---T---	2498
SIVMM32H	---T-A---T---T---	2498
SIVMNE	---T---T---T---	2486
SIVSMH4	---C-C---C---A---	2375
SIVSMH9	---C-C-T---Y---A---R-R---	2384
SIVSMMPBJ	---C-C---C---A-A---	2384
SIVSTM	---G-C-C---A---A---GT---G---	2441
SIVSMHL1B1	---C-T-A---A-C-C---A---T-G-A-C---	656
HIVPOLC12	---C-GA-G---G-A---T---A---G---	372
HIVFORTC2	---C-GA-G---G-A---A---G---	1915

CONSENSUS .A	ACActTAGaaGGAAaaTcATtATAGTAGCAGTcCATGTTCGAAGTGGATT?ATAGAaGCAGaAGTcATCCCA	2579
HIV2ROD	---T-----G-----T-----	2499
HIV2NHZ	---T-----G-----T-----G-----T-----	2496
HIV2ISY	-----T-----	2496
HIV2ST	-----T-----	2556
HIV2BEN	-----C-----	2556
HIV2CAM2	---TC---G-----C-----T-----	2496
HIV2D194	-----G-----G-----C-----	2610
HIV2GH1	-----G-----C-----T-----	2556
HIVFA	-----C-----A-----	445
HIV2S132P1	-----G-----C-----	445
CONSENSUS .B	ACA?TTAGaaGGaAAggTcATAATAGTcGCAGTcCATGTAGCCAGTGGaTTtATAGA?GCAGaGTAATACCC	2497
HIV2D205	---T-----AA-T-----G-----A-----	2556
HIV2UC1	---C-----G-----	2559
HIV2EHOA	---T---G-----T-----T-----C---A---A-----	2547
HIV7312AI	---T---G---A-----A-----G---A-----	445
HIVJA	---C-----G---A-----	445
HIVON	---C-----G---A-----G-----	445
CONSENSUS .C	CCATTTAGaAGGAAAGTTATCATAGTAGCAGTCCATGTGGCCAGTGGCTTTATAGAAGCAGAGTAATCCCA	445
HIV2POLB7	-----	445
CONSENSUS .D	aCAtcTAGaGGaAAaaTaatcATAGTcGCAGTACATGTaGcTAGtGGaTTcATaGaaGCAGaAGTAATcCCa	2485
SIVMM251	C-----A-----G-----	2559
SIVMM142	C-----G-----	2559
SIVMM239	C-----G-----	2571
SIVMM32H	C-----	2571
SIVMNE	C-----	2559
SIVSMNH4	G-C---A-C---T-----G-----T---G-----C-G	2448
SIVSMNH9	---C---C---T-----G-G-----G-----C-G	2457
SIVSMNPBJ	---C---C---T-----G-----G-----G-----G	2457
SIVSTM	---A-C-G---T---A-----G-T---G-----C	2514
SIVSMMLIB1	---T---C---T---T---G-T---C-T---C---G---G	729
HIVPOLC12	---C---G---A-T---G---T---C---C---C---C---C	445
HIVFORTC2	---C---G---C---A-T---G---T---T---T---T---T	1972
CONSENSUS .A	CAGaAATCAGgAAAgCA?ACaGCaCTcTTCc?AcTaAAAcTGGCCAGLAgGTGGCCAAATAACaCActTgCaCa	2651
HIV2ROD	---A---A-----G-----A-----T-----	2572
HIV2NHZ	---A---G---T---C-G-----C-----	2569
HIV2ISY	---A-----	2569
HIV2ST	---A-G-----C-----T-----	2629
HIV2BEN	---G-----G-----	2629
HIV2CAM2	---A-----T---T---C-----A-----	2569
HIV2D194	---G-----	2683
HIV2GH1	---G---G-----A-----	2629
HIVFA	---A---A---A-G-----C-----	518
HIV2S132P1	---G---C-----	518
CONSENSUS .B	CAAGAAACaGGaAGACA?ACAGCTCTCTTCTGTAAAT?TGGCCAGCAGATGGCCtATcACACACcTgCaCa	2568
HIV2D205	---G---G-----AC---G-----A-----	2629
HIV2UC1	---A---A-----A-----	2632
HIV2EHOA	---G-----GC-----	2620
HIV7312AI	---A---A-----GC---A-----C---T---T-----	518
HIVJA	---G---G-----A-----A-----	518
HIVON	---A---A---A-----A-----G-----	518
CONSENSUS .C	CAAGAAACTGGCAGACAAACAGCTCTTTCTCTTAAATTTGGCAGCAAGATGGCCtATTACCCACCTACACA	518
HIV2POLB7	-----	518
CONSENSUS .D	CaaGAAACAGGaAGaCAGACAGCActTaTTcCTGtTaAAATTTgGcaaGCAGATGGCCtATtACaCActcTgCaCa	2558
SIVMM251	---G-----G-----	2632
SIVMM142	-----	2632
SIVMM239	---G-----G-----A-----	2644
SIVMM32H	---G-----G-----G-----A-----	2644
SIVMNE	---G-----G-----G-----A-----	2632
SIVSMNH4	---G-----A---G-----G---A---TG-----C---C-----T	2521
SIVSMNH9	---RG-----A---G---C---A---C-----C---C-----T	2530
SIVSMNPBJ	---G-----A---G---C---A---C-----C---C-----T	2530
SIVSTM	---G-----G---T---G---G---A---C-----C---C-----C	2587
SIVSMMLIB1	---G-----G---T---G---C---A---T-----C---C-----T	802
HIVPOLC12	---G-----A---TT---C-----G---C---C-----C	518

HIV2/SIV POL

CONSENSUS .A	CAGAcAATGG?GCCAACTTCACcTCACAGGAaGTgAAGATGGTgGCATGGTGGaTAGGtATAGaCAATCCTT	2723
HIV2ROD	---T---T---G---A---A---G---A---	2645
HIV2NIH2	---T---T---A---A---G---A---	2642
HIV2ISY	---T---T---G---A---A---G---C---	2642
HIV2ST	---T---T---A---A---G---A---	2702
HIV2BEN	---CC---G---	2702
HIV2CAM2	---T---T---A---A---G---C---A---	2642
HIV2DI194	---CC---G---	2756
HIV2GHI	---CT---G---	2702
HIVFA	---CT---C---A---G---A---	591
HIV25132P1	---CT---G---	591
CONSENSUS .B	CAGACAA?GG?GCCAACTTCACcTCACaAgatGTgAAGATGGTgGCcTGGTGGgTAGG7ATAGAACAACcTT	2638
HIV2D205	---C---T---C---AG---A---A---A---T---	2702
HIV2UC1	---C---C---C---C---A---G---	2705
HIV2EHOA	---C---T---T---A---CA---A---G---A---	2693
HIV7312AI	---T---T---T---T---A---T---A---	591
HIVJA	---T---C---C---A---A---G---	591
HIVON	---T---C---G---	591
CONSENSUS .C	CAGATAATGGTGGCAATTTCACTTCACAGGAAGTAAAAATGGTGGCGTGGTGGGTAGGCCTGGAACATACCTT	591
HIV2POLB7	---	591
CONSENSUS .D	CaGAtAAtGGtGCCAAcTTcaC?TCgCAaGAAgTaAAGATGGTtGCaTGGTGGGcagggATaGAgCA?ACcTT	2629
SIVMM251	---T---TG-C---	2705
SIVMM142	---T---TG-T---	2705
SIVMM239	---T---TG-T---	2717
SIVMM32H	---T---TG-C---	2717
SIVMNE	---TG-T---	2705
SIVSMNH4	-T-----T---A-A---G-A---C-----T-A-G---	2594
SIVSMNH9	-T-----T---A-A---G-A---T-----R-T-A-G---	2603
SIVSMNPBJ	-T-----T---A-A---G-A---T-----T-A-G---	2603
SIVSTM	-G-----C-T---C-A---G-G---A-C---G-A---A-A---	2660
SIVSMNLB1	-----C-A---G---A---AAC---A-G-T---	875
HIVPOLC12	---C-C---T-A---G-G-A---A---T-A---A---	591
CONSENSUS .A	tGGAGTAcCTTACAAcCCACAaAGCCA?GGAGTAgTgAaGCAATGAATCAcCAccTaAaaAaAcCAGATAggt	2795
HIV2ROD	---G---A---T---C---A---	2718
HIV2NIH2	---G---A---T---T---GAC---	2715
HIV2ISY	---G---G---G---GAA---	2715
HIV2ST	C-----A---G-----C---	2775
HIV2BEN	---C---G-----G---	2775
HIV2CAN2	---C---G-----A---	2715
HIV2DI194	-----G-----	2829
HIV2GHI	---C---G---T---G---	2775
HIVFA	C-----A---G-----	664
HIV25132P1	---C---G---G---	664
CONSENSUS .B	cGG?GTaCCCTAtAACCCAcAAAGTCagGGAGTAGTgAaAGCAAT?GAACCATCA?CT?AAAAA?CAaATAGAL	2707
HIV2D205	T-A-----A---G---C---G---T---C---	2775
HIV2UC1	---G-----C---A---C---	2778
HIV2EHOA	---A-G-----T-G---T-G---T-G---C---	2766
HIV7312AI	T-A-----C---T---G---T---	664
HIVJA	---G-----C---A---C---	664
HIVON	---G-----G---T---A---C---	664
CONSENSUS .C	TGGAGTACCTTACAACCCACAGAGTCAGGGAGTAGTAGAGGCCATGAACCATCATCTGAAAAACCAATAGAT	664
HIV2POLB7	---	664
CONSENSUS .D	TggGgTAcCAaTAcAAcCCACAGAGTCaagGAGTACTgGAaGCaATGAATCAcCAcTgAaaAaAcCAaATAGAL	2702
SIVMM251	-----G---C---C---G---	2778
SIVMM142	-----G---C---C---G---	2778
SIVMM239	-----G---C---C---G---	2790
SIVMM32H	-----G---C---C---G---	2790
SIVMNE	-----G---C---C---	2778
SIVSMNH4	---G-T-T---C---C---T---CC-G---	2667
SIVSMNH9	-R-----T-T---C-R-G---G---C---CC---	2676
SIVSMNPBJ	---T-T---C---G---C---C---	2676
SIVSTM	---A-T---C---C---T---G-CC-G---C---	2733
SIVSMNLB1	---A-T---C---A-G---C---G-G-CC---	948
HIVPOLC12	---G---T-C---T---A---T---G---	664

CONSENSUS .A	AgAATtAGAGaACagGCAAAACaTaGAAACaATAGTAcTaATGGCagTTCATTCATGAATTTTAAAGaA	2868
HIV2ROD	-----C-----A-----	2791
HIV2NIH2	-----G-----G-----T-----	2788
HIV2ISY	-----G-----G-----	2788
HIV2ST	-----G-----C-----G-----G-----	2848
HIV2BEN	-----G-----G-----	2848
HIV2CAM2	-----G-----	2788
HIV2D194	-----	2902
HIV2GH1	-----C-----	2848
HIVFA	-----G-----C-----	708
HIV25132P1	-A-----A-----	708
CONSENSUS .B	AGaaTcAG?GAtCagGCAGTATCAaTAGAGACAGTGT?cTAATGGCAaCTCACTGCATGAATTTTAAAGAA	2778
HIV2D205	---C---A---C---A---A---A---A---A---A---G---A---G---	2848
HIV2UC1	---G---C---A---A---A---A---G---	2851
HIV2EHOA	---T---A---A---A---GT---	2839
HIV7312AI	---G---A---A---A---G---	708
HIVJA	---G---A---A---G---	708
HIVON	---G---C---A---A---	708
CONSENSUS .C	AGAATCAGAGAGCAGGCAAAACACAATGGAGACCATAGTCTTGAT	708
HIV2POLB7	-----	708
CONSENSUS .D	AgAATtAGgGAaCAaGCAAAAtCAGTaGaaACcaTAGTAcTaATgGCagTTCATTCaATgAATTTTAAaAGAA	2775
SIVMM251	-----C-----T-----	2851
SIVMM142	-----C-----T-----	2851
SIVMM239	-----C-----T-----	2863
SIVMM32H	-----C-----T-----	2863
SIVMNE	-----C-----A-G-----T-----	2851
SIVSMH4	-----A-----C-----A---G---T-----T-----	2740
SIVSMH9	-----A---G-----A---T-----A-----C-----	2749
SIVSMH9BJ	-----A---G-----A---T-----	2749
SIVSTM	-----A---T-----CA---TG---C-----C-----G-A---	2806
SIVSMMLIB1	-A-----T-----A-----C-----AC-----	1021
HIVPOLC12	-A-----A---C---G-----C---	708
CONSENSUS .A	GGGGAAGGAATAGGGGATATGACcCCAGcAGAAAGACTAaTCAATATGaTcACCaCAGAACAaGAaATACAATT	2941
HIV2ROD	---G-----T---T---T---G---	2864
HIV2NIH2	-----A-----	2861
HIV2ISY	-----T---A---G---	2861
HIV2ST	-----G---TG---G---	2921
HIV2BEN	-----C---T---	2921
HIV2CAM2	-----T---T-----G---	2861
HIV2D194	-----T---C---	2975
HIV2GH1	-----C---T---	2921
CONSENSUS .B	GGGGAGGAaTAGGGGATATGAACCCCTGCAGaaAGAcTAGTtAACATGATAACcACAGaaCAAGAAATACAgTT	2851
HIV2D205	-----G---	2921
HIV2UC1	-----G---	2924
HIV2EHOA	-----A---C-----T-----A---	2912
CONSENSUS .D	GGGGAGGAaTAGGGGATATGACTcCAGCAGAAAGATTaGTTAA7ATGATCAC?ACAGAAACAAGAAATACAATT	2846
SIVMM251	-----A---C-----T-----	2924
SIVMM142	-----A---C-----T-----	2924
SIVMM239	-----A---C-----T-----G---	2936
SIVMM32H	-----A---C-----T-----	2936
SIVMNE	-----C---C-----T-----	2924
SIVSMH4	-----C---T-----C-----	2813
SIVSMH9	-----C---T-----T---C-----	2822
SIVSMH9BJ	-----C---T-----C-----	2822
SIVSTM	-----C-----C-----T-----C-----	2879
SIVSMMLIB1	-----G---C---T-----C-----	1094

HIV2/SIV POL

		5' 3' sj (rare)	
CONSENSUS.A	CCTCCAagcaAAAAATTCAAAATTaaAAaATTTcGgGTCTATTcCAGAGAAGGCAGAGATCagcTGTGGaAA		3014
HIV2ROD	-----C-----G-----T-----		2937
HIV2NIH2	-----		2934
HIV2ISY	-----A---A---C---		2934
HIV2ST	-----C-----		2994
HIV2BEN	-----AG-----T---T---C-A-----A-----		2994
HIV2CAM2	-----C-----P-----		2934
HIV2D194	-----AG-----T---T---A---C-A-----A-----		3043
HIV2GH1	-----AG-----T---T---C---C-A-----A-----		2994
CONSENSUS.B	CcTCCAagCAAAAAATTTAAAAATTTcAAAAATTTCCaGGTCTATTACAGAGAAGGCAGAGATCAACCTCTGGAAg		2924
HIV2D205	-----		2994
HIV2UC1	-----A-----		2997
HIV2EHOA	-C-----A-----C-----G-----		2985
CONSENSUS.D	?CAaCAATCAAAAAATTCAAAATTTAAAAATTTTCGGGTCTATTACAGAGAAGGCAGAGATCA?CTgTGGAA?		2915
SIVMM251	T-----C-----A-----G-----		2997
SIVMM142	T-----C-----G-----G-----		2997
SIVMM239	T-----C-----A-----G-----		3009
SIVMM32H	T-----C-----A-----G-----		3009
SIVMNE	C-----C-----G-----A-----		2997
SIVSMNH4	C-----T-----C---G-----A-----		2886
SIVSMNH9	C-----T-----C---G-----A-----		2895
SIVSMNPBJ	C-----T-----C---G-----A-----		2895
SIVSTM	T-----T-----C-----G-----A---A-----G-----		2952
SIVSMMLIB1	C---G-----T-----A---C-----A-----		1167
		5' 5' sj (rare)	
CONSENSUS.A	GGACCTGGGGAaCTaCTGTGGAAgGGaGacGGAGCAGTCaTAGTCAAGGTAGGgaCaGacATAAAgTaaTAC		3087
HIV2ROD	-----A---A---C-----A---A-----		3010
HIV2NIH2	-----G-----A---G-----A-----G-----		3007
HIV2ISY	-----G-----A---G-----T-----		3007
HIV2ST	-----T-----A-----G---T-----A-----		3067
HIV2BEN	-----T-----A-----G-----		3067
HIV2CAM2	-----A---G-----A---T-----A-----		3007
HIV2D194	-----A---G-----G---G-----G-----		3121
HIV2GH1	-----A-----G-----		3067
CONSENSUS.B	GGACCTGGTGA?CTATTGTGGAAAGGGGAAGGAGCAGTCaTCATAAAGGTAGGgACAGAAATCAAAAGTAaTAC		2996
HIV2D205	-----A-----G-----		3067
HIV2UC1	-----G-----C-----G-----		3070
HIV2EHOA	-----T-----		3058
CONSENSUS.D	GGACCCGGTGAGCTaTTGTgGAAAGGGGAAGGAGCAGTCaTCcTAAAGGTAGGgACAGAcATcAAgGTAGTAC		2988
SIVMM251	-----A-----		3070
SIVMM142	-----A-----		3070
SIVMM239	-----A-----		3082
SIVMM32H	-----A-----		3082
SIVMNE	-----T-----G-----		3070
SIVSMNH4	-----C-----G---C-----		2959
SIVSMNH9	-----C-----G---C-----		2968
SIVSMNPBJ	-----C-----G---C-----		2968
SIVSTM	-----G-----A-----		3025
SIVSMMLIB1	-----G-----		1186
		vif cds ->	
CONSENSUS.A	CAAGaAGgAA?GCCaAGATcATcagaGACTaTGGAGgAAAGgCAaGAg?TGGATAGTggttcCcaactGGagGg		3158
HIV2ROD	-----A-----A-----A-----TT-----		3083
HIV2NIH2	-----A-----T-A-----A-----A-----TT-----		3080
HIV2ISY	-----A---G-----CC-----A-----C-----		3080
HIV2ST	-----A---T-----A-----A-----C-----A---T-----		3140
HIV2BEN	-----G-----T-----G-----A---G---AC-----A---C-----		3140
HIV2CAM2	-----G---A-----C-----CA-----		3080
HIV2D194	-----G-----T-----G-----AC-----A-----		3194
HIV2GH1	-----G-----T-----AC-----TCCCA--TGGA--GT-C		3140
CONSENSUS.B	CCAGaAGaAAAGCAAAgATcATAAGGcACTaTGGAGGgGGAAAAGaATTGGATTGcAGTgCCGACaTGGAGGA		3069
HIV2D205	-----G-----A---T-----G-----G-----		3140
HIV2UC1	-----G-----G-----G---A-----		3143
HIV2EHOA	-----A---A-----C-----G-----		3131
CONSENSUS.D	CcAGgAGaAAgGCTAAaATcATCAAAAGATcATOGAGGAGGgAAAGAGgTGGATAGTcAGTCCcACaTGGAGGA		3061
SIVMM251	-----A-----C-----		3143
SIVMM142	-----A-----A-----		3143
SIVMM239	-----A-----C-----		3155
SIVMM32H	-----A-----G-----C-----		3155
SIVMNE	-----A-----G-----C-----		3143
SIVSMNH4	-A---G---A-----C-----AT-----G-----T-----		3032
SIVSMNH9	-A---G---A-----C-----AT-----G-----T-----		3041
SIVSMNPBJ	-A---G---A-----C-----AT-----G-----T-----		3041
SIVSTM	-A-----C-----G-----C-----		3098

CONSENSUS.A	tgccagGGAGgATGGAGA??TGGCA?GCCCTTGTCAGTACCTGAAATACAGAACAAAAGACCTAGAGSAGGT	3228
HIV2ROD	-----AA-----	3108
HIV2NIH2	-----GA-----	3105
HIV2ISY	-----AA-----	3105
HIV2ST	-----GG-----	3165
HIV2BEN	-----AA-----T-----	3213
HIV2CAM2	-----A-----AG-----	3105
HIV2D194	-----GG-----	3219
HIV2GH1	CAGGGA-----GG-----	3165
CONSENSUS.B	TACCAGGCAGGCTAGAGAgATGGCACAGTCTgAT?AAGTACCTGAAGCA	3117
HIV2D205	-----	3174
HIV2UC1	-----A-----G-C-----	3192
HIV2EHOA	-----T-----G-----A--T--	3168
CONSENSUS.D	TACCgGaGAGGCTaGAGAgGTGGCA	3086
SIVMM251	-----	3168
SIVMM142	-----	3168
SIVMM239	-----	3180
SIVMM32H	-----	3180
SIVMNE	-----	3168
SIVSMH4	-----	3057
SIVSMH9	-----G-----A-----	3066
SIVSMMPBJ	-----G-----A-----	3066
SIVSTM	-----A-G-----G-----	3123
CONSENSUS.A	GCGCTATGTTCCCCACCACAAGGTGGGATGGGCATGGTGGACTTGCAGCAGGGTAATATTCCCACTACAAGGA	3301
HIV2BEN	-----	3286
CONSENSUS.A	AAAAGTCATCTAGAAATACAGGCATATTGGAACCTAACACCAGAAAAAGSATGGCTCTCCTCTCATGCAGTAA	3374
HIV2BEN	-----	3359
CONSENSUS.A	GATTAACCTGGTATACAGAAAAATTCTGGACAGATGTTACCCAGACTGTGCAGACATCCTAATACA	3441
HIV2BEN	-----	3426

HIV2/SIV ENV

CONSENSUS.A	ATG??g?gtggtaggaatCagCTgcttaTTgccattTactAGCtAGtgctTgCtTaataAtatTGC????????	63
HIV2ROD	---ATG-----T-----G-----ACC	60
HIV2NIH2	---AA-G-A-A-A-G-T-G-----A-C-C-----AAG	69
HIV2ISY	---A---AA-T-----G-T-C-G-A-----ACC	66
HIV2ST	---T-----A-AT-G-----GC-G-----GTC	66
HIV2BEN	---GA-CC-----T-G-T-----A-A-----G-----TAGC	69
HIV2CAM2	---GA-CC---C-A-----C-----AGACAA	72
HIV2D194	---GA-CC-----G-T-A-----AAA	69
HIV2GH1	---T-----A-G-TA-ATG-G-----GC-G-----A-G-----ACC	66
HIV2D1024	---G-GC-----A-----A-----C-----AG	59
HIV2D766	---ACAC-A-A-GC-----T-----A-----AA	68
HIV2D808	---GC-CCCCAA-A-C-----T-----A-----T-----GA	68
HIV2D868	---GC-C-AA-C-----AA-C-----G-----CA	68
HIV2CBL21	---T-A-----TA-A-G-----GC-G-----AT-----AC	65
HIV2CBL22	---A-----AA-T-----G-----G-A-----G-----AC	65
HIV2CBL23	---AT-G-----A-----G-----G-----A-----G-----AC	68
HIV2CBL24	---T---C-A-----A-----G-A-T-C-C-----C-C-----G-----C-----AC	65
HIV2CAM1	---GC-TA-AAA-----A-----A-----AAACGCA	74
HIV2CAM2BR	---GG-T---C-A-----C-----AAACA	71
HIV2CAM3	---AC-C-CAAAG-----T-----C-A-----AOGCA	71
HIV2CAM4	---GC-T-----AA-GC-----C-----G-A-A-CA-----C-T-----GC	68
HIV2CAM5	---AC-T-----A-----G-A-A-----A-G-C-----AA	68
HIV2CAM6	---GGTC-AAA-----A-----A-----A-----C-----AA	68
HIV2ALI	---AT-TC-A-----A-----G-A-T-C-----G-----T-----AA	68
CONSENSUS.B	ATGGGCacAt?TAgcaactcaCCTgctTaTT?cgCTCCTgCTTATaAGTaTCLaTGGGTtT?TggGtaA??????	65
HIV2UC1	---CAC-----T-----T-----A-----G-----C-----C-TAAGAA	74
HIV2D205	---T-T-----GC-G-----C-----G-----G-AG-----G-AT-----A	69
HIV2EHOA	---G-A-T-----A-G-A-A-----A-A-----C-G-----G	69
CONSENSUS.D	ATGGgaTGTcTtGGgAATCAgCTGCTTATCGCcaTCTtGcTtLAAGTGtcTa?ggGAtcTaTTGtattT	68
SIVMM251	-----T-----C-----C-----	69
SIVMM32H	-----C-----C-----	69
SIVMM1A11	-----T-----C-----	69
SIVMM142	---T-----A-----C-----T-----T-----C-----	69
SIVMM239	-----T-----C-----C-----C-----	69
SIVMNE	-----T-C-----C-T-----C-----	69
SIVSMH4	-----GC-----AG-----T-TA-A-----T-G-----G-----	69
SIVSMH9	-----GC-C-A-Y-----CT-CA-----T-----G-----	69
SIVSM4PB3	-----GC-----A-----CT-CA-----T-----G-----	69
SIVP209C19	-----GC-----AG-----CT-TAAA-----T-----G-----	69
SIVSTM	---CC-C-C-----A-----A-----T-----A-----CT-GCTT-----C-----C-----	69
CONSENSUS.A	?cAatataGTgACTGTTTctTAtGgcaTaCCcgcgTGGAgaaATGCAtCcATtCCctTctTtTGTGCAACCaaA	136
HIV2ROD	---A-----G-----A-----A-----G-----G-----	133
HIV2NIH2	---T-----C-----G-----G-----	142
HIV2ISY	---A-----AG-----T-----A-----T-----T-----	139
HIV2ST	---G-----G-----P-----T-----T-----	139
HIV2BEN	---G-----A-----T-----T-A-----T-----	142
HIV2CAM2	---G-----A-----T-----T-----	145
HIV2D194	---G-----T-----A-----	142
HIV2GH1	---G-G-----T-----T-----	139
HIV2D1024	A---C-A-----C-----G-----G-----G-----G-----	133
HIV2D766	A---A-C-A-----C-----G-----C-----	142
HIV2D808	A---A-----C-----G-----G-----	142
HIV2D868	A---A-----AG-----	142
HIV2CBL21	C-G-----G-G-----A-----C-----T-----	139
HIV2CBL22	C-----AG-----G-----C-----	139
HIV2CBL23	CA-T-----G-----	142
HIV2CBL24	CA-----G-----	139
HIV2CAM1	A-----A-----G-----T-----	148
HIV2CAM2BR	A-----G-----A-----T-----	145
HIV2CAM3	A-G-----A-----T-----A-----	145
HIV2CAM4	C-----A-A-----G-----T-----T-----	142
HIV2CAM5	G-----A-----G-----T-----C-----	142
HIV2CAM6	A-----A-A-----T-----G-G-----A-----G-----T-----	142
HIV2ALI	A-----T-----G-G-----A-----A-----	142
CONSENSUS.B	?aAcTaTGTcACTGTCTTCTATGGcATACCCGCATGGAggaAACGCAaCagTTCCTCTCTtTtTGTGC8ACCACAA	138
HIV2UC1	A-----T-----T-----C-----	148
HIV2D205	---C-A-----T-----A-----	142
HIV2EHOA	---T-----T-----AA-T-----T-A-----T-----G-----	142
CONSENSUS.D	?CaATATGTcACaGTcTtTtTATGGTgTACCAGctTGGAggaATGCACaaTTCCCTCTCTtTtTGTGC8AACCaagA	141
SIVMM251	-----G-----G-----	142
SIVMM32H	-----G-----	142
SIVMM1A11	---T-----G-----	142
SIVMM142	-----T-----T-----	142
SIVMM239	---T-----T-----	142
SIVMNE	-----T-----G-----	142
SIVSMH4	---A-----A-C-----A-----A-----	142
SIVSMH9	---A-----A-C-----A-----G-----CA	142
SIVSM4PB3	---A-----A-C-----A-----A-----G-----	142
SIVP209C19	---A-----A-C-----A-----A-----G-----	142
SIVSTM	---G-----G-A-----A-----A-----G-----	142

CONSENSUS .A	AtAGaGATACTTGGGGaACCATaCAGTGCTtgCCaGACAATGATGATATaCAGGAaaTA7cttTaAATGTgACA	209
HIV2ROD	-----G-----A-----T-----A-----G-----A-----	207
HIV2NIH2	-----A-----A-----A-----A-----C-----A-----	216
HIV2ISY	-----C-----C-----C-----A-----G-----C-----A-----	213
HIV2ST	-----C-----G-----A-----A-----G-----G-----	213
HIV2BEN	-----C-----G-----T-----A-----C-----G-----AT-----	216
HIV2CAM2	-----G-----G-----T-----A-----C-----G-----C-----	219
HIV2D194	-----G-----C-----C-----A-----A-----A-----C-----	216
HIV2GM1	-C-----A-----AA-----A-----A-----A-----C-----G-----	213
HIV2D1024	-----G-----A-----C-----G-----C-----G-----	207
HIV2D766	-----G-----C-----A-----G-----	216
HIV2D808	-----G-----C-----G-----G-----G-----G-----	216
HIV2D868	-----G-----G-----G-----G-----G-----G-----	216
HIV2CBL21	-----G-----G-----G-----G-----G-----C-----	213
HIV2CBL22	-----A-----GC-----A-----A-----	213
HIV2CBL23	-----G-----G-----A-----G-----	216
HIV2CBL24	-----A-----G-----G-----	213
HIV2CAM1	-----G-----C-----C-----C-----	222
HIV2CAM3BR	-----G-----T-----A-----C-----G-----C-----	219
HIV2CAM3	-----C-----G-----C-----A-----T-----C-----G-----	219
HIV2CAM4	-C-----G-----C-----A-----C-----C-----	216
HIV2CAM5	-----G-----G-----T-----	216
HIV2CAM6	-----C-----C-----G-----	216
HIV2ALI	-----A-----G-----	216
CONSENSUS .B	ACAGaGATACCTGGGGAACTGTACAgTGcCTCCAGACAAATGgtGAcTAcACoGAGATCag7?TAAATATAACA	210
HIV2UC1	-----G-----T-----TG-----	222
HIV2D205	-----C-----T-----T-----T-----GC-----C-----	216
HIV2EHOA	-----A-----T-----A-----T-----A-----CAAT-----	216
CONSENSUS .D	ATAGGGATACTTGGGGAACTGAGTGcCTaCCAGATAAATggTGATTATTCAGAAATGGCCctTAATgtTACA	215
SIVMM251	-----C-----	216
SIVMM32H	-----C-----	216
SIVMM1A11	-----A-----	216
SIVMM142	-----A-----	216
SIVMM239	-----G-----	216
SIVMM6	-----A-----A-----	216
SIVSMH4	-----C-----A-----A-----T-----G-----A-----C-----AA-----C-----C-----	216
SIVSMH9	-----C-----A-----A-----T-----G-----C-----AA-----C-----C-----	216
SIVSMMPBJ	-----C-----A-----A-----T-----G-----C-----AA-----C-----C-----	216
SIVP209C19	-----C-----A-----A-----T-----G-----AA-----C-----AG-----C-----A-----C-----	216
SIVSTM	-----A-----T-----C-----C-----AA-----A-----C-----	216
CONSENSUS .A	GAGGCTTTGATGCATGGaAT??AATACAgTaACAGAAcAAGCAaTaGAAGATGTCTGGaaLCTaTTTGAGAC	280
HIV2ROD	-----T-----C-----C-----C-----	278
HIV2NIH2	-----T-----C-----G-----T-----G-----G-----	287
HIV2ISY	-----T-----C-----G-----T-----G-----G-----	284
HIV2ST	-----C-----C-----G-----G-----G-----G-----	284
HIV2BEN	-----T-----G-----G-----C-----	287
HIV2CAM2	-----G-----A-----G-----G-----C-----	290
HIV2D194	-----A-----T-----G-----G-----GA-----G-----	287
HIV2GM1	-----G-----C-----G-----G-----	284
HIV2D1024	-----T-----C-----G-----G-----	278
HIV2D766	-----A-----C-----	287
HIV2D808	-----C-----C-----	287
HIV2D868	-----C-----C-----	287
HIV2CBL21	-----C-----T-----C-----G-----G-----G-----G-----	284
HIV2CBL22	-----C-----T-----G-----C-----C-----	284
HIV2CBL23	-----C-----G-----G-----G-----G-----	287
HIV2CBL24	-----C-----G-----G-----G-----G-----	284
HIV2CAM1	-----G-----A-----G-----G-----C-----	293
HIV2CAM2BR	-----T-----A-----G-----G-----C-----	290
HIV2CAM3	-----T-----A-----C-----G-----G-----C-----	287
HIV2CAM4	-----T-----C-----AA-----C-----G-----G-----C-----	290
HIV2CAM5	-----T-----G-----G-----G-----G-----	287
HIV2CAM6	-----G-----G-----G-----GA-----	287
HIV2ALI	-----G-----G-----	287
CONSENSUS .B	GAGGCTTTTGATGCATGGaAT...AATACAGTgACAgAACAGGCAgtAgAtGATGTGTGGAGLCTcTTTGAAAC	281
HIV2UC1	-----A-----A-----C-----T-----	293
HIV2D205	-----C-----A-----A-----G-----A-----	287
HIV2EHOA	-----T-----AC-----A-----G-----G-----	287
CONSENSUS .D	GAAagcTTTGATGCTTGGaA?..AATACAGTCACAGAACAGGCAATAGAGGAtGTaTGGcA7CTcTTTGAGAC	284
SIVMM251	-----G-----C-----A-----	287
SIVMM32H	-----G-----A-----A-----	287
SIVMM1A11	-----G-----A-----A-----	287
SIVMM142	-----G-----A-----C-----A-----	287
SIVMM239	-----C-----A-----T-----A-----	287
SIVMM6	-----G-----T-----T-----	287
SIVSMH4	-----GGCT-----T-----A-----G-----A-----C-----A-----	287
SIVSMH9	-----GGCT-----T-----A-----A-----C-----A-----A-----	287
SIVSMMPBJ	-----GGCT-----T-----A-----A-----C-----A-----A-----	287
SIVP209C19	-----GGCT-----A-----T-----A-----G-----A-----C-----A-----	287
SIVSTM	-----GGCT-----C-----C-----A-----C-----A-----C-----T-----A-----	287

HIV/SIV ENV

CONSENSUS.A	ATCAATAAAACCATGTGTCAAAATaACaCCtTaTGTgTagcaATGaa?TGT????a??aa?a??????a	336
HIV2ROD	-----C-----A-----CAGCAGC-CAG-G-GCAGCACAG	352
HIV2NIRZ	-----C-----C-----ACAAGG	345
HIV2ISY	-----G-----A-----C-----	336
HIV2ST	-----C-----C-----C-----CGT-----	336
HIV2BEN	-----GC-----C-----G-----T-----AGCAGGGTTC-AGGG	354
HIV2CAM2	-----A-----	342
HIV2D194	-----G-----G-----G-----CC-----G-----G-----T-----	339
HIV2GH1	-----C-----C-----C-----C-----GT-----	336
HIV2D1024	-----C-----C-----C-----CA-----C-----GG--A-GCAGTGGC-	346
HIV2D1071	-----C-----C-----C-----CA-----C-----GC--A-ACAGTGGC-	25
HIV2D766	-----C-----C-----C-----T-----C-----GC--A-ACAGTGGC-	355
HIV2D808	-----C-----C-----C-----T-----CAGCAAA-ACT-T-ATAACACA-	361
HIV2D868	-----C-----C-----C-----T-----C-----GC--GGACAGTACG-	355
HIV2CBL21	-----C-----C-----C-----G-----A-----T-----CA-CTATGAAA-	349
HIV2CBL22	-----G-----G-----C-----C-----CC--A-CAAGGAGG-	352
HIV2CBL23	-----G-----G-----C-----G-----QA-----AC--C-CAGATGCA-	355
HIV2CBL24	-----C-----C-----A-----AC--T-CA-----	345
HIV2CAM1	-----C-----A-----GATGGCAAC-	355
HIV2CAM2BR	-----A-----AACATA-	349
HIV2CAM3	-----G-----G-----C-----A-----AC--C-CACAGAGT-	358
HIV2CAM4	-----G-----GC-----G-----A-----C-----AG-GG-TGACAGGG-	355
HIV2CAM5	-----G-----G-----C-----A-----AA-----G-----A-----A-AATTAGGG-	355
HIV2CAM6	-----G-----C-----T-----AGCAGATCC-GACCCAGT...	357
HIV2ALI	-----G-----A-----G-----AGCAAC-TA-GC-CAGAGAGC-	361
CONSENSUS.B	CTCcaTAAaACCATGTGTAAaCTaACcCCaCTGTGTGTGgCAATGaa?TGFAaTAA?AC?ggaAcc??????	345
HIV2UC1	-----G-----G-----G-----G-----C-----GA-----C-----A-----	360
HIV2D205	-----C-----C-----C-----G-----A-----C-----A-----AAACCCAG	361
HIV2EHOA	-----A-----T-----A-----G-----G-----GT-G-G-TCA....	357
CONSENSUS.D	cTCAaTaAaGcCtTGTGTAAaATaAcCCcAtTATGcATaCtATGagaTgc	336
SIVMH251	-----T-----	339
SIVMH32H	-----T-----	339
SIVMH1A11	-----T-----	339
SIVMH142	-----T-----	339
SIVMH239	-----T-----	339
SIVMNE	-----A-----	339
SIVSMH4	A--C--T--C--C--C--C--T--AG-A--T--	339
SIVSMH9	A--C--T--C--C--C--C--T--A--A--G--T--	339
SIVSMNPBJ	A--C--T--C--C--C--C--T--A--A--G--T--	339
SIVP209C19	A--C--T--C--C--C--C--T--AG-A--T--	339
SIVHMB670	-----	9
SIVSTM	A--C--T--A--G--T--C--T--G--A--G--T--	339
CONSENSUS.A	??a??a??aca??	341
HIV2ROD	GG-AC-AC-----ACCTCAAGAGGC	375
HIV2NIRZ	..-AC-TG-C-----ACATGG	360
HIV2ISY	..-ACCA-GT-----ACAGAGAGC	354
HIV2ST	..-AC-GC-----ACTGCAAAAAC	357
HIV2BEN	..-AT-CC-G-----ACCCGGAATCCC	381
HIV2CAM2	..-AC-TA-GT-----ACAGTGACACA	363
HIV2D194	..-AT-TA-T-----TCAGGG	354
HIV2GH1	..-AC-GC-C-----ACTAACAAC	354
HIV2D1024	AC-CA-GT-----GGGAACGCA	366
HIV2D1071	AC-CA-GT-T-----	36
HIV2D766	AC-CA-GT-----	366
HIV2D808	GTGCAGGG-ACACAACCACACCCACAGCCGGAATCCATG	399
HIV2D868	AC-CA-GT-----	366
HIV2CBL21	AC-TA-GC-----	360
HIV2CBL22	AC-TG-CT-T-----	363
HIV2CBL23	GG-AC-CA--C-----	366
HIV2CBL24	..GCC-CA-----ACCCAGAGGGG	363
HIV2CAM1	TA-CA-GT-----	366
HIV2CAM2BR	GT-CA-GTGAC-----ACA	363
HIV2CAM3	GT-AC-CAG-T-----	369
HIV2CAM4	AC-CA-CC-----	366
HIV2CAM5	ACGCA-CC-----	366
HIV2CAM6	..-AC-GC-----GGAAATGTAAACC	378
HIV2ALI	CA-CC-CAT-C-----	372
CONSENSUS.B	??aat?c?a?b.....acTA??	355
HIV2UC1	..---A-T-C-----AA	375
HIV2D205	GG---G-C-G-----G--CT	378
HIV2EHOA	-----G--GC	363
CONSENSUS.D	..AAtAAaAgTGAgaCAGatAgatGGGgATtgACaAaatca?ca?????????????????????acaacAaca	386
SIVMH251	..-AC-----A-----T-----	390
SIVMH32H	..-AC-----A-----T-----	390
SIVMH1A11	..-AC-----A-----T-----C	390
SIVMH142	..-AC-----A-----T-----	390
SIVMH239	..-AC-----A-----AT-----	390
SIVMNE	..-C-----A-----T-----	390
SIVSMH4	..-C-----C-----G-----T-----GG-AAAG--GGG--	393
SIVSMH9	..-C-----M--MC-----T--A-----GGAGSACCAGCACCA--	396
SIVSMNPBJ	..-C-----C-----T--A-----GGAACACCAGCACCAACAACA--CA--	402
SIVP209C19	..-C-----G-----T-----AGAGCGG--GAG	384
SIVHMB670	..-G-CA--A--C-A--G-A--GGGAATG--ATA--	63
SIVSTM	..-G-A--A--C-A--GGGA--A--GT--	390

CONSENSUS.A	pc?????ac?ac?ac?ac?c??c?ac?c?????g?????t?acaaatgagactcctccatgcac	380
HIV2ROD	A-AAGCACA-C-A-C-C-CA-AGAC.....CAGGAGCAAGAG---G---GA-A-----GC	440
HIV2NIH2	A-AGGAGG--AGAC--CCA-----AATATCACAA-T-----C---C-AT...GC	413
HIV2ISY	G-AGTTCCA-T-A-GCC-T-T-----GACCTGATA-G-----T---GA-----	413
HIV2ST	A-AACCTCC-AC-A-A-CA-CA-A-AGCA.....AACACACA...-GGA---A---T-----	422
HIV2BEN	T-GAGTCC--A-CT-G-G-C-AC-C-ATCC.....CAGGCTCCA-A-----A---AAC-----	449
HIV2CAN2	A-PATGATA-GG--A-C--C-CAGC--AGCT.....AAGAGGACCA--G---C-A---T-----	431
HIV2D194	A-TACCGCG--CC-G-GTC--C-A-----AATATACAA-A--G---A-A---A-C-T--	413
HIV2GHI	A-AACCACA-AGGG-GC--A-A-----GGATGTCAGAG-----A-----CATA	413
HIV2D1024	A-CACATCC--AC-C-----ATGATAAT-AGACCTCTA-A-----C-----	425
HIV2D1071	GGA-GC--A-CA-AA-C-AGAC.....CAGCCATGG-G-----C---A-----	95
HIV2D766	GGG-GT--A-CA-AC-CG-A-----CCATATCCA-G-----C-----	422
HIV2D808	GGG-AC--AC-CA-AT-C-A-----CCTCTTATG-T-----C---T-----	422
HIV2D868	TG-G-GAC-CA-AA-CGGATCTCGAATACAGAAAGTA-AC---A---A---T-----	425
HIV2CBL21	G-CCAG-GCATAA-C-ACTGCAAG-CAGTAGTGA-A-ATG--ACTGA--	422
HIV2CBL22	A-ACCACA--AG-AT-CC-GCCTA-A-TAAACCCGTGACAGAG.....-G---A---CT-----	434
HIV2CBL23	GAAGCAGAG-C-GT-C--A-TG-TG-TACAGCAACA-GGGCAATGGAG-----G-----	437
HIV2CBL24	GGG-AC--A-C--T-CA-T-GAGCAGCATCATCCGAGGAA--G---A---T-----	434
HIV2CAM1	A-TATGATA-GG--A-C--C-CAGC--AGCTAAGAG--ACCAATAAGTGAC--ATTTC--C---T--	437
HIV2CAM2BR	AGAGGCC-AG-CAATCCATG-AGAPAAATGAG-----A-----	428
HIV2CAM3	G-C-C-AG-CAGCA-A-CCTGTAAAT-ACACACCC-AC-----G-AG-A-----	431
HIV2CAM4	CTCTCAAT-AA--C-C--T-CT-T-CGCAATCTA-CGATAGATGAT-AT-C-.....	431
HIV2CAM5	A-AACATAC-CT-A-C--G-CG-A-CACCACACC-GGATGACAG-G-----	452
HIV2CAM6	C-GAGCC-AGGAGCAGACTCAAAACCC-G-----GCGA-----	425
HIV2ALI		
CONSENSUS.B	ac?a??actac?cctactaCaaC?cCtc?c??a?c?????c?cagAAC?tac?CAACGAaaacagcccatGtAT	411
HIV2UC1	C-T-TT--A--T--TA-AAACC-----T-AG-A--T-----C---T-----	443
HIV2D205	-C-CC--AG-----C--GT-----GGG-T--AACGA-T-----GA---C---	443
HIV2EHOA	-AAGAG--C--TA-GT-CT-G-G--C-TCAG-TCCTCTA-T--C-C-----GAT--AA---	437
CONSENSUS.D	gcaacacagca?caaca?Cagca?caa?c?ca?c??c?aaag?AaAcgt?gT?AATGagactagt?CtTGtAT	445
SIVMM251	A---G---C---T---C-GT-T---GAA---AT-G--A-G-C-----T-----	461
SIVMM32H	-C---C---GG-----T-----AT-G--A-G-C-----T-----	452
SIVMM1A11	A---G---C---AGG-ATGT-GA-----AT--A-GA-C-----T-----	452
SIVMM142	-T---A-A--A-A-G-A-T-GTAGAG-C-AG-G-A-A-C-----C---G---	464
SIVMM239	-T---A-T---GA---T-GC-----T-G-A-G-C-----T-----	455
SIVMMNE	-C---AT-C--A-AA-G-GAG.....GC-AT--A-G-C-----A---C-----	455
SIVSMH4	A-G--T-A-A--G-A-C--GTGT-----GC-A--T-TA-A--A-G-A-C---C---	461
SIVSMH9	CA---T-A-A-C-C-----T-CCAATA-T-C--G-T-A--C-G-GA-C---	461
SIVSMH9B	A---CA-G-T-A-A-C-C-T-CCAATA-C-C--G-T-A--C-G-GA-C---	476
SIVP209C19	A---A---C---CG-C-C-AGTGA--A--G-TA-A--AGG-GA-C---C---	458
SIVHUMB670	G-C-T--AG-C-C-AAAGTA-C-C--T-A--A---T---G---	137
SIVSTM	-TG--C-A-G-G-G--G-C-----GCC-G-GT-A-A--A-A-CT-C--G---	455
CONSENSUS.A	acgc?ca?ACaactGctCa??Ggatt?ggggagGagAgatggtCaatTGTcagTTcaatATGacaGGattAG	448
HIV2ROD	---G-G-----G-A-----A-C-A-----C-----	511
HIV2NIH2	G-TG-G-----A---AAA-----A-A-G-C-----T-G-----	484
HIV2ISY	-AATIGA--T-----C-GA-----C---G-G-----C-----	484
HIV2ST	-A-G-----A---G-G-A-A-----G-C-----G-----	493
HIV2BEN	-GAAAACA--CA-G-----A-T-T-G-----A-GC-A--G-----AG-G---	520
HIV2CAN2	-A-A-----A-A-A-A-A-A-----G-C-----C-----	502
HIV2D194	-G-GACA-----A-----A-A-A-G-----TG-G--G-----G-GC---	484
HIV2GHI	C..T-G-----A-----G-A-----A-----T-----	481
HIV2D1024	-A-A-----A-----GA-A-----G-----G-----	496
HIV2D1071	-G-A-----GA-A-----C-A--G-----G-----A---	166
HIV2D766	-A-G-----AA-A-----G-----G-----G-----A---	493
HIV2D808	-G-A-----G-AA-----A-----G-----G-----C---	493
HIV2D868	-A-A-G-C-----GA-A-----A--G-----G-----	493
HIV2CBL21	-A-A-G-C--TG-----A-A-C-----CA-G-----G-----	496
HIV2CBL22	-AAAA-A--TA-----C-G-A-----C-A-----G--G-----	493
HIV2CBL23	-G-A-----G-A-A-----G-----A-----	505
HIV2CBL24	-A-A-A--T-A-----A-A-C-----A-TC--G-----C---	508
HIV2CAM1	-A-TG-G-----C-----GAA-----TG-----	505
HIV2CAM2BR	-G-A-----A-A--A--A--A--GG--C-----	508
HIV2CAM3	G--A-G-----AAA-----G-----C-----	499
HIV2CAM4	-G-AG-A-----A-----G-AA--G--T-A-GG-G-----G--C---	502
HIV2CAM5	-AAAAATG--C-----A-A--T-----C-----T-----	502
HIV2CAM6	G--TG-G-----G-ACA-AG-----G-G-----G--G-----	523
HIV2ALI	-AAGG-G-----C-CAGG--C-A--T-----G-----C-----	499
CONSENSUS.B	taAAAATGACaCTGc?Ca...GGaATAGGactAGAagAaAaattTgAcTGT?AATTTAA?ATGACaGGaCTAA	477
HIV2UC1	-----C--G-----A-C-C-G-T-----T-C-----C-----	514
HIV2D205	A-----A-----C-----GA-G-----GC-A--A-T--GT--G---	514
HIV2EHOA	-C-----TG-----G-----G-A-----C-A-C-A-----T---	508
CONSENSUS.D	Aa7c?ataAatTGcaCA...GGCTTgGAacAaGagccaaTgaTAaGcTGTAATTCaAcATGACAGGgtTAA	514
SIVMM251	-GC-C-G-----A-----C-----	532
SIVMM32H	-C-C-G-----A-----G-----	523
SIVMM1A11	-GT-C-G-----A-----G-----	523
SIVMM142	-GT-C-G-----A-----	535
SIVMM239	-GCCG-GG-----A-----	526
SIVMMNE	-A-C-G-----A-----	526
SIVSMH4	-AAA--G--TG-----G-----C-----G-T-----T-----	532
SIVSMH9	-GAAG-----T-----G-----C--G--T-----T-----	532
SIVSMH9B	-AAT-----T-----G-----C--G--T-----T-----	547
SIVP209C19	-GAA-----G--TG-----G-----G-T-----T-----	529
SIVHUMB670	-AAA-C-----T-----A-G-G-A-----G-----T-----	208
SIVSTM	-G-A-C--C--C-----G-----AT--C-AG-G-T-----T-----AC---	526

HIV2/SIV ENV

CONSENSUS.A	agagaGAtaAgaaaaAaacagTATAaTGAaCaTGGTACtcaaaaGAtGTgGTLTgtGaa??????????????	507
HIV2ROD	-A-----GACA-----AAT	576
HIV2NIH2	-----C-----G-----C-----G-----C-----	543
HIV2ISY	--TT-----G-----C-----TCA-----GAT	549
HIV2ST	-G-----TA-----A-----C-----TCA-----	555
HIV2BEN	-ACA-----GGAG-----G-----C-----T-T-G-----C-----	579
HIV2CAM2	-A-----C-----G-----GC-----T-----	561
HIV2D194	-ACA-----GGA-----C-----CG-----G-----CAAG-----ACA	549
HIV2GH1	-----A-----TCA-----	543
HIV2D1024	C-----CC-----GG-----G-----GT-----C-----ACA-----GAAACA	564
HIV2D1071	C-----ACC-----AGA-----GC-----T-----CGCA-----CCAAC	573
HIV2D766	C-----CC-----AG-----GT-----ACAACAGGAGCAACT	567
HIV2D808	C-----ACCG-----TG-----CGC-----AAT-----	555
HIV2D868	CA-----CC-----AG-----GT-----GGCAACAGGAACAAAT	567
HIV2CBL21	-----C-----TCG-----	558
HIV2CBL22	T-----CC-----GC-----G-----C-----GC-----	552
HIV2CBL23	-----G-----A-----G-----G-----A-----	564
HIV2CBL24	-A-----TA-----G-----C-----AGCACACCAGGT	579
HIV2CAM1	-A-----C-----C-----CA-----CCAATGGA	573
HIV2CAM2BR	-A-----C-----C-----OC-----TAAT-----	570
HIV2CAM3	-----C-----G-----G-----C-----T-----C-----GCA-----GAT	564
HIV2CAM4	-----C-----GT-----GAGA-----A-----CACA-----	564
HIV2CAM5	CACCT-----CC-----AG-----C-----A-----T-----	561
HIV2CAM6	C-----CC-----C-----C-----AAGCG	588
HIV2ALI	-----CC-----	558
CONSENSUS.B	gAAGAGATGAGT?AAAACAATATAAGGACACCTGGTAc?aa?AAGATTTAGAGTGT?a?????????????g?	531
HIV2UC1	-----AA-----G-----A-----CA-----C-----G-----C-----	576
HIV2D205	-----CT-----A-----TA-----G-----	570
HIV2EHOA	A-----TC-----TA-----C-----T-----G-----AAAGGGGACAAAGGA-----T	582
CONSENSUS.D	AaAGaGACaa?Aa?AaaGagTACATGAAACCTGGTACtClacaGAtTggT?TGtGaaCaAa?????????ggg	576
SIVM251	-----G-----CA-----G-----T-----	597
SIVM32H	-----G-----CA-----G-----T-----	588
SIVM41A11	-----G-----A-----G-----A-----	588
SIVM4142	-----G-----A-----G-----T-----	600
SIVM4239	-----G-----A-----G-----A-----	591
SIVM4E	-----G-----GA-----G-----G-----T-----	591
SIVSMNH4	-T-----G-----A-----G-----A-----T-----A-----T-----A-----G-----AA-----C-----G-----GTCA-----C-----	600
SIVSMNH9	-----A-----G-----G-----A-----T-----A-----T-----A-----G-----A-----G-----AATAGCAAT-----A-----	606
SIVSMNPBJ	-----A-----G-----G-----A-----T-----A-----T-----A-----G-----A-----T-----AAT-----A-----C-----	615
SIVP209C19	-----G-----GA-----GG-----A-----T-----A-----T-----A-----G-----AA-----C-----G-----GTCAGCAAT-----	600
SIVHMB670	-----G-----A-----G-----A-----	228
SIVSTM	-----T-----A-----G-----G-----G-----G-----GC-----C-----AA-----C-----	588
CONSENSUS.A	aa?aa?caca????????????????????????????caga?ca?gTGTtacATGAacCATtGCAaACatCagT	550
HIV2ROD	-T-----G-----AAT-----C-----CA-----	626
HIV2NIH2	-T-----A-----C-----TCAAGT-----G-----A-----	596
HIV2ISY	-C-----G-----GAC-----GA-----AA-----GA-----	599
HIV2ST	-TGA-----C-----AAGAAA-----G-----AA-----CA-----C-----	608
HIV2BEN	-C-----CA-----GCTGG-----CA-----GA-----A-----A-----	626
HIV2CAM2	-T-----G-----GATCAGACCACAAAT-----G-----C-----C-----	623
HIV2D194	-CGGA-----GG-----CA-----GA-----	593
HIV2GH1	-T-----AT-----C-----AAAGATGGG-----A-----A-----GA-----	599
HIV2D1024	-T-----A-----C-----TCAAAC-----G-----G-----	617
HIV2D1071	-C-----C-----T-----AACCAGAGCCTCAAATAACACCTCAGAC-----G-----G-----G-----	308
HIV2D766	-C-----C-----T-----AAC-----GA-----G-----	617
HIV2D808	-----AACAACCCCAACC-----A-----A-----GA-----	605
HIV2D868	G-----C-----C-----TT-----AAC-----G-----G-----	617
HIV2CBL21	-GT-----CA-----GT-----GTGACGAGCCAAACAAATTTGGACAGG-----AA-----G-----	626
HIV2CBL22	TCAGGT-----AGTAACAACACAAAC-----G-----GA-----T-----G-----	614
HIV2CBL23	-----GGAAATGGCACCACACAGAT-----CA-----G-----	614
HIV2CBL24	-C-----G-----AAC-----G-----CGA-----	629
HIV2CAM1	-T-----G-----CACA-----A-----CCAG-----C-----A-----	623
HIV2CAM2BR	-----G-----GAT-----CAGACCACAAATG-----C-----C-----	629
HIV2CAM3	-T-----A-----C-----TCAAGCAAGAGTAACACCTCA-----AA-----G-----T-----	629
HIV2CAM4	-T-----A-----C-----AACACAAATAAGACTTGACAAT-----G-----C-----T-----	632
HIV2CAM5	-C-----G-----A-----CCAA-----CC-----A-----T-----	608
HIV2CAM6	-TGG-----C-----TCAG-----CCAG-----A-----	638
HIV2ALI	-----CCATTTAACACCACCACAAAC-----C-----G-----	617
CONSENSUS.B	AAtaa?AcC-----?????????a?ca?ca?ATGtTAtATAgAACcTGtAAtACatCAaT	574
HIV2UC1	-C-----GC-----G-----C-----T-----C-----G-----T-----C-----A-----G-----	623
HIV2D205	-T-----AGGAAGTAT-----C-----G-----G-----C-----C-----A-----	626
HIV2EHOA	-G-----A-----G-----A-----	623
CONSENSUS.D	AATagcact-----g??aaTGAaAGLAgATGtTAcATGaatCACTGEAAACACCTtCtGt	627
SIVM251	-----AT-----C-----C-----	650
SIVM32H	-----AT-----C-----C-----	641
SIVM41A11	-----GT-----	641
SIVM4142	-----GT-----T-----	653
SIVM4239	-A-----GT-----C-----	644
SIVM4E	-----AA-----	644
SIVSMNH4	-GAA-----G-----A-----C-----T-----CAG-----	647
SIVSMNH9	-GAA-----C-----A-----C-----T-----CAG-----	653
SIVSMNPBJ	-GAA-----C-----A-----C-----T-----CAG-----	662
SIVP209C19	-TGAG-----G-----A-----C-----G-----T-----CAG-----	650
SIVSTM	-GT-----GCG-----A-----G-----C-----GA-----T-----C-----T-----CAG-----	641

CONSENSUS .A	CATcAcAGagTCATGTGATcAAGCACTATTGgGATGctATgaggTTTAgATAcTGTgCACCcCCgGgtTttGcc?	623
HIV2ROD	-----A-----C-----A-----A-----C-----	700
HIV2NIH2	-----C-----A-----T-----	670
HIV2ISY	-----C-----A-----T-----	673
HIV2ST	-----C-----A-----T-----	682
HIV2BEN	-----A-----C-----C-----	700
HIV2CAM2	-----C-----A-----	697
HIV2D194	-----A-----C-----A-----	667
HIV2GH1	-----A-----A-----A-----	673
HIV2D1024	-----A-----A-----A-----T-----G-----T-----	691
HIV2D1071	-----A-----A-----A-----A-----T-----T-----A-----T-----	382
HIV2D766	-----A-----A-----A-----T-----G-----A-----TT-----	691
HIV2D808	-----A-----A-----A-----T-----A-----A-----T-----	679
HIV2D868	-----A-----A-----AC-----T-----T-----A-----T-----	691
HIV2CBL21	-----A-----C-----A-----	700
HIV2CBL22	-----A-----G-----C-----A-----	688
HIV2CBL23	-----C-----C-----A-----	688
HIV2CBL24	-----C-----C-----C-----	703
HIV2CAM1	-----A-----AG-----A-----	697
HIV2CAM2BR	-----A-----A-----	703
HIV2CAM3	-----A-----A-----A-----	703
HIV2CAM4	-----T-----A-----A-----TA-----	706
HIV2CAM5	-----G-----C-----A-----C-----	682
HIV2CAM6	-----G-----C-----A-----T-----AC-----T-----	712
HIV2ALI	-----A-----A-----T-----AC-----T-----	691
CONSENSUS .B	tATCCAGAgTCATGTGATcAAACATTATGGGAcAGcTTAAGaTTTAGGTATGTGTCTCCCCAGGAtTgCTc	648
HIV2UC1	G-----T-----C-----C-----A-----	697
HIV2D205	-----G-----G-----G-----G-----G-----TT-----	700
HIV2EHOA	-----A-----T-----T-----T-----	697
CONSENSUS .D	tATCCAGAgTcTGTGACAAcATTATTGGGATgCctAtcAGaTtctAGgTATGTGTGACCtCCAGGTATAGCTT	701
SIVMM251	-----C-----A-----	724
SIVMM32H	-----C-----G-----G-----A-----	715
SIVMM1A11	-----C-----G-----	715
SIVMM142	-----C-----G-----G-----G-----A-----	727
SIVMM239	-----C-----G-----	718
SIVMNE	-----C-----G-----A-----C-----G-----	721
SIVSMNH4	-----C-----G-----Y-----A-----G-----	727
SIVSMNH9	-----C-----G-----A-----G-----	736
SIVSMMPBJ	-----G-----C-----G-----A-----C-----C-----	724
SIVP209C19	A-----C-----CT-A-----G-----A-----T-----	715
SIVSTM	-----A-----T-----	697
CONSENSUS .A	TgCTAAGaTgCAATGATAcCaATtATtCAGGCTTtgggCCcAAATGCTcCTAAAGTAgTAGCTgCtAcATGcACc	774
HIV2ROD	-----T-----CA-----C-----T-----	744
HIV2NIH2	-----A-----CA-----T-----C-----	747
HIV2ISY	-----G-----C-----A-----T-----G-----T-----T-----A-----	756
HIV2ST	-----AT-----T-----C-----A-----T-----G-----A-----T-----A-----	774
HIV2BEN	-----A-----G-----C-----CA-----C-----T-----G-----T-----A-----	771
HIV2CAM2	-----A-----C-----A-----T-----G-----T-----T-----A-----	741
HIV2D194	-----A-----C-----C-----A-----G-----T-----T-----A-----	747
HIV2GH1	-----A-----C-----C-----T-----C-----	765
HIV2D1024	-----AT-----C-----C-----T-----C-----	456
HIV2D1071	-----A-----AT-----C-----C-----T-----C-----	765
HIV2D766	-----A-----T-----C-----C-----C-----T-----	753
HIV2D808	-----A-----T-----C-----T-----C-----T-----	765
HIV2D868	-----A-----T-----C-----T-----C-----T-----	774
HIV2CBL21	-----A-----A-----TA-----T-----A-----	762
HIV2CBL22	-----A-----C-----G-----	762
HIV2CBL23	-----A-----TA-----A-----A-----A-----	777
HIV2CBL24	-----A-----T-----A-----A-----G-----A-----A-----	771
HIV2CAM1	-----A-----G-----C-----CA-----C-----T-----G-----T-----	777
HIV2CAM2BR	-----A-----G-----C-----CA-----C-----T-----G-----T-----	777
HIV2CAM3	-----A-----TG-----C-----A-----A-----T-----T-----T-----	780
HIV2CAM4	-----A-----A-----CA-----T-----T-----G-----T-----	756
HIV2CAM5	-----A-----A-----A-----A-----	786
HIV2CAM6	-----A-----CA-----A-----	765
HIV2ALI	-----A-----CA-----A-----	721
CONSENSUS .B	TgCTAAGATGtAATGATAcCaAAcTATTCAGGCTTcATGCCcAAcTGTAGTAAgGTAGTAGtTCTc?TgCaca	721
HIV2UC1	-----C-----T-----A-----A-----A-----T-----	771
HIV2D205	-----A-----C-----C-----C-----	774
HIV2EHOA	-----C-----T-----A-----A-----TCTG-----A-----	768
CONSENSUS .D	TGCTtAGaTGLAATGAcCaAATTATtCAGGCTTtAgCCTAA?TgTtCTAAGGTAgTgTgTCTCTCaTGcACa	774
SIVMM251	-----A-----A-----G-----	798
SIVMM32H	-----A-----A-----G-----	789
SIVMM1A11	-----A-----A-----G-----	789
SIVMM142	-----C-----C-----C-----	801
SIVMM239	-----A-----G-----G-----	792
SIVMNE	-----G-----TT-----T-----GCT-----C-----G-----	792
SIVSMNH4	-----G-----T-----C-----GCT-----C-----A-----T-----	795
SIVSMNH9	-----G-----TT-----T-----GCT-----C-----A-----T-----	801
SIVSMMPBJ	-----G-----TT-----T-----GCT-----C-----A-----T-----	810
SIVP209C19	-----G-----TT-----T-----GCT-----C-----A-----T-----	798
SIVSTM	-----C-----C-----T-----A-----GCA-----T-----C-----A-----C-----G-----T-----	789

HIV2/SIV ENV

CONSENSUS .A	AGGATGATGGAAACgCAaActTCLAC?TGGT?TTGGC?TTaaTGGCaCtAGagCAGAgAAAtAGaACaTataTcTa	770
HIV2ROD	-----C-A-----	848
HIV2NIH2	--A-----A-----A-----	818
HIV2ISY	--A-----T-C-----G-A-----	821
HIV2ST	-----C-C-T-----G-A-----	830
HIV2BEN	-----T-----A-----	848
HIV2CAM2	--A-----G-----G-----C-----	845
HIV2D194	-----T-----A-----	815
HIV2GH1	-----C-C-T-----G-----	821
HIV2D1024	-----G-----G-----T-----	839
HIV2D1071	-----G-----G-----G-----T-----	530
HIV2D766	-----G-----G-----TG-----G-----A-----T-T	839
HIV2D808	-----A-----G-----C-----C-T-----	827
HIV2D868	-----A-----A-----CG-T-----	839
HIV2CBL21	-----C-C-T-----G-----A-----G-----	848
HIV2CBL22	-----T-----A-----	836
HIV2CBL23	--A-----A-----A-----T-----	836
HIV2CBL24	-----T-----C-----GT-----T-----	851
HIV2CAM1	-----C-A-----	845
HIV2CAM2BR	--A-----G-----A-----C-----	851
HIV2CAM3	-----A-----A-----A-----	851
HIV2CAM4	--A--A-----T-----T-----G-----G-C-----	854
HIV2CAM5	--A-----G-----G-----A-----A-----	830
HIV2CAM6	-----A-----G-----A-----	860
HIV2ALI	-----G-T-----T-----G-----A-----	839
CONSENSUS .B	AGAATGATGGAAACACAGaCCTCTACATGGTTTGGCTTCAATGGTACaAGGgCAGAsAACAGGACATATA?TA	794
HIV2UC1	-----G-----G-----A-----T-----G-----	845
HIV2D205	-----T-----G-----A-----	848
HIV2ENOA	-----C-----	842
CONSENSUS .D	AGaATGATGGAgACaCagACTTCTACTTGGTTTgGcTTtAATGGaACTAGaGCAGAAAATAGAAC?TAtATLTA	847
SIVMM251	--G-----T-----	872
SIVMM32H	--G-----T-----	863
SIVMM1A11	--G-----T-----	863
SIVMM142	--G-----C-G-----C-----	875
SIVMM239	--G-----T-----	866
SIVMNE	-----T-----	866
SIVSMH4	-----G-A-C-----C-----T-----G-----A-C-----	869
SIVSMH9	--A-----A-C-----T-C-----T-----A-C-----	875
SIVSMMPBJ	--A-----A-C-----T-C-----T-----A-C-----	894
SIVP209C19	-----G-A-C-----C-----T-----G-----A-C-----	872
SIVSTM	-----A-----A-----A-----A-----C-----	863
CONSENSUS .A	TTGGCAtggtAgagataacAGaaCTaTcATcAGcttAaACaaatattatAatCTcactaT?cattTGTaaGAGGc	843
HIV2ROD	-----C-----T-----G-T-G-----	922
HIV2NIH2	-----A-----T-----T-C-T-----G-----	892
HIV2ISY	-----G-----T-----C-A-T-----G-A-----	895
HIV2ST	-----G-----T-----T-----G-T-----CG-A-----	904
HIV2BEN	-----C-----T-G-----T-----G-----A-G-G-----A-----	922
HIV2CAM2	-----A-----T-----T-----GC-----C-----T-----GT-----G-----	919
HIV2D194	-----AG-----T-G-----T-----G-----A-G-----A-----	889
HIV2GH1	-----C-----T-G-----G-----C-----T-----A-----A-----	895
HIV2D1024	-----A-----A-----A-----C-----T-----A-----	913
HIV2D1071	-----A-----A-----A-----C-----T-----AT-----G-----	504
HIV2D766	-----A-----A-----GG-----GG-----A-----GT-----C-----G-----T-----AT-----G-----	913
HIV2D808	-----A-----A-----A-----C-----C-----T-----AT-----	901
HIV2D868	-----A-----A-----A-----C-----T-----A-----	913
HIV2CBL21	-----G-----T-G-----T-----G-----A-----	922
HIV2CBL22	-----A-G-C-T-----T-----G-----C-----C-A-----	910
HIV2CBL23	-----G-----A-----	910
HIV2CBL24	-----C-----T-----T-----T-----A-----	925
HIV2CAM1	-----CA-----AG-----G-----G-G-----G-----	919
HIV2CAM2BR	-----A-----T-----G-----C-----T-----GT-----G-----	925
HIV2CAM3	-----C-A-----T-----C-C-----G-G-G-----A-----	925
HIV2CAM4	-----A-----AT-----AG-----C-----T-----G-----GC-----TC-----C-----G-----A-----	928
HIV2CAM5	-----A-----AT-----AG-----C-----T-----G-----GC-----TC-----C-----G-----A-----	904
HIV2CAM6	-----A-----A-----G-----T-----G-----	934
HIV2ALI	-----T-----C-G-----T-----G-----	913
CONSENSUS .B	TTGGCAtggtAAAGAcAATAGGACcATCATAAGCTTAAATcCATATATATAATcTgaCAATgCACTGtAaaaAGGC	868
HIV2UC1	-----A-----T-----G-----AG-----A-----T-----C-G-----	919
HIV2D205	-----AA-----T-----C-----T-----T-----A-----G-----	922
HIV2ENOA	-----T-----T-----	916
CONSENSUS .D	cTGGCAtgGcAgagataAATAGGACtAaATTAGcttAaAatAAGcATtATaATcTAAcAATgAaatGTAGAAAGAC	921
SIVMM251	-----T-----G-----	946
SIVMM32H	-----T-----G-----	937
SIVMM1A11	-----T-----G-----A-----A-----	937
SIVMM142	-----T-----C-----C-----	949
SIVMM239	-----T-----G-----	940
SIVMNE	-----A-----A-----G-----	940
SIVSMH4	T-----A-----A-AG-----A-----C-----C-----G-----	943
SIVSMH9	T-----R-----AGC-----A-----C-----C-----R-----G-----	949
SIVSMMPBJ	T-----AGC-----A-----C-----C-----G-----	958
SIVP209C19	T-----A-AG-----A-----C-----C-----C-----T-----GC-----	946
SIVSTM	T-----G-----G-----T-----C-----G-----C-----C-----T-----GC-----	937

CONSENSUS.A	CaGGaAaTAAgAcAGT?gtacCcaATAacacTtATGTcAGGattagtgTTTCacTCc77777CaGccaaTc777	907
HIV2ROD	-----G-----T-----GAA-----A-----TG-----CAT-----CACTAC-----G-----	993
HIV2NIH2	-G-----A-----GT-----T-----G-----TAA-----G-----ATC	960
HIV2ISY	-----A-----T-----C-----CCGCAGA-----AAG-----ATC	963
HIV2ST	-----C-----T-----C-----G-----	969
HIV2BEN	-----TT-----T-----	987
HIV2CAM2	-----G-----G-----GCA-AGA-----G-----ATT	987
HIV2D194	-----T-----T-----GCG-AG-----T-----G-----G-----TAC	957
HIV2GH1	-----C-----T-----C-----G-----	960
HIV2D1024	-----A-----A-----T-----	978
HIV2D1071	-----A-----A-----T-----	669
HIV2D766	-----A-----A-----G-----G-----G-----T-----	978
HIV2D808	-----A-----A-----T-----	966
HIV2D868	-----A-----G-----T-----	978
HIV2CBL21	-----C-----T-----C-----CACAG-----G-----G-----ATC	990
HIV2CBL22	-----A-----T-----C-----G-ACAAA-----G-----G-----ACTC	978
HIV2CBL23	-----G-----G-----AG-----GG-----ATC	978
HIV2CBL24	-----T-----G-----A-----T-----G-----	990
HIV2CAM1	-G-----A-----G-----	984
HIV2CAM2BR	-----G-----G-----GCA-AGA-----G-----ATT	993
HIV2CAM3	-----A-----A-----G-----G-----	990
HIV2CAM4	-----A-----T-----C-----G-----G-----	993
HIV2CAM5	-----T-----T-----G-----	969
HIV2CAM6	-----A-----A-----T-----G-----	999
HIV2ALI	-----T-----G-----A-----A-----	978
CONSENSUS.B	CAGGaAACAagAcgGTtGTaCCAATAAGAAcGtGTTCAGGacT?cttTTCATTCa.....CAGCCTaTC...	932
HIV2UC1	-----G-----A-----CA-----C-----T-----A-----GAA-----G-----C-----	984
HIV2D205	-----A-----T-----G-----A-----	987
HIV2EHOA	-----A-----T-----G-----TA-----T-----C-----	981
CONSENSUS.D	CAGGaAATAAgACAGTtTTACCAGTCACCAATtATGTCTcggaTTGGTtTTCACtCa.....CAaCCaaTc...	986
SIVM251	-----	1011
SIVM32H	-----	1002
SIVM1A11	-----CA-----	1002
SIVM142	-----C-----G-----	1014
SIVM239	-----	1006
SIVMNE	-----C-----	1005
SIVSMH4	-----A-----A-----G-----C-----T-----G-----G-----C-----A-----	1008
SIVSMH9	-----C-----A-----G-----C-----T-----G-----G-----C-----A-----	1014
SIVSMFBJ	-----C-----A-----G-----C-----T-----G-----G-----C-----A-----	1023
SIVP209C19	-----G-----A-----G-----C-----T-----G-----G-----C-----A-----	1011
SIVSTM	-----G-----C-----A-----C-----A-----C-----C-----T-----G-----C-----T-----	1002
CONSENSUS.A	AATAAaAgaCCcaggCAaGCATgGTGctGGTTCaaaGGcaaaTGGaagGaaGCCATGcagGAGGTgAagcaAAC	981
HIV2ROD	-----A-----A-----G-----G-----A-----C-----G-----	1067
HIV2NIH2	-----A-----A-----T-----G-----C-----G-----G-----G-----G-----	1034
HIV2ISY	-----A-----A-----C-----G-----G-----G-----A-----A-----	1037
HIV2ST	-----G-----G-----G-----G-----A-----T-----	1043
HIV2BEN	-----C-----C-----G-----T-----G-----C-----TGG-----G-----G-----A-----	1061
HIV2CAM2	-----G-----G-----C-----CA-----A-----	1061
HIV2D194	-----C-----A-----TG-----G-----T-----TC-----C-----TA-----G-----	1031
HIV2GH1	-----C-----T-----G-----G-----A-----	1034
HIV2D1024	-----C-----T-----A-----C-----A-----	1052
HIV2D1071	-----CC-----A-----C-----C-----C-----	743
HIV2D766	-----C-----T-----A-----C-----C-----GA-----	1052
HIV2D808	-----C-----T-----A-----G-----G-----	1040
HIV2D868	-----G-----T-----A-----C-----T-----G-----G-----G-----	1052
HIV2CBL21	-----G-----A-----T-----G-----T-----G-----G-----G-----	1064
HIV2CBL22	-----C-----A-----A-----C-----G-----G-----	1052
HIV2CBL23	-----C-----A-----A-----T-----TG-----G-----G-----	1052
HIV2CBL24	-----C-----C-----A-----A-----	1064
HIV2CAM1	-----G-----T-----G-----T-----C-----	1058
HIV2CAM2BR	-----G-----G-----C-----CA-----A-----	1067
HIV2CAM3	-----C-----T-----A-----G-----A-----G-----	1064
HIV2CAM4	-----TG-----T-----G-----G-----G-----A-----A-----	1067
HIV2CAM5	-----G-----G-----C-----G-----G-----A-----A-----	1043
HIV2CAM6	-----GG-----T-----A-----G-----C-----GA-----G-----A-----	1073
HIV2ALI	-----A-----A-----G-----G-----G-----	1052
CONSENSUS.B	AATAa?AGACCCcAgACaAGCTTGGTGTGGTTTAAGGGAAACTGGACAGAAAGCCATAaAGGAGGTgAAagaGAC	1005
HIV2UC1	-----CC-----G-----T-----A-----G-----	1058
HIV2D205	-----G-----A-----T-----A-----C-----AG-----	1061
HIV2EHOA	-----A-----A-----T-----A-----C-----	1055
CONSENSUS.D	AATGAgAGgCCAAAaCagGCATGGTGTtGGPTTgAGGaaaaTGGaaggagGCaATAaaaGaggTyAAGcagc	1060
SIVM251	-----T-----G-----G-----T-----T-----A-----	1085
SIVM32H	-----T-----G-----G-----T-----T-----A-----	1076
SIVM1A11	-----T-----G-----G-----A-----T-----T-----	1076
SIVM142	-----T-----G-----A-----A-----T-----	1088
SIVM239	-----T-----G-----A-----A-----T-----	1079
SIVMNE	-----T-----A-----A-----A-----T-----A-----	1079
SIVSMH4	-----A-----C-----C-----A-----GC-----AA-----C-----CC-----G-----A-----G-----A-----	1082
SIVSMH9	-----A-----C-----C-----G-----A-----C-----CCGG-----A-----GNAM-----	1088
SIVSMFBJ	-----A-----C-----C-----G-----A-----AA-----C-----CC-----G-----A-----G-----A-----	1097
SIVP209C19	-----A-----C-----C-----CC-----G-----C-----GC-----C-----CC-----G-----A-----G-----A-----	1085
SIVSTM	-----A-----A-----T-----C-----G-----GA-----GA-----T-----G-----A-----	1076

HIV2/SIV ENV

CONSENSUS A	CcTtTgcAAaACatCCcAGgTAlaagga...??accatga?Acaa??????aa?aftaactTtacagcaC	1039
HIV2ROD	-----G-----G-----C-----GG-----T-----G-----G-----G-----	1129
HIV2NIH2	-----G-----A-----G-----CAGG-GCCGCACAGAG-T-----A-----A-----	1099
HIV2ISY	-----T-----C-----A-----C-----C-----AT-----A-----A-----	1099
HIV2ST	-----T-----C-----A-----C-----C-----GAA-----A-----CGT-----T-----G-----	1105
HIV2BEN	-----T-----C-----A-----C-----T-----GGG-----A-----GAA-----	1123
HIV2CAM2	-----GG-G-----ACC-AA-CATA-G-----G-T-----C-----C-A-----	1123
HIV2D194	-----CGG-----A-----T-----GGA-----A-----GAAG-----	1093
HIV2GH1	-----AT-----C-----C-----AG-----T-----A-----AA-----	1096
HIV2D1024	-----T-----C-----G-----G-----A-----G-----GAA-----G-T-----C-----G-----	1114
HIV2D1071	-----C-----G-----G-----A-----T-----GGA-----G-----G-----	805
HIV2D766	-----T-----C-----A-----GG-----G-----GG-----G-----G-----	1114
HIV2D808	-----A-----C-----C-----GG-----T-----GGA-----G-----G-----	1102
HIV2D868	-----A-----G-----C-----GG-----A-----T-----GGA-----T-----A-----GTG-----	1114
HIV2CBL21	-----T-----G-----G-----A-----G-----GAA-----T-----C-----G-----G-----	1126
HIV2CBL22	-----G-----CA-T-----GAG-----GT-----C-----G-----	1114
HIV2CBL23	-----A-----AG-----T-----A-----G-----	1114
HIV2CBL24	-----G-----G-----C-----C-----GC-----G-----T-----G-----	1126
HIV2CAM1	-----T-----G-----A-----GG-----C-----GGA-----C-----G-----	1120
HIV2CAM2BR	-----GG-G-----AA-C-T-----CG-----G-T-----C-----C-A-----	1129
HIV2CAM3	-----C-----C-----AG-----G-----G-----	1126
HIV2CAM4	-----CT-G-----A-----G-----CAA-C-T-----AC-----C-----CCG-C-----G-G-AG-----	1129
HIV2CAM5	-----G-----G-----G-----GAACCCA-T-A-----CA-----G-----T-----	1102
HIV2CAM6	-----AA-G-----GAA-----TT-----TT-----AT-----T-----C-----GA-----	1135
HIV2ALI	-----T-----C-----AC-----C-A-----AA-----	1114
CONSENSUS B	CATCAtAAAaCATCCcAGgTATAaAGGA????cAcAAATa?cac?????aaGaFAag?TtagcAGagC	1064
HIV2UC1	-----A-----CGAG-----G-----A-----G-T-----G-----	1120
HIV2D205	-----GGTG-----A-----T-----AGCG-A-----T-----G-----T-----A-----	1129
HIV2EHOA	-----T-A-----T-----A-----TC-----A-----A-CTCAC-----G-----	1117
CONSENSUS D	CaTtGTCAAAATCCcAGGTATAActGGA...ACTAAcaAeActgaE...aAAATcAaEtTaaagGCTC	1122
SIVMM251	-----G-----	1147
SIVMM32H	-----G-----	1138
SIVMM1A11	-----G-----	1138
SIVMM142	-----G-----	1150
SIVMM239	-----G-----	1141
SIVMNE	-----G-----	1141
SIVSMH4	-----T-G-----G-----TG-----AGG-----T-----C-----A-----	1144
SIVSMH9	-----T-G-----AG-----TR-K-Y-R-A-----C-----T-----GC-----A-----	1150
SIVSMH9BJ	-----T-G-----G-----T-G-----A-----C-----T-----GC-----A-----	1159
SIVP209C19	-----T-G-----G-----TG-----A-G-----T-----C-----A-----	1147
SIVSTM	-----T-G-----A-----C-----TG-C-----G-CA-----A-GGA-----GT-----	1138
CONSENSUS A	CaGgaaaagggcTCaGAccCaGaaGtggcATacATGTGGACTAacTGcAGaGGAaTtctTCTAcTGcAAaATG	1113
HIV2ROD	-----G-----A-----G-----	1203
HIV2NIH2	-----G-G-----AA-----C-----	1173
HIV2ISY	-----A-----A-----T-----C-----T-----	1173
HIV2ST	-----G-C-----A-----T-----T-----	1179
HIV2BEN	-----G-GC-----G-----TT-----T-----	1197
HIV2CAM2	-----A-G-----T-----A-----T-----T-----	1197
HIV2D194	-----T-----T-----A-----T-----T-----	1167
HIV2GH1	-----G-----G-----CA-----G-----G-----	1170
HIV2D1024	-----G-----A-T-----G-----CA-----T-----G-----G-----	1188
HIV2D1071	-----G-----A-T-----G-----CA-----T-----G-----G-----	879
HIV2D766	-----G-----G-----C-----G-----G-----	1188
HIV2D808	-----G-----G-----C-----G-----T-----	1176
HIV2D868	-----G-----G-----C-----G-----	1188
HIV2CBL21	-----G-C-----C-----C-----	1200
HIV2CBL22	-----G-G-----T-----A-----C-----T-----	1188
HIV2CBL23	-----G-----G-----G-----T-----	1188
HIV2CBL24	-----A-----T-----T-----T-----	1200
HIV2CAM1	-----T-----T-----T-----	1194
HIV2CAM2BR	-----A-G-----T-----A-----T-----T-----	1203
HIV2CAM3	-----G-----G-----G-----	1200
HIV2CAM4	-----T-----T-----T-----	1203
HIV2CAM5	-----G-----A-----CT-----T-----	1176
HIV2CAM6	-----G-----A-----G-----G-----	1209
HIV2ALI	-----G-----TG-----A-----T-----T-----T-----	1188
CONSENSUS B	acggcga?AgGtTCAGAtCCAGAggtAaGAtATATGTGACACtAAcTGTAAGGAGAA?TtctT?TA?TGTAATATG	1135
HIV2UC1	C-T-GC-----C-----G-----C-----C-----P-----	1194
HIV2D205	-T-GA-A-----AC-C-----C-----T-----A-----A-----C-----	1203
HIV2GH2	-----G-A-C-----G-----TC-C-----C-----	18
HIV2EHOA	-----G-A-C-----G-----TC-C-----C-----	1191
CONSENSUS D	CtggaggGAGGA...GATCCGGAaGtAcCTtCATGTGACAAATTCaGAGGAGA?TtctTcTA?TG?AAAAATG	1190
SIVMM251	-----G-----C-----T-----	1218
SIVMM32H	-----A-----G-----C-----T-----	1209
SIVMM1A11	-----A-----G-----T-----C-----T-----	1209
SIVMM142	-----A-----G-----T-----C-----T-----	1221
SIVMM239	-----G-----C-----T-----	1212
SIVMNE	-----G-----T-----C-----T-----	1212
SIVSMH4	-A-C-----A-----C-----T-----T-----TC-----A-----T-----A-----C-----	1215
SIVSMH9	-A-C-----A-----T-----T-----TC-----A-----T-----A-----T-----C-----	1221
SIVSMH9BJ	-A-C-----A-----T-----T-----TC-----A-----T-----A-----T-----C-----	1230
SIVP209C19	-A-C-----A-----C-----T-----T-----T-----A-----A-----T-----C-----	1218
SIVSTM	-----G-----A-----G-----C-----C-----T-----A-----T-----T-----C-----	1209
SIVSTM22579	-----A-----P-----T-----T-----C-----	42

CONSENSUS.A	ACtTGGTTCTCAATTgGgTaGaa.....aAcAaaac????ataPg???	1149
HIV2HOD	-----A-----G-----T-G-A-----	1236
HIV2NINZ	-----G-----GG-GGTC-G-AA-----	1215
HIV2ISY	-----T-----G-GGTC-A-----	1212
HIV2ST	-----G-GAATC-G-CA-----	1221
HIV2BEN	-----G-G-AAAC-A-CA-----	1236
HIV2CAM2	---C-----T-C-TAATACG-CA-----	1239
HIV2D194	-----T-G-GAACC-A-CA-----	1209
HIV2GH1	-----G-C-GAATC-G-CA-----	1212
HIV2D1024	-----A-----G-----G-CATGA--A-----	1230
HIV2D1071	-----A-----G-----G-CGCGA-C-G-----	921
HIV2D766	-----A-----G-----G-TGCA--G-----	1230
HIV2D808	-----G-----TGCAA--CA-ACA-----	1221
HIV2D868	-----A-----G-----TGAGA--GCAACA-----	1233
HIV2CBL21	-----G-----C-A--A-CA-ACA-----	1242
HIV2CBL22	-----T-G-G--GG-CA-ACA-----	1230
HIV2CBL23	-----A-----T-G-A-----	1221
HIV2CBL24	-----G-A--A--CAACT-----	1242
HIV2CAM1	-----T--GG-GAA-----	1233
HIV2CAM2BR	---C-----T-C-T--C-ACA-----	1245
HIV2CAM3	-----T-----G-TGAGA--G-ACA-----	1245
HIV2CAM4	-----G-A--GGAGG-CCA-----	1245
HIV2CAM5	-----G-G--GGCCA-GAA-----	1218
HIV2CAM6	-----G-A--A-CA-ACA-----	1251
HIV2ALI	-----G--GG-CA-GAA-----	1230
CONSENSUS.B	AcTggTTcTTAAACTGGGTaGAg.....AAcA?AACa???aatacaact.	1175
HIV2UC1	---C-----T-----G-----G-----GG-----	1236
HIV2D205	---C-----A-----	1245
HIV2GH2	---AT-----A-----AGA-----	60
HIV2EHOA	---T-TT-----A-----T-G--TGGG-----	1227
CONSENSUS.D	AaTtGGTTtCtAaATtGGgTAGAagatagg???aata?aacTa?c?????????a?g?????????cA?Aagcc	1239
SIVMM251	-----G-----G-GT--C-----G-G-----	1268
SIVMM32H	-----G-GT--A-----G-G-----	1259
SIVMM1A11	-----T-----C-G--A-----G-----	1259
SIVMM142	-----G-CT--C-----G-----	1271
SIVMM239	-----C-G--A-----G-----	1262
SIVMNE	-----A-----CTG--GGAACCTACC-----G-----	1268
SIVSRR4	-----T-----G-C--AGACC-A-AGGG-GG-----G-TGGAAACAA--A--TAG	1280
SIVSMMH9	-----T-----Y-TAAAA--GGTT-----G-TGGACAAGT--A--C-A	1283
SIVSMMPEJ	-----T-----A-TACAA--GGTT-----G-TGGACAAGT--A--C-A	1292
SIVP209C19	---G-----T-----G-C-AGACC-A-AT-G-A-----G-TGGAAACAA--A--AA	1283
SIVSTM	-----CT-----A-----C-G--G-CATCAGAAATG-G-GATTGGACAA-A--CAA	1274
SIVSTM22579	-----CT-----C-G--G-CATTAGAAATG-A-AATTGGAACA-A--C-A	107

HIV2/SIV ENV

CONSENSUS.A	??cagcacAAtTATgtgccATGcCaTATAaagCAaATAATtAAcaCcTGGCAaAagGTAGggA	1210
HIV2ROD	-C-G-----CA-G-----A-----	1297
HIV2NEHZ	-G-----CA-G-----G-----T-----C-G-----	1276
HIV2ISY	-T-C-----G-----G-----T-----	1273
HIV2ST	-G-----C-----G-----T-----C-----	1282
HIV2BEN	-G-G-C-----G-----T-----A-----	1291
HIV2CAM2	A-G-----CA-----G-----	1300
HIV2D194	-CGG-C-----G-G-----	1270
HIV2GH1	-C-G-----G-----T-----C-----	1273
HIV2D1024	ACATG-----	1294
HIV2D1071	ACATG-T-----	975
HIV2D766	ACATG-----G-----	1294
HIV2D808	TG-----	1282
HIV2D868	TG-----	1294
HIV2CBL21	-A-----G-----T-----C-----	1303
HIV2CBL22	-G-C-----G-----T-----	1291
HIV2CBL23	-C-G-----A-G-----GA-----	1282
HIV2CBL24	TG-----G-----A-G-C-T-----	1303
HIV2CAM1	A-----	1294
HIV2CAH2BR	A-G-----CA-----G-----	1306
HIV2CAM3	TG-----G-----A-----	1306
HIV2CAM4	-C-----G-----G-G-----T-----A-----	1306
HIV2CAM5	-C-----CA-G-T-----G-----TG-T-----AA-----	1279
HIV2CAM6	-G-----A-----	1312
HIV2ALI	-CA-G-----T-----C-A-----	1291
CONSENSUS.B	??c?GAgcAATTTATG?GcCaTGcCA?ATCagACAGATA?TcAACACGTGGCACAaggTTGG?A	1231
HIV2UC1	-A-AG-----T-A-G-----C-----A-----G-----A-----	1297
HIV2D205	-G-----C-----T-----T-----A-----G-----	1306
HIV2GH2	-G-----T-----T-----A-T-----A-----G-----	121
HIV2EHOA	CTCAG-----C-T-----C-----G-----A-----A-----	1291
CONSENSUS.D	aaagGAaCgacAtasaaAggAATTA7GTgCCaTGtCAtATtaGACaaTAATcAACACtTGGCAlAaAGTAGGcA	1312
SIVMM251	-G-----G-----C-----G-----	1342
SIVMM32H	-G-----G-----C-----	1333
SIVMM1A11	-A-----C-----	1333
SIVMM142	-G-----C-----A-----	1345
SIVMM239	-AG-----C-----	1336
SIVMNE	-C-----G-----C-----	1342
SIVSMNH4	G-A-G-A-G-G-AA-----T-----G-----C-----	1354
SIVSMNH9	G-A-----T-CC-G-AA-----T-----R-----G-----G-----C-----	1357
SIVSMNPBJ	G-A-G-----A-GG-A-----T-----G-----G-----C-----	1366
SIVP209C19	GCCA-G-A-G-G-A-----T-----G-----G-----C-----	1357
SIVSTM	-----A-A-G-----T-A-----C-C-C-----G-T-T-A-----G-----	1348
SIVSTM22579	-----AC-A-----T-----C-C-C-----G-T-T-----G-----G-----	181

CONSENSUS .A	aaaatgtATATTTCCTCCcAGGGAAGGggagtTgacCTGcaActCaACaGTgACcAGcaTaaTtGCTAACATt	1284
HIV2ROD	G-----C-----C-----T-----A-----A-----	1371
HIV2NIH2	-----C-----T-----A-----T-----G-----A-----C-----	1350
HIV2ISY	-----T-----AC-----A-----T-----T-----C-----	1347
HIV2BT	-----T-----G-----TG-----A-----A-----	1356
HIV2BEN	-----G-----T-----A-----A-----	1365
HIV2CAM2	-----C-----T-----A-----T-----A-----	1374
HIV2D194	-----T-----A-----AC-----	1344
HIV2GH1	-----C-----T-----A-----T-----	1347
HIV2D1024	-----C-----T-----T-----	1368
HIV2D766	-----C-----T-----	1368
HIV2D808	-----C-----A-----A-----T-----	1356
HIV2D868	-----C-----T-----	1368
HIV2CBL21	-----T-----ACT-----T-----	1377
HIV2CBL22	-----T-----A-----T-----C-----C-----	1365
HIV2CBL23	-----T-----A-----T-----A-----T-----	1356
HIV2CBL24	-----T-----A-----T-----C-----	1377
HIV2CAM1	-----G-----	1368
HIV2CAM2BR	-----G-----	1380
HIV2CAM3	-----T-----A-----GT-----G-----A-----C-----G-----	1380
HIV2CAM4	-----GA-----T-----A-----A-----T-----T-----A-----	1353
HIV2CAM5	-----G-----	1386
HIV2ALI	-----T-----A-----A-----T-----G-----	1365
CONSENSUS .B	2AaAt?TaTATTTCCTCCaAGaGA?GGTgagCTcTC?TGTAATTCCaCTGTgAC?AGccTcATAGCCAACAT?	1299
HIV2UC1	A-T-CG-----G-----G-----AC-----C-----C-----T-----T-----	1371
HIV2D205	A-----A-----A-----A-----A-----T-----C-----	1380
HIV2GH2	G-----A-----A-----CT-----A-----A-----T-----A-----C-----	1395
HIV2EHOA	G-----G-----G-----A-----C-----T-----C-----T-----	1365
CONSENSUS .D	aaaATGTtTaTTTGCTCCaAGaGA?GGAGACCTcACgTGTaa?TCCAC?GTgACCAGtcTcATAGCaaAcATe	1383
SIVMM251	-----G-----C-----A-----	1416
SIVMM32H	-----G-----C-----A-----	1407
SIVMM1A11	-----G-----TC-----A-----	1407
SIVMM142	-----G-----C-----A-----	1419
SIVMM239	-----G-----C-----A-----	1410
SIVMNE	G-----G-----G-----T-----A-----	1416
SIVSMNH4	-----A-----T-----G-----A-----G-----A-----C-----T-----A-----T-----G-----G-----	1428
SIVSMNH9	-----G-----R-----T-----G-----A-----G-----A-----C-----T-----A-----T-----C-----G-----A-----	1431
SIVSMNPBJ	-----G-----T-----G-----A-----G-----A-----T-----T-----A-----T-----C-----G-----A-----	1440
SIVP209C19	-----A-----T-----G-----A-----G-----A-----T-----T-----A-----T-----G-----G-----	1431
SIVSTM	-----C-----A-----T-----T-----T-----T-----A-----A-----	1422
SIVSTM22579	-----G-----G-----A-----A-----T-----T-----T-----T-----A-----A-----G-----T-----	255
CONSENSUS .A	gAc?????????g7c77?aatcagaCa?????aAtATTACCTTTAGTGCAGAGGcGGcaGAaCTaTAcogAtT	1339
HIV2ROD	-----TGGCAAAACAAAT-----C-----A-----	1436
HIV2NIH2	-----T-----CGGGAG-----A-----C-----	1409
HIV2ISY	-----T-----GTTGAT-CAGAT-----C-----G-----	1412
HIV2BT	-----GGA-GAGAG-----C-----	1418
HIV2BEN	-----ATA-A-AAA-----G-----TCATACC-----C-----A-----G-----	1433
HIV2CAM2	-----GAGAGG-A-----ACA-----T-----G-----	1439
HIV2D194	-----TCA-A-GGA-----C-----C-----A-----G-----	1406
HIV2GH1	-----GTAATAG-----C-----	1409
HIV2D1024	-----GGA-G-----G-----C-----	1427
HIV2D766	-----GGA-G-----C-----	1427
HIV2D808	-----GGA-G-----G-----	1415
HIV2D868	-----GGA-G-----	1427
HIV2CBL21	-----T-----ACATAC-G-----C-----G-----	1439
HIV2CBL22	-----T-----GTAAAT-G-----A-----	1427
HIV2CBL23	-----T-----GCAAT-GA-----A-----T-----G-----	1418
HIV2CBL24	-----GTTAAAA-----C-----	1439
HIV2CAM1	-----ACTGACATG-----CGGT-----TGAGACA-----	1436
HIV2CAM2BR	-----GAGAGG-A-----ACA-----T-----G-----	1445
HIV2CAM3	-----GTGGAA-CA-----	1442
HIV2CAM4	-----ATGCCG-----CG-PT-FACA-----T-----	1442
HIV2CAM5	-----GACAAATATGCACAAG-C-----T-----AA-----	1421
HIV2CAM6	A-----GTGGAC-----GTCA-----A-----G-----T-----	1448
HIV2ALI	-----ACGGATGCC-----C-----G-----	1427
CONSENSUS .B	?AC?????LcaGataa?AA?g?TAcCa?????AATATT?C?gTGAGTGCAGAAGTgTcAGAAGCTGTATAgAtT	1357
HIV2UC1	G-GTG-AT-AT-----GGC-----T-----A-----AGACC-----A-----CA-----C-----	1445
HIV2D205	A-----C-----CAG-----CC-----T-----T-----	1442
HIV2GH2	A-----CA-----T-----T-----CC-----T-----C-----C-----	257
HIV2EHOA	G-----GGAT-----G-----CCT-----T-----A-----T-----A-----	1430
CONSENSUS .D	gAt...TGGATt?Atgggaa?CA?ACT...AatATCACCcatgAGTGCAGAGGTGGCAGAAGCTGTATCGATT	1444
SIVMM251	-----C-----G-----C-----A-----G-----	1481
SIVMM32H	-----C-----G-----C-----A-----	1472
SIVMM1A11	-----C-----G-----C-----A-----	1472
SIVMM142	A-----C-----G-----C-----A-----G-----	1484
SIVMM239	-----T-----G-----C-----A-----	1475
SIVMNE	-----T-----G-----C-----A-----	1481
SIVSMNH4	-----TCA-----A-----C-----TG-----G-----C-----	1493
SIVSMNH9	-----TCA-----C-----TA-----G-----C-----	1496
SIVSMNPBJ	-----TCA-----C-----TG-----G-----C-----	1505
SIVP209C19	-----TCA-----AAC-----TG-----G-----C-----	1496
SIVSTM	-----C-----C-----A-----CAAT-----TG-----G-----C-----TGCA-----	1487
SIVSTM22579	-----C-----C-----A-----CAAT-----TG-----G-----TGCA-----	320

HIV2/SIV ENV

CONSENSUS .A	gGaactTgGGaGAtTAcAAaTfGgTAGAaaTaACAcCAaTTgGcTTcGCaCCTACa?cagaaaaaaGatActcct	1412
HIV2ROD	---G-----	1510
HIV2NIH2	---G---C---A---C---T---T---G---	1483
HIV2ISY	---G---A---G---C---	1486
HIV2ST	A---G---A---G---T---C---T---	1492
HIV2BEN	A---C---C---C---AA---GATC-G-G-G---	1507
HIV2CAM2	---G---T---C---G---C---	1513
HIV2D194	A---G---C---C---A---G---CCG---AA-G---T---	1480
HIV2GH1	A---AA---G---GAG-G---T---	1483
HIV2D1024	A---T---T---C---T---	1501
HIV2D766	---C---C---T---CC---	1501
HIV2D808	--G-A-T---C---T---G---T---C-GG---	1489
HIV2D868	A---C---C---T---C---C---T---	1501
HIV2CBL21	---C---A---G---C---T---T---	1513
HIV2CBL22	---A---A---G---	1501
HIV2CBL23	A-G---G---	1492
HIV2CBL24	A---G---A---T---C---CT---	1513
HIV2CAM1	---A---A---C-G---	1510
HIV2CAM2BR	---G---G---T---C---G---C---	1519
HIV2CAM3	---G---T---GC---	1516
HIV2CAM4	---C---AA---G---T---GATC-G---C---	1516
HIV2CAM5	---C---A---T---T---	1495
HIV2CAM6	---C---C---GA-C---C---	1522
HIV2ALI	A---G---A---G---T---GG---	1501
CONSENSUS .B	GGAActgGGGGACTA?AAAtTAGTAGAaATAACACCAaTTGGCTTtGCACCTACAgel?TAAaAAGATA?TCCT	1428
HIV2UC1	---C---G---GA---T---	1519
HIV2D205	---T---C---G---G---C---	1516
HIV2GH2	---CT---T---C---AG---C---	331
HIV2EHOA	---C---G---AG-A---T---	1504
CONSENSUS .D	GGAAaTTGGGAGATTA?AAATfGgTAGAgATCCTCCaATfGgCTTgGCCCCcACaaTGTgAagAGgTACACtA	1517
SIVMM251	---G---T---G---	1555
SIVMM32H	---T---G---	1546
SIVMM1A11	---T---A---G---	1546
SIVMM142	---T---A---	1558
SIVMM239	---T---G---	1549
SIVMNE	---T---	1555
SIVSMNH4	---G---C---A---T---G---A-GA---C---	1567
SIVSMNH9	---G---C---T---G---A-A---C---	1570
SIVSMMPBJ	---G---C---T---C---C---G---A-A---C---	1579
SIVP209C19	---G---C---T---G---A-GA---C---	1570
SIVSTM	---C---A---A---A---G---A---A---	1561
SIVSTM22579	---C---G---A---A---A---G---A---A---	394

CONSENSUS.A	Ct...gctCc?gggAGacataaaagaGGTgttccGTGCTAGGGTTCtTgGGTTTCTCgCaaCagcaGGttcT	1482
HIV2ROD	---AC---C-----T-----C-A-----	1581
HIV2NIH2	---ACCA---C-----T-----C-A-----	1554
HIV2ISY	---A---G-----C-T-----C-A-----A-G-----G-	1557
HIV2ST	---A-T---GA-----A-----A-G-----AG-	1563
HIV2BEN	-A..A---A-T---GA-C-----G-----	1578
HIV2CAM2	---AC---C-----C-T-----A-----AGT-	1584
HIV2D194	-G...A-T---GA-C-----G-----	1551
HIV2GH1	-G...A-T---GA-C-----T-----G-----	1554
HIV2D1024	---AC---CG-----	1521
HIV2D766	---AC---CG---	1524
HIV2D808	---AC---G-GAG---	1512
HIV2D868	---ACA---OCG---	1525
HIV2CBL21	---A---GA-----CC-----A-----AG--	1584
HIV2CBL22	---A---AAT---GA---G-----A-G-----G-----	1572
HIV2CBL23	---A---AAT---GA-C---G-----C-----	1563
HIV2CBL24	---T---AAA---GA-----A-----A-----G-	1584
HIV2CAM1	---AT---G-----T-----A-----GT-	1581
HIV2CAM2BR	---AC---C-----C-T-----A-----AGT-	1590
HIV2CAM3	---ACAA---C-----A-----	1587
HIV2CAM4	---A---AACA---GAG-----G---C---A-	1587
HIV2CAM5	---G---A-----G---CG---	1565
HIV2CAM6	---A---AT---A-----A-----G---	1593
HIV2ALI	---A---AA---GA-----C-----G-----	1572
HIV2GB1222	---C-T-----G-----	48
CONSENSUS.B	CA...gtgA?accgAGgAATAAAAGAGGTGTA?TgGTGCTAGG?TTCTTGGGATTCCTTGCaatGGCAGGTCTCT	1496
HIV2UC1	---AC---C-----G-----C-T-----	1590
HIV2D205	---A---A-----A-----A-----C-T-----	1587
HIV2GH2	---AT---T---A---C-----C-----G---A-----	402
HIV2EHOA	---C-----C-T-----G-T-----G-C-----	1575
CONSENSUS.D	CtgggtgG?acCTCAGGAaTAAaAGAGGGGTCTTGTGCTAGGGTTCTTgGGTTTCTCGC?AC?GCAGGTCTCT	1588
SIVMM251	---C-----A-G-----	1629
SIVMM32H	---C-----A-G-----	1620
SIVMM1A11	---C-----A-G-----	1620
SIVMM142	---C-----A-G-----	1632
SIVMM239	---C-----A-G-----	1623
SIVMM6	---C-----A-G-----	1629
SIVSMH4	-AAC---TG-----G-----A-----	1641
SIVSMH9	-AAC---TG-----G---A-----	1644
SIVSNMFB7	-AAC---TG-----G---A-----	1653
SIVP209C19	-AAC---TG-----G---A-----	1644
SIVSTM	---A-T---C---G-----C-----G---A-----	1632
SIVSTM22579	---A-T---C---G-----C-----G---A-----	465
CONSENSUS.A	GCAATGGGcgcgggcgtCctTgACgcTGTGcGCTCAGTCTCCGACTTTAcTggccGGGATAGTGCAGCAACAGCA	1556
HIV2ROD	---C---CG-----A-----T-----	1655
HIV2NIH2	---C-----TC-----A-----T-----	1628
HIV2ISY	---G-----TC-----T-CCGT-	1631
HIV2ST	---CG---C-----A-C---C-----T-----	1637
HIV2BEN	---CG---C-----A-C---C-----T-----	1652
HIV2CAM2	---A-----A-C---C-----T-----	1658
HIV2D194	---GC-----C-----C-----	1625
HIV2GH1	---C-----C-----C-----	1628
HIV2CBL21	---TC-----C-----T-----	1658
HIV2CBL22	---TC-----C-----T-----	1646
HIV2CBL23	---A-----C-----T-----	1637
HIV2CBL24	---A-A-----T-----	1658
HIV2CAM1	---A-A-----A-C---C-----T-----	1655
HIV2CAM2BR	---A-----A-C---C-----T-----	1664
HIV2CAM3	---A-----C-A-----A-C---C-----	1661
HIV2CAM4	---A-----C-A-----A-C---C-----	1661
HIV2CAM5	---T-----T-----T-----	1640
HIV2CAM6	---A---AG-T---A-----T-----C-----T-----	1667
HIV2ALI	---A---AG-T---A-----T-----T-----	1646
HIV2GB1222	---CG-----C-----	122
CONSENSUS.B	GCAATGGGCGCcaagTC?TTGACGcTGTcAGCTCCCGACTTTACTGGCtGGGATAGTGCAGCAGCAGCA	1569
HIV2UC1	---C-----C-----	1664
HIV2D205	---T-----C-----	1661
HIV2GH2	---GC---T-----	476
HIV2EHOA	---GG---C---T---G-----	1649
CONSENSUS.D	GCAATGgGCGCGCGGTCTgTACgcTGAc?GCTCAGTCCCGACTTTAcTGGCtGGGATAGTGCAGCAGCAGCA	1661
SIVMM251	---C-----C-----A-----	1703
SIVMM32H	---C-----C-----A-----	1694
SIVMM1A11	---C-----C-----	1694
SIVMM142	---CG---C-----	1706
SIVMM239	---C-----A-----	1697
SIVMM6	---C-----C-----	1703
SIVSMH4	---CG---T-G-----G-----	1715
SIVSMH9	---CG---T-G-----G-----	1718
SIVSNMFB7	---A---CG---T-G-----G-----	1727
SIVP209C19	---C---CG---T-G-----G-----	1718
SIVSTM	---A---C---G-----A-----	1706
SIVSTM22579	---A---C---G-----G-----	539

HIV2/SIV ENV

CONSENSUS .A	ACAGCTGtTgGAcgTgGTCaaGAGACAACaAGAAaTGTTGCGACTGACCgTCGCGGGaaCaAAAAaCTCCAGg	1630
HIV2ROD	-----C-----	1729
HIV2NIH2	-----T-----	1702
HIV2ISY	-----T-----C-----	1705
HIV2ST	-----	1711
HIV2BEN	-----A-----G-----C-----	1726
HIV2CAM2	-----A-----T-----T-----	1732
HIV2D194	-----T-----G-----	1699
HIV2GH1	-----A-----G-----	1702
HIV2CBL21	-----A-----	1732
HIV2CBL22	-----G-----C-----	1720
HIV2CBL23	-----O-----	1711
HIV2CBL24	-----A-----	1732
HIV2CAM1	-----A-----G-----T-----	1729
HIV2CAM2BR	-----A-----T-----	1738
HIV2CAM3	-----T-----A-----	1735
HIV2CAM4	-----G-----	1735
HIV2CAM5	-----G-----	1714
HIV2CAM6	-----A-----G-----	1741
HIV2ALI	-----A-----G-----G-----	1720
HIV25132E2	-----A-----G-----	64
HIV27312A	-----	64
HIV2FA	-----	64
HIV2GB1222	-----G-----A-----	196
HIV2C1171	-----G-----	64
HIV2C19	-----A-A-----G-----	64
HIV2IS053	-----A-----A-----	64
HIV2VI390	-----A-----T-----G-G-----	64
HIV2VI905	-----C-----	64
CONSENSUS .B	ACAGCtg?TGGACGTGGTCAAAAGACAACAGGAAGTGTgCGGCTGACCGTCTGGGGaaCgAAaAACCTCCAGa	1642
HIV2UC1	-----C-----A-----	1738
HIV2D205	-----CTG-----G-----	1735
HIV2GH2	-----C-----GT-----A-G-----	550
HIV2EHOA	-----CG-----A-----G-----	1723
HIV2JA	-----A-----	64
HIV2ON	-----	64
CONSENSUS .C	GACGTGGTCAAAAGACAACAGGAAGTGTGCGGCTGACCGTCTGGGGAACTAAAAACCTCCAGA	64
HIV2238	-----	64
CONSENSUS .D	ACaGCTGTTGGACGtGGTCAaGAGACAACAAGAAaTGTTGCGACTGACCGTCTGGGGaaC?AAgAACCTCCAGA	1734
SIVMM251	-----A-----	1777
SIVMM32H	-----A-----	1768
SIVMM1A11	-----T-----A-----	1768
SIVMM142	-----A-----	1780
SIVMM239	-----A-----	1771
SIVMNE	-----A-----	1777
SIVSMBH4	-----T-----T-----	1789
SIVSMBH9	-----T-----T-----	1792
SIVSMBPB3	-----T-----G-T-----	1801
SIVP209C19	-----T-----T-----	1792
SIVSTM	-----C-----	1780
SIVSTM22579	-----C-----C-----	613
HIV2P0784	-----G-T-A-----	64
SMC18	-----C-----C-----T-A-----	64
SMC12	-----C-----C-----T-A-----	64

CONSENSUS .A	CAAGAGTCACCTGCTATCGAGAAGTACTTAAaggAcCAGGCgCagcTAAATTCaTGGGGaTGtGcgTtLAGaCaa	1704
HIV2ROD	-----A-----C-C-----G-----	1803
HIV2NIW2	-----A-----C-----	1776
HIV2ISY	-----T-----C-GCA-----GA-----	1779
HIV2ST	-----A-----C-----A-----	1785
HIV2BEN	-----C-----C-T-----A-----G	1800
HIV2CAM2	-----A-----C-----T-----A-----	1806
HIV2D194	-----A-----A-----A-----G-G	1773
HIV2GHI	-----A-----A-----T-----G-----G-G	1776
HIV2CBL21	-----T-----A-----C-----G-----	1806
HIV2CBL22	-----A-----G-----	1794
HIV2CBL23	-----A-----A-----	1785
HIV2CBL24	-----A-----G-----	1806
HIV2CAM1	-----A-----C-----TA-----A-C-----	1803
HIV2CAM2BR	-----A-----C-----T-----A-----	1812
HIV2CAM3	-----A-----C-----A-A-----C-----	1809
HIV2CAM4	-----G-----A-----A-----C-T-----G-----	1809
HIV2CAM5	-----T-----A-----A-G-----A-----G	1788
HIV2CAM6	-----A-----C-----A-----	1815
HIV2ALI	-----A-----G-----A-----	1794
HIV25132E2	-----A-----A-----G-G	138
HIV27312A	-----T-----A-----A-----G-----	138
HIV2FA	-----T-----A-----A-----C-----G-----	138
HIV2GB1222	-----A-----C-----CG-----	270
HIV2CI171	-----A-----G-----G-G	138
HIV2CI9	-----A-----A-----G-----G-G	138
HIV2I8053	-----GA-----C-----	138
HIV2VI390	-----A-----A-----A-----G-G	138
HIV2VI905	-----A-----A-----C-----A-----C-----G	138
CONSENSUS .B	CAAGAGTCACCTGCCATCGAGAATACCTaAaggaacCAOGCAG?ACTAAATTCgTGGGGATGTGCTTT?AGaCAG	1714
HIV2UC1	-----T-----T-----T-----G-----	1812
HIV2D205	-----T-----A-----A-----	1809
HIV2GH2	-----T-----A-----C-----	624
HIV2EHOA	-----T-C-A-----A-----A-----C-----	1797
HIV2JA	-----T-----T-----C-----T-----	138
HIV2ON	-----T-----T-----T-----	138
CONSENSUS .C	CTAGAGTCACCTGCCATCGAGAAGTACCTAAAGGACCAAGCAAAACTAAATTCATGGGGATGTGCCTTTAGACAG	138
HIV2238	-----	138
CONSENSUS .D	CTAGAGTCaCTGCCATCGAGAAGTACCTaAAGGATCAGGCgCagcTaAATTCtTGGGGATGtGCCgTtLAGaCaa	1808
SIVMM251	-----G-----T-----C-----G-----	1851
SIVMM32H	-----G-----T-----C-----G-----	1842
SIVMM1A11	-----G-----A-----T-----C-----G-----	1842
SIVMM142	-----G-T-----T-----C-----G-----	1854
SIVMM239	-----G-----T-----C-----G-----	1845
SIVMNE	-----T-----T-----C-----G-----A-----	1851
SIVSMH4	-----T-----G-----A-----T-----G-----	1863
SIVSMH9	-----T-----CRN-----A-----T-----G-G	1866
SIVSMPEJ	-----T-----T-----A-----T-----G-G	1875
SIVP209C19	-----T-----G-----A-----T-----G-----	1866
SIVSTM	-----T-----A-----A-----G-----	1854
SIVSTM22579	-----T-----A-----A-----G-----	687
HIV2F07B4	-----T-----C-----A-----C-----C-----G	138
SMCI8	-----A-----A-----A-G-----A-----	138
SMCI7	-----A-----A-G-----G-----A-----	138

HIV2/SIV ENV

CONSENSUS.A	GTCTGCCACACTaCTGTaCCaTgGgtAAATgActccTTaacaCCtgatTGGaacAAcATGACaTGGCAggAaTG	1778
HIV2ROD	-----T-----T-----G-----C-----G-----G-----	1877
HIV2NIM2	-----T-----TA-----G-----G-----G-----	1850
HIV2ISY	-----A-----A-----G-----C-----A-----	1853
HIV2ST	-----A-----G-----C-----C-----	1859
HIV2BEN	-----G-----T-----T-----C-----A-----G-----	1874
HIV2CAM2	-----C-----G-----T-----	1880
HIV2D194	-----C-----	1847
HIV2GH1	-----T-----T-----T-----G-----C-----	1850
HIV2CBL21	-----A-----CA-----C-----A-----	1880
HIV2CBL22	-----G-----AGA-----G-----C-----	1868
HIV2CBL23	-----TA-----G-----G-----G-----A-----	1859
HIV2CBL24	-----T-----CA-----C-----C-----C-----	1880
HIV2CAM1	-----T-----GC-----	1877
HIV2CAM2BR	-----C-----G-----T-----	1886
HIV2CAM3	-----T-----A-----	1883
HIV2CAM4	-----C-----T-----T-----A-----A-----G-----	1883
HIV2CAM5	-----AC-----T-----T-----C-----C-----	1862
HIV2CAM6	-----T-----G-----	1889
HIV2ALI	-----A-----A-----G-----C-----G-----A-----G-----	1868
HIV25132E2	-----T-----G-----C-----G-----A-----	212
HIV27312A	-----AG-----G-----G-----C-----G-----AC-----	212
HIV2FA	-----AG-----G-----C-----G-----C-----	212
HIV2GB1222	-----T-----T-----	344
HIV2CI171	-----T-----T-----G-----C-----A-----A-----G-----	212
HIV2CI9	-----T-----G-----G-----C-----	212
HIV2ISO53	-----T-----A-----C-----C-----C-----G-----	212
HIV2VI390	-----AG-----C-----	212
HIV2VI905	-----A-----CA-----C-----G-----C-----T-----C-----	212
CONSENSUS.B	GTCTGcCACACgACTGTACCATGGccAaATGaaACtCTcaaaCCAgActTGGaA?AACATGACATGGCAaCAGTG	1787
HIV2UC1	-----T-----C-----C-----A-----T-----G-----A-----	1886
HIV2D205	-----C-----C-----A-----T-----T-----	1883
HIV2GM2	-----G-----C-----T-----C-----G-----	698
HIV2ENOA	-----GT-----T-----C-----T-----G-----T-----A-----	1871
HIV2JA	-----A-----A-----C-----T-----A-----C-----A-----G-----	212
HEV2OM	-----C-----A-----G-----A-----G-----C-----G-----A-----	212
CONSENSUS.C	GTTTGCCACACTACTGTGCCATGGGTGAACGACAGTATGCAGCCAGAGTGGCAAAACATGACTTGGCAACAATG	212
HIV22238	-----	212
CONSENSUS.D	GTCTGcCACACTACTGTAccatGGccAaATGa?acttTgacaCC?gatTGGaAcAAATatgActTGGCAagAgTG	1880
SIVMM251	-----CA-----G-----C-----A-----A-----C-----GAT-----	1925
SIVMM32H	-----CA-----G-----C-----A-----A-----C-----GAT-----	1916
SIVMM1A11	-----CA-----G-----C-----A-----AA-----G-----GA-----	1916
SIVMM142	-----T-----CA-----G-----C-----A-----A-----GA-----	1928
SIVMM239	-----CA-----G-----C-----A-----AA-----G-----GA-----	1919
SIVMNE	-----T-----CA-----A-----C-----A-----AA-----GA-----	1925
SIVSM4H4	-----T-----A-----A-----GTG-----TA-----GA-----	1937
SIVSMOH9	-----K-----C-----A-----TARC-----Y-----	1940
SIVSM4PB1	-----A-----C-----A-----TA-----C-----	1949
SIVF209C19	-----T-----CT-----A-----GTG-----TAT-----G-----C-----	1940
SIVSTM	-----C-----TT-----GT-----G-----A-----G-----A-----	1928
SIVSTM22579	-----C-----CT-----GT-----G-----A-----G-----A-----	761
HIV2F0784	-----C-----GA-----AGT-----C-----C-----T-----C-----G-----	212
SMCI8	-----G-----T-----GAG-----C-----CA-----T-----T-----A-----	212
SMCI2	-----G-----T-----GA-----A-----C-----CA-----T-----T-----A-----	212

CONSENSUS .A	GGaacaacsagCCGctacCTaGAgGCAAAATACAGT?aaagttTaGAacAgGCaCAaATtCagCAaGaaAaAgA	1851
HIV2ROD	---A---G---A---G---A---	1951
HIV2NIH2	---A---G---C---C---A---	1924
HIV2ISY	---CA---A---T---G---C---	1927
HIV2ST	---G---G---A---A---G---C---	1933
HIV2BEN	---GA---G---G---G---A---C---A---	1948
HIV2CAM2	---A---G---G---G---C---C---A---	1954
HIV2DL94	---A---G---A---C---C---A---	1921
HIV2GH1	---A---G---C---C---A---	1924
HIV2CBL21	---A---AC---A---	1954
HIV2CBL22	---GA---GGCCAG---C---	1942
HIV2CBL23	---GGCA---A---G---G---C---CA---A---	1933
HIV2CBL24	---G---C---G---CTG---C---	1954
HIV2CAM1	---A---G---G---C---G---C---A---G---	1951
HIV2CAM2BR	---A---G---G---C---G---C---A---	1960
HIV2CAM3	---A---A---G---G---TG---C---G---	1957
HIV2CAM4	---T---G---AC---T---G---T---C---CT---C---G---	1957
HIV2CAM5	---CA---A---TG---G---C---C---A---C---	1936
HIV2CAM6	---G---GA---G---G---C---C---G---	1963
HIV2ALT	---T---T---G---CAG---G---	1942
HIV25132E2	---GG---G---C---C---A---G---A---	286
HIV27312A	---A---A---G---G---G---C---C---A---	286
HIV2FA	---A---G---G---G---C---C---A---	286
HIV2GB1222	---G---A---G---AC---A---G---A---	418
HIV2CI171	---GATG---C---CAG---A---	286
HIV2C19	---GG---G---G---TC---A---G---	286
HIV2IS053	---C---A---G---CA---G---G---	286
HIV2VI1390	---CGA---G---T---C---CAG---G---G---C---A---G---A---	286
HIV2VI905	---A---T---GC---	286
CONSENSUS .B	GGAAaagcaagTC?atTtCTTGGATCCAAATATAACag?a?TatTAGAAGagGC?CAGATACAaCA?GAAAaGA	1856
HIV2UC1	---G---A---CCC---G---T---A---G---	1960
HIV2D205	---C---C---G---CTC---G---T---G---	1957
HIV2GR2	---G---A---A---CTC---A---T---A---	772
HIV2EHOA	---G---G---COC---AA---T---C---A---A---	1945
HIV2JA	---C---A---G---A---C---A---CA---T---A---G---G---	286
HIV2ON	---GCAA---G---C---A---CA---T---A---A---G---G---	286
CONSENSUS .C	GGAAAAACAGATTGCTTCTTGGAGGATAATATCACAGAGCTGTTAGAGCAGCTCAAAATACAGCAGGAGAAAA	286
HIV22238	---	286
CONSENSUS .D	GGAgagaaAGgtTtgacttccTgGAgGcAAATATaaCagaaactattTaGAagA?GCaCAaATtCaaCAaGaaAaAgA	1953
SIVMM251	---C---T---A---CC---CC---G---G---	1999
SIVMM32H	---C---T---A---CC---CC---G---G---	1990
SIVMM1A11	---C---T---A---CC---CC---G---G---	1990
SIVMM142	---C---G---T---A---G---CC---CC---G---G---	2002
SIVMM239	---C---T---A---A---CC---CC---G---G---	1993
SIVMNE	---C---T---A---G---CC---T---G---G---	1999
SIVSM044	---A---C---A---TC---T---A---A---G---	2011
SIVSM049	---A---A---C---GA---A---TS---TV---G---A---G---H---	2014
SIVSMMPBJ	---A---A---C---GA---A---TC---TC---G---A---G---	2023
SIVP209C19	---A---C---C---A---TC---T---A---A---G---	2014
SIVSTM	---T---T---C---G---A---T---G---G---	2002
SIVSTM22579	---T---T---C---A---T---G---G---	835
HIV2F0784	---A---C---T---A---C---T---G---A---G---G---G---A---	286
SMCI8	---A---T---A---CAGAGAT---CT---AA---G---GA---C---A---	286
SMCI2	---A---T---A---CAGAGAT---CT---TC---G---G---G---A---	286

HIV2/SIV ENV

CONSENSUS . A	AtAtgTATGAacTaCAAAaatTaAAtAgCTgGGATgTtTTggcAAcTgGtTTGactTaaCCtCCTgGaTcaa?	1924
HIV2ROD	-----A-----T-----G-----G	2025
HIV2NIH2	-----G-----AC-----T-----C-----G-----G	1998
HIV2ISY	-----G-----G-----GC-----T-----F-----G-----G	2001
HIV2ST	-C-----T-----A-----C-----T-----G-----A	2007
HIV2BEN	-----T-----A-----C-----G-----G	2022
HIV2CAM2	-----C-----A-----AC-----G-----G	2028
HIV2D194	-----C-----C-----C-----T-----G-----A	1995
HIV2GH1	-----T-----GC-----C-----T-----G-----G	1998
HIV2CBL21	-C-----T-----T-----G-----GA	2028
HIV2CBL22	-----G-----G-----C-----G-----A-----G	2016
HIV2CBL23	-----T-----G-----G-----G-----G	2007
HIV2CBL24	-----G-----G-----G-----G-----G	2028
HIV2CAM1	-----G-----G-----G-----G-----G	2025
HIV2CAM2BR	-----A-----AC-----G-----G	2034
HIV2CAM3	-----CA-----T-----G-----G	2031
HIV2CAM4	-----G-----G-----G-----A-----T-----T-----G-----G	2031
HIV2CAM5	-C-----G-----C-----A-----T-----G-----GA	2010
HIV2CAM6	-C-----GT-----C-----G-----G-----G	2037
HIV2ALI	-CA-----AC-----C-----G-----G-----G	2016
HIV2S132E2	-----T-----C-----C-----A	360
HIV2T312A	-C-----T-----T-----G-----G-----A	360
HIV2PA	-C-----T-----C-----T-----G-----A	360
HIV2GB1222	-----GT-----C-----C-----G-----A	492
HIV2CI171	-----T-----C-----C-----A-----T-----A-----A	360
HIV2CI9	-----T-----C-----TC-----A	360
HIV2IS053	-C-----G-----G-----T-----A	360
HIV2VI390	-----T-----C-----T-----G-----A	360
HIV2V1905	-C-----G-----A-----A-----C-----G-----GGC	360
CONSENSUS . B	AcATGTATGA?tTaCAGAAAtTaAAt??CTGGGATgTTTTgGTAATTGGTTtGACTTCACCTCCTGGATggcA	1927
HIV2UC1	-----A-----AG-----C-----AAA-----	2034
HIV2D205	-----A-----A-----AG-----C-----AAA-----	2031
HIV2GH2	-----A-----G-----C-----AG-----C-----AAA-----	846
HIV2EHOA	-----G-----C-----CAA-----A-----CA-----AAA-----	2019
HIV2JA	-T-----GC-----CCAA-----C-----A	360
HIV2ON	-T-----GC-----G-----CAG-----A-----G-----G	360
CONSENSUS . C	ATATGTATGAGTTGCAGAAATTAAATAGCTGGGATgTTTTTGGCAATTGGCTTGATCTCACCTCTTGGGTAAAA	360
HIV2238	-----A-----G-----G-----G-----G-----G	360
CONSENSUS . D	AcAtGTATGAATtgCAAAAtTaAATAgctGGGATgTgTtGGCAAttGGtTTGAcctTaCtTcTTGGATaAaa	2027
SIVMM251	-----A-----G-----G-----G-----G-----G	2073
SIVMM32H	-----A-----G-----G-----G-----G-----G	2064
SIVMM1A11	-----A-----G-----G-----G-----G-----G	2064
SIVMM142	-----A-----G-----G-----G-----G-----G	2076
SIVMM239	-----A-----G-----G-----G-----G-----G	2067
SIVMNE	-----A-----G-----G-----G-----G-----G	2073
SIV5MMH4	-T-----A-----C-----G-----G	2085
SIV5MMH9	-T-----SK-----N-----T-----Y-----K-----Y-----	2088
SIV5MMFBJ	-T-----C-----A-----T-----G-----G	2097
SIVP209C19	-T-----T-----C-----A-----C-----G-----G	2088
SIVSTM	-T-----G-----C-----A-----C-----G-----G	2076
SIVSTM22579	-T-----G-----C-----A-----A-----G-----C-----G-----G	909
HIV2F0784	-----G-----A-----C-----A-----A-----C-----T-----T-----G	360
SMCI8	-----G-----G-----C-----TT-----C-----G-----C-----	360
SMCI2	-T-----G-----CC-----TT-----C-----G-----C-----	360

CONSENSUS.A	TATaTTCAaTATGGaGTTtataTAGTAgtaGgaatAaTAGcttTaAGhaATagtaaTatAtgTAgTaCAaaTGtT	1998
HIV2ROD	-----GCT-----A-----C-G-----G-----A-----G-----	2099
HIV2N1HZ	-----G-----G-----G-----A-----G-----	2072
HIV2ISY	-----CATG-----G-----C-C-----C-----	2075
HIV2ST	-----G-----CATG-----T-----	2081
HIV2BEN	-----GC-----G-----C-----C-----G-----T-----	2096
HIV2CAM2	-----T-----C-----C-----G-----T-----C-----	2102
HIV2D194	-----G-----CC-----A-----G-----T-----	2069
HIV2GH1	-----G-G-----T-----C-----A-----G-----T-----	2072
HIV2CBL21	-----C-T-----A-----G-----G-----C-----T-----A-----	2102
HIV2CBL22	-----C-T-----A-----G-----G-----C-----G-----	2090
HIV2CBL23	-----A-----A-----G-----T-C-T-----A-----G-----	2081
HIV2CBL24	-----G-----G-A-A-----C-T-----A-G-----G-G-----	2102
HIV2CAM1	-----C-----T-----CA-----G-----C-----	2099
HIV2CAM2BR	-----T-----C-----C-----	2108
HIV2CAN3	-----C-----C-----T-----C-----A-----	2105
HIV2CAM4	---C-----G-----T-A-A-----G-G-----C-T-----G-----A-----A-----	2105
HIV2CAM5	---G-----GC-----C-----TC-----G-----C-----	2084
HIV2CAM6	---G-----GC-----C-----TC-----G-----C-----	2111
HIV2ALT	-----A-----G-----C-T-----G-----	2090
HIV2S132E2	-----C-----C-----G-----C-----CA-----G-----C-----	434
HIV27312A	---G-----G-----C-C-----G-A-----A-----	434
HIV2EPA	---G-----G-----C-T-----A-----	434
HIV2GB1222	---C-----GTC-----G-T-----C-----	566
HIV2CI171	---G-----T-----G-----T-C-T-----	434
HIV2CI9	-----TA-----A-A-----G-T-----C-----A-----G-----T-----	434
HIV2IS053	---G-----A-----C-----T-C-T-----	434
HIV2VI390	-----G-----C-----G-----C-----G-----	434
HIV2VI905	---G-----G-----C-T-----T-----A-----	434
CONSENSUS.B	tAcaTcaTgTTAGGAcTATATgTAGTgcAGGatTAaTAGTgTAAAGaATAGCaATATACATTaTgCAGATGCT	2000
HIV2UC1	-----G-----	2108
HIV2D205	---T---CAT-----A-----G-----G-----G-----G-----	2105
HIV2GH2	---T---CAT-----GC-----C-----G-----G-----	920
HIV2EHOA	---G-----T-----A-----AT-----A-G-----A-----C-----A-----	2093
HIV2JA	---AA-----G-----A-----	434
HIV2ON	C---C-T-G-----C-----A-----G-----	434
CONSENSUS.C	TATATCTATTTAAGGATTTTATATAGTAGCAGGAGTAATAGTGTAAAGaATAGCTATCTATGTAGTACAAATGTT	434
HIV223B	-----	434
CONSENSUS.D	TATATAcAaatatGGagPttataTAGTtgTAGGagTaaTA?tgTTAAGaATagtaaTCTAT?TAGTaCAaaTgctT	2099
SIVMM251	-----A-----G-----C-----G-----A-----C-----	2147
SIVMM32H	-----C-----G-----A-----C-----	2138
SIVMM1A11	-----C-----A-----C-----	2138
SIVMM142	-----A-----A-----C-----G-----A-----C-----	2150
SIVMM139	-----C-----C-----G-----A-----C-----	2141
SIVMNE	-----C-----G-----A-----C-----	2147
SIVSMH4	-----T-ACT-----C-----GG-----A-G-----G-----	2159
SIVSMH9	-----T-ACTG-----C-----GGA-----G-A-G-----G-G-----	2162
SIVSMFBJ	-----T-ACTG-----C-----GGA-----G-A-G-----G-G-----	2171
SIVP209C19	-----T-ACTA-----T-----G-GG-----A-G-----G-G-----	2162
SIVSTM	---C-----A-----T-----A-----T-G-A-----G-C-----A-A-G-----T-----	2150
SIVSTM22579	---C-----G-A-----T-----A-----A-G-A-----G-C-----G-----G-----	983
HIV2F0784	---G-----C-TC-----A-----TA-----C-----A-----G-----	434
SMC18	---CT-TATA---T-----G-----A-----A-T-----GGA-----G-----CT-A-----T-A-----	434
SMC12	---CT-TATA---T-----G-----A-----A-T-----GG-----A-G-----CT-A-----T-A-----	434

HIV2/SIV ENV

CONSENSUS.A	AaGTAGacTtAGaAagGGCTATAGGCTGTPTTCTCtCCcCoCCcGgTTataTcCAACAGaTCCAtATCCaca	2072
HIV2ROD	-----G-----	2173
HIV2NIMZ	-----C-----	2146
HIV2ISY	-----T-----	2149
HIV2ST	-----C---CT-----	2155
HIV2BEN	-----T-----	2170
HIV2CAM2	-----G-----C---T-----	2176
HIV2D194	-----C-----	2183
HIV2GH1	-----T-----	2146
HIV2CBL21	-----T---G---C---T---C---G---	2176
HIV2CBL22	-----G-----G-----	2164
HIV2CBL23	-----CC-----	2155
HIV2CBL24	-----CC-----TC	2176
HIV2CAM1	-----G-----T-----	2173
HIV2CAM2BR	-----C-----	2182
HIV2CAM3	-----GT-----	2179
HIV2CAM4	-----T-----C-----C---T-----	2179
HIV2CAM5	-----C-----	2158
HIV2CAM6	-----C-----	2185
HIV2ALI	-----G---C---G---C---T---C-----	2164
HIV25132E2	-----	450
HIV27312A	-G-----GA	450
HIV2FA	-----G---	450
HIV2GB1222	-----G-----C-----	640
HIV2CI171	-----G---	453
HIV2CI9	-----	453
HIV2IS053	-G-----G---	453
HIV2VI390	-----T-----	453
HIV2VI905	-G-----G---	453
CONSENSUS.B	AGCAAgacTTAGGAAGGGCTATAGGCCAGTATtTCCTCCCTCCCTCTTATACTCAACAGATCCCTATCCGCA	2074
HIV2UC1	-----	2182
HIV2D205	-----	2179
HIV2GH2	-----T-----	994
HIV2EHOA	-----G-----	2167
HIV2JA	-----A-----	450
HIV2ON	-----G-----	450
CONSENSUS.C	AATGAGACTTAGGAAGGGC	453
HIV2238	-----	453
<- premature stop in some SIVs		
CONSENSUS.D	aGcTAaGttaAGgcAgGG?TATcGCCcAGTGTCTCTCCCG?CCGctCTTaTtt?cAGcAGA?cCaTAtCcaa?	2168
SIVMN251	-----G-----A-----C---T---CT---C---C	2221
SIVMN32H	-----A---G-----A-----C-----C-----C	2212
SIVMN1A11	-----A---G-----A-----C---T---C-----C	2212
SIVMN142	-----G---A---G-----A-----C---T---C---C---C	2224
SIVMN239	-----G-----A-----A-----C-----C-----C	2215
SIVMNE	-----A---G-----A---T-----C---T---C-----G-C	2221
SIVSMH4	-----G---A---T-----T---G---G---T-----T---C-----CA	2233
SIVSMH9	-----G---A---T-----T---G---R---G---T---T---C---Y-RA	2236
SIVSMMPBJ	-----G---A---T-----T---G---G---T-----T---C---GA	2245
SIVP209C19	-----G---A---T-----T---G---G---T-----T---C---CA	2236
SIVSTM	-----G-C-T---A-A-T---C---T---T---G-CGT---T---C---CA	2224
SIVSTM22579	-----G-C-T---A-A-T---T---T---CAT---T---T---CA	1057
HIV2P0784	-----TA---	450
SMCI8	G-G---CAT---A---	450
SMCI2	-G---C-T---A---	450

	1/3' s3	
CONSENSUS.A	aGgAcCGGgACAGCcagcCagaGaAGaAACAGAAgAGACGttGg?aCaaCg?TGGagaCagaT?cTggCct	2143
HIV2ROD	-----G-----AC-----G-----AG-----G-----A-----C	2247
HIV2NIH2	-----A-----G-----AC-----G-----AG-----G-----CT-----	2220
HIV2ISY	-----T-----A-----GA-----G-----T-----AG-----C-----	2223
HIV2ST	-----T-----A-----G-----T-----AT-----GG-----C	2229
HIV2BEN	-----G-----AC-----G-----C-----GG-----G-----TT-----GAC-----TG-----	2244
HIV2CAM2	-----C-----C-----G-----GG-----G-----C-----	2250
HIV2D194	C-----A-----G-----AC-----CC-----G-----G-----AG-----TTT-----G-----C-----TG-----	2217
HIV2GH1	C-----G-----AC-----G-----ACA-----G-----G-----A-----TT-----GAC-----TG-----A-----	2220
HIV2CBL21	-----C-----G-----C-----C-----G-----G-----T-----C-----C-----	2250
HIV2CBL22	-----T-----A-----G-----G-----C-----C-----AG-----C-----C-----	2238
HIV2CBL23	-----T-----A-----T-----A-----C-----T-----C-----TG-----A-----	2229
HIV2CBL24	G-----C-----T-----A-----G-----G-----T-----GAC-----C-----TG-----	2250
HIV2CAM1	-----T-----T-----AG-----T-----T-----C-----	2247
HIV2CAM2BR	-----C-----C-----G-----G-----C-----	2256
HIV2CAM3	-----A-----C-----A-----CAT-----G-----C-----	2253
HIV2CAM4	C-----AA-----G-----G-----T-----C-----TG-----	2253
HIV2CAM5	-----TT-----AC-----G-----G-----A-----C-----AC-----C-----C	2232
HIV2CAM6	-----T-----AC-----G-----A-----G-----T-----CT-----C	2259
HIV2ALI	-----A-----A-----G-----G-----T-----G-----TG-----C	2238
HIV2GB1222	-----A-----A-----AG-----T-----T-----C	714
CONSENSUS.B	AggaCOGGGGACAGCCAGCCAAAGAAAGAAACAGAAAGaGgAGGTGGAAaAcgAcGgGGgcTACAGATCtTGGCCC	2148
HIV2UC1	-----AC-----A-----CT-----	2256
HIV2D205	-----G-----A-----C-----	2253
HIV2GH2	-----G-----A-----	1068
HIV2EHOA	-----A-----A-----	2241
CONSENSUS.D	aGG?cCaGG?acTGCcAACCAaAGAAAGGaaaAgAAggaGacGGTGGAGacaGcGGTGGcCaCAgcTCcTGCCCT	2240
SIVMM251	-----A-----C-----C-----G-----C-----AG-----	2295
SIVMM32H	-----A-----C-----C-----G-----C-----AG-----	2286
SIVMM1A11	-----A-----C-----C-----G-----C-----AG-----	2286
SIVMM142	-----AT-----C-----CT-----C-----A-----G-----	2298
SIVMM239	-----A-----C-----C-----G-----C-----A-----AG-----	2289
SIVMM2	-----A-----C-----C-----G-----C-----AG-----	2295
SIVSMH4	-----G-----A-----C-----G-----A-----G-----A-----T-----	2307
SIVSMH9	C-----G-----A-----G-----G-----R-----A-----A-----T-----	2310
SIVSMHBJ	C-----G-----A-----G-----G-----A-----A-----T-----	2319
SIVP209C19	-----G-----A-----C-----G-----A-----A-----G-----A-----T-----	2310
SIVSTM	-----G-----A-----A-----C-----A-----GA-----A-----A-----T-----A-----	2298
SIVSTM22579	-----G-----A-----A-----T-----C-----A-----GA-----	1098
	<- tat	
CONSENSUS.A	TggCcgatagcatatatacatcttctgatccgctgctgattcgccctcttgacgggaactatacaacAtCTGCAG	2217
HIV2ROD	-----A-----A-----G-----	2321
HIV2NIH2	-----A-----G-----	2294
HIV2ISY	-----A-----C-----A-----G-----	2297
HIV2ST	-----AG-----A-----A-----	2303
HIV2BEN	-----A-----AAC-----G-----G-----A-----A-----C-----T-----G-----	2318
HIV2CAM2	-----A-----A-----A-----A-----C-----	2324
HIV2D194	-----AC-----AAC-----C-----A-----A-----A-----C-----G-----G-----	2291
HIV2GH1	-----AAC-----A-----A-----A-----C-----G-----G-----	2294
HIV2CBL21	-----A-----C-----TT-----	2324
HIV2CBL22	-----A-----AG-----A-----AG-----TT-----G-----G-----G-----	2312
HIV2CBL23	-----AG-----A-----A-----	2303
HIV2CBL24	-----A-----A-----T-----	2324
HIV2CAM1	-----C-----A-----G-----	2321
HIV2CAM2BR	-----A-----A-----	2276
HIV2CAM3	-----C-----C-----A-----G-----	2327
HIV2CAM4	-----CC-----G-----A-----A-----G-----G-----	2327
HIV2CAM5	-----A-----AG-----A-----G-----	2306
HIV2CAM6	-----C-----T-----A-----AGC-----T-----G-----G-----G-----	2333
HIV2ALI	-----C-----T-----A-----AGC-----T-----G-----G-----G-----	2312
HIV2GB1222	-----A-----A-----G-----T-----G-----G-----	788
CONSENSUS.B	TGGCAGATaGAATACattCAcTTtCtg?TtCGCCA?CtGAGGaaACCTCTTgATTPTGGCTaTACAAcGcCTGCAG	2220
HIV2UC1	-----GC-----A-----A-----G-----G-----	2330
HIV2D205	-----T-----AC-----G-----G-----A-----	2327
HIV2GH2	-----C-----C-----C-----A-----G-----C-----A-----	1142
HIV2EHOA	-----C-----C-----C-----A-----G-----	2315
CONSENSUS.D	TgGCAGATAGAAATATATCATTTCTTgATcCGCCaacTgaTACGCCCTCTGACTTGGcTaTtCAgCAacTgCAG	2314
SIVMM251	-----A-----	2369
SIVMM32H	-----A-----	2360
SIVMM1A11	-----A-----	2360
SIVMM142	-----A-----	2372
SIVMM239	-----A-----	2363
SIVMM2	-----A-----T-----	2369
SIVSMH4	-----GT-----G-----	2381
SIVSMH9	-----R-----G-----GT-----	2384
SIVSMHBJ	-----GT-----GT-----	2393
SIVP209C19	-----A-----T-----G-----G-----A-----A-----T-----	2384
SIVSTM	-----A-----C-----G-----AG-----G-----A-----A-----T-----	2372

HIV2/SIV ENV

CONSENSUS.A	ggacCTTAcTATCCAgga?cttccc?gacccCTCcaacc?atcttccaga?cttc?????????aca?c??tca	2272
HIV2ROD	-----G-----T-----TC-----A-----	2374
HIV2NIH2	-----T-C-C-T-----A-----G-----CAGAGAGCACTA-----G-AA-----	2369
HIV2ISY	-----A-----T-C-TG-C-----ACCA-----	2338
HIV2ST	-----G-----A-----TA-C-----G-----TCGAGAGCATTG-----G-AG-----	2377
HIV2BEN	A-----A-A-C-C-----G-G-TG-C-----G-----A-----G-AA-----	2380
HIV2CAM2	-A-----T-----A-----C-----A-----	2377
HIV2D194	-G-----A-A-C-C-----G-G-TG-C-----G-----A-----G-AA-----	2393
HIV2GH1	A-----C-A-C-C-----A-G-TG-C-----A-----A-----G-AA-----	2356
HIV2CBL21	AAC-----A-----A-G-----A-----A-----C-----GG-----TCAGAGAGCACTG-----G-AA-----	2398
HIV2CBL22	-----G-----T-----T-----G-----A-----	2353
HIV2CBL23	-A-----T-C-T-----G-AG-T-----G-----TCAGAGAGCACTG-----G-AA-----	2377
HIV2CBL24	-----G-C-T-T-----A-C-----G-----CAGAGAGCACTA-----G-AA-----	2398
HIV2CAM1	-----T-----T-----C-----A-----	2374
HIV2CAM2BR	AA-----T-----A-----C-----A-----	2329
HIV2CAM3	-----G-----T-----A-----A-----	2380
HIV2CAM4	-----G-----C-----AGAGCA-----CTG-----G-GA-----	2380
HIV2CAM5	-A-----A-G-----T-----TA-----	2347
HIV2CAM6	-----G-----T-----C-----A-----	2386
HIV2ALI	-----T-----TC-T-----G-----G-----TCAGAGAGCGCTG-----A-AA-----	2386
HIV2GB1222	-A-----T-----T-----G-C-----A-----	841
CONSENSUS.B	aAcCTTAcTgTgAAGACCTtCCaAA?CCTCCAAcCA?ta?????????ctccaa	2364
HIV2UC1	G-A-----C-----T-----GC-----	2376
HIV2D205	-----T-----C-A-A-CTCAACCAAT-----	2382
HIV2GH2	C-----G-----C-----A-----CTCCAAACCG-----	1191
HIV2EHOA	-----T-A-CA-----C-----G-----	2361
CONSENSUS.D	asccTtgCTATcgAGag?aTaCCAgatCCTCCAAcCAatacTCCAgAGgcTCTC?acGacccTaCagAGaattc	2386
SIVMM251	-----C-----TG-----GA-GG-----	2443
SIVMM32H	-----C-----TG-----G-----	2434
SIVMM1A11	-----T-----TG-----GA-G-----	2434
SIVMM142	-----C-----T-----TG-----G-----	2446
SIVMM239	-----T-----TG-----G-----	2437
SIVMNE	-----C-----AT-----C-----C-----	2443
SIVSMNH4	GGAT-G-----T-----GAGC-----G-G-----C-----A-----GT-G-A-----G-C-----	2455
SIVSMNH9	GGAT-G-----T-----GANC-G-----A-----G-----G-----C-----A-G-----G-----GCC-----	2458
SIVSMMPB3	-GAT-G-----T-----GAAC-G-----A-C-----G-----G-----C-----A-G-----G-----GCC-----	2467
SIVP209C19	GGAT-G-----T-----GAGC-----G-G-----A-----A-----G-----G-----A-----G-----	2458
SIVSTM	GG-T-GT-----C-----GATC-----C-----C-----CCT-----A-----A-----C-G-----TA-----G-----A-----	2446
<- rev cds		
nei cds ->		
CONSENSUS.A	GaGACTGGctgAGaacttaagcaGCctaccTCCAaTATGGGtGCgAGTGGATCCAAGaAGCgtTCCagGCcttc	2346
HIV2ROD	-----G-A-----T-T-----A-----GC-----	2448
HIV2NIH2	-----G-----G-----G-----A-----C-T-----	2442
HIV2ISY	-----C-G-----G-----C-----	2412
HIV2ST	-----T-----CA-----G-----GA-----	2451
HIV2BEN	-G-----G-G-----C-A-----T-----A-A-----	2454
HIV2CAM2	-----G-C-----T-----GCT-----	2451
HIV2D194	-G-----G-G-----GA-A-----	2427
HIV2GH1	-G-----G-----G-----A-----A-A-----	2430
HIV2CBL21	-----C-OGG-----C-----A-----	2472
HIV2CBL22	-G-----G-----T-----G-----G-----	2427
HIV2CBL23	-----A-----AC-G-C-----T-----G-----G-----	2451
HIV2CBL24	-----G-----G-----G-----CC-----A-C-----	2472
HIV2CAM1	-----T-----OC-GTA-----T-----A-----	2448
HIV2CAM2BR	-----G-C-----T-----GCT-----	2403
HIV2CAM3	-----G-A-----T-----C-C-----	2454
HIV2CAM4	-----T-----G-----G-----A-----T-----	2454
HIV2CAM5	-G-----A-----CA-----T-T-----	2421
HIV2CAM6	-----G-A-----T-----	2460
HIV2ALI	-G-----G-----G-----T-----	2460
HIV2GB1222	-----C-----G-A-----T-T-----A-----GC-----	915
CONSENSUS.B	ccgCTCAGGCT?cc??tTGC?TACCTCCagTATGGGATcagCTGGTCCCAAGaAGC??TCC?AGCAGCA	2326
HIV2UC1	TT-AC-----A-----AA-----A-----	2445
HIV2D205	-A-----C-CC-----G-----GC-----G-----G-----	2451
HIV2GH2	-C-TGC-----G-----CA-----G-----GC-----G-----	1257
HIV2EHOA	-T-ACC-----A-----GA-----AA-----A-----	2430
CONSENSUS.D	GaGAAGTCcTCAGgattgaactaaCCTAccTACAaTATGGGTGgagcTattTCCAaGAaGCgGtcCAagccggc	2460
SIVMM251	-----C-----G-----T-----G-----	2517
SIVMM32H	-----C-----G-----G-----T-----	2508
SIVMM1A11	-----C-----G-----A-----T-----G-----G-T-A-----	2508
SIVMM142	-----C-----G-----C-----	2520
SIVMM239	-----C-----G-----T-----G-----T-----	2511
SIVMNE	-----C-----C-----G-----T-----C-----	2517
SIVSMNH4	-T-----A-----A-----G-----C-----A-----A-----GT-G-----	2529
SIVSMNH9	-T-----A-----AG-----GA-----G-----C-T-CC-----A-CG-----RRG-T-G-----	2532
SIVSMMPB3	-T-----A-----AG-----C-GA-----G-----G-----C-T-CC-----A-CG-----G-GT-G-----	2541
SIVP209C19	-T-----A-----A-----A-----G-----G-----A-----G-----G-GT-G-----	2532
SIVSTM	-----G-----AC-A-G-GC-G-----TT-----G-----C-T-GGA-----A-CG-----A-CA-----	2520

CONSENSUS.A	gc?agggcTAc?agAGAGACTcCTTgCgggCgctGgaGgggctTgTgpggggca?tggtacgaaTcGGgAGggg	2416
HIV2ROD	--G-----A-----A-----C-----A-----T-T-A-----	2522
HIV2NIH2	--A-----A-----A-----A-----G-----A-A-----A-----GC-CA-----	2516
HIV2ISY	--G-----T-----A-----A-----A-----G-----A-----TC-----G-----A-----	2486
HIV2ST	--G-----GG-----A-AAA-----C-----A-----C-----A-C-G-A-----T-----	2525
HIV2BEN	--G-----A-----G-----A-----G-AT-G-A-----AA-----GC-CG-----C-----	2528
HIV2CAM2	--A-----G-A-----A-----C-----A-G-----AATGT-ACA-----C-T-A-----	2525
HIV2D194	--A-----A-G-G-----A-----A-----G-A-T-AA-----GC-CA-----C-----	2501
HIV2GH1	--A-A-A-----G-----A-----A-----G-----G-A-T-CA-----G-CA-----CG-----	2504
HIV2CBL21	--G-----G-----A-AA-----TC-----AAA-----C-T-----A-A-----GG-A-----T-----	2546
HIV2CBL22	--A-AATT-----G-----A-----A-A-----G-C-----G-----A-----A-----	2501
HIV2CBL23	--A-----G-----A-----A-----A-----C-CAG-----G-----	2525
HIV2CBL24	A-A-----AC-----A-----A-----AA-----C-CAG-----C-----	2546
HIV2CAM1	--GG-----A-----G-A-A-----A-A-----C-----A-A-----T-----T-----	2522
HIV2CAM2BR	--A-----G-A-----A-----C-----A-----AA-TGT-ACA-----C-T-A-----	2477
HIV2CAM3	--G-----A-----A-----A-AA-----C-----A-A-----T-----ACAATT-T-----	2528
HIV2CAM4	--A-----A-G-G-----A-----AA-----A-----A-----G-----C-----A-----	2528
HIV2CAM5	--G-----G-----A-----A-----AG-C-----A-C-----A-----T-----G-----T-----T-----	2495
HIV2CAM6	--G-----ATT-G-A-----A-----C-----A-----A-----C-----G-----T-----T-----	2534
HIV2ALI	--A-----AT-G-G-----G-----A-AAA-A-C-----A-A-----A-CAGT-GG-----A-----	2534
HIV2GB1222	--G-----A-----C-----C-----A-----T-----G-A-----T-----	989
CONSENSUS.B	gCaAGGcT?CG?GAGAGACTCTTGGGag?gCgGga?gacCcttaTGGga?g??CTCAGgAG?cgCGgggAGc	2390
HIV2UC1	A-G-----G-A-----ACA-A-----AGAG-----C-----A-A-CT-----A-GA-A-----A-----	2519
HIV2D205	--A-----A-G-----A-----T-----C-----GA-----G-CT-----AG-----A-----	2525
HIV2GH2	--T-----A-----A-----T-----GA-----C-----GATC-----GAT-----	1331
HIV2EHOA	--C-----G-G-----C-----CGAG-----CG-----GA-TC-----AG-----A-----A-----	2504
CONSENSUS.D	tGgAGAtcTGcgacAGAGACTcCTTGGCGGGCGGAGACTATGgGAGaCTCTTggaAGgGtTGGaAGatG	2534
SIVMH251	-----A-----A-----A-----A-----A-----A-----A-----A-----	2591
SIVMH32H	-----A-----A-----A-----A-----A-----A-----A-----A-----	2582
SIVMH1A11	-----A-----A-----A-----A-----A-----A-----A-----A-----	2582
SIVMH142	CA-----G-----A-----G-----CA-----G-----	2594
SIVMH239	-----A-----A-----A-----A-----A-----A-----A-----A-----	2585
SIVMNE	-----G-----A-----	2591
SIVSMH4	--A-CT-----CG-----AA-----A-----G-----G-----G-----G-----	2603
SIVSMH9	--A-T-----YRCG-----AA-----A-----G-----G-----G-----G-----	2606
SIVSMH9B3	--A-T-----TACG-----AA-----A-----A-----G-----G-----G-----G-----	2615
SIVP209C19	--A-T-----CG-----A-----A-----A-----A-----G-----G-----G-----G-----	2606
SIVSTM	--G-----G-----G-----A-----G-----A-----A-----A-----G-----A-----	2594
CONSENSUS.A	aaTAcTtgcagTcCCaAGaAGgATCAGgCAGGGAGCAGAAaTcgCcCTCCTgTGA	2471
HIV2ROD	-----C-----G-----T-----A-----	2574
HIV2NIH2	G-----C-----T-----C-----A-----C-----	2569
HIV2ISY	G-----C-----T-----C-----A-----T-----	2538
HIV2ST	-----G-----	2577
HIV2BEN	--C-----T-----A-----C-----	2580
HIV2CAM2	--C-----G-----A-----C-----T-----	2577
HIV2D194	--C-----C-----	2553
HIV2GH1	--C-----C-----A-----G-----	2556
HIV2CBL21	--C-C-----A-----A-----GC-----	2601
HIV2CBL22	--C-----T-----C-----A-----G-----	2556
HIV2CBL23	-----C-----A-----A-----G-----A-----	2580
HIV2CBL24	-----C-----A-----A-----C-----	2601
HIV2CAM1	G-----G-----G-----C-----T-----	2577
HIV2CAM2BR	-----TG-----A-----A-----C-----T-----	2532
HIV2CAM3	-----T-----A-----A-----G-----	2583
HIV2CAM4	G-----C-----T-----G-----A-----C-----	2583
HIV2CAM5	-----C-----T-----G-----A-----C-----	2550
HIV2CAM6	-----C-----GA-----A-----A-----G-----	2589
HIV2ALI	G-----C-----G-----G-----G-----T-----	2589
HIV2GB1222	--C-----G-----A-----G-----A-----	1044
CONSENSUS.B	?ATCaTcGCgATcCCCAGGAG?ATCAGACAAGGactTGAACtTgCCCTCTTGTGA	2443
HIV2UC1	A-----A-----G-----	2571
HIV2D205	A-----A-----A-----A-----	2577
HIV2GH2	G-----G-----A-----A-----	1383
HIV2EHOA	G-----T-----A-----A-----G-----GGC-----	2559
CONSENSUS.D	gATactcGCaATCCCTaGgaG7aTcAGaCAAGGGCTTgAGCTcactCTCTTGTGA	2588
SIVMH251	-----C-----G-----T-----G-----G-----	2643
SIVMH32H	-----G-----T-----	2634
SIVMH1A11	--C-----A-----T-----	2634
SIVMH142	--C-----G-----T-----	2646
SIVMH239	-----C-----G-----T-----	2637
SIVMNE	-----G-----C-----T-----	2643
SIVSMH4	-----C-C-----G-----	2655
SIVSMH9	A-----HR-AC-C-----G-----	2658
SIVSMH9B3	A-----A-AC-C-----G-----	2667
SIVP209C19	-----C-CG-----G-----	2661
SIVSTM	---CGGG---C-----A-----C-----A-----A-----	2646

HIV2EHOA

LOCUS HIV2EHOA 8636 bp ds-DNA VRL 11-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) complete genomic sequence, isolate HIV2EHOA.
 ACCESSION U14545
 SOURCE Human immunodeficiency virus type 2 DNA derived from the peripheral blood lymphocytes of a male patient from the Ivory Coast with full-blown AIDS.
 REFERENCE 1 (bases 1 to 8636)
 AUTHORS Rey-Cuille,M.A., Galabru,J., Laurent-Crawford,A., Krust,B., Montagnier,L. and Hovanessian,A.G.
 TITLE HIV-2 EHO isolate has a divergent envelope gene and induces single cell killing by apoptosis
 JOURNAL Virology 202, 471-476 (1994)
 COMMENT The HIV-2 EHO sequence codes for a viral strain which is highly virulent, binds soluble CD4 receptors with high affinity, down-regulates CD4 receptors with accelerated kinetics and does not form syncytia. The divergent genotype of HIV-2 EHO is most apparent in its cell surface glycoprotein, which is 25 kDa smaller than other currently identified HIV-2 strains. However, this discrepancy in protein size is not evident when comparing predicted molecular weights. Rey-Cuille et al. report that the EHO envelope sequence contains fewer N-linked glycosylation sites than other HIV-2 sequences and further suggest that the presence of fewer oligosaccharide chains on the modified protein may in part explain the difference in protein size. The fact that HIV2 EHO does not induce syncytia and is highly virulent is unusual since non-syncytium-inducing isolates are usually associated with decreased virulence. HIV-2 EHO clusters with the B subtype HIV-2 viruses.

BASE COUNT 2972 a 1759 c 2066 g 1839 t

ORIGIN 549 bp downstream from the start of gag

```

1 cacccttat gatattaatc agatgttaaa ttgtgttagga gaacatcagg cagccatgca
<- gag cds
61 aattattagg gagataatta atgaggaagc agcagactgg gaccaacage atccatgccc
121 aggcccaatg ccagcagggc aactcagggg accaagaggg tcagacatag cgggaaccac
181 cagcacagta gaagaacaga tacagtggat gtacagaccc caaaatcctg tcccagtggg
241 aaacatctat agaaggtgga ttcagttagg gctacagaaa tgtgtccgga tgtacaatcc
301 tactaatata ctggacataa agcaagggcc aaaagaaccc ttccaaagct atgtagatag
361 attctacaaa agcttacggg cagagcaaac agaccagca gtaaaaaatt ggtgacaca
421 aacactgctg attcagaatg ctaaccacga ttgcaaatg gtgcttaagg gcttgggaat
481 gaaccccaca ttgaagaaa tgctaacagc ctgccagggg ataggaggcc cagggcagaa
541 ggcaaggcta atggctgagg ctttaaaagc agccttgaca ccttcacca atccgtttgc
601 cgccgctcaa ccaagagcag ggaagaggac agtgacatgc tggaaactgc gcaaggcggg
661 acatacagcc aggcagtgc aagcccttag aaggcagggA TGctggaaat gtggacagca
pol cds start ->
721 aggacacatc atgtcaaaat gccagaaaag acaggcgggt tttttagggt tcggccctctg
781 gggaaagaag cctcgcaatt tccccgtcca ggcacccag gggatagtg catctgcgcc
841 cccgatgaac ccagcattcg gcatgacacc tcagggtgcg attccatctg caccocctgc
901 agatccagca gaggagatgc taaagaacta catgcaacta gggaaagaagc agaaggagaa
961 cagagagaga cctacaagg aggtgacaga ggatttgctg cactcgaatt ctctctttgg
1021 agaagaccag TAGtcaaagc aactattgag ggtcaatcag tagaagtatt actagacaca
<- gag cds end
1081 ggagctgatg actcaatagt agcagggata gaattaggca gcaattacac cccaaaaata
1141 gtaggtggga taggaggatt tataaatacc aatgaataga aaatgtaga aatagaagta
1201 gtaggaaaaa gagtaagagc aacagtaatg acaggggaca cccaataaa catttttggc
1261 agaaatatat taaatagctt aggcagact ctaaatctcc cagtagcaag gatageacca
1321 gtaaaagtcc agttaaagcc tgaaaaagat gggccaaaaa tcagacaatg gcccctatcc
1381 aaagagaaaa tactagocct caaagaaatc tgtgaaaaaa tggaaaaaga gggacagtta
1441 gaagaggcgc ctctactaa tccatacaat tcgccaccc tcgccataaa aaagaagac

```

```

1501 aaaaacaaat ggaggatgct aatagatttc agagaactaa acaaggtaac ccaagaattt
1561 acagaggtcc agctgggtat tcttcaccca gcaggactgg catcaaagaa aagaataaca
1621 gtactagatg taggagatgc ctacttcagt gtcccactag atccagactt cagacaatat
1681 acagcattta ctttgccagg agtaataaat gcagaaccag gaaagagata tctttacaaa
1741 gtccctaccac agggatggaa gggatcccca gcaattttcc agtacaccat ggcaaaggta
1801 ctagaccctt tcagaaaagc caacaatgat gtcactataa tccagtacat ggatgacatt
1861 ctcggtggcaa gtgacaggag cgatctggag catgacaggg tagtgtctca actaaaagag
1921 ctattaaata acatgggatt ctctactcca gaagaaaagt tccaaaaaga cctccattc
1981 aaatggatgg ggtatgagct ctggccaaag aaatggaaac tgcaaaaaat acagctacca
2041 gaaaaagagg tttggacagt aaatgacatt cagaagttag tgggagtatt aaattgggca
2101 gctcaacttt tccgggggat taagaccagg catatatgta aactaataag gggaaagatg
2161 acctaacag aagaggtaca atggactgaa ttggcagagg cagaattcca ggaacacaaa
2221 atcatcctag aacaagagca ggaaggatcc tattacaaag aaggggtacc tttagaagca
2281 acagtgcaga aaaatctagc aaatcagttg acatacaga ttcatcaggg agataaaatc
2341 ctaaaagtag gaaaatctgc aaagggttaa aacactcaca ccaatggagt aagactattg
2401 gctcatgtag tccaaaaaat aggaaaggaa gcattggtca tctggggaga gataccaatg
2461 ttccatctac cagtagaaag agagacatgg gatcagttgt ggacagatta ctggcaagta
2521 acctggatcc cagaatggga tttgtctca accccaccat taataagggt agcctataac
2581 ctggtcaaaag accccctaga aggagtagaa acttactaca cagatggatc ctgtaacaaa
2641 gcctcaaaag aagggaagc aggatatgtc acagacaggg gaaaggataa agttaaacca
2701 ttagaacaaa caacaaatca gcaagcagag cttgaagcat ttgcaactgc actacaggac
2761 tcaggaccac aggtcaatat catagtagat tcacaatatg tcatgggaat agtagctgca
2821 caaccaacag aaacagaatc accgatagta agagaaataa ttgaagaat gatcaaaaag
2881 gaaaaaatat atgtaggatg ggtaccagct cacaagggac tgggtggtaa tcaggaaagta
2941 gaccacctag tgagccaagg aattagacaa atcctatttc tagaaaaaat agaaccagct
3001 caagaagaac atgaaaaata tcataataat gtaaaagaac tagtccataa atttgggatt
3061 ccacaattag tggcaagaca aatagtaaat tctgttgata aatgccaca aaaaggggaa
3121 gctattcatg gacaggtaaa ttcagaacta gggacatggc aaatggactg tacacattta
3181 gaggggaaag ttataatagt ggcagttcat gtagccagtg gattcataga agcagaagta
3241 ataccocaaag aaacaggaag acagacagct ctcttctgtg taagctggc cagcagatgg
3301 cctatcacac acctgcacac agacaacggt gccaaactca cttcacaaga tgtgaaaatg
3361 gcagcctggt ggtatgggat agaacaaaaca ttgggagtgc cctataatcc agaaagtcag
3421 ggagtagtag aagcaatgaa ccatcctctg aaaaatcaga tagacagsat tagagatcag
3481 gcagtatcaa tagagacagt tgtgttaatg gcaactcact gcatgaattt taaaagaagg
3541 ggaggaatag gggatatgac cctgcagaa agaatagtca acatgataac tacagaacaa
3601 gaatacaaat tctccaaaac aaaaaattta aaattccaaa atttccgggt ctattacaga
3661 gaaggucagag atcaactctg gaagggaacct ggtgatctat tgtggaaagg ggaaggagca
3721 gtcacataaa aggtagggac agaaatcaaa gtaataccca gaagaaaagc aaagatcata
3781 aggaactATG gaggggaaa agaattggat tgcagtgccg acgtggagga taccatgcag

vif cds start ->
3841 gctagagagg tggcacagtc taatTAAata cctgaatat aggacaaaag acttgcaaca
      <- pol cds end

3901 ggtctcttat gttcctcacc ataaggtagg atgggcttgg tggacttgca gtagagtaat
3961 atttcccttg aaagaaggag cacatctaga agtccaagga tactgggaacc tgaccccaga
4021 aaggggattc ttgagttcct atgctgtaag actaacatgg tatgagagga gcttttatac
4081 agatgtaact cctgatgtag cagaccgatt actgcatggg tcttatttct ctctgtttac
4141 agctaataaa gtaaggagag ccatcagggg agaaaagata ttgtccact gcaactaccc
4201 atcagctcat acagggcagg taccaagttt acagtttcta gccctaagag ttgtacaaga
4261 aggaaaagAT Ggatcccagg gagagagtag caccaggaaa cagcggcgaa gaaacagtag

vpX cds start ->
4321 gagaggcatt cgaatggcta gagacaacat tagaacatct caacagagta gcagtcaatc
4381 acttggccag ggaacttatt ttccagggtct ggcagaagtc ttgggcatat tggcgtGAgg
      <- vif end

4441 aacagggcat gtcaattagt tataccaaat ataggatttt gtgtctaatt caaaaagcaa
4501 tgtttataca ctttgcaaag ggttggtggAT Gtctgcggga aggccatggg ccagggggat
      vpR cds start ->
4561 ggagatcagg acctcctect cctcctcccc caggtctage cTAAtggcag aagcagtcct
      <- vpX cds end
4621 agagattcct ccagaggata aaaaccacaa aagagaaccg tgggaacagt gggtagtgga

```

HIV2EHOA

```

4681 cgctcctggag gaaataaaac aagaagcctt aaagcatttt gacccccgtt tattaactgc
4741 acttggcaat tttatctaca ataggcATGg aaataccctt gaaggagcag gagagctcat
      tat cds start ->
4801 taaactectc cagcggggc acctectcca ctccagaggg ggttgccaac actcaaggat
4861 tggacaaccg gggggaggaa atccctctctc agctataccg cccctctTAAG gcctgcagta
      <- vpR cds end
4921 atacatgcta ctgtaaaaaa tgttcttacc attgccagct ttgctttctt aaaaagggcc
4981 tggggatatg ttATGaacgc tcgagaaaga gatcttcaaa aagggtctag actactacat
      rev cds start ->
5041 cttctgcacc aaacgagtga gtATGgcaca tgttaataat tacctacttg ttacactcct
      env cds start ->
5101 gcttataagt atctatgggt atatgggcaa gaactttgtc actgtcttct atgggtatacc
5161 cgcattgaaa aatgcacaa tccccctctt ttgtgtctac agaaacagag atacctgggg
5221 aactgtacaa tgcctcccag ataattgatg ctataccgaa atccaattaa atataacaga
5281 ggcctttgat gcctgggata atacagtgc acatcaggca acaaaggatg tgtggagctc
5341 ctttgagacc tcaataaaac catgtgttaa attaacccca ctgtgtgtga caatgaagtg
5401 taataagacg tggagctcag ctagcaaaag gaccactacg tctcccggt cctcagatc
5461 ttctactcag accctactca acgaagatag caaatgtatt caaaatgaca gctgtgcagg
5521 gataggacta gaagaaatga tagactgtca attcaaaatg acaggattaa aaagagatga
5581 gtcaaaaaca tataaggaca cctgggtataa acaagattta gtgtgtgaaa aggggacaaag
5641 gagtaatgaa agcaaatgtt atataaaaac ctgtaataca tcaattatoc aagaatcatg
5701 tgacaaacat tattgggata gtttaagatt taggtattgt gctccccag gatttgcctt
5761 gctaagatgc aatgatacta aatattcagg ctctatgccc aactgtagta aggttagtagt
5821 atctctgtac agaattgatg aaacacagac ctctacatgg ttggcttca atgggtacaag
5881 ggcagaaaaac aggcacatata tctattggca tggtaaaagc aataggacta tctaaagctt
5941 aaattcatat tataatctga caatgcactg taaaaggcca ggaacaaaaa tgggtgtgccc
      HIV-1 V3 loop equivalent ->
6001 aataagaacc gtgtcaggta ttctcttcca ttccagacct atcaataaaa gacctaaaca
6061 agcttgggtc tgggttaaag gaaactggac agaagccata caggagggtga aagagaccat
      <-
6121 taaaaatcat cccagatatt caggaacaac aaatatctca cagataagggt tagcagagca
6181 cgcgagaagc tcagatccag aagtaagata tatgtggact aactgtaggg gagagtttct
6241 ctactgtaat atgacttttt tcttaaactg ggtagaaaat agaactgggc tcaagagaaa
6301 ttatgctca tggcacatca gacagatagt caacacgtgg cacaagattg gaagaaatgt
6361 gtatttgctt ccaagagagg gtgaactctc ctgtaatcc actgttacca gctcatagc
6421 caacattgac ttgatagata agaactctac taatattact gtgagtgcag aagtgtcaga
6481 actgtataaa ttggaaactg gggactacaa attagttagg ataacaacca ttggctttgc
6541 acctacaagt ataaaaagat attcctcagt gacaccgagg aataaaaagag gtgtacttgt
6601 gctaggggtt ttgggattcc ttgcgacggc aggtttctga atgggcggcg cgtccttgac
6661 gttgtcggct cagtcctcga ctttactggc tgggtagtg cagcagcagc aacagctcgt
6721 ggacgtggtc aaaagacaac aagaactgtt gcggctgacc gtctggggga cgaaaaacct
6781 ccaggcaaga gtcactgcca tcgagaaata tctcaagac caggcacaac taaattcatg
6841 gggatgtgct ttccagacagg tctgtcacac gactgtacca tgggtaaatg aatcccttaa
6901 gccagactgg aataacatga catggcaaca atgggagagg caagtccgt tcttggatgc
6961 aaatataaca aaattactag aagaggcaca gatacaacaa gaaaagaaca tgtatgagtt
7021 acagaaatta aatcaatggg atattttcag taattgggtt gacttcacct cctggatggc
7081 atacatcagg ttaggattat atatagtat aggaatagta gtattaaaga tagcaatata
7141 cattatacag atgctagcaa ggcttaggaa gggctatagg ccagtattct cctccctccc
7201 ctcttatact caacagatcc ctatccgcaa ggaacgggga cagccagcca acgaagaaac
7261 agaagaaggc ggtggaaaca acgagggcta cagatcttgg ccttggcaga tcgaatacat
7321 ccacttcccg attcgcacaac TGAgggacct cttgatttgg ctatacagcg gctgcagaac
      <- tat cds end
7381 cttattatca aagaccttcc aaaccttcca accagtactc caacogctca ggcttccacc
7441 tgcatacctc cgatATGgga tcagctggtt ccaagaagca atccaagcag cagccagggc
      nef cds start ->
7501 tgcgggagag actcttgca ggcggggcag gacctcgtgg ggagtcttca ggagagcggc
7561 aggagagatc attgcaatc ccaggaggat cagacaaggg gcTGAactcg cctcttTG
      <- rev cds end

```

```

7621 Atggacaaaa gactttggga gcagaagggg gaggaaca agattcagat gaggatgatg
    <- env cds end
7681 aggacaatga agtaggggtc cgtgtaagac cgggggtccc actgaggcca atgacattca
7741 aactagcagt agacatgtct catTTTTTaa aagaaaaggg ggaactggaa gggattttct
    LTR ->
7801 atagtgagag aaggcataaa atactagaca catacttaga aatgaagaa ggcatttgtt
7861 ctggatggca gaactacaca catggggccag gagtaagata tcccaagttc ttggctggc
7921 tatggaagct agtgccaata aatatgatag cagaaccaga ggacgaggaa accattgtc
7981 tggatcatcc agcacagacc tcggcatggg atgaccccca cgaagagacc cttgtctggc
8041 agtttgactc cctcctagca tatgactatg tggccttcag cagggtccca gaggagtttg
8101 ggtatcagtc aggaatgcca gagaaagagt ggaaggctaa actgagagca agaggaatac
8161 ctacagagTA Gacaggaant aacagcaca agaggaacta gctaacactg catagagaag
    <- nef cds end
8221 gaaactagct gatactGCAG GACTTTCAG AAGGGGCTGT aacagtGGGA GGgacatggg
    -> enhancer <- Spl ->
    enhancer -> <-
8281 AGGAGCaGGG AGGaagccc atcaagactc tgTATAAATg tacccgcttt gtgcattgta
    Spl -> Spl -> signal ->
8341 ttcagtggct ctggggagag gctggcagat cgagccctgg gaggttctct ccagcactag
    mRNA -> | -> tar core I <- | | ->
8401 caggtagagc ctgggtgttc cctgctagac tctcaccagc acttggccgg tgctgggcag
    tar core II <- |
8461 acggctccac gcttgcttgc ttaaaacctt cttAATAAAA ctgectatta gaagcaagac
    signal ->
8521 aagtgttgtt tccatctct cctagtcgcc gcttggctcat tgggtgctcc actagaaacc
8581 ctggtctgtt aggaacctt ctgctttggg aaaccaaggc aggaatacc ctagca
    <- LTR

```


HIV222381

LOCUS HIV222381 781 bp ss-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample 2238, clone 1, partial cds.

ACCESSION U33077

SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA derived from a 47 year old asymptomatic male from rural Liberia.

REFERENCE 1 (bases 1 to 781)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 2238 was a 47 year old asymptomatic male from rural Liberia who had no history of sexually transmitted disease, chronic illness, blood transfusion or homosexuality. Primer pairs failed to amplify the pol (reverse transcriptase) region. Sample 2238, a highly divergent strain, defines a new subtype which has been designated as subtype C [1].

NCBI gi: 532129

BASE COUNT 280 a 155 c 190 g 156 t

ORIGIN 111 bp downstream from the start of gag

1 agcgaacgaa ttggacagat ttgcattagc agaaagcctg ttggagacca aagaaggatg

gag cds ->

61 ccasaaaatt ttaacagtgt tagaaccatt agttccgaca gggtcagaaa atttaagag

121 cctttataat actacctggc tcatttgggtg ttacacgca gaacagaaag tgaacatac

181 agaggaacga aacgaggtag tagagagaca tctagcggca gaaactaaaa atgcagaaaa

241 aatgccagcc acaagtagac caacagctcc atctagtggc ggcagaggaa attaccagc

301 gcagcaagta gcaggcaatt atgtccacca tctctaatg cctagaacac tgaatgcatg

361 ggtaaaagta gtagaagaaa agaagtttgg ggcagaggta gtaccaggat ttcaggcact

421 gacagaaggg tgcacaccct atgatattaa tcaattgcta aattgtgtag gagatcatca

481 ggctgctatg cagattatta gagaaattat caatgaggaa gcagccgatt gggatgaaca

541 ccacccagta ccaggccccc tgcagcagg acagcttagg gatccaaggg gctctgacat

601 tgcaggaaca accagcacgg tagaagaaca gatacagtg atgtatagag cacagaaccc

661 agtaccggta ggcaacatct atcggagatg gatccaacta ggctgcaaa agtgtgtcag

721 aatgtataat ccaactaaca ttctagacgt gaaacagggt ccaaaagaac catttcagag

781 t

gag cds ->

LOCUS HIV222382 708 bp ss-RNA VRL 24-SEP-1992
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) pol integrase sequence, sample 2238, partial cds.
 ACCESSION M87138
 SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA derived from a 47 year old asymptomatic male from rural Liberia.
 REFERENCE 1 (bases 1 to 708)
 AUTHORS Gao, F., Yue, L., White, A.T., Pappas, P.G., Barchue, J., Hanson, A.P., Greene, B.M., Sharp, P.M., Shaw, G.M. and Hahn, B.H.
 TITLE Human infection by genetically-diverse SIVm-related HIV-2 in West Africa
 JOURNAL Nature 358, 495-499 (1992)
 REFERENCE 2 (bases 1 to 781)
 AUTHORS Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT Patient 2238 was a 47 year old asymptomatic male from rural Liberia who had no history of sexually transmitted disease, chronic illness, blood transfusion or homosexuality. Virus isolation was attempted without success. Sample 2238, a highly divergent strain, defines a new subtype which has been designated as subtype C [2].
 NCBI gi: 325740
 BASE COUNT 260 a 131 c 168 g 149 t
 ORIGIN 2184 bp downstream from the start of pol
 1 gttagggatt atagcagggc agccaacaga aacagagtc cactagtaa acagaatcat
 pol cds ->
 61 agaagaatta ataaggaagg aagctgttta catcagctgg gtgccagctc acaagggtct
 121 cggaggaaat caagaagtag accacctagt aagtcaggga attaggcaag tattattctt
 181 agaaaaaata gaaccagcac aggaagagca tgaagattt catagcaatg ccaaagagct
 241 gacacacaaa ttgggtatac ctaacttagt ggctaaacag atagtaaata cttgtcatgt
 301 gtgtcaacag aaaggagagg ccatacatgg gcaagtaaat gcagatttag gaacatggca
 361 aatggaactgt acccathtag aaggaaaagt tatcatagta gcagtcctatg tggccagtgg
 421 ctttatagaa gcagaggtaa tcccacaaga aactggcaga caaacagctc tttttctgtt
 481 aaaattggca gcaagatggc ctattaccca cctacacaca gataatggtg ccaatttcac
 541 ttacacaggaa gtaaaaatgg tggcgtggtg ggtaggcgtg gaacatacct ttggagtacc
 601 ttacaacca cagagtcagg gagtagtaga ggocatgaac catcatctga aaacaccaat
 661 agatagaatc agagagcagg caaacacaat ggagaccata gtcttgat
 pol cds ->

HIV260415K1

LOCUS HIV260415K1 781 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample 60415K, clone 8, partial cds.

ACCESSION U33076

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 34 year old heterosexual asymptomatic man from urban Senegal.

REFERENCE 1 (bases 1 to 781)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F0784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 60415K was a heterosexual 34-year old asymptomatic man from urban Senegal, who was living in the Washington, D.C. area. He traveled extensively before coming to the United States where he was identified as HIV-2 positive after seeking medical care for a STD. He was married to an American IV drug user. Virus isolation was successfully performed at the University of Alabama at Birmingham. Virus 60415K clusters with the subtype A HIV-2 viruses. NCBI gi: 532131

BASE COUNT 277 a 165 c 194 g 145 t

ORIGIN 111 bp downstream from the start of gag

1 agcgaaatgaa ttggacagat tgggattagc agagagcctg ttggagtcaa aagaagggtg

gag cds ->

61 ccaaagaatt cttacagttt tagatccatt agtacogaca ggttcagaaa atttaaaaag

121 cctctttaat actgtctgcy tgccttgggtg catcacgca gaagagaaag tgaaagatac

181 tgaaggagca aaacaaatag tacagagaca tctagtggca gaaacaggaa ctgcagagaa

241 aatgccaaagc acaagtagac caacagcacc acctagcggg gaaggaggaa acttccccgt

301 gcaacatgta ggcggcaact ataccatcat accgctgagc ccccgaaact taaatgcttg

361 ggtaaaatta gtagaggaas aaaagttcgg ggcagaagta gtgccaggat ttcaggcact

421 ctccagaaggc tgcacgccct atgatattaa ccaaatgctt aattgtgttg gcgaccatca

481 ggcagctatg caataaatca gggaaattat taatgaagaa gcagcagaat gggatgtgca

541 acatccaata ccaggccctt taccagcggg gcagcttaga gagccaagag ggtctgacat

601 agcaggggaca acaagcacag tagatgaaca gatccagtgg atgttttaggc caaaaaatcc

661 cgtaccagta ggaacatct atagaagatg gatccagata ggactgcaga agtgcgtcag

721 gatgtacaac ccgaccaaca tctagacat aaaacaggga ccaaagagc cgttccaaag

781 c

gag cds ->

LOCUS HIV260667K1 766 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample 60667K, clone 1, partial cds.
 ACCESSION L33078
 SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 37 year old heterosexual female originally from Sierra Leone who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 766)
 AUTHORS Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBNCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 60667K was a heterosexual 37-year old female from Sierra Leone, who was living in the Washington, D.C. area. Originally she lived in an urban area in West Africa before traveling extensively and coming to the United States. After seeking medical care for a STD, patient 60667K was identified as HIV-2 positive. She had reduced CD4 counts but was otherwise asymptomatic. Her two teenage children were seronegative. Western blot analysis of patient 60667K's peripheral blood cells showed HIV dual reactivity. Virus isolation was successfully performed at the University of Alabama at Birmingham. Sample 60667K clusters with the subtype B viruses.
 NCBI gi: 532133
 BASE COUNT 275 a 160 c 187 g 144 t
 ORIGIN 111 bp downstream from the start of gag

1 agtaaatgaa ttagaagat ttggattggc agaaaacctg ttgggatcaa aagaaggatg
 gag cds ->
 61 ccataaaatt ctgcagctct tagagccatt agtaccacaa ggatcagaaa atttaagag
 121 ccttttttaat accgtctgcy tcattttattg cctgcacgca ggagagaaaag tgaaagatac
 181 tgaggaagca aaaaagatag cacagcgaca tctagtggga aacacagaaa aaatgccagc
 241 tacaagtaga ccaacagcac catctagcgg cggaatttat ccagtgcagc aaatagctgg
 301 aaattattcc cacctgcgcg taagccctcg gaccttaaat gcctgggtaa aattggtaga
 361 ggaaaagaag ttctggggaga acgtggtaac aggtattccag gcactatcag aaggatgcac
 421 cccctatgat attaatcaga tgtaaattg tgtaggagac catcaggcag ccatacgaat
 481 cattagggaa- ataattaatg aggaagcggc ggaatgggac caacagcatc catcgccagg
 541 cccaatgccg gcaggacaac tcaggggaacc aagagggtca gacatagcgg gaaccaccag
 601 cacagtagaa gaacagatac aatggatgtg cagggcccaa aatcctgtcc cagtaggaaa
 661 catttataga agatggattc agatagggct gcaaaaatgt gtccggatgt acaacctac
 721 caatatactg gacataaagc aagggccaaa ggagcccttc caaaac
 gag cds ->

HIV27312A1

LOCUS HIV27312A1 766 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, isolate 7312A, clone JK, partial cds.

ACCESSION L33079

SOURCE Human immunodeficiency virus type 2 DNA from cocultured PBMCs derived from a 32 year old male from urban Cote d'Ivoire with lymphadenopathy and who showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 766)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Rui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F0784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.

Patient 7312A was a 32 year old male from urban Cote d'Ivoire who had reported frequent visits to urban prostitutes. His sera was HIV dually reactive and he had lymphadenopathy. Unlike other sequences in this study, gag, pol, and env sequences for patient 7312A were not PCR derived. Instead, they were obtained from a full length proviral genome cloned from cultured PBMC DNA. Virus isolation was successfully performed at the University of Alabama at Birmingham. Virus 7312A clusters with subtype B viruses for gag and pol(RT and int) and with the A viruses for env. Gao et al. conclude that virus 7312A has arisen from the recombination of subtype A and B viruses.

NCBI gi: 532137

BASE COUNT 269 a 157 c 188 g 152 t

ORIGIN 111 bp downstream from the start of gag

1 agcaaatgaa ttagacagat tccgattggc agaaaacctg ttggaatcaa aagaaggatg
gag cds ->

61 tcagaaaatt cttacggtct tagctccatt aatgccaaac gggtcagaaa atttaaaaag
121 cctttttaat actgtctgcg tcatttgggtg cttacacgca gaaaagaaag tgaagatcac
181 cgagggaagca aaaaagctag cacagagcca tctagtggta gacactgaac aaacacctec
241 tgcagacaaa ccagcagcga catctagtgg cagaaattac ccagtgcac aagtagctgg
301 caattatgtc cactgtccat taagcccccg aacctaaat gcctgggtaa aattagtggg
361 ggaaaagaaa ttcggggcag aagtagtgcc aggatctcag gcactatcag agggatgtac
421 ccttatgat attaatcaga tgctaaattg ttagggagaa catcaggcag ccattgcaagt
481 cattagggaa ataataatg aggaggccgc agactgggac cagcagcacc cggtgccagg
541 tccactgcca gcaggacaac ttagggaacc cagaggatca gatatagcgg gaaccaccag
601 cacagtagag gaacagatac agtggatgta caggggtcaa agtccatccc cagtagggaa
661 catttataga agatggattc agctaggatt gcagaagtgt gtcgggatgt acaatctac
721 taatatacta gatgtaaaac aaggggccaa agaacccttc caaagc
gag cds ->

LOCUS HIV27312A2 995 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus 2 (HIV-2) pol reverse transcriptase sequence, isolate 7312A, clone JK, partial cds.
 ACCESSION L93092
 SOURCE Human immunodeficiency virus type 2 DNA from cocultured PBMCs derived from a 32 year old male from urban Cote d'Ivoire with lymphadenopathy and who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 995)
 AUTHORS Gao, P., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 6041SK, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 7312A was a 32 year old male from urban Cote d'Ivoire who had reported frequent visits to urban prostitutes. His sera was HIV dually reactive and he had lymphadenopathy. Unlike other sequences in this study, gag, pol, and env sequences for patient 7312A were not PCR derived. Instead, they were obtained from a full length proviral genome cloned from cultured PBMC DNA. Virus isolation was successfully performed at the University of Alabama at Birmingham. Virus 7312A clusters with subtype B viruses for gag and pol(RT and int) and with the A viruses for env. Gao et al. conclude that virus 7312A has arisen from the recombination of subtype A and B viruses.
 NCBI gi: 532141

BASE COUNT 390 a 200 c 218 g 187 t

ORIGIN 806 bp downstream from the start of pol

1 aaaggagggg gcagctagaa gaagcatctc ccactaatcc atacaacaca cccacctttg

pol cds ->

```

61 ctataaagaa aaaggataaa acaaatgga gaatgcta atagacttcagg gaattaaata
121 aggtaaccca agaatttaca gaagtccaac taggtattcc ccacccggca ggctggccag
181 aaaaaaggag aataacagtg atagatgtag gagatgcta cttcagtgtc ccactagacc
241 cagacttcag gcaatataca gcattcactt tggcatcaat aaataatgca gagccaggaa
301 aaaggtagat ttataaagtc ctaccacagg gatggaaggg atcccagca atttttcaat
361 actccatgag aaaggtagta gacccttcca gaagggccaa caatgatgtc accataatcc
421 agtagatgga tgacatcctt gtggcaagcg atagaagtga tctagaacat gataaggtag
481 tgtctcaact aaagagagta ctaaatgaca tgggattttc caccacagaa gaaaaattcc
541 aaaggagccc tccattccag tggatgggtt atgagctctg gccaaagaag tggagctgct
601 aaaaaataca aataaccaga agggaggttt ggacagtaaa tgacattcag aaactagtgg
661 gagtatttaa ctgggcagct caacttttcc cgggaatcaa aacaaggcac atatgcagac
721 tgatcaaggg aaagatgacc ctgacagagg aggtgcagtg gacagaacta gcagaggcag
781 agatgcagga aataaaatc atcctaggac aggaacagga aggatcttac tacaaagaag
841 gggtagctct agaagcaaca gtgcagaaaa acttagcaaa ccagtggaca taaaaaatc
901 atcagggaaa taaagtcta aaagtaggaa aatatgcaa ggtaaaaac acccacacca
961 atggagtaag aatactagca catgtagtc aaaaa

```

pol cds ->

HIV27312A3

LOCUS HIV27312A3 708 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus 2 (HIV-2) pol integrase sequence, isolate 7312A, clone JK, partial cds.

ACCESSION L33088

SOURCE Human immunodeficiency virus type 2 DNA from cocultured PBMCs derived from a 32 year old male from urban Cote d'Ivoire with lymphadenopathy and whose showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 708)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F0784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.

Patient 7312A was a 32 year old male from urban Cote d'Ivoire who had reported frequent visits to urban prostitutes. His sera was HIV dually reactive and he had lymphadenopathy. Unlike other sequences in this study, gag, pol, and env sequences for patient 7312A were not PCR derived. Instead, they were obtained from a full length proviral genome cloned from cultured PBMC DNA. Virus isolation was successfully performed at the University of Alabama at Birmingham. Virus 7312A clusters with subtype B viruses for gag and pol(RT and int) and with the A viruses for env. Gao et al. conclude that virus 7312A has arisen from the recombination of subtype A and B viruses.

NCBI gi: 532139

BASE COUNT 275 a 121 c 167 g 145 t

ORIGIN 2184 bp downstream from the start of pol

1 catgggaata gtagctggac agccaacaga aacagagtc cggatagtaa ataagataat
pol cds ->

61 tgaagaaatg atcaaaaagg aagcaatata tgtaggatgg gtgccagctc acaaaggact
121 aggtggtaat caggaagtag accacctagt aagtcaggga attaggcagg ttctattcct
181 agaaagata gaaccagccc aggaggaaca tgaaaaatat catggcaatg taaaagagct
241 gggtcataaa ttgggaattc cacaactagt ggcaaaacaa atagtaaaact cctgtgataa
301 atgtcaacaa aaaggagaag ctgttcattg acaggtaaat gcagaactag ggacatggca
361 aatggactgt acacatttag aagggaaagt cataatagta gcagtccatg tagccagtgg
421 gtttatagaa gcagaggtaa taccccaaga aacaggaaga caaacagctc tcttctgtt
481 aaagctggca agcagatggc ccattacaca cttgcacaca gacaatgggtg ccaactttac
541 ttcacaagat gtaagatgg tggtctgggt ggtaggaata gaacaaacct ttggagtacc
601 ctacaacca caaagtcagg gtagtagtag agcaatgaat catcatctga aaaatcaaat
661 agataggatc agagatcagg cagtatcaat agaaacagtt gtgctaet
pol cds ->

LOCUS HIV27312A4 453 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp41 env sequence, isolate 7312A, clone JK, partial cds.
 ACCESSION L33094
 SOURCE Human immunodeficiency virus type 2 DNA from cocultured PBMCs derived from a 32 year old male from urban Cote d'Ivoire with lymphadenopathy and who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 7312A was a 32 year old male from urban Cote d'Ivoire who had reported frequent visits to urban prostitutes. His sera was HIV dually reactive and he had lymphadenopathy. Unlike other sequences in this study, gag, pol, and env sequences for patient 7312A were not PCR derived. Instead, they were obtained from a full length proviral genome cloned from cultured PBMC DNA. Virus isolation was successfully performed at the University of Alabama at Birmingham. Virus 7312A clusters with subtype B viruses for gag and pol(RT and int) and with the A viruses for env. Gao et al. conclude that virus 7312A has arisen from the recombination of subtype A and B viruses.
 NCBI gi: 532135
 BASE COUNT 162 a 76 c 110 g 105 t
 ORIGIN 1861 bp downstream from the start of env
 1 gacgtggtca agagacaaca agaaatgttg cgactgacgg tctggggaac aaaaaatctc
 env cds ->
 61 caggcaagag tcaactgctat tgagaataac ttaaaggacc aggcgcaact aaattcatgg
 121 ggatgtgcgt ttaggcaagt ctgccacact actgtaccat gggtaaatga cagcttgaca
 181 cctgattggg acaacatgac gtggcaacaa tgggaaaaac aaatccgga cctggaggca
 241 aatatcagtg aaagtctaga acaggcacaa atccagcaag aaaagaacat gtatgaatta
 301 caaaaattaa atagctggga tgtttttggc aactggtttg atttagcctc ctgggtcaaa
 361 tatattcagt atggagtta tatagtagta ggaatagtag ctctcagagt aataatatat
 421 gtagtacaaa tgataggtag acttagaaga ggc
 env cds ->

HIV27810A1

LOCUS HIV27810A1 766 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample 7810A, clone 2, partial cds.

ACCESSION L33080

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from an asymptomatic 31 year old heterosexual man originally from Cote d'Ivoire.

REFERENCE 1 (bases 1 to 766)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,W., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 7810A was a heterosexual 31-year old student from Cote d'Ivoire, who was living in the Washington, D.C. area. Originally he lived in an urban area in West Africa before traveling extensively and coming to the United States. After seeking medical care for a STD, patient 7810A was identified as HIV-2 positive. He had a heterosexual partner of 3 years who was HIV-2 seronegative. He was asymptomatic with normal CD4 counts. Sample 7810A clusters with the subtype B viruses.

NCBI gi: 532143

BASE COUNT 270 a 168 c 183 g 145 t

ORIGIN 111 bp downstream from the start of gag

1 agcaaacgaa ttggacagat ttggactagc agagaccctg ttgggatcaa aagaaggatg

gag cds ->

61 tcacaaaatt cttacggcct tagctccatt agtaccacaa ggatcagaaa atttaaaaag

121 ccttttttaat actgtctgcg tcggtttattg cttacacgca gaagagaaag tgaagatac

181 tgaggaagca aaaaagatag cacagagcca tctagcggcg aacacagaaa aaatgccagc

241 tacaagtaaa ccaacagcac cacctagcgg cagaaattat ccagtgcagc aggtggccgg

301 caattacgtc catctgccgc taagcccccg aaccttaaat gcctgggtaa aattagtaga

361 ggataagaaa ttcggggcag aagtagtacc aggattccag gcactatcag aagggtgtac

421 cctttatgat attaatcaaa tgctaaattg tgtaggagaa catcaagcag ccatgcaaat

481 tattagggaa ataactaatg aggaagcagc ggactgggac cagcagcacc cgtcaccggg

541 cccactacca gcaggacagc ttagggaccc aagagggtcg gacatagcgg ggaccaccag

601 cacagtggag gaggagatac aatggatgta cagggcccaa aatcctatcc cagtaggaaa

661 catttatagg agatggatcc aattaggact gcagaagtgt gtccggatgt ataaccctac

721 taatatactg gacataaaac aaggacccaa agaacccttc caaagc

gag cds ->

LOCUS HIV27924A1 781 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample 7924A, clone 1, partial cds.

ACCESSION L33081

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 48 year old heterosexual male originally from urban Guinea Bissau who showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 781)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient 7924 was a heterosexual 48-year old teacher from Guinea Bissau, who was living in the Washington, D.C. area. Originally he lived in an urban area in West Africa before traveling extensively and coming to the United States. After seeking medical care for a STD, patient 7924 was identified as HIV-2 positive. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity. Virus isolation was successfully performed at the University of Alabama at Birmingham. Sample 7924A clusters with the subtype A HIV-2 viruses.

NCBI gi: 532145

BASE COUNT 279 a 157 c 195 g 150 t

ORIGIN 111 bp downstream from the start of gag

1 agcgaatgaa ttggacagat tcggattagc agagagcctg ttggagtcaa aagaaggttg
gag cds ->

61 tcaaagaatt atttcagttt tagatccatt agtaccgaca ggctcagaaa atttaaaaag
121 cctttttaat actgtctgcg tagtttggtg catcacgaca gaagagaaag tgaagacac
181 tgaagaagca aaaaagatag tacagagaca tctagtggca gaacaggaa ctgcagagaa
241 aatgccaaat acaagtagac caacagcacc acctagcggg aaaggaggaa actacccgt
301 gcaaagcata ggcggcaact ataccatgt gccgtgagt cccgaaccc taaatgcttg
361 ggtgaaatta gtagaagaga agaagttcgg gccagaagta gtgccaggat tccaggcgct
421 ctcaagaaggc tgcacgccc atgacattaa tcaaatgctt aattgtgtgg gcgaccatca
481 agcagctatg caaataatca gggaaatcat taatgaagta gcagcagatt gggatgttgc
541 acatccaata ccaggccctt taccagcagg gcagcttagg gaaccaagag ggtctgcacat
601 agcaggaaca acaagcacag tagaagagca gatccagtgg atgttttagag cacaaaaccc
661 tgtaccagtg ggaacatct atagaagatg gatccagata gggctgcaga aatgcgtcag
721 gatgtacaat ccaaccaaca tectagatgt aaacaggga ccaagggaat catttcaag
781 c

gag cds ->

HIV2ALI

LOCUS HIV2ALI 2589 bp ss-RNA VRL 07-MAR-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) envelope glycoprotein sequence, non-syncytium-inducing isolate HIV-2ALI, complete cds.
 ACCESSION L25445
 SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from cocultured PBMCs derived from a male patient from Guinea-Bissau who had AIDS-related complex (ARC).
 REFERENCE 1 (bases 1 to 2589)
 AUTHORS Costa Taveira,N., Bex,F., Burny,A., Sharp,P., Santos Ferreira,M. and Moniz-Pereira,J.
 TITLE Genetic analysis of the env gene from a non-syncytium-inducing HIV-2 isolate (HIV-2 ALI)
 JOURNAL Aids Res. Hum. Retroviruses 10, 223-224 (1994)
 COMMENT HIV2ALI is a non-syncytium-inducing isolate with a deduced V3 loop amino acid sequence that is similar to other HIV-2 V3 loop regions. Costa Taveira et al. note that the most significant difference between this isolate and others which induce syncytium is the absence of positively charged amino acids within the flanks of the V3 loop. This difference suggests, as it does for HIV-1, that the syncytium inducing phenotype may be due to the hydrophilic nature of the V3 loop. In addition, this isolate possesses a leucine at residue 7 with the TM protein fusion domain. This characteristic is unique to all human retroviruses and may play a part in its inability to form syncytium. In phylogenetic analysis, HIV2ALI grouped with HIV-2 subtype As from Guinea-Bissau, even though it is phenotypically different from others in the group with respect to ability to form syncytium.
 BASE COUNT 846 a 530 c 611 g 602 t
 ORIGIN env cds start

```

1 ATGatgtcta gtagaatca gctgcttgtt actatcttac tagctagtgc ttgcttagta
env cds ->
61 tattgtaaac aatatgtgac tgttttttat ggcgtgccag catggaaaaa tgcattccatt
121 cccctctttt gtgcaaccaa aaatagagat acttggggaa ccatacagtg cttaccagac
181 aatgatgatt atcaggaaat agctttaaat gtgacagagg ctttcgatgc atgggataat
241 acagtaacag aacaagcagt agaagatgtc tggagactat ttgagacatc aataaaacca
301 tgtgtcaagt taacaccttt atgtatagca atgaagtgtg gccacataag cacagagagc
361 acaaccacat ccccgagccc agggagcaca ctcaaacccc tgataaatga gagcgatcca
421 tgcataaagg cagacaactg ccccagggga ctaggggatg aagagatggt caattgtcgg
481 ttcaacatga caggattaca gagagataag ccaaaacagt ataattgaaac atgggtactca
541 aaagatgtgg tttgtgaacc atttaacacc accacaaacc agaccagggtg ttacatgaac
601 cattgcaaca ctcagtcac cccagagtaa tgtgataagc actattggga tgcataaaga
661 tttagatact gtgcaccacc tggttacgcc ttgctaagat gcaatgatac caattatcca
721 ggctttgcac ccaattgctc taaagtagta gctgctacat gcacaaggat gatggagatg
      v3 loop equivalent ->
781 caaacttcta cttggtttgg ctttaattggc actagggcag aaaatagaac atatatctat
841 tggcatggtg gagataatag aactatcctc agcttaacaa aacagtataa tcttactatg
901 cattgTAAga ggccaggaaa taagacagtt gtaccaataa cacttatgtc agggtaata
      <- v3 loop equivalent
961 tttaactccc agccaatcaa taaaagaccc agacaagcat ggtgctggtt caaaggcgaa
1021 tggaggggag coattgcagga ggtgaaggaa acccttgtaa aacatcccag gtataaagga
1081 accaatgaca caaaccaaat taactttaca aaaccaggaa gaggtcaga tgcagaagag
1141 gtatatatgt ggactaactg cagaggagaa tttctctatt gcaacatgac ttggttcttc
1201 aattgggtgg aaaaacaaac gggtcaggaa cagcacaatt atgcaccgtg ccatataaag
1261 caaataatta atacctggca caaagtggg aaaaatgtat atttgctccc taggggaagg
1321 gagttgacct gcaactcaac agtaaccagc ttgattgcta acattgacac ggaatggcaac
1381 cagacaaata ttacctttag tgcagagggtg gcagagctat accgattaga attgggggat
1441 tacaatttag tagagataac accaattggc ttgcaccta catcagaaag gagatactcc
1501 tctactccaa ggaggaataa aagagggtgtg ttctgtctag ggtcttggg ttttctcgcg
1561 acagcaggtt ctgcaatggg cagggcagct ttaacgctgt ctgctcagtc tggacttta

```

```

1621 ttggccggga tagtgcagca acagcaacag ctgttggaca tggccaagag acaacaggaa
1681 atgttgcgac tgaccgtctg gggaaacgaaa aatctccagg caagagtcac tgctatcgag
1741 aaataacttaa aggaccaggc ggggctaaat tcatggggat gtgcatttag acaagtctgc
1801 cacactactg taccatgggt aaataactcc ttaaaacctg attgggacaa catgacgtgg
1861 caagagtggg aacaacaaat ccgttaccta gaggcacata tcagtgaaca gttagaacag
1921 gcacaattc agcaagaaaa gaatacatal gaactacaaa aattaaatag ctgggatgtt
1981 ttaccaaact ggcttgactt aaccgcctgg gtcaagtata ttcaatatgg agtttatata
2041 atagtaggaa tagtagctct tagaatagta atatatgtag tgcaaatgtt aagtaggctc
2101 aggaagggtt ataggcctgt ttctctctcc cctcccggtt acatccaaca gatccatctc
2161 cacaaggacc aggeacagcc aaccagagga gaaacagaag aagacgttgg agacaacgtt
2221 ggggacagat tgtggccctg gcgatcgca tatttacatt tctgatcca cctgctagct
2281 cgcctcttga tcgggctgta cagcatctgc agggacttac tatccaggat cttctcgatc
2341 ctccaacoga tcttcagag tcttcAGaga gcgctgacaa caatcagggg ctggctgaga
      3' sj (tat.rev) /\
2401 cttaaagcag cctacctgca gtatgggtgc gagtggatcc aagaagcgtt ccaggccttt
2461 gcaaggattg cgagagagac gcttacaaac acctggagag acttgtgggg ggcaatgcag
2521 tgggtcggga ggaggatact cgcagtccca aggaggatca ggcagggggc agaaattgcc
2581 ctctctTGA

```

```
<- env cds end
```

HIV2CI171

LOCUS HIV2CI171 453 bp ss-RNA VRL 28-OCT-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp 36 envelope sequence, isolate HIV2CI171, partial cds.

ACCESSION X78505

SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from cocultured PBMCs derived from a patient from Cote d'Ivoire with AIDS.

REFERENCE 1 (bases 1 to 453)

AUTHORS Peeters,M., Janssens,W., Fransen,K., Brandfull,J., Hendrickx,L., Koffi,K., Delaporte,E., Piot,P., Gershy-Damet,G., Van der Groen,G.

TITLE Isolation of simian immunodeficiency virus from two wild-captured sooty mangabeys in Cote d'Ivoire: virological and genetic characterization and relationship to either HIV-2 and SIVsm/mac strains.

JOURNAL AIDS Res. Hum. Retroviruses 10, 1289-1294 (1994)

COMMENT This sequence, HIV2CI171, was obtained through a study designed to compare genetic isolates of HIV-2 and SIV from the same region of Africa, specifically Cote d'Ivoire. During the study, blood samples were collected from five human patients (CI171, CI9, ISO53, VI390, and VI905) and two sooty mangabey monkeys (SMCI2 and SMC18). PBMCs from the animals and human patients were cocultured with those from a healthy human donor, and the viral sequences were PCR amplified, cloned, and sequenced.

Both HIV-2 strains CI9 and CI171 were obtained from patients from Cote d'Ivoire with AIDS. HIV-2 strain VI390 was isolated from an asymptomatic Ghanaian patient living in Belgium. HIV-2 strain VI905 was derived from an asymptomatic Senegalese patient also living in Belgium. Finally, strain ISO53 was obtained from an asymptomatic Belgium homosexual who had never had contact with Africans. Both strains SMC18 and SMC12 were derived from sooty mangabeys from Cote d'Ivoire.

Through phylogenetic analysis, Peeters et al. determined that the SMC12 and SMC18 sooty mangabey sequences were equidistant from both the HIV-2 sequences and from other SIVsm and SIVmac sequences. Peeters et al. also determined that the HIV-2 isolates obtained in their laboratory cluster with the HIV-2 subtype A viruses. Sequence CI171 and the other four HIV-2 viral sequences determined in this study are shown in alignment below: CI9 (X78506), ISO53 (X78507), VI390 (X78510), and VI905 (X78511).

BASE COUNT 161 a 73 c 105 g 114 t

ORIGIN 1861 bp downstream from the start of env

1 gacgtggtca agagacaaca agaaatgttg cgactgaccg tctggggaac
env cds ->

51 gaaaaatctc caggcaagag tcactgctat cgagaaatac ttaaaggacc
101 aggcgcagct aaattcatgg ggatgtgcgt ttaggcaggt ctgccacact
151 actgtaccat tggtaaataa ctctttaacg cctgactgga aaaatatgac
201 atggcaagag tgggagatgc aagtcgceta cctagaggca aatatcagtc
251 aacagttaga acaggcaca attcaacaag aaaagaatat gtatgaatta
301 caaaaactaa atagctggga tgtctttagt aacTAGtttg atttaacctc
premature termination of env <-

351 cTAGatcaaa tatattcagt atggagtttt tatagtagta ggaatagtag
<- premature termination of env

401 ttcttagaat agtaatatat gtagtacaaa tgtaagtag acttagaagg
451 ggc
env cds ->

CONSENSUS	GACGTGGTCAAGAGACAACAAGAAATGTTGCGACTGACCGTCTGGGGAACGAAAAATCTC	60
HIV2CI171	-----	60
HIV2CI9	---A-A-----	60
HIV2ISO53	-----A-----A-----	60
HIV2VI390	-----A-----T-----G-----	60
HIV2VI905	-----C-----	60
CONSENSUS	CAGGCAAGAGTCACTGCTATCGAGAAATACTTAAAGGACCAGGCGCAGCTAAATTCATGG	120
HIV2CI171	-----	120
HIV2CI9	-----G-----A-----	120
HIV2ISO53	-----G-----	120
HIV2VI390	-----A-----	120
HIV2VI905	-----A-----C-----A-----	120
CONSENSUS	GGATGTGCGTTTAGGCAGGTCTGCCACACTACTGTACCATGGGTAAATGACTCCTTAACG	180
HIV2CI171	-----T-----T-----	180
HIV2CI9	--G-----T--G--	180
HIV2ISO53	---C-----A--A---T-----A-----A	180
HIV2VI390	-----A-----	180
HIV2VI905	---C-----A-----A-----A	180
CONSENSUS	CCTGACTGGAACAATATGACATGGCAGGAATGGGAG?A?CAAGTCCGCTACCTAGAGGCA	238
HIV2CI171	-----A-----A--G-----ATG-----	240
HIV2CI9	-----G--A-----	240
HIV2ISO53	---T-----C-----C--G---AC--C---A-----	240
HIV2VI390	-----C--A--G-----T-----	240
HIV2VI905	--CA--T---C--G--C---T---C---AC--A---A---T-----	240
CONSENSUS	AATATCAGTGAACA?TTAGAACAGGCACAAATTCACAAGAAAAGAATATGTATGAATTA	297
HIV2CI171	-----C-----G-----	300
HIV2CI9	-----ATC-----G-----	300
HIV2ISO53	-----A-----G-----G--G-----C-----GC--G	300
HIV2VI390	-----C---G--G--G-----C-----G--A-----	300
HIV2VI905	-----C--AGT-----G-----C-----C---	300
CONSENSUS	CAAAAATTAAATAGCTGGGATGTCCTTGGCAACTGGTTTGATTTAACCTCCTGGATCAAA	357
HIV2CI171	-----C-----A--T---A-----A-----	360
HIV2CI9	-----C-----	360
HIV2ISO53	-----T-----	360
HIV2VI390	-----G-----	360
HIV2VI905	-----G-----A---A-----C--G-----GGC	360
CONSENSUS	TATATTCAGTATGGAGTTT?TATAGTAGTAGGAATAATAGTTCTTAGAATAGTAATATAT	416
HIV2CI171	-----T-----G-----	420
HIV2CI9	-----A-----TA--A--A---G--T---T--A---C-----	420
HIV2ISO53	-----A---A---C-----	420
HIV2VI390	-----A-----A-----C---G-----C--T--A-----C-----	420
HIV2VI905	-----G-----C-----	420
CONSENSUS	GTAGTACAAATGTTAACTAGACTTAGAAGGGGC	449
HIV2CI171	-----	453
HIV2CI9	A---G---T-----A---	453
HIV2ISO53	-----G-----	453
HIV2VI390	---G-----T---A---	453
HIV2VI905	--T-----A--G-----	453

HIV2D1024

LOCUS HIV2D1024 1521 bp ss-RNA VRL 15-APR-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp105 envelope glycoprotein sequence, sample D1024, partial cds.
 ACCESSION U07104
 SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from uncultured PBMCs derived from a patient from Bombay, India whose sera was dually reactive (HIV1/HIV2).
 REFERENCE 1 (bases 1 to 1521)
 AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Mahambre,G., Delwart,E.L., Mullins,J.I. and Ruebsamen-Waigmann,M.
 TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses
 JOURNAL J. Virol. 68, 2161-2168 (1994)
 COMMENT This sequence was obtained as part of a study to determine the genetic breadth of HIV-2 variation in India. In this study, blood samples were obtained from four patients from Bombay, India (D766, D808, D868, and D1024) and from one patient from Goa, India (D1071); each of the five patients was serologically dually reactive. The mean genetic divergence of the set of four sequences from Bombay was 7.3%, while the sequence from Goa was from 7 to 10.5% different from each of the Bombay sequences. The entire set of Indian sequences, including both the Bombay and Goa sequences, diverged from eleven non-Indian HIV-2 sequences by 23%. Because of the similarity of the Indian sequences, Grez et al. suggest that similar HIV-2 strains might be found elsewhere in India. All five of these sequences clustered with HIV-2 subtype A viruses. HIV-1 env sequences from Indian patients D766, D808, D868, and D1024 clustered with HIV-1 C viruses, and as a set showed a similar degree of interpatient genetic similarity as did their HIV-2 counterparts.
 BASE COUNT 526 a 305 c 336 g 354 t
 ORIGIN envelope cds start
 1 ATGggggctc agctgcttat taccatttta ctaactagtg cttgcttaat atactgcaga
 env cds ->
 61 caatacgtaa ctgtttccta tggcatatccc ggggtggagga atgcattccat tccccctcgtt
 121 tgtgcaacca agaataagga tacttgggga accatacagt gcttaccaga caatgatgac
 181 tatcaggaaa tagccttgaa tgtgacagag gcttttgatg catggaacaa tacagtaaca
 241 gaacaagcaa tagaagatgt ctggaatcta ttgagacat caacaaaacc atgtgtcaaa
 301 ttaacaccct tatgtgtagc aatgacatgc agggaaaagca gtggcaacac aagtacaggg
 361 aacgcaacca catccacacc catgataaat gagacctcta taataaatga gaacctctca
 421 tgcataogca caaacaactg ctccaggattg agagaggaag agatggtcag ttgtcagttc
 481 aatatgacag gattagcgag agataagcca aaagggtata atgaaacgtg gtactcaagt
 541 gacgtgggtt gtgaacaga aacaaataac acctcaaaac agagcaggtg ttacatgaac
 601 cattgcaaca catcagtcac cacagagtca tgtgataagc actattggga tgatatgagg
 661 tttagatatt gtgcacogcc gggttttgcc ttgctaagat gcaatgatac caattactca
 721 ggctttgagc ccaattgttc taaagttagta gctgccacat gcaaccaggat gatggagacg
 781 caaacttcta cgtggttttg ctttaattggc actagagcag agaatagaac atatatttat
 841 tggcatagta aagataacag aactataatc agcttaaaca aacattataa tctctctata
 901 cattgtaaga ggccaggaaa taaaacagta gtaccaataa cacttatgtc aggattagt
 961 tttcattccc agccaatcaa taacagacct agacaagcat ggtgctgggtt caaaggcaaa
 1021 tggaaagaag ccattgcagga ggtgaagcaa acccttgtaa aacaccccag gtatgagggg
 1081 accaataaga cagaagatat tcactttgca gcaccgggaa aaggctcaga cccggaagcg
 1141 gcatacatgt ggactaactg caggggagaa tttctctact gcaacgtgac ttggttcctc
 1201 aattggatag agaacagaac catgaataag acatggcaca attatgtgcc atgccatata
 1261 aagcaataaa ttaacacctg gcataaggtg gggaaacatg tatatttggc tcccagggaa
 1321 ggggagttga cctgcaactc aactgtgacc agcataattg ctaacattga cggaggtaat
 1381 cagacaaata ttacctttag tgcagaggtg gcggaactat accgactaga attgggagat
 1441 tataaattgg tagaataaac accaattggc ttgcaccta catcagtaaa acgatactcc
 1501 tctgctcagc ggagacatac g
 envelope cds ->

LOCUS HIV2D1071 975 bp ss-RNA VRL 15-APR-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp105 envelope glycoprotein sequence, sample D1071, partial cds.

ACCESSION U07105

SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from uncultured PBMCs derived from a female prostitute from the state of Goa, India whose sera was dually reactive (HIV-1/HIV-2).

REFERENCE 1 (bases 1 to 975)

AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Nahambre,G., Delwart,E.L., Mullins,J.I. and Ruebsamen-Waigmann,H.

TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses

JOURNAL J. Virol. 68, 2161-2168 (1994)

COMMENT This sequence was obtained as part of a study to determine the genetic breadth of HIV-2 variation in India. In this study, blood samples were obtained from four patients from Bombay, India (D766, D808, D868, and D1024) and from one patient from Goa, India (D1071); each of the five patients was serologically dually reactive. Patient D1071 was a female prostitute who was identified as HIV-2 positive in April, 1992 and dually seroreactive in July of 1992. The mean genetic divergence of the set of four sequences from Bombay was 7.3%. The sequence from Goa was from 7 to 10.5% different from each of the Bombay sequences. The entire set of Indian sequences, including both the Bombay and Goa sequences, diverged from eleven non-Indian HIV-2 sequences by 23%. Because of the similarity of the Indian sequences, Grez et al. suggest that similar HIV-2 strains might be found elsewhere in India. All five of these sequences clustered with HIV-2 subtype A viruses. HIV-1 env sequences from Indian patients D766, D808, D868, and D1024 clustered with HIV-1 C viruses, and as a set showed a similar degree of interpatient genetic similarity as did their HIV-2 counterparts.

BASE COUNT 355 a 208 c 203 g 209 t

ORIGIN 340 bp downstream from the start of envelope

1 ATGacatgca gcaaaaacag tggcaacaca agtataggaa gcacaaccac aaccacagac
 env cds ->

61 gcagccatgg tgataaatga gacctctcaa tgcatacggc caaacaactg ctcaggattg
 121 agagaggaag agctagtcag ttgtcagttc gatatgacag gattaacgag agataaacca
 181 aaaagatata atgaacatg gtactcaagc gatgttgttt gtgacgcacc aactaacacc
 241 tcaaacccaga cctcaaataa cactcagac cagagcaggt gttgcatgaa ccattgcaac
 301 acatcagtc tccagaatc atgtgaaaag cactatTAGg atgctataag gtttagatat
 premature termination of env cds <-

361 tgtgcaccac caggttatgc cttgctaaga tgcattgata ccaattactc aggettttgag
 421 cccaattgcc cttaagtagt agctgccaca tgcaccagga tgatggaaac gcaaacttct
 481 acgtggtttg gctttaatgg cactagagca gagaatagga catatattta ttggcatagt
 541 agagataaca gaactataat cagcttaaac aaacattata atctctctat atattgtaag
 601 aggccaggaa ataaaacagt agtaccata acacttatgt caggattagt gtttcattcc
 661 cagccaatca ataccagacc cagacaagca tctgtctggc tcaaggcca atggaaggaa
 721 gccatgcagg aggtgaagca aaccttgca aaacacccga ggtataaggg aaccaataat
 781 acaggaaaga ttaactttac agcacgggga aaaagttcag acccggaagc agcatacatg
 841 tggactatct gcagaggaga atttctctac tgcaacatga cttggtttct caatTAGgta
 premature termination of env cds <-

901 gaagacagaa ccgcgaacag gacatggcat aattatgtgc catgccatat aaagcaaata
 961 attaacacct ggcac
 envelope cds ->

HIV2D766

LOCUS HIV2D766 1524 bp ss-RNA VRL 15-APR-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp105 envelope glycoprotein sequence, sample D766, partial cds.

ACCESSION U07106

SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from uncultured PBMCs derived from a patient from Bombay, India whose sera was dually reactive (HIV1/HIV2).

REFERENCE 1 (bases 1 to 1524)

AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Mahambre,G., Delwart,E.L., Mullins,J.I. and Ruebsamen-Waigmann,H.

TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses

JOURNAL J. Virol. 68, 2161-2168 (1994)

COMMENT This sequence was obtained as part of a study to determine the genetic breadth of HIV-2 variation in India. In this study, blood samples were obtained from four patients from Bombay, India (D766, D808, D868, and D1024) and from one patient from Goa, India (D1071); each of the five patients was serologically dually reactive. The mean genetic divergence of the set of four sequences from Bombay was 7.3%, while the sequence from Goa was from 7 to 10.5% different from each of the Bombay sequences. The entire set of Indian sequences, including both the Bombay and Goa sequences, diverged from eleven non-Indian HIV-2 sequences by 23%. Because of the similarity of the Indian sequences, Grez et al. suggest that similar HIV-2 strains might be found elsewhere in India. All five of these sequences clustered with HIV-2 subtype A viruses. HIV-1 env sequences from Indian patients D766, D808, D868, and D1024 clustered with HIV-1 C viruses, and as a set showed a similar degree of interpatient genetic similarity as did their HIV-2 counterparts.

BASE COUNT 512 a 307 c 346 g 359 t

ORIGIN envelope cds start

1 ATGacacgta gaagggctca gctgcttatt tccattttac taactagtgc ttgcttaata

env cds ->

61 tattgcaaac aaaacgtaac tgtttttctat ggcatacccg cgtggagaaa tgcattccatt

121 cccctctttt gtgcaaccac aaatagggat acttggggaa ccatacagtg cttgccagac

181 aatgatgact atcaggaat aactttgaat gtgacagagg otttcgatgc atggaacaat

241 acagtaacag aacaagcaat agaagatgtc tggaaatctat ttgagacatc aacaaaacca

301 tgtgtcaaat taacaccctt atgtgtagca atgaattgca gcaaaaacag tggcaacaca

361 agtacaggga gtacaaccac acccgagcc atatccatga taaatgagac ctctccatgc

421 atacgcacag acaactgtc aggattaaga gaggaagaga tggtcagttg tcagttcagt

481 atgacaggat taacgagaga taagccaaaa aggtataatg aaacatggta ctcaagtgat

541 gtgggtttgtg aaacaacagg agcaactaac acctcaaac aggacaggtg ttacatgaac

601 cattgcaaca catcagtcac cacagagtca tgtgataagc actattggga tgatatgtgg

661 tttagatatt gtgcaccgcc gggtrtatgtc ttgctaagat atactgatac caattactca

721 ggcctttgagc ccaattgttc taaagttagta gctgccacat gcaccaggat gatggaaacg

781 cagacttcta cgtgggtttgg cttttgtggc gctagagcag agaaaagaac atatattttt

841 tggcatagta aagatggcag ggctataatc agctgtaaca catatgataa tctctctata

901 cattgtgaga ggccaggaaa taaaacagta gtgccgataa cacttatgtc gggattagtg

961 tttcattccc agccaatcaa taacagacct agacaagcat ggtgctggtt cacaggcaca

1021 tggagagaag ccattgcagga ggtgaagcaa acccttgtaa aacaccccag atatagggga

1081 accaatgaga caaggaagat taactttaca gcaccgggaa aaggctcaga cccggaagcg

1141 gcatacatgt ggaactaactg caggggagag tttctctact gcaacatgac ttgggttctc

1201 aattggatag agaacagaac tgtgaatagg acatggcaca attatgtgcc atgccatata

1261 aagcagataa ttaacacctg gcataaggta gggaaacatg tatatttgcc tcccaggga

1321 ggggagttga cctgcaactc aactgtgacc agcataattg ctaacattga cggaggtaat

1381 cagacaaata ttacctttag tgcagaggtg gcagaactat accgactgga attgggagat

1441 tacaacttgg tagaaataac accaattggc ttgcaccta catcagaaaa ccgatactcc

1501 tctgctcacg ggagacatac gaga

envelope cds ->

LOCUS HIV2D808 1512 bp ss-RNA VRL 15-APR-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gpl05 envelope glycoprotein sequence, sample D808, partial cds.

ACCESSION U07107

SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from uncultured PBMCs derived from a patient from Bombay, India whose sera was dually reactive (HIV1/HIV2).

REFERENCE 1 (bases 1 to 1512)

AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Mahambra,G., Delwart,E.L., Mullins,J.I. and Ruebsamen-Waigmann,H.

TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses

JOURNAL J. Virol. 68, 2161-2168 (1994)

COMMENT This sequence was obtained as part of a study to determine the genetic breadth of HIV-2 variation in India. In this study, blood samples were obtained from four patients from Bombay, India (D766, D808, D868, and D1024) and from one patient from Goa, India (D1071); each of the five patients was serologically dually reactive. The mean genetic divergence of the set of four sequences from Bombay was 7.3%, while the sequence from Goa was from 7 to 10.5% different from each of the Bombay sequences. The entire set of Indian sequences, including both the Bombay and Goa sequences, diverged from eleven non-Indian HIV-2 sequences by 23%. Because of the similarity of the Indian sequences, Grez et al. suggest that similar HIV-2 strains might be found elsewhere in India. All five of these sequences clustered with HIV-2 subtype A viruses. HIV-1 env sequences from Indian patients D766, D808, D868, and D1024 clustered with HIV-1 C viruses, and as a set showed a similar degree of interpatient genetic similarity as did their HIV-2 counterparts.

BASE COUNT 536 a 304 c 319 g 353 t

ORIGIN envelope cds start

1 ATGgcgcccc aaagactca gctgcttatt tccattttac taactagtgc ttgcttaata

env cds ->

61 tattgtgaac aatatgtaac tgttttctat gccataccgc cgtggaggaa tgcattccatt

121 cccctctttt gtgcaaccag aaatagggat acttggggaa ccatacagtg cttgccagac

181 aatgatgact atcaggaaat agctttaaat gtgacagaag ctttcgatgc atggaacaat

241 acagtaacag aacaagcaat agaagatgtc tggaaatctat ttgagacatc aacaaaacca

301 tgtgtcaaat taacaccctt atgtgttagca atgaattgca gcaaaaacta taataacaca

361 agtgcaggga acacaaccac acccacagcc gaatccatga taatgagaa ctcttcacgc

421 atacgcgcaa acaactgttc aggattggaa gaggaagaga tagtcaattg tcagttcaat

481 atggcaggat tacogagaga taaccggaaa tggataatg aaacatggta ctccagcgat

541 gtgggtttgtg aaataacaa cccaacccaa aacagatgtt acatgaacca ttgcaacaca

601 tcagtcacatc cagagtcacg tgaaaagcac tattgggatg atataaggtt tagatattgt

661 acaccacogg gttatgcctt gctaaaatgc aatgatacca attactcagg ctttgagccc

721 aattgctcta aagtagtagc tgcacacgc actaggatga tggaaacaca aacttctacg

781 tggtttggct ttaatggcac tagagcagag aacagaacat acatttattg gcatagttaa

841 gataacagaa ctataatcag cctaaacaca cattataatc tctctatata ttgtaagagg

901 ccaggaaata aaacagtagt accaataaca cttatgtcag gattagtgtt tcattcccag

961 ccaatcaata acagacctag acaagcatgg tgctggttca aaggcagatg gaaggaagcc

1021 atgcaggagg tgaaggaaac ccttacaaaa caccocagggt ataggggaac caatgataca

1081 ggaagagatta actttgcagc accgggaaaa ggctcagacc cgggaagcggc atacatgtgg

1141 actaactgca ggggagaatt tctctattgc aacatgactt ggttcctcaa ttgggtggaa

1201 aacaaaactg caaatcagac atggcacaat tatgtgccat gccatataaa gcaataaatt

1261 aacacctggc ataaggtagg gaacatgta tatttgctc ccagggaagg ggaatttaac

1321 tgcaactcaa ctgtgaccag cataattgct aacattgacg gaggtaatca gacaaatatt

1381 accttttagtg cagaggtggg agaactatac cgattgggaa ttggagatta caaattggta

1441 gaaataacat caattgggtt cgcacctaca tcagaaaaac gggactctc tgctcacggg

1501 agacatagag ag

envelope cds ->

HIV2D868

LOCUS HIV2D868 1525 bp ss-RNA VRL 15-APR-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp105 envelope glycoprotein sequence, sample D868, partial cds.
 ACCESSION U07108
 SOURCE Human immunodeficiency virus type 2 from PCR-amplified DNA extracted from uncultured PBMCs derived from a patient from Bombay, India whose sera was dually reactive (HIV1/HIV2).
 REFERENCE 1 (bases 1 to 1525)
 AUTHORS Grez,M., Dietrich,U., Balfe,P., Von Briesen,H., Maniar,J.K., Mahambre,G., Delwart,E.L., Mullins,J.F. and Ruebsamen-Waigmann,H.
 TITLE Genetic analysis of human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2) mixed infections in India reveals a recent spread of HIV-1 and HIV-2 from a single ancestor for each of these viruses
 JOURNAL J. Virol. 68, 2161-2168 (1994)
 COMMENT This sequence was obtained as part of a study to determine the genetic breadth of HIV-2 variation in India. In this study, blood samples were obtained from four patients from Bombay, India (D766, D808, D868, and D1024) and from one patient from Goa, India (D1071); each of the five patients was serologically dually reactive. The mean genetic divergence of the set of four sequences from Bombay was 7.3%, while the sequence from Goa was from 7 to 10.5% different from each of the Bombay sequences. The entire set of Indian sequences, including both the Bombay and Goa sequences, diverged from eleven non-Indian HIV-2 sequences by 23%. Because of the similarity of the Indian sequences, Grez et al. suggest that similar HIV-2 strains might be found elsewhere in India. All five of these sequences clustered with HIV-2 subtype A viruses. HIV-1 env sequences from Indian patients D766, D808, D868, and D1024 clustered with HIV-1 C viruses, and as a set showed a similar degree of interpatient genetic similarity as did their HIV-2 counterparts.
 BASE COUNT 526 a 301 c 335 g 363 t
 ORIGIN envelope cds start

```

1 ATGgcgcgtg aaaggactca gctaattatt gccacattac tagctagtgc ttgcttaaga
env cds ->
61 tattgccaac aatatgtaac tgttttctat ggcatacccg cgtggaagaa tgcattccatt
121 cccctctttt gtgcaaccaa aaatagagat acttggggaa ccatacagtg cttgccagac
181 aatgatgatt atcaggagat agctttgaat gtgacagagg ctttcgatgc atggaacaat
241 acagtaacag aacaagcaat agaagatgtc tggaaatctat ttgagacatc aacaaaacca
301 tgtgtcaaac taacaccctt atgtgtagca atgaattgca gcaaggacag tagcaacaca
361 agtacaggga acacaccac atccacagcc tcttatgtta taatgagac cctctcatgc
421 atacacacag acacctgctc aggatgtgag gaggaagaga tagtcagttg tcagttcagt
481 atgacaggat tagcaagaga taagccaaaa aggtataatg aaacatggtc ctcaagtgat
541 gtggtttgtg aggcaacagg acaaatgac acctaaacc agagcaggtg ttacatgaac
601 cattgcaaca catcagtcac cacagagtca tgtgataagc actattggaa tgacatgagg
661 tttagatatt gttcaccacc gggttatgcc ttgctaagat gcaatgattc caattactca
721 ggctttgagc ccaattgttc taaagtagta gctgccacat gcaccaggat gatggaaacg
781 caaacttcta catggtttgg ctttaattgg actagagcag agaatagaac atacgtttat
841 tggcatagta aagataacag aactataatc agcttaaca aatactataa tctctctata
901 cattgtaaga ggccaggaaa taaaacagtg gtaccaataa cacttatgtc aggattagtg
961 tttcattccc agccaatcaa tagaagacct agacaagcat ggtgctggtt cacaggcaaa
1021 tgggaaggag ctatgcggga ggtgaggcaa accattgcag aacaccccag gtataggggg
1081 accaataata caggaaatat taaatttgtg gcaccgggaa aaggctcaga cccggaagcg
1141 gcatacatgt ggactaaactg caggggagaa tttctctact gcaacatgac ttggttcctc
1201 aattggatag agaacaaaac tgagaatgca acatggcaca attatgtgcc atgccatata
1261 aagcaataaa ttaacacctg gcataaggta gggaaacatg tatatttgcc tcccagggaa
1321 ggggagttga cctgcaactc aactgtgacc agcataattg ctaacattga cggaggtaat
1381 cagacaaata ttaccttag tgacagagtg gcagaactat accgattaga attgggagat
1441 tataaattgg tagaataaac accattggc ttgcaccta caccagaaaa acgatactct
1501 tctgctcaca ggagacatcc gagag
envelope cds ->

```

LOCUS HIV2FA1 784 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample FA, clone 2, partial cds.

ACCESSION L33082

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 36 year old female living in urban Ghana.

REFERENCE 1 (bases 1 to 784)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient FA was a 36-year old female who was identified at the University of Ghana Medical Center in Accra as HIV-2 positive and who was admitted with pulmonary tuberculosis. She had lived in both Ghana and Cote d'Ivoire where she suffered from AIDS dementia complex. Virus isolation was successfully performed at the National Cancer Institute. Virus FA clusters with one of the subsets of subtype A in gag and pol(int) and with the other subset in env and pol(RT). Gao et al. conclude that it may be a recombinant virus derived from two subtype A viruses, one from each of the two subsets. NCBI gi: 532149

BASE COUNT 271 a 164 c 201 g 148 t

ORIGIN 111 bp downstream from the start of gag

1 agcgaacgaa ttggacagat ttggattggc agagagcctg ttggagtcaa aagaaggttg
gag cds ->

61 ccaaaagatt atcacagttt taaatccatt agtaccgaca gggtcagaaa atttacaag
121 tctttttaat acagtctgcg tcatctggtg cttgcacgca gaagagaaag tgaagatac
181 tgaagaagca agsaaactag tacagagaca tctagtggca gaaacaggaa ctgcagagaa
241 aatgccaaat acaagtagac caacagcacc acctagcggg agtggaggaa acttcctgt
301 gcaacagaca ggcggtggca actacgtcca tgtgccacta agtccccgaa ctctaaatgc
361 ttgggtaaaa ttaatagagg ataagaagtt cggggcagaa gtagtgccag gatttcaggc
421 actctcagaa ggcctgcacg cctatgatat caatcaactg cttaattgtg tggcgatca
481 ccaagcagct atgcaataaa tcagagaaat tatcaatgaa gaagcagcag attgggatgc
541 acagcaccga ataccaggcc ccttgccagc agggcagctt agagacccaa ggggatctga
601 catagcagga acaacaagca cagtagatga acagatccag tggatgtata gggcacaana
661 tcccataccg gtagggaaca tctacagaag gtgggtccag atagggtac agaagtgtgt
721 caggatgtac aacccaacta acgtccctaga cgtgaagcag ggaaccaagg aatcgttcca
781 gagc
gag cds ->

HIV2FA2

LOCUS HIV2FA2 708 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) pol integrase sequence, sample FA, clone 1, partial cds.
 ACCESSION L33089
 SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 36 year old female living in urban Ghana.
 REFERENCE 1 (bases 1 to 708)
 AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient FA was a 36-year old female who was identified at the University of Ghana Medical Center in Accra as HIV-2 positive and who was admitted with pulmonary tuberculosis. She had lived in both Ghana and Cote d'Ivoire where she suffered from AIDS dementia complex. Virus isolation was successfully performed at the National Cancer Institute. Virus FA clusters with one of the subsets of subtype A in gag and pol(int) and with the other subset in env and pol(RT). Gao et al. conclude that it may be a recombinant virus derived from two subtype A viruses, one from each of the two subsets.
 NCBI gi: 532151
 BASE COUNT 290 a 135 c 153 g 130 t
 ORIGIN 2184 bp downstream from the start of pol

```

1 aatgggaata gtagtaggcc aaccaacaga gtcagaaaat agaatagtaa accaaatcat
pol cds ->
61 tgaggacatg ataaagaag aagcaatcta tgttgcattg gtcccagccc ataaaggcat
121 aggaggaaac cagggaagtag accacttagt aagtcagggc atcagacagg tattattcct
181 ggaagaata gagcccgctc aagaagaaca tgaataatat catagcaata taaaagagct
241 aactcataaa ttgtgaatac cccaactagt agcaagacag atagtaaaac catgtgcccc
301 atgccaacag aaaggagagg ctatacatgg gcaagtaaat gcagaaatag gtgtctggca
361 aatggactgc acacacttag aaggaaaaat cattatagta gcagtacatg ttgcaagtgg
421 attcatagaa gcaaaagtca tcccacagaa atcagaaaga cagacagcac tcttctatt
481 aaaactggcc agcaggtggc caataacaca cttgcacaca gacaatggct ccaacttcac
541 ctacacaggaa gtgaaaatgg tggcatgggt gataggtata gagcaaacct tcggagtacc
601 ttacaatcca caaagccaag gtagtagtag ggcaatgaat caccacctaa aaaatcagat
661 aagtagaatt agagagcagg caaacacaat agaaacaata gtactaat
pol cds ->

```

LOCUS HIV2FA3 453 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp41 env sequence, sample FA, clone 15, partial cds.

ACCESSION L33095

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 36 year old female living in urban Ghana.

REFERENCE 1 (bases 1 to 453)

AUTHORS Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F0784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient FA was a 36-year old female who was identified at the University of Ghana Medical Center in Accra as HIV-2 positive and who was admitted with pulmonary tuberculosis. She had lived in both Ghana and Cote d'Ivoire where she suffered from AIDS dementia complex. Virus isolation was successfully performed at the National Cancer Institute. Virus FA clusters with one of the subsets of subtype A in gag and pol(int) and with the other subset in env and pol(RT). Gao et al. conclude that it may be a recombinant virus derived from two subtype A viruses, one from each of the two subsets.

NCBI gi: 532147

BASE COUNT 161 a 78 c 108 g 106 t

ORIGIN 1861 bp downstream from the start of env

1 gacgtggtca agagacaaca agaatgttg cgaatgacg tctggggaac aaaaatctc

env cds ->

61 caggcaagag toactgctat tgagaataac ttaaaggacc aggcgcaact aaattcatgg

121 ggatgcgcgt ttaggcaagt ctgccacact actgtaccat gggtaaatga cagcttgaca

181 cctgattgga acaacatgac gtggcagcaa tgggaacaac aaatccgcta cctggaggca

241 aatatcagt agagtttaga acaggcacaa atccagcaag aaaagaacat gtatgaacta

301 caaaaattaa atagctggga tctctttggc aactggtttg atttaacctc ctgggtcaaa

361 tatattcagt atggagttta tatagtagta ggaatagtag ctottagaat aataatatat

421 gtagtacaaa tgtaagtag acttagaagg ggc

env cds ->

HIV2FO784

LOCUS HIV2FO784 781 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample FO784, clone 103, partial cds.
 ACCESSION L33083
 SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 46 year old asymptomatic man from rural Liberia who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 781)
 AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient FO784 was a 46 year old asymptomatic male from rural Liberia who had no history of sexually transmitted disease, chronic illness, blood transfusion or homosexuality. His serum was HIV dually reactive. Virus isolation was attempted twice at the University of Alabama at Birmingham without success. This sample defines a distinct subtype, which has been designated as subtype D by Gao et al. Interestingly, virus FO784 clustered with SIVsm and SIVmac, rather than with other HIV-2 strains. This and other strong evidence suggest that human and macaque contact with infected sooty mangabeys may be the source of HIV-2 and SIVmac infection [1].
 NCBI gi: 532153
 BASE COUNT 278 a 158 c 196 g 149 t
 ORIGIN 111 bp downstream from the start of gag

1 agcaaatgaa ttggacagat ttggattagc agaaagcctg ttggaaaata aagaagggtg
 gag cds ->
 61 tcaaaaaatt ctctcggtct tagctccatt agtaccgaca gggttcagaaa atttaaaaaag
 121 cctttacaat actgtctgcg tcatctggtg ttacacgcga gaagagaaaag tgaagacac
 181 tgaggagagca aaacagatag tgcagagaca tctagtgtg gaaacaggaa ctgcagacaa
 241 aatgccatca acgagtagac caacagcacc acctagtac agaggagaa attaccagt
 301 gcagcaagtc ggcggaaatt atgtccacct accgctgagc cccagaacat taaatgcatg
 361 ggttaagtta gtggaggaca aaaaattcgg ggcagaggta gtgccagggt ttcaggcact
 421 atcggaaggc tgcactccgt atgacatcaa tcagatgcta aattgtgtag gagaacatca
 481 ggcagccatg cagatcataa gggaaataat caatgatgag gcagcagatt gggatcagca
 541 gcatccacaa ccaggccacac taccagcagg acagctcaga gatccacgag gatctgatat
 601 agcaggaacc actagcacag tggaggaaca aatacagtggt atgtacaggc agcagaatcc
 661 cataccagtt ggaaatatct ataggagatg gatccagcta gggttacaga aatgtgtcag
 721 aatgtacaac ccaactaaca ttctggatat aaaacaaggg ccaaaagaga cgttccagag
 781 c
 gag cds ->

LOCUS HIV2JA1 766 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample JA, clone 1, partial cds.

ACCESSION L33085

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 27 year old female living in urban Ghana who showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 766)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Naequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, PA, FO784, PT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient JA was a 27-year old female, married to a Nigerian man. She had lived in Nigeria for five years before she became ill with chronic diarrhea, severe weight loss, anemia and alopecia and returned to Ghana. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity. Sample JA clusters with the subtype B viruses.

NCBI gi: 532159

BASE COUNT 265 a 163 c 193 g 145 t

ORIGIN 111 bp downstream from the start of gag

1 agtgaacgaa ttggacagat ttggattagc ggagagcctg ttggaatcaa aagaaggatg
gag cds ->

61 ccacaaaatt cttacggtct tagctccatt agtaccaaca ggatcagaaa atttaaaaag
121 tctttttaat actgtctgog ttgtttattg cttgcacgca gaagagaaag tgaagatac
181 tgaggaagca aaaaagatag cacagagcca tctagcgggtg aacacggaaa aaacyccagc
241 tacaagtaga ccaacagcac cgcttagcgg cagaaactat ccagtgcagc aggtggccgg
301 caattacgtc cacctgcgcg taagcccccg aacottaaat gcctgggttae aactagtaga
361 gganaagaaa ttccggggcag aagtagtacc aggattccag gcactatcag aagggtgcac
421 ccottatgat attaatcaaa tgcataattg tgtaggggaa catcaagcag ccattgcaaat
481 tattagggaa ataataaatg aggaagcagc agactgggac cagcagcacc cgggtgccagg
541 tccactacca gcaggacagc ttagggaacc aagagggtca gacatagcgg gaaccaccag
601 cacagtggag gaggagatac agtggatgta cagggcccaa aatcctgtcc cagtaggaaa
661 catttatagg agatggattc aattagggct gcagaagtgt gtccggatgt ataacctac
721 taatatactg gacataaaaag aaggaccaaa agaacccttc caaagc
gag cds ->

HIV2JA2

LOCUS HIV2JA2 708 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) pol integrase sequence, sample JA, clone 9, partial cds.

ACCESSION L33090

SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 27 year old female living in urban Ghana who showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 708)

AUTHORS Gao,P., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F07B4, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.

Patient JA was a 27-year old female, married to a Nigerian man. She had lived in Nigeria for five years before she became ill with chronic diarrhea, severe weight loss, anemia and alopecia and returned to Ghana. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity. Sample JA clusters with the subtype B viruses.

NCBI gi: 532161

BASE COUNT 275 a 130 c 167 g 136 t

ORIGIN 2184 bp downstream from the start of pol

1 catgggaata atagctggac agccaacaga aacagagtca ccaatagtaa ataaaataat

pol cds ->

61 tgaagaaatg atcaaaaagg aagcaatata tgtaggatgg gtgccagctc acaaaggact

121 aggtggtaat caggaaagtag accacctagt aagtcaagga attaggcagg ttctattcct

181 agaaaagata gaaccagcac aagaagaaca tgaaaagtat catggcaatg taaaagaact

241 agttcataaa tttggaatcc acagtttagt ggcaaaacaa atagttaaact cctgtgataa

301 atgccaacaa aaaggggaag ctgttcctag acaggtaaat gcagaactag gaacgtggca

361 gatggactgt acacacttag aaggaaaggt cataatagtg gcagtccatg tagccagtgg

421 atttatagag gcagaagtaa tcccccaaga aacgggaaga cagacagctc tcttcctgtt

481 aaaattggcc agcagatggc ctatcacaca cctacacaca gacaatggcg ccaacttcac

541 ctcacaagac gtgaagatgg tggcctgggt ggtaggaata gaacgaacct tcggcggtacc

601 ctataaccca caaagtcagg gagtagtaga agcaatgaac catcacctaa aaaaccaaat

661 agatagaatc agggatcagg cagtatcaat agagacagtt gtgctaata

pol cds ->

LOCUS HIV2JA3 453 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp41 env sequence, sample JA, clone 12, partial cds.

ACCESSION U33096

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 27 year old female living in urban Ghana who showed HIV dual seroreactivity.

REFERENCE 1 (bases 1 to 453)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient JA was a 27-year old female, married to a Nigerian man. She had lived in Nigeria for five years before she became ill with chronic diarrhea, severe weight loss, anemia and alopecia, and returned to Ghana. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity. Sample JA clusters with the subtype B viruses.

NCBI gi: 532157

BASE COUNT 162 a 85 c 105 g 101 t

ORIGIN 1861 bp downstream from the start of env

1 gacgtggtca aaagacaaca ggaactgttg cggctgaccg tctgggggac aaaaaacctc
 env cds ->

61 cagacaagag tcactgocat cgagaaatac cttaaaggacc aggcactact aaactcgtgg
 121 ggtgtgtgctt ttagacaggt ctgtcacaca actgtaccat ggccaaatga aactctccat
 181 ccaactggc aaaacatgac atggcagcag tgggaacaga aggtcaacta cttggatgca
 241 aatataacac aattattaga agaggcacag atacagcagg aaaagaatat gtatgagcta
 301 cagaaattaa accaatggga tgttttttgt aattggtttg acttcaactc ctggatggca
 361 tacatcaaat taggactata tgtagtagca ggattaatag tgttaaggat agtaatatat
 421 attatacaga tgctagcaaa acttaggaag ggc
 env cds ->

HIV2ONFT1

LOCUS HIV2ONFT1 766 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample ON, clone 18, partial cds.
 ACCESSION L33086
 SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 48 year old fisherman living in urban Ghana who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 766)
 AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type. Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.
 Patient ON was a 48 year old male fisherman who had sexual encounters with prostitutes in Ghana and in Cote d'Ivoire. He was hospitalized at the University of Ghana Medical Center in Accra where he was identified to be HIV-2 positive and suffered from end-stage AIDS. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity.
 Patient FT, patient ON's wife, was also identified at the University of Ghana Medical Center in Accra as HIV-2 positive. She was clinically asymptomatic. Samples ON and FT both cluster with the subtype B viruses.

NCBI gi: 532165

BASE COUNT 267 a 166 c 188 g 145 t

ORIGIN 111 bp downstream from the start of gag

1 agtgaacgaa ttggacagat ttggattggc agagcgccctg ttggaatcaa aagaaggatg
 gag cds ->

61 tcacaaaatt cttacggctt tagctccatt agtaccacaa gggttcagaaa atttaaaaag
 121 ccttttttaat actgtctgcg tcctctactg cttgcacgca gaagagaaaag tgaaagatac
 181 tgaggaagca aaaaagatgg cacagagcca tctagtgggtg gacgcagaaa aaacgccagc
 241 tacaagtaga ccaacagcac cacctagogg cagaaactat ccagtgcac aggtagccgg
 301 caattatgtc cagtgccgc taagcccccg aacctaaat gctggggtaa aattaataga
 361 agaaaagaag ttccggggcag aagtagtgcc aggatccaa gcactatcgg aagggtgcac
 421 ccttatgat atcaatcaga tgctaaattg tgtaggagaa caccaggcag ccatgcaaat
 481 tattagggaa ataatcaatg aggaggcagc agactgggac cagcagcaac ccataccagg
 541 tccactgccg gcaggacaac ttagggaacc cagagggtca gatatagcag gaaccaccag
 601 cacagtagag gaacaaatac agtggatgta caggcagcca aatcctgttc cggtagggaa
 661 catttacaga agatggattc aaatagggct gcagaaatgt gttcggatgt acaattccac
 721 caatatatta gatgtaaaac aagggccaaa ggaacccttc caaage

gag cds ->

LOCUS HIV2ONFT2 708 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) pol integrase fragment, sample ON, clone 6, partial cds.
 ACCESSION U33091
 SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 48 year old fisherman living in urban Ghana who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 708)
 AUTHORS Gao, F., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ko, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, PA, PO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.
 Patient ON was a 48 year old male fisherman who had sexual encounters with prostitutes in Ghana and in Cote d'Ivoire. He was hospitalized at the University of Ghana Medical Center in Accra where he was identified to be HIV-2 positive and suffered from end-stage AIDS. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity.
 Patient FT, patient ON's wife, was also identified at the University of Ghana Medical Center in Accra as HIV-2 positive. She was clinically asymptomatic. Samples ON and FT both cluster with the subtype B viruses.
 NCBI gi: 532167

BASE COUNT 277 a 130 c 164 g 137 t

ORIGIN 2184 bp downstream from the start of pol

1 catgggaata atagctggac aaccaacaga aacagagtca ccaatagtga ataaaataat
 pol cds ->

61 tgaagaaatg atcaaaaagg aagcaatata tgtaggatgg gtgccagctc acaaaggact
 121 aggttgtaat caggaagtag accacctagt aagtcaagga attaggcagg ttctattctt
 181 agaaaagata gaaccagcac aagaagaaca tgaaaagtac catggcaatg taaaagaact
 241 agttcataaa tttggaatcc cacagttagt agcaaaacaa atagtaaact cctgtgataa
 301 atgccaacaa aaaggggaag ctgttcattg acaggttaat gcagaactag gaacatggca
 361 gatggactgt acacacttag aagggaagat cataatagtg gcagtccatg tagccagtgg
 421 atttatagag gcagaggtaa taccocaaga aacaggaaga caaacagctc tcttctatt
 481 aaaattggcc agcagatggc ctatcacaca cctgcgcaca gacaatggcg ccaacttcac
 541 ctcaacaagt gtgaagatgg tggcctgggt ggtaggagata gaacaaacct tgggggtacc
 601 ctataacca caaagtcagg gagtatgga agcaatgaac catcatctaa aaaaccaaat
 661 agatagaatc agggatcagg cagtatcact agaaacagtt gtactaat

pol cds ->

HIV2ONFT3

LOCUS HIV2ONFT3 453 bp ds-DNA VRL 23-AUG-1994
 DEFINITION Human immunodeficiency virus type 2 (HIV-2) gp41 env sequence, sample ON, clone 12, partial cds.
 ACCESSION L33097
 SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from a 48 year old fisherman living in urban Ghana who showed HIV dual seroreactivity.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
 JOURNAL J. Virol. 68, 7433-7447 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, F0784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation.
 Patient ON was a 48 year old male fisherman who had sexual encounters with prostitutes in Ghana and in Cote d'Ivoire. He was hospitalized at the University of Ghana Medical Center in Accra where he was identified to be HIV-2 positive and suffered from end-stage AIDS. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity.
 Patient FT, patient ON's wife, was also identified at the University of Ghana Medical Center in Accra as HIV-2 positive. She was clinically asymptomatic. Samples ON and FT both cluster with the subtype B viruses.

NCBI gi: 532163

BASE COUNT 155 a 83 c 113 g 102 t

ORIGIN 1861 bp downstream from the start of env

1 gacgtggtca aaagacaaca ggaactgttg cggctgaccg tctgggggac gaaaaacctc
 env cds ->

61 cagacaagag tcaactgctat cgagaaatac ctaaaggacc aggcactact aaattcgttg
 121 ggatgtgctt ttagacaggt ctgccacacg actgtaccat ggccaaatga aactatcgaa
 181 ccagactgga acaacatgac atggcagcaa tgggaagcaa aggtccaatt ctgggatgca
 241 aatataacac aattcattaga agaagcacag atacagcagg aaaagaatat gtatgagcta
 301 cagaaattga atcagtggga tgttttttagt aattggtttg acttcacctc ctggatggca
 361 caccttaggt taggactata tgtagtgcga ggaataatag tgttaaggat agtaatatat
 421 attatgcaga tgctagcaag gcttaggaag ggc

env cds ->

LOCUS HIV2ONFT4 766 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample FT, clone 1, partial cds.

ACCESSION L33084

SOURCE Human immunodeficiency virus type 2 PCR-amplified DNA recovered from an asymptomatic female living in urban Ghana.

REFERENCE 1 (bases 1 to 766)

AUTHORS Gao, P., Yue, L., Robertson, D.L., Hill, S.C., Hui, H., Biggar, R.J., Neequaye, A.E., Whelan, T.M., Ho, D.D., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient FT, patient ON's wife, was identified at the University of Ghana Medical Center in Accra as HIV-2 positive. She was clinically asymptomatic. Patient ON, patient FT's husband, was a 48 year old male fisherman who had sexual encounters with prostitutes in Ghana and in Cote d'Ivoire. He was hospitalized at the University of Ghana Medical Center in Accra where he was identified to be HIV-2 positive and suffered from end-stage AIDS. Subsequent Western blot analysis of peripheral blood cells showed HIV dual reactivity. Samples FT and ON cluster with the subtype B viruses.

NCBI gi: 532155

BASE COUNT 269 a 162 c 185 g 150 t

ORIGIN 111 bp downstream from the start of gag

1 agtgaatgaa ttggacagat ttggattagc agagaccctg ttggagtcaa aagaaggatg
gag cds ->

61 tcacaaaatt cttacggctct tagctccatt agtaccaaca ggatcagaaa atttaaaaag
121 ccttttttaatt actgtctgag tcatttttctg cttgcacgca gaacagaaaag tgaagatgac
181 tgagggaagca aaaaagatag cacagagcca tctagtgtg gacgcagaaa aagtgccagc
241 tacaagtaaa tcaacagcac cacctagtgg cagaaattat ccagtgcagc aggtagccgg
301 caattatgtc caagtgcgc taagcccccg aaccctaaat gcctgggtta aattaataga
361 agaaaagaag ttcggggcag aagtatgcc aggattccaa gcactatcgg aagggtgcac
421 ccttatgat attaatcaga tgctaaattg tgtaggagaa caccaggcag ccattgcaaat
481 tattagggaa ataataatg agggggcggc agactgggac cagcagcatc caataccagg
541 cccactgcca gcaggacaac ttagggaacc cagagggtca gatatagcag gaaccaccag
601 cacagtagag gaacaaatac agtggatgta caggcaacca aatcctgttc cggtagggaa
661 catttacaga agatggatcc aatatgggct gcagaaatgt gtccggatgt acaatccac
721 taatatatta gatgtaaaac aagggccaaa ggaacccttc caaagc
gag cds ->

HIV2PA1

LOCUS HIV2PA1 781 bp ds-DNA VRL 23-AUG-1994

DEFINITION Human immunodeficiency virus type 2 (HIV-2) p16/p28 gag sequence, sample PA, clone 1, partial cds.

ACCESSION U33087

SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 35 year old asymptomatic male U.S. immigrant from rural Sierra Leone.

REFERENCE 1 (bases 1 to 781)

AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology

JOURNAL J. Virol. 68, 7433-7447 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient PA was a 35 year old male U.S. immigrant from rural Sierra Leone who apparently acquired HIV-2 infection prior to his move to the United States, roughly twenty years ago, and thus represents a long-term survivor of HIV-2. He had normal CD4 levels and was otherwise asymptomatic. Sample PA defines a distinct subtype, which has been designated as subtype E [1].

NCBI gi: 532169

BASE COUNT 276 a 167 c 194 g 144 t

ORIGIN 111 bp downstream from the start of gag

1 agccagagaa ttagacagat ttggattggc agaaagcctg ttggaaaaca aggaaggatg
gag cds ->

61 tcacagaata ctgacagtac tagaaccatt aatgccaaaca ggctctgaaa atttaaaaag
121 cttgtttaac gctgtctgcg tagtttactg cctgcacgca gcagtgaag tgaagatatac
181 agaggaagcg aaaaaacatg tacagcagca ccacctagtg gcaggaggta aaactacaga
241 gaaactgcct cctcagagta gaccaacagc gccacctagt ggagggtggag ggaattatcc
301 agtgcaacag ataggaaata actatgtgca ctctccactg tccccaaaga cattgaatgc
361 atgggtcaaa ttagtagaag aaaagaaatt tggagcagag gtagtgccag gottccaggc
421 attatcagaa ggatgcaccc cgtatgacat caaccagatg ctttaattgag tgggggaaca
481 tcaggtcagc atgcaaatga tcagagagat aatcaatgaa gaagcagcag attgggacgt
541 acagcatcca agaggggaac cgcacagaca gggcctaaga gacccatcag gatcagacat
601 agcagggaca accagtaccc ccgcagaaca aatagagtgg atgtacagga atccaaatcc
661 aatccctgtg ggagacatct atagaagatg gatccagcta gggctccaga aatgtgtcag
721 aatgtataat ccaacaaaca ttctggacgt caaacagggg cccaaagaat cttttcagag
781 c

gag cds ->

LOCUS HIV2PA2 995 bp ds-DNA VRL 23-AUG-1994
DEFINITION Human immunodeficiency virus type 2 (HIV-2) pol reverse transcriptase sequence, sample PA, partial cds.
ACCESSION L33093
SOURCE Human immunodeficiency virus type 2 PCR-amplified PBMC DNA recovered from a 35 year old asymptomatic male U.S. immigrant from rural Sierra Leone.
REFERENCE 1 (bases 1 to 995)
AUTHORS Gao,F., Yue,L., Robertson,D.L., Hill,S.C., Hui,H., Biggar,R.J., Neequaye,A.E., Whelan,T.M., Ho,D.D., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
TITLE Genetic diversity of human immunodeficiency virus type 2: Evidence for five distinct sequence subtypes with differences in virus biology
JOURNAL Unpublished (1994)
COMMENT This sequence was obtained as part of a study designed to determine the in vivo extent of HIV-2 genetic and biological variation. Up to now, HIV-2 sequences have been derived almost solely from cultured isolates. Since cultivation may select against viruses which can not grow in a particular cell type, Gao et al. argue that a greater variation may be elucidated by sequencing virus obtained from uncultured PBMCs. With this aim in mind, peripheral blood samples were collected from 12 HIV-2 seropositive patients (2238, 60415K, 60667K, 7312A, 7810A, 7924A, FA, FO784, FT, JA, ON, and PA) from six different West African countries, living in both urban and rural areas. Uncultured PBMC DNA was PCR amplified, and the recombinant clones were manually sequenced. Another portion of the cells was used for virus isolation. Patient PA was a 35 year old male U.S. immigrant from rural Sierra Leone who apparently acquired HIV-2 infection prior to his move to the United States, roughly twenty years ago, and thus represents a long-term survivor of HIV-2. He had normal CD4 levels and was otherwise asymptomatic. Sample PA defines a distinct subtype, which has been designated as subtype E [1].
NCBI gi: 532171
BASE COUNT 392 a 182 c 218 g 203 t
ORIGIN 806 bp downstream from the start of pol
1 agaaggaggg acaattagaa cgagccccc cactaatcc ttataataca ccaacctttg
pol cds ->
61 ctattaggaa aaaagataag agtaaatgga ggatgcttat agatttcagg gagctaaata
121 aggcacaca ggacttcaca gaagccaac tgggtattcc acaccagca ggtctaaaag
181 aaatggagca gataacagtg ttagatatag gggatgcta tttttoggta cccctagacc
241 cagaatttag acagtataca gcattcacca taccctcggt aataatcaa gaaccaggaa
301 aaagatacat ctacaaggtc ctaccacaag ggtggaaggg atccccagca gtcttcagg
361 agaccatgag acaggtacta gagcccttca gaaaagcaaa ttgggacgtt ctactcatc
421 agtacatgga cgacttacta attggcagta acaggagcat gacagagcat gacaggatgg
481 tagtacaatt aaaagacatg ctgaacaacc tggggttcag caccocagag gataagtttc
541 agagagaacc acccttaaaa tggatggggt atttactata cccaagaaa tggaaactac
601 aaaaaataga gcttctgaa aaagaaagat ggacagttaa tgacattcag aagttagtag
661 gaataactaa ttgggcagca caagtataac caggaatcaa aacaaaaaat ttatgcaagc
721 tcatcagagg aaaaatgacc ctaacagaag aggttcagtg gacagaattg tcagaggcag
781 aattggcaga aaataagatt attctaagtc agggacagga gggaaagatac tatagggaag
841 aagaaaactt agaagcaaca gtactaaaga accaagataa tcagtggagc tataagattc
901 atcagggaga cagaattctt aaagtaggaa aatttcccc aataaaaaat acccataca
961 atggagtaag actattagct agttagtagt agaaa
pol cds ->

SIVCPZGAB2

LOCUS SIVCPZGAB2 280 bp ss-RNA VRL 15-JUN-1994

DEFINITION Simian immunodeficiency virus (SIVcpz) partial pol sequence, sample SIVcpzGAB2.

ACCESSION U11495

SOURCE Simian immunodeficiency virus from uncultured PCR-amplified PBMC DNA derived from a chimpanzee who lived in a tropical rainforest in northeastern Gabon.

REFERENCE 1 (bases 1 to 280)

AUTHORS Janssens,W., Fransen,K., Peeters,M., Heyndrickx,L., Motte,J., Badjabaga,L., Delaporte,E., Piot,P. and van der Groen,G.

TITLE Phylogenetic analysis of a new chimpanzee lentivirus SIVcpz-gab2 from a wild-captured chimpanzee from Gabon

JOURNAL AIDS Res. Hum. Retroviruses 10, 1191-1192 (1994)

COMMENT This is the third chimpanzee lentivirus that has been isolated and molecularly characterized. The first isolated was derived from a chimpanzee from Gabon (SIVcpzGAB) and the second isolated was derived from a chimpanzee from Zaire (SIVcpzANT still under study). This 280 bp pol sequence was derived from a wild chimpanzee from the tropical rainforest of northeastern Gabon. SIVcpzGAB2 clusters with the HIV-1 sequences and SIVcpzGAB.

BASE COUNT 117 a 47 c 60 g 56 t

ORIGIN approximately 2588 bp downstream from beginning of pol cds

1 ggtagtagaa tccatgaaca aagaattaaa aaagattata gggcaagtta gagaccaagc

pol cds ->

61 agaacactta aagacagcag tacaaatggc tgtattcatt cacaatttta aaagaaaagg

121 ggggattggg gggtagactg caggagaaaag aatagtagac atgctagcaa cagacttaca

181 aacacaagaa ctacaaaaac aaattctaaa aattcaaaaa ttctggggtt attacagagg

241 cagcagagac ccaatttggg aaggaccagc ctctctcctg

pol cds ->

LOCUS SIVHUB670 228 bp ss-RNA VRL 17-FEB-1994
 DEFINITION Simian immunodeficiency virus (SIVhu) envelope sequence, human isolate SIVhuB670, partial cds.
 ACCESSION S67406
 SOURCE Simian immunodeficiency virus from PCR-amplified DNA from cocultured PEMCs derived from an asymptomatic human laboratory worker.
 REFERENCE 1 (bases 1 to 228)
 AUTHORS Khabbaz,R.F., Heneine,W., George,J.R., Parekh,B., Rowe,T., Woods,T., Switzer,W.M., McClure,H.M., Murphey-Corb,M. and Folks,T.M.
 TITLE Brief report: infection of a laboratory worker with simian immunodeficiency virus
 JOURNAL N. Engl. J. Med. 330, 172-177 (1994)
 COMMENT This partial envelope sequence was isolated from a human laboratory worker who had been working with SIV, specifically with the strain SIVsmB670. This strain was originally isolated from a sooty mangabey and subsequently introduced experimentally into a macaque. The researcher was seroreactive to both HIV-2 and SIV antibodies. However, the infecting virus was determined to be SIV. The researcher is asymptomatic with normal CD4 counts. Khabbaz et al. report difficulty in obtaining virus by both isolation and PCR and conclude that it may be due to a low viral burden or its distribution in tissue. Khabbaz et al. also report a premature termination in the nef gene (preliminary data) not present in the presumably infecting SIVB670 strain.
 NCBI gi: 455953
 BASE COUNT 107 a 36 c 48 g 37 t
 ORIGIN 340 bp downstream from the start of envelope
 1 atgagatgca ataagacaga aacagacaaa tgggggttaa caggggaatgc aataacaaca
 env cds ->
 61 acagcaacaa cagcagcacc atcagcaaga acaccaaaag taacagcaaa cggtgtaaat
 121 gaaactagtt cttgtgtaaa aaacaataat tgtacaggct tagagcagga accaatggta
 181 agctgtaaat ttaacatgac aggggttaaaa agggacaaaa agaaagaa
 env cds ->

SIVSMCI2

LOCUS SIVSMCI2 453 bp ss-RNA VRL 06-OCT-1994
 DEFINITION Simian immunodeficiency virus (SIVsm) env sequence, sooty mangabey isolate SMC12, partial cds.
 ACCESSION X78508
 SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from cocultured PBMCs isolated from an asymptomatic sooty mangabey captured in its natural habitat in Cote d'Ivoire.
 REFERENCE 1 (bases 1 to 453)
 AUTHORS Peeters,M., Janssens,W., Fransen,K., Brandful,J., Heyndrickx,L., Koffi,K., Delaporte,E., Piot,P., Gershy-Damet,G.M., and van der Groen,G.
 TITLE Isolation of simian immunodeficiency viruses from two sooty mangabeys in Cote d'Ivoire: Virological and genetic characterization and relationship to other HIV type 2 and SIV sm/mac strains
 JOURNAL AIDS Res. Hum. Retroviruses 10, 1289-1294 (1994)
 COMMENT This sequence, SMC12, was obtained through a study designed to compare genetic isolates of HIV-2 and SIV from the same region of Africa, specifically Cote d'Ivoire. It was isolated from a seropositive sooty mangabey in good health, with no signs of AIDS-like disease. This animal was captured roughly ten years ago in its natural habitat in Cote d'Ivoire and subsequently housed in isolation at the University of Abidjan, Cote d'Ivoire. PBMCs from the animal were cocultured with those from a healthy human donor, and the viral sequences were PCR amplified, cloned, and sequenced. Through phylogenetic analysis, Peeters et al. determined that the SMC12 and SMC18 sooty mangabey sequences were equidistant from both the HIV-2 sequences and from other SIVsm and SIVmac sequences. Peeters et al. also conclude from serological data that these SIVsm sequences are closer to some HIV-2s than SIVcpz is to HIV-1. Based on several pieces of evidence, the evolutionary history of SIV and HIV-2 infection may be explained by cross-speciation from sooty mangabeys to humans [1]. Both <SMC12> and <SMC18>, accession number X78509, are shown in alignment below.

BASE COUNT 159 a 71 c 108 g 115 t
 ORIGIN 1861 bp downstream from the start of env
 1 gacgtggtc cagagacaaca agaactgttg cgactgaccg tctggggaac taaaaacctc
 env cds ->
 start
 61 cagactagag tcaactgccat cgagaaatac cttaaaggatc aggcacggct aaatgcttgg
 121 ggatgtgcat ttagacaagt ctgccacact acggtacatt gggaaaataa cactttgcaa
 181 cctgattgga ataatatgac ttggcaagaa tgggaaataa agatcagaga tctggaggca
 241 aatattctcag aatcattgga ggaggcaca atacaataag aagaagaat atgtatgagttg
 301 caaaaattaa atagctggga tgtgtttggc aaccgggttg attttacctc ttgggtcaaa
 361 tatattctata taggatttta ttagtagtagt ggaattatag ggtaagaat aatagtctac
 421 ttaatacaat tattaggtaa gcttaggaag ggc
 env cds end ->

CONSENSUS	GACGTGGTCAAGAGACAACAAGAAGTGTGCGACTGACCGTCTGGGGAACATAAAACCTC	60
SMC12	-----	60
SMC18	-----	60
CONSENSUS	CAGACTAGAGTCACCTGCCATCGAGAAATACCTAAAGGATCAGGCACGGCTAAATCTTGG	119
SMC12	-----G-----	120
SMC18	-----T-----	120
CONSENSUS	GGATGTGCATTTAGACAAGTCTGCCACACTACGGTACCTTGGGA?AAT?ACACTTTGCAA	177
SMC12	-----A--A-----	180
SMC18	-----G--G-----	180

CONSENSUS	CCTGATTGGAATAATATGACTTGGCAAGAATGGGAAATAAAGATCAGAGATCTGGAGGCA	237
SMCI2	-----	240
SMCI8	-----	240
CONSENSUS	AATATCTCAGAA?ATTGAG?A?GCACAAATACAACAAGAAAAGAA?ATGTATGAGTTG	292
SMCI2	-----TC-----G-G-----T-----	300
SMCI8	-----AA-----A-C-----C-----	300
CONSENSUS	CAAAA?TTAAATAGCTGGGATGTGT?TGGCAAC?GTTTGATTTTACCTCTTGGGTCAAA	350
SMCI2	-----A-----C-----	360
SMCI8	-----G-----T-----	360
CONSENSUS	TATATCTATATAGGATTTTATGTAGTAGTAGGAATTATAGG?TTAAGAATA?TAGTCTAC	408
SMCI2	-----G-----A-----	420
SMCI8	-----A-----G-----	420
CONSENSUS	TTAATACAATTATT?GGTAAGC?TAGGAAGGGC	439
SMCI2	-----A-----T-----	453
SMCI8	-----G-----A-----	453

CONSENSUS .VERVET	TGGATGGGATTTATTACTCCGATAGGAGGaataagATccT?AaccTgTatGCCcT?AATGAatGGGG	65
SIVAGM3	-----C-----A-----C-----C-----G-----	67
SIVAGMTYO	-----A-----G-----T-----T-----	67
SIVAGM155	-----A-----T-----T-----C-----	67
SIVAGM691	-----A--A--G-A-----TT-G-----AT-A-----	67
SIVAGM90	-----A-----GA-----C-----A-----T-----T-----	67
SIVVER2L	-----A-----T-----T-----C-----C-----G-----	42
SIVAGTYAB	-----T-----A--GC-----A--A--C--T-----T-A-----G-----	67
CONSENSUS .GRIVET	TGGA?GG??T?TATTACTC??AtAGaaGaGAgaaagATccTaaALcTgTATgccc?c?aatgagtgggg	59
SIVAGM677	---T--GA-T-----CG-A-----A-----G-----T-----T-G--C-----	67
SIVGRI3L	A-----G-----A--GT-----C--C--C--C-G--A-----	47
SIVGRI2L	A-----GCT-----TC-C-----A-----	47
SIVAGTYAA	-----CC-T-----AT-----A-A-----T-A-----A-----	43
SIVAGTYAC	---A--AG-A-----TG-----G-----A-----G-----	49
CONSENSUS .SABAEUS	TGGAAGGGATTTATTACAGTGAAAGgAGaaAGAAGATcTTgGatcTcTATGCTcTcAATGAatGGGG	67
SIVSAB1L15	-----G-----	42
SIVSAB1L14	-----	42
SIVSAB2L16	A-----T--A--CT-A-----CA-G-----	42
SIVSAB2L17	A-----A--CT-A-----CA-G-----	42
SIVSAB3L18	-----	42
SIVSAB3L19	-----	42
SIVSAB1C	-----G-----	67
SIVSAB4L10	A--G-----A-----T-G-----G-----	42
SIVSAB4L4	A--G-----A-----T-G-----G-----	42
CONSENSUS .SYKES	TGCAGGGCATGAATTATTGTGAAAAGAGAGATGAAATTCTTCATCTCTACCTACAGAAATGAGCATGG	67
SIVSYK	-----	67
CONSENSUS .VERVET	aATcattGATGATGGaALGcATggtc??aaGGaCCaGGaATcAGaTaccCaaag?TGcTTtGGcTTC	129
SIVAGM3	-----T-----AG-----A-----C-----	134
SIVAGMTYO	G--A-----C-----GA-G-----A-----T--T-AA-----G--	134
SIVAGM155	-----C-----AA-----T--G--A-----G--G-----	134
SIVAGM691	-----A-----C-A--T-ATA-TCC-----T-----C-----T-----A-----G--	134
SIVAGM90	-----AGC-----C--A-----T--CC-C-----T	134
SIVVER2L	-----TA-----AG-----	109
SIVAGTYAB	-----A-----C--C--G--ACA-TCC-----G-----G-----AG--T-----G--	134
CONSENSUS .GRIVET	aat?atagatgatcggcaag??taccacccc?ggcccaag?ataagatacccg??ac?lctggccttc	117
SIVAGM677	-----A-----CT--T-A--A-----G-----G-----AGAGTC-----	134
SIVGRI3L	-----A-----GA-----A-----T-----CTG--A-----	114
SIVGRI2L	C--T-----GA-----GC-----A-----A--G-----CTA--G-----	114
SIVAGTYAA	-----C-----C-----CT--T-----G--A-----A--C-----AAGG--T-----G--	110
SIVAGTYAC	-----	49
CONSENSUS .SABAEUS	cATAGTgGATGGaTGGCAGaAtTATACaGATGGgCCAGGTACcAGgTAcCCaAAgtgcTTTGGaTgq	134
SIVSAB1L15	-----	109
SIVSAB1L14	-----	109
SIVSAB2L16	A-----G-----C-----C-----T-----A-----A--T--T--ACAA-----T--	109
SIVSAB2L17	A-----G-----C-----C-----A-----A-----A--T--T--ACAG-----T--	109
SIVSAB3L18	-----	109
SIVSAB3L19	-----A-----	109
SIVSAB1C	-----	134
SIVSAB4L10	-----C-----C--C--T-----A-----G--CAG-----G--TC	109
SIVSAB4L4	-----C-----C--C--T-----A-----G--CAG-----G--TC	109
CONSENSUS .SYKES	AATCATAGA...CAGGATCAATTACACCTCAGGGCCAGGAACCCGGTACCCGTTAATCTTTGGATGG	131
SIVSYK	-----	131

AGMLTR

CONSENSUS. VERVET	TGcTTtAaGcTaGtPaCCaGTgGaccT?caTGAgGAaGCagAaacaTGTGAaAGaCAtTGcTtGgTcC	194
SIVAGM3	-----A--G-----B-----G-----G--T-----G-----C-----G-	201
SIVAGMTYO	-----C-----G-----T-A-----C-----C-A-----	201
SIVAGM155	-----C-----G-----T-C-----G-----A-----	201
SIVAGM691	--T--G-A-----G-----TAG-----GC--GG-----G--C--TC-----	201
SIVAGM90	-----C-----G--A--GA-G-----A--G-----C-----G-----G-	201
SIVVER2L	-----AC-----G-----A--T-A-----G--A--CAT-----G-----C-----	176
SIVAGTYAB	--T--C-----G-----T-T--C-----G-----G-----	201
CONSENSUS. GRIVET	tgccttAagctagt?cc?gt?gacct??tgaggaggcac?caactgtga?agacactg?ct?ctgc	174
SIVAGM677	-----C-----C--A--G-----GCA-----G-----G-----T--GA-----	201
SIVGR13L	-----C-----A--G--AAC-----CTC--A-----AG-----T--A-----	181
SIVGR12L	-----G--A--G-----CTCC-----GC-----T-T-----T--C--G-----	180
SIVAGTYAA	--T-----T-G--A--G--A-----GCA-----A--C--G--T--C--G--G-----CA--A--C-	177
SIVAGTYAC		49
CONSENSUS. SABAEUS	TGcTTtAaAcTGgTaCCtGTgGaccTgTCAGAgGagGCTaAgaAtagtGAgaacCacTGcCTgCTgC	201
SIVSAB1L15	-----	176
SIVSAB1L14	-----	176
SIVSAB2L16	--T--C--G-----G--A--A--A-----A--C--TG--C--A--T--GA--T--T--C--T-	176
SIVSAB2L17	--T--C--G-----G--A--A--A-----A--C--TG--C--A--T--GA--T--T--C--T-	176
SIVSAB3L18	-----A--G-----	176
SIVSAB3L19	-----T-----A--G-----	176
SIVSAB1C	-----	201
SIVSAB4L10	-----G-----CT--A-----A--CA--C-----GA--C--T-----C-----	176
SIVSAB4L4	-----G-----CT--A-----A--CA--C-----GA--C--T-----C-----	176
CONSENSUS. SYKES	TTATGGGAAC TAGTTCCAAATGAGATAGAGGGATGCTTGGAGTATGAGGAGAATACACTCCTCCTCC	198
SIVSYK	-----	198
CONSENSUS. VERVET	AtCCaGCaCaAa?Tg?g?GAaGAtCC?GATGgtatcaa?CATggaGagatcctGGtctGGAagTTtga	256
SIVAGM3	-----G--A--G-----C--T-----A-----AG-----	268
SIVAGMTYO	-----G--GA--G--A-----A-----GC-----G-----	268
SIVAGM155	-C-----C--CAT-----C--T-----A--T-----A--A--CA-----	268
SIVAGM691	-----T--T--T--ACAG--G-----A-----TGG-----A--CA-----GA--CA--	268
SIVAGM90	-----T-----G--AAA--G--C--C-----GT-----C-----	268
SIVVER2L	-----G--G--A-----A-----C-----G-----	243
SIVAGTYAB	-C-----T-----A--A--G-----T-----A--CCAG--A--AAT--G--GC--A-----	268
CONSENSUS. GRIVET	a?cc?gc?cag???gg?g?gat???gat?aa???g???catggaga?gtccttggtctggaagtktga	222
SIVAGM677	-T--A--A--ATG--G--AA--CCT--GG--TA--AT-----A-----	268
SIVGR13L	-T--T--T--GCG--A--GA--GGG--CC--GT--GA-----G--GAC-----A-----	248
SIVGR12L	-C--T--T--GCA--G--GGAG--GAA--CC--GC--GC-----G-----	247
SIVAGTYAA	-C--A--C--TTA--A--AG--CCCA--CGG--TA--AT-----A-----	244
SIVAGTYAC		49
CONSENSUS. SABAEUS	AtCCTGCTCAAGTgGCCf...atGAGGATGaccCCTGGA...gGAGaccctGGTgTGGAAGTTTGA	262
SIVSAB1L15	-----G-----A-----	237
SIVSAB1L14	-----G-----A-----	237
SIVSAB2L16	-----A--A--A--GG-----T--A-----GT--T-----	237
SIVSAB2L17	-----A--A--A--GG-----T--A-----GT--T-----	237
SIVSAB3L18	-C-----A-----	237
SIVSAB3L19	-C-----	237
SIVSAB1C	-----G-----A-----	262
SIVSAB4L10	-C-----A--A-----A-----T-----T-----	237
SIVSAB4L4	-C-----A--A-----A-----T-----T-----	237
CONSENSUS. SYKES	ATCCAGC...CTCTGGGACAGGGGAGTTCGAA...CATGGGAGAGACCTCATGTGGAAGTTCAA	256
SIVSYK	-----	256

CONSENSUS .VERVET	tCCC?ATGCTaGcag?Cagta?GACCC?accag?gagTa?tt?acaGACaTgcAT????ca?t?gg	310
SIVAGN3	-----C--A--T-----T-A--AA-A--TC-C--T-----...GATC-T--	331
SIVAGMTYO	---T---G---A-A---C---C-ATC-G---C-T-T-----...GGGC-G-T	331
SIVAGM155	---A---G---T---T---C---CT-A-G---C-T---T-AT---TCAA---G-T--	334
SIVAGM691	-----C-----G---A-G---ATG-----G---C-AT-----GCAA---G-G--	328
SIVAGM90	C---T-----G---T---A-A---AC---T-AGA-----...G---C-G-T	332
SIVVER2L	-----C---A---C---T---A---A---T---TGA-----TCAA---G-T--	309
SIVAGTYAB	-----C-----T---A-G---TCG---G.G---C-----C-----	317
CONSENSUS .GRIVET	ccc?.aagtt?gCaGTggagTaccgcCC??7atG?ata?ggatAtgcAcg?gcatgc??g??ta	273
SIVAGM677	---G---G---G-----GGAC---TT---A---C-----AA-----AAA-CGC--	334
SIVGRI3L	---T---A-----TTAAG---ATCG---C---AA-----G-ATG-AAAA-AAG--	314
SIVGRI2L	---T---AA-----A-TAAG---ATGG---C---GA---C---T-AA-----CTGA...	310
SIVAGTYAA	---GG---C-G---AC-A-----G-G-T-TT---CTG---AT---C---G-CATGG	300
SIVAGTYAC	T---C-----G-G-T-TT---C-G-TAT-----	90
CONSENSUS .SABAEUS	ccCa.?TaCTAGCAGTGGACTATGttGCTTGGAGaCTGCACCCAGAgCAGGTGCCGtCgtctG?	326
SIVSAB1L15	---C-----G-----G-----G-----CACAA-G	303
SIVSAB1L14	---C-----G-----G-----G-----CACAA-G	303
SIVSAB2L16	T---G-GT-----A-----C-----C	303
SIVSAB2L17	T---G-GT-----A-----C-----C	303
SIVSAB3L18	-T---A---T-----C-----A-----A-----A	303
SIVSAB3L19	-T.G---T-----A-----G-----A	303
SIVSAB1C	---C-----G-----G-----CACAA-G	328
SIVSAB4L10	T---A-----C-----C	303
SIVSAB4L4	T---A-----C-----C	303
CONSENSUS .SYKES	CCCC...CACCTGGCCTATACACCAGGCTGGGAGATGGCCCGCTT.GCAGTTAGAGAGACAAACAGG	319
SIVSYK	-----	319
<- nef		
CONSENSUS .VERVET	?Aa??ggaagtagc??7accgcagGctTGcGGTTa?gc?cgTc?cc?aggaGargACattaaaaacT	367
SIVAGM3	C---GA-----T.A-----T-----A---A---A---AT---T-----G-----	397
SIVAGMTYO	G---GA-----CAG-----A---C-----GAA---A---A---AT-----G-----	397
SIVAGM155	T-CAG-A---C---C.G---A-----TCCT.G---TG---T-----G-----	399
SIVAGM691	G-----GTA-----G.-G---CGGG-CCT-TGT---CAACC-GCA-	376
SIVAGM90	G---GA-----T.A-----T-----A---TG---GG-----TG-----	397
SIVVER2L	T-CAG-A---C---C.T-A-CGCA-----T-----A---TG---T-----G-----	374
SIVAGTYAB	---TGT---A---TA.GGAT-TG---T-----G---A---CGGG---CT-TGT---CAACC-TCA-	382
CONSENSUS .GRIVET	?a?agc?????????g?L?a?c?ctT?g??c????.....????????????????????	286
SIVAGM677	GTCTCAGCACTTTGCCG-T-GGGACT--CGGC-A-----	367
SIVGRI3L	A-C---AAGAGGAAGT-G-TTC-G-C-A-CT-GGTT.....ACCATGGAGATGACATTAAGAACT	374
SIVGRI2L	A-C---AGGAGGAAGT-G-T-C-G-C-A-CT-GGTT.....ACCATGGAGATGACATTAAGAACT	370
SIVAGTYAA	C-A-----GCTAACGC-G-A---T-TGGTT.....	328
SIVAGTYAC	GCTGACGC-A-A---TGGTT.....	111
CONSENSUS .SABAEUS	CTgACCACAAAACACATCCTATgAGTt.....GTCATGGTGTATGACATTAAGAACT	379
SIVSAB1L15	---A-----	356
SIVSAB1L14	---A-----	356
SIVSAB2L16	-----	356
SIVSAB2L17	-----	356
SIVSAB3L18	-----	356
SIVSAB3L19	-----	356
SIVSAB1C	---A-----	381
SIVSAB4L10	-----A-----	356
SIVSAB4L4	-----A---C-----	356
CONSENSUS .SYKES	AAAACACAAGAGC.....TGCACTCAGCACTTAGTAAGAACA	357
SIVSYK	-----	357

AGMLTR

	enhancer	enhancer	
CONSENSUS.VERVET	gctgac????????????gggacttccagc?aaGGGACTTTCCAgggcggg?cat		409
SIVAGM3	-----T-----		440
SIVAGMTYO	-----G-----A-----A-----		441
SIVAGM155	-----C-----A-----A-----		442
SIVAGM691	ACCA--CATGCTAATGAGCTA		409
SIVAGM90	-----AA-----C-----		439
SIVVER2L	-----P-GA-----		417
SIVAGTYAB	AGCA--GGGTC-----CT-----		411
CONSENSUS.GRIVET	??????GCGACTTTCCa?? ??GGGacttTCca?gGcGGt?cgg???????		317
SIVAGM677	-----CA.GT--TGGG--GGA--A-A-----		404
SIVGR13L	GCTGAC-----GG.CT--G-----G-----G-----		417
SIVGR12L	GCTGAC-----GC.AA--G-----C-----C-----		413
SIVAGTYAA	-----GCT..G---T-----TA-T--GCG-ACTATGGGA		372
SIVAGTYAC	-----GTT..G-----CA-T--GAG-ACCTGAGGT		155
CONSENSUS.SABAEUS	GCTGACTg.....?????????AGAGGGACTTTCCAGGGTGGAGACT		412
SIVSAB1L15	-----		389
SIVSAB1L14	-----		389
SIVSAB2L16	-----CTGCTGACTG-----		399
SIVSAB2L17	-----A-----CTGCTGACTG-----		399
SIVSAB3L18	-----		389
SIVSAB3L19	-----		389
SIVSAB1C	-----		414
SIVSAB4L10	-----		389
SIVSAB4L4	-----		389
CONSENSUS.SYKES	TCPCCTAGGAGACTCCATGGTGACAAGCTCGCCCCAGGGGAGGGACCTGGGCGGGGAA		416
SIVSYK	-----		416
	signal ->		
CONSENSUS.VERVET	gggcggtacGGGGAGTGGcTTtCCCTCAGA?gcTGcATAaAAGCAGATGCTCgctGGC		467
SIVAGM3	-----C-----		498
SIVAGMTYO	-----		499
SIVAGM155	-----		500
SIVAGM691	-----A-----T--A-----TAT--T--T-----TTG---		461
SIVAGM90	-----T-----		497
SIVVER2L	-----		475
SIVAGTYAB	-----A--TCC--T--G-----AT--T-----T-----		467
CONSENSUS.GRIVET	????????ggGGcGgtacGgGGAGTGGcTTtCCCTCAGA.GCTGCATAAAAGCAGATGCTcGCTGGC		376
SIVAGM677	-----T-----		462
SIVGR13L	-----A-----T--G-----		476
SIVGR12L	-----A--T-TGG-----A-----		472
SIVAGTYAA	GGGGCTG-----T-----		437
SIVAGTYAC	GGGGTTT-----A-----G-----		221
CONSENSUS.SABAEUS	GGCGGTACTGGGACTGGCTTtCCCTCAGTc.GCTCAATATAAAGCAGCTGCTTtGCGCC		470
SIVSAB1L15	-----G-----		447
SIVSAB1L14	-----G-----		447
SIVSAB2L16	-----		457
SIVSAB2L17	-----		457
SIVSAB3L18	-----		447
SIVSAB3L19	-----		447
SIVSAB1C	-----G-----		472
SIVSAB4L10	-----T-----		447
SIVSAB4L4	-----C-----C-T-----		447
CONSENSUS.SYKES	GGCTCACCCCTGCTGCATAAAAGCCGATGCTCTGTAGC		453
SIVSYK	-----		453

putative mRNA start ->		
CONSENSUS.VERVET	TTGTAAC?cAGc?CTCTTACT. .AGGAGaCCAGCTcGAGCCTGGGTGTTGGCTGGTcAgCCTAACc	530
SIVAGM3	-----	561
SIVAGMTYO	-----	562
SIVAGM155	-----	563
SIVAGM691	-----G-----A-----	523
SIVAGM90	-----	560
SIVVER2L	-----C-----	538
SIVAGTYAB	-----T-GT-----G-----G-A-----T	531
CONSENSUS.GRIVET	TTGTAAC?cAGT.CTCTTACT. .AGGAGACCAGCTcGAGCCTGGGTGTTGGCTGGT?TAGCCTAACc	439
SIVAGM677	-----	525
SIVGRI3L	-----	539
SIVGRI2L	-----	535
SIVAGTYAA	-----T-----	500
SIVAGTYAC	-----T-----	284
CONSENSUS.SABAEUS	.TGtAAAA.CAGT.CTCTGcPTGGAGGCTgCCAGATTGAGCCTGGGTGTTCTCTGGTAAGcCTGAAC	534
SIVSAB1L15	-----	511
SIVSAB1L14	-----G-----C-----	511
SIVSAB2L16	-----	521
SIVSAB2L17	-----	521
SIVSAB3L18	-----A-----C-----	511
SIVSAB3L19	-----A-----	511
SIVSAB1C	-----	536
SIVSAB4L10	-----T-----	511
SIVSAB4L4	-----	511
CONSENSUS.SYKES	TATCCGOGGAGTTGCAACTAGCTCTCTCCAGCGGCCCTGGGAGCCCTGAGAGACTGGTAGAGCGGCCc	520
SIVSYK	-----	520
CONSENSUS.VERVET	TGgt?TGCCC?.....acCagGGGTAAGga?CTCcttGGCTT?agaaagc?????.....	571
SIVAGM3	-----	602
SIVAGMTYO	-----	603
SIVAGM155	-----	604
SIVAGM691	---AC---C---T---AGC---AC---T---T---	567
SIVAGM90	-----G-----	601
SIVVER2L	-----	579
SIVAGTYAB	---ACT---A---GT---A---AGC---AC---G-ATGCTATGCTT.	581
CONSENSUS.GRIVET	cGGT.TGCCC.....ACCGGGGGTAAGGA.CTCTTGGCTtcATATAGCT.	482
SIVAGM677	-----	568
SIVGRI3L	-----	581
SIVGRI2L	-----	577
SIVAGTYAA	-----G-----	543
SIVAGTYAC	-----CGG-----	327
CONSENSUS.SABAEUS	CAGCTTGAGCCTGGGTGTTGGTGGGGAAGGCaAGCCTaAGAGCACcCTGGTGAGGTCTcGGTAGAt	601
SIVSAB1L15	-----C-----	578
SIVSAB1L14	-----C-----	578
SIVSAB2L16	-----G-----	588
SIVSAB2L17	-----C-----	588
SIVSAB3L18	-----C-----	578
SIVSAB3L19	-----C-----	578
SIVSAB1C	-----C-----	603
SIVSAB4L10	-----T---A---	578
SIVSAB4L4	-----T---A---	578
CONSENSUS.SYKES	GGGAGCCCTTGCCAGTAAAGGCTAGTGAGGTCCGGT?TTGGCCGCC	566
SIVSYK	-----	566

AGMLTR

	signal ->	
CONSENSUS.VERVET	CAATAAAA	578
SIVAGM3		609
SIVAGMTYO		610
SIVAGM155		611
SIVAGM691		574
SIVAGM90		608
SIVVER2L		586
SIVAGTYAB	C-----	588
CONSENSUS.GRIVET	CaATAAAA	489
SIVAGM677		575
SIVGRI3L		588
SIVGRI2L		584
SIVAGTYAA	-G-----	550
SIVAGTYAC		334
CONSENSUS.SABAEUS	CCTCacCTGGcgACTGGCCATTGGccAGTAGCAGaGACTCCGCTTGCTTGCTTGATTGCCCTCAATAAAA	668
SIVSAB1L15	---G-----	645
SIVSAB1L14	---G-----	645
SIVSAB2L16	-----A-----T-----G-----	655
SIVSAB2L17	-----A-----G-----	655
SIVSAB3L18	-----G-----	645
SIVSAB3L19	-----G-----	645
SIVSAB1C	---G-----	670
SIVSAB4L10	---A---T-----	645
SIVSAB4L4	---A---T-----	645
CONSENSUS.SYKES	CGGTCTPCGGCAGCTCTCCGTTGCTTTGCTTTAGCTGATCAGCCCAATAAAA	622
SIVSYK		622
CONSENSUS.VERVET	.?cTTgcCTGCattAGagctt..aTctGAGTCaaGTGccCtCATTTgacGcCTCact????gaaacggg	637
SIVAGM3	.T---CG-----C-----	672
SIVAGMTYO	CTCTT-----	674
SIVAGM155	A-T-----CTT-----	673
SIVAGM691	.C-----AG-A.C-TC-----GT--GT-C-----C-----CG-TCAGG-G-AT..	637
SIVAGM90	.CA-----C-----CAAGC-GG--A	672
SIVVER2L	.A-----CGAG-G-----	650
SIVAGTYAB	.TT-----T-----AC...TC---GT--GT-C---AC--T---C...GTGCA--G--A	650
CONSENSUS.GRIVET	.cCTGctc.GCTTAGTtGCTATaTTgGAGTCAAGTGC..TCATTGCTgCocCGaa?..cTC??GAAG	547
SIVAGM677	C-----GC-----TA-----	636
SIVGRI3L	CT-----G-----GT...T-----	648
SIVGRI2L	.T---C-----A--T--C-T...--AC-----	645
SIVAGTYAA	C-----C-----G...--AC-----	611
SIVAGTYAC	C-----G...T--CA-----	395
CONSENSUS.SABAEUS	GTCTTGCC...TTAGTTTACAAGTAAAGCAAGTGTGTGCCCCATTtATTCTCA.....GG	720
SIVSAB1L15		697
SIVSAB1L14		697
SIVSAB2L16		707
SIVSAB2L17		707
SIVSAB3L18		697
SIVSAB3L19		697
SIVSAB1C		722
SIVSAB4L10	C-----	697
SIVSAB4L4	C-----	697
CONSENSUS.SYKES	GTGTGTGTAACGCAAGTAAAGTGTGCTCTCATCTGTCTGTGCTGCGTCACTCCCAAGGTGCTCAGG	689
SIVSYK		689

CONSENSUS. VERVET	?????cTccTtACTGGGtTcTCTCp?gaCCcAGGcGAGAgAAA?CTCCAGCAGT	686
SIVAGM3	TAAAA-----A-----	724
SIVAGMTYO	..AAT-----T-----	725
SIVAGM155	.AGAAG-----AA-----	725
SIVAGM691	-C-AA-----T-----G-----	683
SIVAGM90	ACC..G-----T-----	723
SIVVER2L	TCCT..G-----A-----	701
SIVAGTYAB	AT...-C-T-A-----G-T-----A-----A-----C-----	701
CONSENSUS. GRIVET	TGAACCTCTCTTACTGgGTTCTCCTgt..?aCCcAGG?aGGAGAAActcCa???AGT	597
SIVAGM677	-----TG-----	682
SIVGRI3L	-----A-----T-----CG-----GCA	700
SIVGRI2L	-----T-----C-----GCA	697
SIVAGTYAA	-----T...G-----G-----CT-GAGC---	664
SIVAGTYAC	-----T...G-----G-----CT-CAGC---	448
CONSENSUS. SABAEUS	AaaACCCCTGGTTACT.....AAGGATCCCTGttAgaaaaaTCCTTAGCA	765
SIVSAB1L15	-----A-----	742
SIVSAB1L14	-C-----A-----	742
SIVSAB2L16	-----T-----	751
SIVSAB2L17	-----T-----	751
SIVSAB3L18	-----	741
SIVSAB3L19	-G-----A-----	742
SIVSAB1C	-C-----A-----	767
SIVSAB4L10	-----G-AG-----	742
SIVSAB4L4	-----G-AG-----	742
CONSENSUS. SYKES	GAGCECGCCTTCAGTGGCGCCGAACAGGGACTTGACTAGGCCAGGGACCCGAGAGGAGGTGACGCCA	756
SIVSYK	-----	756
CONSENSUS. SYKES	GTCAAAGGGAGTCCGGATCCAGAGGGATCGGCTGTCCGTCTAGGCCGACCCGGTGCTGGTGGCTCG	823
SIVSYK	-----	823
CONSENSUS. SYKES	TGGTGTACACACGAGACGAGGTAAGTGATCCAGCG	858
SIVSYK	-----	858

AGMENY

VERVET.COM	ATGA?g?agacatTA?taa??tTaggaaTA?PtagtaAFAGGgATAGGaaTAGtG?TaAgTacaAagca?CAaT	65
SIVAGM3	---A-CT---C-G...A---G---C-T---G---C-T-A---G---A---	70
SIVAGMTYO	---G-T-T---A---A---CC---A---G---A---G---T---T---G---A---	67
SIVAGM155	---CAA---TTC---GG---TT---TAT-G-T---G---G---A---A---G---G---	73
GRIVET.COM	ATGGGGAGATTGCTTATAAAAACTACTAATAATAGCAATAGGGATAAGTATAGGAATAGGT.....AACCTGT	67
SIVAGM677	-----	67
SABAEUS.COM	ATGAAGCTTCTTACAGTACTGCTTTGGCTAAAGTGGGTGTTGGAGCTTAGTGTGTTA.....GTCCAGT	64
SIVSAB1C	-----	64
SYKES.COM	ATGGCAGCTTTTGAACCTACATAGTTTGTCTTTTATAGCTTAATAAGTTTAGGATTATGGAAAAACAACAAT	73
SIVSYK	-----	73
VERVET.COM	GGaTaACAGT?TtTATGGAgTACCAGTATGGAAAAACAGCTCaGT?CAgGCTTTcTGCATGACaCCcACcAC	136
SIVAGM3	---G-C---A---A---G---G---T---T---G---A---	143
SIVAGMTYO	---C---A---G---G---T---T---G---A---	140
SIVAGM155	---G-C---A---G---G---T---T---G---A---	146
GRIVET.COM	ATGTGACAGTGTTTTATGGAATCCAGTATGGAAAAATTCACAGTTCAGOCATTTTGCATGACGCCCAATAC	140
SIVAGM677	-----	140
TANTALUS.COM	GCAGCATTTTGTATGACGCCCAATAC	27
SIVAGMB14	-----	27
SABAEUS.COM	ATCTTACAGTGTTTTATGGTATCCAGTATGGAAAAATAGTTTCAAGCCTTTTGTAGACCCCCAATAC	137
SIVSAB1C	-----	137
SYKES.COM	ATGTAACAGTATTTTATGGGATACCAGTATGGGAAGATGCATATGCACCTCTCTTTTGTACCACTCCCATAA	146
SIVSYK	-----	146
VERVET.COM	?AG?tTaTGGGCaACTACTAAeTgcATACCAGATGATCATGACTAtACA.....GAaGFACCacTa...	195
SIVAGM3	C---AC---C---CG---C---G---T---T---G---A---	204
SIVAGMTYO	A---C---T---G---G---T---T---G---A---	201
SIVAGM155	T---G---G---G---G---G---G---G---G---G---	207
GRIVET.COM	CAATATGTGGGCAACCACCAACTGCATACCAGATGATCATGATAATACA.....GAGGTGCTCTTA...	201
SIVAGM677	-----	201
TANTALUS.COM	CAATTTATGGGCCACACAAATTTGTATACCAGATGATCATATATACACA.....GAAGTCAATTA...	88
SIVAGMB14	-----	88
SABAEUS.COM	GAACCTATGGGCTCTACTAATTTGCATCCCGATGATGAACCAGAAGGAACAATAGCAGAAGTACCCATACCT	210
SIVSAB1C	-----	210
SYKES.COM	AGGAGGATGGGCTACTAAGAATTTGTGTGCGG...TCAGCAGATCAATA...GAAGTAAGGGTC...	204
SIVSYK	-----	204
VERVET.COM	AALATcAct...GAaCCATTTGAaGCATGGGctGAcAGAAaCCcTTAgTAGCACAAGCAGGaAGTAaCAT?C	264
SIVAGM3	---C---G---G---G---G---G---G---G---G---	274
SIVAGMTYO	---A---A---G---G---G---G---G---G---	271
SIVAGM155	---A---A---G---G---G---G---G---G---	277
GRIVET.COM	AACATTAACA...GAAGCTTTGGAGGCTTGG...GAT...AATCCGCTGGTAAACCAAGCAGAGATTAATATAC	265
SIVAGM677	-----	265
TANTALUS.COM	AATGTCTCA...GAGAAATTTGAAGCGTGGAAAGACAGGAACCCCTTTGGTGGCACAAGCAGAAAGCAATATAC	158
SIVAGMB14	-----	158
SABAEUS.COM	AATATCACA...GAAAAATTTGATGCTTTGGAAAAATCGGAACCCCTCTAGTAGGCCAAGCAGAGATTAACATAC	280
SIVSAB1C	-----	280
SYKES.COM	AACATTAACAGGAGAATACTTTCCAGCATGGAAATAGCAGTTCACATGATAAGGCAGCAATAATTAGAGGACATGT	277
SIVSYK	-----	277
VERVET.COM	AccTgCT?TTTGAaGcAGAct?TgAaGcC?TGTGTAaAgTATCACCTcTaTGCATCAaATGaatTGTGTAGA	334
SIVAGM3	---C---A---C---A---A---A---A---A---A---	347
SIVAGMTYO	---T---A---C---A---A---A---A---A---A---	344
SIVAGM155	---G---A---A---A---A---A---A---A---A---	350
GRIVET.COM	ATCTACTCTTTGAACAAACGATGAGGCTTGTGTTAAGCTCTCCCCCATATGTTATTAATAATGTCCTGTGTAGA	338
SIVAGM677	-----	338
TANTALUS.COM	ATCTGCTTTTGAAGTACCTTTGAAGCCTTGTGTCAAGTTAACACCAATGTGCATAAGAATGAATTGCATAG	231
SIVAGMB14	-----	231
SABAEUS.COM	ACCTTTGTTTGAAGTACTTTTAAAGCCATGTGTTAAGCTCTCACCAATGTGCATAAAAAATGAAGTGTCTATCG	353
SIVSAB1C	-----	353
SYKES.COM	CAGCTTTATTCTTGCAGGCCAATAGACCCCTGTGTGAAGCTAGCACCCATGTGTATTAGGATGTTGTGTACATT	350
SIVSYK	-----	350
VERVET.COM	gTTaAA?lcctC?g?a?c?ac?acCa?Caac?acEAC?ACGgC??AaC?Acc????????????????	376
SIVAGM3	A---G---C---T---AGC---T---C---TC---AA---G---C---CTC---A---AATATCACAGCCTCAACA	420
SIVAGMTYO	---T---A---AAG---GAA---GGG---G---A---C---T---A---G---C---GAA---T---T---GGC---	402
SIVAGM155	---AGG---C---C---A---CT---T---C---G---CT---T---AGG---C---AA---	405
GRIVET.COM	GCTGAATGGTACAGCCACGACAAAGGCCACCCTACTTGAACATACAACAATGACTACCCCTGT.....	402
SIVAGM677	-----	402
TANTALUS.COM	GCTTCCCTCCCTACTCCCTCCTCTAGTACAACAAGGACAAAGAAATCCCTGCCCAGGCCACAACAAAC	304
SIVAGMB14	-----	304
SABAEUS.COM	CTTAGAAGGAGGAGCAGCAACAACAACATCACCATCAACATCAACAGCACGCCCCGAGGTAGTGTGTGGGC	426
SIVSAB1C	-----	426
SYKES.COM	GGACAACAGCCCTGCTACAAGTACCCCTACTACAAGTCCCTTACTACACCTCCA.....	405
SIVSYK	-----	405

VERVET.COM	?????cTACCcTGTGT??g7ae?AagaCaggt?c?aa?cTAcAgTCaTGTAAaGaaAccATcATAGAAAaGG	436
SIVAGM3	ACCACCTT-G--G-----CCAG--C-----A--A-TGTGT--G-A-----A-----G--	493
SIVAGMTYO	-AG-COOG-C-T-----GAA--T-----C-----C--G--T-----G--	469
SIVAGM155	-TA-A--T--A-----ACT-C--C-----C--C--C-----	472
GRIVET.COM	CAGAATTGCAGTACAGACGAGATAGAAGGAG	433
SIVAGM677	GATGAGTGCATAGCCACTTGGTGACAAATA	335
TANTALUS.COM	AATGACTCAGTGATAGAACAAG	451
SIVAGMB14		335
SABAEUS.COM	TTT.....	451
SIVSAB1C	AATGAAACG...TGGTGGGGAG	424
SYKES.COM		424
SIVSYK		424
VERVET.COM	AgATGaAtGA?GAGcC?GC?TctAAtTGTACaTtTGCaATGGCTGGGTATgTAgTAgAGaGATCagAAGaAGaAt??	504
SIVAGM3	-AT-A-----A-----T--T-----A-----A-----A-----G--	564
SIVAGMTYO	---G--T-----C--C-----C-----C-----C-----A-----TA	542
SIVAGM155	---C--G--A--G--A--C--C--C-----T-----CA--T--G--C--A-----	543
GRIVET.COM	AAATGGCAGAGGAACCGCATCCAACTGCACCTTTTGCAATTGCAGGATATCAAGAGATGTAAAAAAGAAT..	504
SIVAGM677	GCATGGAAATTGAAAACAGTTCATATGTGCACCTTTTGCAATTGCAGGATACATGAGGGATCAGAAAAAGACA..	406
TANTALUS.COM	AGATGGAAAAAGAACAGGCCATGAACCTGTTCCTTTTGCTATGGCAGGATATAGAAGGGATGTAAAGAAAAAT..	522
SIVAGMB14	ACAACAGTACAGAGCCAGATTCaATTGCTCTTTTAACTTGACAGGGGGTTCAGAGGATAAAAAACAGCAA..	495
SABAEUS.COM		495
SIVSAB1C		495
SYKES.COM		495
SIVSYK		495
VERVET.COM	?TactCaGtgTgTGGAAATGATGCAGAAATCcttTGTAAg?????a?ta?????aAtAat?ctA?CaAGAG	562
SIVAGM3	A-G-----AGGGT--C-----T--A--G--A	627
SIVAGMTYO	T-----T-----A-----C--AAATAAG-C-.....GCA--G--	609
SIVAGM155	---C-----A--A-----CGT-G--CATCGC-----GGG-CA-----	612
GRIVET.COM	.TATAGCATGACCTGGTATGATCAGGAGTTFAGTCTGCAATAATAAAACAGGAAGTGAAAAGGGAGTAAGGAT	576
SIVAGM677		576
TANTALUS.COM	.TATAACAGCACATGGTATGATGCAGAACTCATGTGTGAACCGAGAGCAAAAAG...AACTCA...AGGGGG	472
SIVAGMB14		472
SABAEUS.COM	.TACTCAACAGTATGGGATGATCAAGAGGTGGTTGTGAAGAGGGAGGGAAAAATCAAAACGCAACACATACT	594
SIVSAB1C		594
SYKES.COM	.TATAGAGCTTTCTTCTATAAGGATGACTTGATG.....AAAGAGGAAGGCCAATAGTAGCTAT...	552
SIVSYK		552
VERVET.COM	TGTtALATGATfCALtGTATATGALTCAgTtATAaaGAAAGC?TGTGatAAaACATATTGGGATgAgC	628
SIVAGM3	C-----C-----A-----C-----C-----	694
SIVAGMTYO	C-----C-----A-----C-----C-----	676
SIVAGM155	C-----C-----G-----T-----G-----A-----	679
SIVVER1E	--A-----	4
SIVVER2E	---	4
SIVVER266E	---	4
SIVVER385E	--C-----	4
GRIVET.COM	TGTTACATGATACATTOTATATGATTCAGTGATAAAAGAGCTTGTGATAAAACATATTGGGATACtt	643
SIVAGM677		643
SIVGR13E	--A-----	4
SIVGR12E	--C-----	4
TANTALUS.COM	TGTTACATGATCCATGTGAACGATTCAgTCATAAAGGAAGCTTGCAGAAAAACTTACTGGGATACct	539
SIVAGMB14		539
SIVTAN1E	---	4
SIVTAN17E	---	4
SIVTAN40E	--A-----	4
SIVTAN49E	--AC-----	4
SABAEUS.COM	GTGGGATGTTACATGATCCATTGTAAATACCTCAGTAATCAAGAGAGGCTTGTGATAAAACTTTATTGGGATACct	667
SIVSAB2E	--A-----	4
SIVSAB3E	---	4
SIVSAB4E	--T-----	4
SIVSAB1C		667
SYKES.COM	TATTACCTGTTACATTGTAAATACAAGTGTATATCCGAGCATGTGAGAAAACAGACCTTTCAACCCCT	619
SIVSYK		619

AGMENV

VERVET.COM	TaAGatTaAGaTAcTGTGCTCCaGCAGGaTatGCTtTgcTaAAATGtAAAGat?a?gAtTATgatGGgTttaa	699
SIVAGM3	-----G-----C--G-----G-T-----AT-----C--C--T-T-----CA-----	767
SIVAGM155	-G--G-----G--T-----T-----G-A-----A-----A-----	749
SIVVER1E	-----C-----T-----C-----T-----TGG-----CA--A-----	752
SIVVER2E	-----GC-----A--T-----C-----AGAA-----ACA--C---GC	77
SIVVER266E	-----G-----C-----C--T-G-----CC-T-----A-----AC--	77
SIVVER385E	-----GC-----T-----C-----T--CC-----GCA-----T-----	77
GRIVET.COM	-----G-----G-TC--CC-AT-----CT-T--C-----	77
SIVAGM677	TaAGAgTaAGATAcTGTGCACCAGCAGG?TaTGCTT?gtTaAaATGcaatGATAAGgATTATAa?GGCTTTGC	714
SIVGRI3E	-----G-----T-----C-----T-----GA-----	716
SIVGRI2E	-----A-----A-----T-----T-----	77
TANTALUS.COM	-T--T-G-----T-----C-T-----G-G-----TTA-----A-----C-----	77
SIVAGMB14	TaAGgtTaaGATAcTGTGCACCAGCaGgctTTGCTaTcTTaAAATGtAagGATACcAATTataCAGGctTtGG	612
SIVTAN1E	-T-----T-----G-----A-----C-----	612
SIVTAN17E	-A-----C-----C-----A-----	77
SIVTAN40E	-G--C-T-----T-----T-----C-----A-----	77
SIVTAN49E	-A--C-----T-----C-T-----C-G-----G-----	77
SABAEUS.COM	-C-T-----A-----A-----	77
SIVSAB2E	TPAG??TtAG?TATTGTGCTCCaGCAGGtTATGCTtTgcTaAGATGtgcAGATaCAGAtTAIttCAGGcATAA	737
SIVSAB3E	-----AT-G--A-----T-----	77
SIVSAB4E	-----GC-----G-----T-----	77
SIVSAB1C	-----AT-A--A-----T-----C-----G--A-----AG---T---C---CAG-----	77
SYKES.COM	-----GC-----G-----	740
SIVSYK	TTCCGATTCaATACTGTGCTCCaCCGGGTACTCATGTGTTGAATGCAATGATACAAATTTTGAAGGAGACGA	692
	-----	692
VERVET.COM	accaaa?TGcttc?AATGT?tcaGTaGtgcATtGtACagg?tTaATgAATACACAGT?ActACaGG?tTgtT?	765
SIVAGM3	GA-----C-----T-----T-----G-----TAAC--G--A-----G-----T--AC---G	840
SIVAGM155	--A--T--C--A-----A-----C-----G--A-----G-----G	822
SIVVER1E	G--G-A-----T-----T-----C--ACT-----A--C--T--TC---A	825
SIVVER2E	T-----GA--AGA-----TAGT-----C-----G--G-----C-G-----C--C-A	150
SIVVER266E	--AG--C--C--T-----C-----G-----A-----T--C---G	150
SIV2010E	-A-----C--C--C-----A-----G--A-----CAACC-----G-----G-----A	150
SIVVER385E	--A--T-----C-----A-----GC-----A-----G---C-G	15
GRIVET.COM	TCCaAa?TGtAAGAATGTtCAGTagt?CAGTGTACTAGatTAATaAATACtACtaTAaccACAGG?ATAGGa	784
SIVAGM677	-----G--C-----G--T-----C-----T-----G-----	789
SIVGRI3E	-----A-----A-----G-----A-----G	150
SIVGRI2E	-----T-GT-----GA-----TACT-----C-----C-----A--AG--G-----T-----	150
TANTALUS.COM	?...attTGtAgaaATGTtTCAGTgGTaAGcTGCACaGG?tTaATGAATACtACAGTgAGcTcAGCTTTTGG?	679
SIVAGMB14	C...-CA-----G-----T-----C-----C-----T	682
SIVTAN1E	T...G-----T-----T-----T-----A--G-----G--C-----A-----C	147
SIVTAN17E	G...-G-----G-----G-----GC-----C-----C	147
SIVTAN40E	T...-----AC-----T-----T--T--AC-----A-----T-----G	147
SIVTAN49E	G...-A-----A-----T-----C-----C-----A	147
SABAEUS.COM	G...??eTGTaagAATGTaACAGTatTCAGCTTCAGCTAGgtTaAaAATACTACTgTAAGcACAGG?ATAGGt	804
SIVSAB2E	...GC-----A-----G-----	147
SIVSAB3E	...AA-----A-----G-----	147
SIVSAB4E	...AAT--C-A-----T-----T-----C-T--T-----A-----T-----A-----	147
SIVSAB1C	...GC-----G-----C-----A-----A-----	810
SYKES.COM	T...GTCTGTACAAATGTAACTGCAGTATCATGCACACAGGAATTCAATACATTGGCTAGCACATGGTTCCAA	762
SIVSYK	-----	762

VERVET.COM	LTGAATGGAGCtAttctGAGAAAtcGAACccAgAT?TGGCAaAAaCATagagtaAgcAATgActCaGtgcTaa	837
SIVAGM3	-----G-----C-----A-----A-----G-----G-----	913
SIVAGMTYO	C-----CA-----A-----G-----G-----AT---A---A---A---G---	895
SIVAGM155	-----A---A-----A-----G-----G-----G-----G-----	898
SIVVER1E	--A---T---T---A---A-----A---AG-A---T-----A---G---T---A---T---TA-T-	223
SIVVER2E	C-A---C-----CA-----C-----G---G---G-----A-----	223
SIVVER266E	C-----A-----A-----A-----G-----GACG-----C-CAC-T-	223
SIV2010E	-----G-----C-----CAA-----T-----A-----T---G---AC-G---	88
SIVVER385E	-----C-----C-----G-----T-----T-----T-----	223
GRIVET.COM	cTAAATGGLAGTAGAgcaGAAATAGAACaGAGATATGGCAagAAAagGGaAATagTAAATgATACAGTcATaA	857
SIVAGM677	T-----T-----T-----GA---CGA---GA-----	862
SIVGRI3E	-----G-----T-----G-----A-----A-----A-----	223
SIVGRI2E	-----T---G-----A-----G-----C---A-----T-----	223
TANTALUS.COM	?TAAATGGCAGtCAAdCAGAAaAtAGAACAGAAaT?TGGCAAAAGCATgGggttAg?AATgAttctGT?AT?A	747
SIVAGMB05	-----A---C-----G---A---T-----	31
SIVAGMB14	C-----T-----G-----G-----A-----A-----A---A---T---C---C---T---	755
SIVTAN1E	A-----G-----C-----A-----A-----A---G---C---A-----G---A---	220
SIVTAN17E	T-G---C---G---C-----G-----GG-T-----T-----A---A---	220
SIVTAN40E	T---C---G---C-----T-----A-----CAA---T-----A---C---T---	220
SIVTAN49E	C-----G-----G---C-----A---A---AC-----AG---G---A---	220
SABAEUS.COM	AT?AATGGtAGTTATGTAGCAAAATAGAACAGAAATcTGGCAGaAAAAATGGaAATccCAATgAT?CtGtTATcA	875
SIVSAB2E	--C-----G-----C---A-----	220
SIVSAB3E	--C-----G-----T-----	220
SIVSAB4E	--T-----A---AG-----AGT---A---A---A---C---A---	220
SIVSAB1C	--T-----C-----T-----	883
SYKES.COM	CTGAATGGGAOCtACAAAGCAAAAGATAAAAGTTAGATTATCAAAACAAAAGGATAAAATGAGTCAGTAATTA	835
SIVSYK		835
VERVET.COM	T7ttg?TtAAaAagcatTs?AatcTaaCagT?ActTG?aaaAGsCCAGGsaAatAA?ACAGTctTaCCaGTaAC	904
SIVAGM3	-G---AT-----A-----C-----T-----C-----C---A-----	986
SIVAGMTYO	-C---T---C---C---C---T-----T---C---C---T---GG-----C---G-----C---G-----	968
SIVAGM155	-C---C---C---T---C---G---T---A---C---G---G---G---G---G---G---G---G---	971
SIVVER1E	-TAG-A-----T---T---A---A---GG---TGT-----G---A-----A-----G---	296
SIVVER2E	-A---A-----GC---TT-----C---C---G---G-----A-----T-----	296
SIVVER266E	-T---AC-----C---T-----G---T-----G-----C---C-----	296
SIV2010E	-A---AC-----A---C---G---C---C---T---G-----G---A---G---T---C---	161
SIVVER385E	-ACA-A-----A---C---C---A---T---C---G-----C---A-----T---G---	296
GRIVET.COM	TaAAGTTaAATAAGTTTTacAACTTgACAGTgAGATGccGaAGACCAGG?AAtAAAACAGTgtTGCCAGTaaC	929
SIVAGM677	-----G-----T---T-----C-----C---C-----A---AC---T-----	935
SIVGRI3E	-T-----T---C---T---A-----TA-----A-----G-----	296
SIVGRI2E	-T-----C---A-----CC-A-----G---G---C---G-----G---A---G---C-----	296
TANTALUS.COM	T?AAGcTtAAaAA?catTA?aa?Ttaac?atAat?TGTAg?AGACCAGGaaAacAAGACaGTC?TACCAGTAAC	812
SIVAGMB05	-T-----C---A-----CC-A-----G---G---C---G-----G---A---G---C-----	104
SIVAGMB14	-A-----G---C---TC-G---A---T---G---A-----C---A---T-----	828
SIVTAN1E	-A---A---A-----A---T---G---G---AG---G---G---A-----T-----	293
SIVTAN17E	-T-----GTT---T---C---A---AA---G---T-----C-----	293
SIVTAN40E	-A---T---G---A---TC-G---AA---AA---TA---T---G---G---C-----	293
SIVTAN49E	-T-----GT-C---C---A---GGGG---CA-----G-----T-----	293
SABAEUS.COM	TcAGGCtAAATAGa?ATPtTAATtTaACcATcAGATGTAGaAgcCCAGGsaAatAAGACAGTatTgCCtGTaAC	947
SIVAGMD29	-----T-----C-----G---A---G---	38
SIVAGMD30	-----G-----C-----G---A---G---	38
SIVAGMD45	-----C-----C-----C-----A-----	32
SIVSAB2E	-----C-----	293
SIVSAB3E	-C-----C-----G-----G-----	293
SIVSAB4E	-A-----GT---A---C---T---C---A-----G---G-----G-----	293
SIVSAB1C	-----C---T-----	956
SYKES.COM	TTTTAGTTCCAGAAGCACTTAGATTGCAGATTATATGTGAAGAGACCAGGAAATGAGTCTATCAAAAATATTCA	908
SIVSYK		908

AGMENV

VERVET.CON	aATaATGGCaGGacTaGTgTT?CA?TCtCA?AagTatAAcacaAg?cTgAGa.....CAa	955
SIVAGM3	---C-----G-----T-T-----G-G---C-T-----G-----	1041
SIVAGMTYO	G-----G-----G-T-C-C-----A-A-C-----TG-AG-T-----G	1023
SIVAGM155	G-----T-----C-C-C-A-G-----T-----A-A-G-----G	1026
SIVVER1E	C-C-----T-G-----C-C-C-A-A-C-----G-AAT-A-----	351
SIVVER2E	-----T-----T-----T-----G-A-----G-AT-----	351
SIVVER266E	-----T-T-----T-T-T-----G-----TT-----GT-A-G-----	351
SIV2010E	T-C-----A-G-----C-T-C-A-----T-C-A-----G-----	216
SIVVER385E	-----T-T-C-----T-T-C-----A-----G-AT-----	351
GRIVET.CON	AATCATGGCaGGattaGTgTTTCAttcTCAGAAaTatAATAC?AgaTTAAAA.....CAA	983
SIVAGM677	-----G-----A-----C-----C-----C-G-----	990
SIVGRI3E	-----CG-----C-----G-A-----	351
SIVGRI2E	-----C-G-----AG-----G-----T-----	351
TANTALUS.CON	CATcATGGCaGGcTgGTgTTCCACTCtCAaaatAcAAtACaAg?TgAGa.....CA?	865
SIVAGMB05	---G-----G-----GC-----C-AT-----G-----	159
SIVAGMB30	A-----CA-----G-A-----C-A-----A-----	48
SIVAGMB53	-----T-C-----C-----C-----A-----	24
SIVAGMB14	-----AA-----A-----AT-----G-----	883
SIVTAN1E	-----C-----A-----C-----T-----AT-A-----A-----	348
SIVTAN17E	-----G-----T-----C-----G-C-C-----G-----	348
SIVTAN40E	-----T-CT-----A-----A-----G-----T-A-G-----A-----	348
SIVTAN49E	-----T-----T-----G-----C-T-----G-----	348
SABAEUS.CON	catcatGGctGGctTggTcttccCA?AGcCAGAAATATAACACCAGGtTgaga.....CAa	1001
SIVAGMD29	---TGC-----A-----T-C-----A-----T-----AGT-----G-----	93
SIVAGMD30	-----C-----C-----T-A-A-----	93
SIVAGMD37	AA-----T-----A-----T-----	36
SIVAGMD45	G-----A-----C-T-G-----T-C-----A-----C-A-G-----	87
SIVSAB2E	-----T-----A-----	348
SIVSAB3E	-----T-----A-----	348
SIVSAB4E	-----T-C-----A-----A-----	348
SIVSAB1C	-----T-----A-----	1013
SYKES.CON	ATTG...GCTGCAGGCTATTTC...CTCCCAGTAATACAAGGG...AAATTGAAAACCCGGAGGGAGCGCAAAACGT	975
SIVSYK		975
VERVET.CON	GC?TGGTGTCACTTccaagGcaAtTGGagaGGaGCCTGGaagGAAGTaaaacaagAaATagtaaAgtTaCCaa	1027
SIVAGM3	--T-----G-----C-----A-----A-T-----A-----	1114
SIVAGMTYO	--T-----G-----T-----CG-----G-----	1096
SIVAGM155	--C-----C-----G-A-----A-----T-----C-G-G-----A-----	1099
SIVVER1E	--A-----A-----TG-C-----GA-----A-----T-----G-----C-----	424
SIVVER2E	--T-----G-----T-----G-----A-CGG-----ACC-----TC-----	424
SIVVER266E	--A-----C-----C-A-----A-----T-----GG-----CC-G-----C-----	424
SIV2010E	--C-----T-----TA-T-----G-----G-----AG-----G-----	289
SIVVER385E	--C-----CAT-----T-----A-----T-----C-----	424
GRIVET.CON	GC?TGGTGTCACTTCCAAGGagATTGGcAAAGGcCTGGAAAGAAAGTa...AAGsAAG?AA	1052
SIVAGM677	-----C-----G-----A-----G-----A-----C-----TG-----	1060
SIVGRI3E	--A-----C-----G-----A-----T-----G-----A-----A-----AG-----A-----	421
SIVGRI2E	--A-----C-----T-----A-C-----T-----C-----	421
TANTALUS.CON	GC?TGGTGTCACTTccaagGGgAtTGGAAAGgaGCCTGGaagGaaGTgAg?gA?AC?AT?gt?aa?T?CCaa	930
SIVAGMB05	--T-----T-----T-----G-----G-----T-----TACA-----T-----A-----	232
SIVAGMB30	--C-----TC-----A-----G-----T-----C-A-----T-----GC-G-----A-CA-----C-C-----CC	121
SIVAGMB53	--C-----TC-----A-G-----G-A-----T-----G-----A-AA-----A-C-----T-----TGCCT-G-----	97
SIVAGMB14	--A-----C-----T-----G-----A-----A-----A-----T-----AT-G-----	956
SIVTAN1E	--A-----G-----T-----G-----AA-A-----A-----GG-----C-T-----	421
SIVTAN17E	--T-----C-----T-----TA-A-----AG-----A-----G-----A-----A-----C-----T-----G-----AT-A-----C-----	421
SIVTAN40E	--T-----T-----C-----T-----A-----G-----G-----A-----A-----T-----AC-T-----	421
SIVTAN49E	--C-----T-----C-----T-----C-----G-----G-----C-----T-----G-----C-----A-----	421
SABAEUS.CON	GCtTGGTGTCTGGTTTGGaOGtaAtTGGagAggaGCCTGGcAaGaaGrCAaAGaa?AaTaGtgaatTacCac	1073
SIVAGMD29	-----T-----A-----A-----CA-----GA-----C-----A-----	166
SIVAGMD30	--C-----A-----A-----A-T-----A-----G-C-----A-----GC-----A-----	166
SIVAGMD37	-----G-----AAT-----G-----GGA-----T-----T-----A-----	109
SIVAGMD45	--A-----T-----C-----C-----A-----A-G-----T-G-----C-----CAGCCA-G-----A-----	160
SIVSAB2E	-----C-----G-----G-----A-----GCT-----	421
SIVSAB3E	-----GCTC-----T-----	421
SIVSAB4E	--C-----T-----G-----G-----T-----T-----C-----GA-----T-----T-----	421
SIVSAB1C	-----A-G-----C-----G-C-----	1084
SYKES.CON	GCCTTCTGTGCACTAACAGGAAATTGGACAGAATTPTTAAAGCAGGTACATGAACAGGCCACGAAAACCTGGA	1048
SIVSYK	-----	1048

VERVET.COM	aAgA?ag?TAc?e?GG?ACCAAtgAtaCagAacaaATcT?t?TgcaaAGaCa?T??GGaGA?CCAGAagCagC	1089
SIVAGN3	---T-A---C-A-A---A---T---G-G-T-A-C---G---TA-TT---T---T---	1187
SIVAGNTYO	---C-G---TA-A-A---A---C---A-ACC---A-GG---C---T---	1169
SIVAGN155	---AC-G---C-A-C---A---CA---T-T---A-TT---C---T---	1172
SIVVER1E	---A-G---TC-G---G-T---C---A-G---TCC-A---G---GG---T---G---T---	497
SIVVER2E	C-C-T-AA---AGG---G-T---GA-G---A-T-C---A-TT---C---G---	497
SIVVER266E	---A-A---AGA-C---CC-A---G---TCT-C---G-GG---C---A---	497
SIV2010E	---C-A---TC-G---A---C---G---C-G-G---A-T-A---G-TT---T---T---	362
SIVVER385E	---C-G---AGG---G---C---C---T-A-T-AC---G-G-GG---T---G---	497
GRIVET.COM	Aa.....AATcT?ACagaagtaa7?AtAGAAAacATAcattTGAGAAAGCaATGGGGaGatCCAGAagCaGC	1116
SIVAGN677	---T---T---GC---T---C---AT---T---	1127
SIVGRI3E	-G.....C-T---AA---T---C---T---	488
SIVGRI2E	---T-G---C---TG---A-A---G---T---	479
TANTALUS.COM	aagaaheaTatcaaGGGACaAataAtaCaAagCaAAT?TGGC-TaAGGaGACaATGGGGaGatCCAGA?GCAGC	1001
SIVAGMB05	---TC---C---G---C---T---T---G---A---	305
SIVAGMB30	C---T---C---C---A---G---T---AC---C---G---	194
SIVAGMB53	---G---A---C---T---T---G---G---A---	170
SIVAGMB14	-C-T---T---CG---T---A---A---C---G---	1029
SIVTAN1E	---G---G---C---G---G---T---T---A---	494
SIVTAN17E	---G---G---A---G---A---A---T---G---C---T---	494
SIVTAN40E	-G---C---G---C---CAG-G---A---A---A---	494
SIVTAN49E	G---C-G---CTCC---T---A---G---	494
SABAEUS.COM	caga?AggTatagaGGCAC7Aatga7Acaaa?aAaATtTTcct?CAAAGaCaATGGGGaGatCCaGaagCaGA	1141
SIVAGMD29	---G-A---CA---A---T---GT---G---G---	239
SIVAGMD30	A---T---A---C---T---C---T---GP---A---	239
SIVAGMD37	A---A-AA---CA---C---C-T---G-GC---G---G---	182
SIVAGMD45	ACTCC-CT---C---A---GAA-T-TCTCCC-G---A---TT-G---G---C---GT---	233
SIVSAB2E	---T---C---A---C---C---G---C---T---G---G---T---	494
SIVSAB3E	---T---CA---A---A---C---T---C---A---A---T---	494
SIVSAB4E	---A---C---G---C---AA-C---TCAC---A---G---G---	494
SIVSAB1C	--A-A-A---T---A---C---C---G---T---A---T-T---	1157
SYKES.COM	AA.....AATGTAAACCAACACCATGGAGGTCCCAGCCAGGGGAGATCTAGAGGTACC	1103
SIVSYK	---T---T---T---T---T---T---T---T---	1103
VERVET.COM	aAattT?TGGTTtAatTgtCAaGGaGaaTtTcTtATtTgtAAaATGGatTGGtTtTtTaAatTatcTaAatAAt	1161
SIVAGN3	---A---A---G---G---T---C---G---	1260
SIVAGNTYO	C---C-G---T---T---G---CT---	1242
SIVAGN155	---C-A---C---C---G---C---C---G---	1245
SIVVER1E	T---G---T---G---T---C---G---A-GC---C---CA-T---C---C	570
SIVVER2E	---C-A---C---G---	570
SIVVER266E	G---G---C---G---T---C---G---	570
SIV2010E	---A---C---G---CC---A---	435
SIVVER385E	---G---C---G---T---CC---C---G---C---T---	570
GRIVET.COM	aAattTtTGGTTtAaCTGtCAaGGaGAATtTtCTATtGTAAgATGGatTGGTtTtAATtATcTaAatAAT	1189
SIVAGN677	G---T---T---T---T---C---C---C---	1200
SIVGRI3E	---C---C---C---C---C---T---	561
SIVGRI2E	---C---T-G---C---C---A---A---	552
TANTALUS.COM	?AacataTGGtT?AATtGtCAaGGGAGA?TtCTtCTA?TgtACaCCtGATTGGtTtT?AATt?gcTaAacAAc	1068
SIVAGMB05	C---T---T---T---T---C---A---AT---T---	379
SIVAGMB30	A---T---T---A---T---C---C---T---C---G---	267
SIVAGMB53	T---ACA---C---A---T---C---A---C---ATT-G---C	243
SIVAGMB14	G-G---T---TG---A---T---C---A---ATT-G---	1102
SIVTAN1E	T---T---C-C---G---T---T---T---T---G---G---T---	567
SIVTAN17E	T---T---T---C---G---T---T---A---ACT---T---	567
SIVTAN40E	C---T---T---C---T---A---C---C---A-G---G---T---	567
SIVTAN49E	C---T---C---T---C---G---C---A---C---G---	567
SABAEUS.COM	attctTcTtTtTtAatTgtCAAGG?GAaTT?Tt?AtTGTAAaATGGatTGGTtTcT?AA?TaTtTgAAtAAt	1210
SIVAGMD29	G---G---GG---C---G---T---C---T---A---T---C---	312
SIVAGMD30	---G---C---G---A---C---T---G---T---T---C---	312
SIVAGMD37	---TA---C---C---G---G---C---T---G---T---T---	255
SIVAGMD45	---AA---G---G---A---C---T---G---T---A---C---	306
SIVSAB2E	---T---T---T---C---C---C---C---	567
SIVSAB3E	---T---T---C---C---C---C---	567
SIVSAB4E	---C---C---A---T---T---A---T---A---C---	567
SIVSAB1C	---T---C---C---C---C---	1230
SYKES.COM	AACACATGGTTCCAATGTGGGGAGAGTCTTCTACTGT.....AAT	1146
SIVSYK	---T---T---T---T---T---T---T---	1146

AGMENY

VERVET.CON	?t?ACA.....gtagATgC?ga?CAT?????AAt?atTgc???aatagtac??g?a?aaaa	1200
SIVAGM3	CGT---.....C-G-C-.....CCG---...-G---G-.....-GG	1303
SIVAGMTYO	AAA---.....TGG---AT-C-.....TT---.....A-C-GC---	1282
SIVAGM155	T-A---.....-G---T-T-.....C---.....-AAAC-ACGC-G	1288
SIVVER1E	C-C-T-.....-GA-T-.....A---.....-C---ACAGGGA....	615
SIVVER2E	C-T-.....-T-C-.....AG--C-.....-G--T-AA-G-CG-GCC	619
SIVVER266E	T-A-G-.....-T-T-.....-G-A--AAG---CA-GTG-T-CA---	622
SIV2010E	T-G-.....-T---AA-T---AACACT--CAGA--C-.....AG--GG-G-AA-C--	490
SIVVER385E	AAG---.....TGGA---A-C-.....-C-A--GAT-----T-.....-G-	616
GRIVET.CON	casACA.....GtaGATGCAGatGG?????Aa?A?cTgc???.....?????Aa?aaaa	1221
SIVAGM677	-G---.....-A-----A---TACTAAT-GG-C---GAC.....-G	1243
SIVGRI3E	-G---.....-A-----A---A-G-----ACTTCGC-C---	604
SIVGRI2E	G-----G-----G-----G-----T-AA-C-.....AACA-ACCG-	592
TANTALUS.CON	Gagctcgtcga?gGaa?ata??tagatGTaGaaGga.....AA?aa?Tgt.....?c??a?A?ac?A???	1113
SIVAGMB05	-----AA-T---C-A-CC-T---G-----G-.....-C-G---.....C-AG-C-GG--GTCAG	439
SIVAGMB30	-C-----G---GC-C-GAA-----TG-C-.....C-CA-T-CG--A-TGG	328
SIVAGMB53	-----C-AG--T---C-TAC-----T-A-----G-CA-P-CG--T-CAC	304
SIVAGMB14	-----CAG--A---GC-TCACT--C-----C-.....T-GA-----T-AA-C-TA--A-GTG	1163
SIVTAN1E	-----AA-AGT--G-A-TG-----T-----T--T--C-----	612
SIVTAN17E	-----C--A--CCGGA-G-G-A-AGAG---G---G-.....CG-C-----	612
SIVTAN40E	-----AG--A-TCA---A--CAG-TA-C---G-----C-CT-----	612
SIVTAN49E	-----A--GA---AC--CAC-----G-----CCCT-----TCC-AT-OC-CTA	625
SABAEUS.CON	caqtca.....GTtagacCCagatCAT.....Aac??cTgt.....?	1238
SIVAGMD29	-TAA-----G--T--C---C-.....AAA-----	342
SIVAGMD30	-----T-----C-.....AA-----	342
SIVAGMD37	A---T-.....-G-GT---C---CCA-----	285
SIVAGMD45	-AAGT-.....TA-T--T-----TCAG--C-----	336
SIVSAB2E	-----C-----TCC-----G	598
SIVSAB3E	-----TCC-----	597
SIVSAB4E	-A--G-.....-T-----TA-----ACA-----	597
SIVSAB1C	A-----C-----AA-----	1260
SYKES.CON	GTATCTAAGCTCTTTGCTAACATTACGAATGGAAACGCAAGTAAGAATAATTATGCCAGTAACCTTCGT...	1215
SIVSYK		1215
VERVET.CON	gt??a?????agggcaCCaGG?CC?TGTGtACaAG?ACTTATGTTGctTgcCAtATcaGtCtGTcTaAa	1262
SIVAGM3	-AAA-OGT...-A-----A-C---C-----A-A-----AC-----T--	1373
SIVAGMTYO	AGAA-GGA...CAC-----A-A-----G-G-----A-----G-----A-----	1352
SIVAGM155	-GAA-OGT...C-AAGT---T-C-----G-A-----C-----C-----C-----	1358
SIVVER1E	AAA--A-.....G-G---C---G-A---A-GG--T-----A---C--	677
SIVVER2E	C-----G--T--C---CC---G---G-----A---A-----	683
SIVVER266E	-GGTAACAAG-.....G-T-----G-A-----G-C-----TC-----T	695
SIV2010E	-GG-.....-G---T-T---C---A-----C-----G-T--	554
SIVVER385E	C-GG-.....A-T-----G-----C-----	683
GRIVET.CON	gggg?Cca.....OGACCAGGcCcaTGTGt?cAgagaACaTATGTTGCCTGcCAtATAaGACAAgTAGTaaa	1286
SIVAGM677	G-AAG-----T-----T-----C-----	1310
SIVGRI3E	-----T-A-----CGA--A-----	671
SIVGRI2E	-A-A-----T-T-----A-AC-----T-C-----G---C--	659
TANTALUS.CON	7agga?t?ac?gggtc????agaaaaTGtlt?AaaAGaC?TAtGTAgccTgcCAtATcaGGTCagTtGtCAA	1176
SIVAGMB05	G---C-C--A--T--T...C-C---C-G--G--A--C-----T--T-----C-----	509
SIVAGMB30	G---C-C--A--A...C-A-----C-A-----G-----T-----A-----	398
SIVAGMB53	C-CC-T-A--T-----TACC-----G---C-A-----T-C-----T--T-----A-----A-----	377
SIVAGMB14	G---C-G--A---CACC-----G---C-T-----C-C--T-G--TG--T--AC-----	1236
SIVTAN1E	ACC--T--AAAGGAT-A-COC--C-AC-----A---CC-----T--T-----A-----	680
SIVTAN17E	AAAGATCCA--G-----P-----AT-G--C--TC-----A-----G-----	674
SIVTAN40E	ACTC-T-A-AAAGAT--GCT---AAG--G---A-----C-----A--T--	680
SIVTAN49E	A--GGGA-TGACAGGAAGC-----G---A---A---T-G-----T-----G-----	698
SABAEUS.CON	????a?taaaaa?aatgatAaaaaATGtTGGcagAGaActTAtGtCtCCt?GcCacAT?Ag?CAagTaGTAaA	1302
SIVAGMD29	-T--GGGG-----C-----G-----T--A--G-----G--	407
SIVAGMD30	CA--C-A-----C-----A--G--A-----T--T--TA--O-----G--	410
SIVAGMD37	G--G--C-----G-----A---C---C-----G-----G-----	353
SIVAGMD45	AA-C--C-----G-----GCA--A---G---T--T--C---A---G--C--	404
SIVSAB2E	GAGGA-A-G-C--G---A-C-GG-----C-----C-----A--G--A-----	671
SIVSAB3E	A--C-CT-CC--C-----AC-----C-----A--G-----	665
SIVSAB4E	C-CC--A-GC--C---G-----G-----T--A-----	665
SIVSAB1C	GCA-----T---ACA--GCC-----C-----C-----A--G-----	1328
SYKES.CON	CTGTCA?GTGCCATTCGCCAAATAATTAA	1244
SIVSYK		1244

VERVET.COM	tgatTggtAtAcaaTatcaAagaaaAcATatGctCCaCCaAGAGAaGG?CattTgcAaTGcac?TCcACgGT?	1312
SIVAGN3	-----C-----C-----G-----C-----A-----G-----G-----C-----A-----A	1446
SIVAGMTYO	-----CC-----CC-----T-----G-----A-----G-----A-----G-----A-----C	1425
SIVAGM155	-----T-----A-----G-----G-----A-----T-----A-----T-----C-----A-----T	1431
SIVVER1E	-----C-A-T-GG--GA-G-T-----A-T-C-----T-----G-G-T-C-T-A-C	750
SIVVER2E	-----C-----C-----A-----G-----A-----G-----A-----G-----A-----C	756
SIVVER266E	ATGG-T-G-G-----T-----G-T-C-AQ-----A-----T-----G	768
SIV2010E	-----G-TAGT-ATTG-----C-----G-T-----G-----CGC-----G	627
SIVVER385E	-----G-----C-----C-----C-----A-----A-----A-----A	756
GRIVET.COM	tGAtTGGTAcAcTcTcaCtAagAAaAcATatGctCCgCCAAGaGAAGG?CAtTTGGAGTGtAAcTC?cCAQTC	1357
SIVAGM677	-----G-T-----A-GGT-----A-----G-----T-----C-----A	1383
SIVGRI3E	C-C-----A-C-----C-----C-----G-----GA-----	744
SIVGRI2E	-----T-C-----CCG-----C-G-----C-----T-----	732
TANTALUS.COM	tGAtTGGTATAcCctTgTgtaAgAA?AccTATGCaCCgCCaAGaGAaGGaCActTgGA?TGcAcaTCcACaGTG	1246
SIVAGMB05	-----T-----GT-TCA-----G-----A-----C-----A-----G-----A-----	582
SIVAGMB30	-----ATCAC-----A-----C-----A-----A-----T-----	471
SIVAGMB53	-----C-T-CTT-C-T-A-----A-----G-----A-----G-----	450
SIVAGMB14	-----A-ATTGC-A-A-GA-----G-----G-----G-----G-----	1309
SIVTAN1E	-----C-----CA-A-----A-----A-----G-----A-----A-----A-----	753
SIVTAN17E	C-----T-A-----G-----A-G-G-----A-----A-----	747
SIVTAN40E	-----T-----C-A-A-A-G-----A-G-----A-G-T-----	753
SIVTAN49E	-----G-T-A-A-G-----G-----G-----G-----G-----	771
SABAEUS.COM	TGAtTGGTATAcCctTgtCcAaaAAGACaTatGc?CCtCCcAGaGAAGGcCActTGGAgTgtAacTC?ACAGCc	1373
SIVAGMD29	-----C-----T-----T-----A-T-----A-----CA-T-----T	480
SIVAGMD30	-----T-----T-----T-----T-----A-----C-----T-----A	483
SIVAGMD37	-----T-----C-----T-----A-----GC-----C-T-A-----T	426
SIVAGMD45	-----C-----TG-A-G-----C-A-----T-GC-----CT-T-----	477
SIVSAB2E	-----C-T-----T-----C-C-----G-----G-----	744
SIVSAB3E	-----C-T-----T-----C-C-----C-----A-----	738
SIVSAB4E	-----C-----T-----A-----G-----A-----A-----A	738
SIVSAB1C	-----C-T-----T-----C-C-----G-----G-----	1401
SYKES.COM	TGAtTGGAGATATGTAAGAAATGATATACCTGCCCCCTACGGCAGGGCATATTAAATGTACTTCCAATGTC	1317
SIVSYK	-----	1317
VERVET.COM	ActGGgATGcAGTgGAgcTaAATATaAa?aaTaagAacAGaAC?...AATGTgACatTgagtCCcCagATaG	1400
SIVAGM3	-----G-T-----T-----G-----C-T-G-----G-T-----A-----A-----	1516
SIVAGMTYO	-----T-----T-----T-----T-G-A-----C-----C-----A-----	1495
SIVAGM155	-----C-----C-----T-CC-----G-A-----A-----A-----	1501
SIVVER1E	-----G-T-----G-----TCC-----G-A-----T-----TCA-----	820
SIVVER2E	-----C-----AT-G-----T-G-CGC-----A-----A-G-----TC-----	826
SIVVER266E	-----A-----G-A-----TCCCC-A-T-----C-----A-----	838
SIV2010E	-----A-----C-----CC-----G-----T-----T-----C-A-A-T-----	697
SIVVER385E	-----G-----TC-----G-C-T-----C-----C-----	826
GRIVET.COM	ACgGCACtATatGTGGaaATaaATTatAaCaAaAgctTGGcCCaATaAATGTgACccTatctcCTCaggTat	1430
SIVAGM677	-----C-----C-----G-----C-----A-----T-----A-----A-----GA	1456
SIVGRI3E	-----A-----G-----G-----C-T-----G-AG-----T-----T-----AT-G-A-----A	817
SIVGRI2E	-----A-----G-----G-----C-T-----G-AG-----T-----T-----AT-G-A-----A	805
TANTALUS.COM	ACaTC?ATgaTgGtGTcActTgGatTAcacac?cAAaaAcAtGAct...AATGTgACatTgACAGCaAAAtTgG	1314
SIVAGMB05	-----C-T-----A-----T-A-----TTCA-----T-----A-----	652
SIVAGMB30	-----C-----G-A-----T-----G-----A-----C-----T-----A-----	541
SIVAGMB53	-----C-C-----C-----C-----C-----A-----C-----T-----G-C-----	520
SIVAGMB14	-----A-C-----G-----T-----GA-----G-----A-----C-----A-----	1379
SIVTAN1E	-----T-----A-----T-----G-----G-A-G-----T-----T-----	823
SIVTAN17E	-----C-----A-----A-C-----T-----AT-----T-----CC-A-----A-----	817
SIVTAN40E	-----C-A-A-A-C-----T-----C-----AT-----T-----C-T-----CA-T-----	823
SIVTAN49E	-----C-A-----A-----T-----TTCG-----T-----C-----CC-A-----	841
SABAEUS.COM	ACaGC7cTgTATGTgGAacTgaa?TA?AaCaqtaAaaAaCaGgACc...AATGTgACatT?tcCCTCagATta	1439
SIVAGMD29	-----T-----T-----T-----C-----T-----A-----A-----C-----C-----A-----T-----C	550
SIVAGMD30	-----TA-----T-----G-----T-----TTCG-----G-----T-----C-C-TG-----A-----	553
SIVAGMD37	-----A-C-----T-----T-----G-----G-A-----C-----G-----C-----	486
SIVAGMD45	-----G-A-----G-C-----T-----TTCAC-T-----T-----C-----GAGT-----C-----	547
SIVSAB2E	-----G-----A-----CTC-----T-----A-----A-----C-----C-----G-----	814
SIVSAB3E	-----G-----A-----C-----T-----A-----A-----A-----G-----	808
SIVSAB4E	-----T-----T-----T-----C-----C-----T-----G-----C-----C-----G-----	808
SIVSAB1C	-----G-----A-----T-----C-----T-----G-----A-----G-----G-----	1471
SYKES.COM	ACTGCAGTATGACGGACATAGAATACTACCCAGGATCTACGTTG...AATTTACTCCAACAGCAAAATGTGG	1387
SIVSYK	-----	1387

AGMENV

VERVET.CON	AAac?ATcTGGGcagCaGAATTGGGcaGaTACAAATTGtAGa?AT?ACaCCAATTGGcTTtGCACCcACaga	1470
SIVAGM3	---C-----G-----A---A---T-----C-----	1589
SIVAGMTYO	--T-T-----G-T-----A-G-A-C-----	1568
SIVAGM155	---C-----G-T-----A-----	1574
SIVVER1E	--G-C-A---T-----T-G-----A-T-G-----T-----A-CCC	893
SIVVER2E	---GT-----A---A-T-----	899
SIVVER266E	--T-T-----A-C---T-----G-C-----	911
SIV2010E	---GT-A---T-T-----T-----G-C-----A-----	770
SIVVER385E	---C-----G-----T-----G-C-T-----	899
GRIVET.CON	G?AGtATaTGGGc?TAtGAaLTGGGgATAcAAATTAGTgAGaATaACACCAATTGG?TTTGCCTCCtACaGA	1500
SIVAGM677	---C-----G-C---C-----C-T-----G-----C-----	1529
SIVCRI3E	---G-----A---G-----T-----G-----	890
SIVGRI2E	---A---T---C-----C-----G-----T-----C-G-	878
TANTALUS.CON	AaAA?ATaTggGcTtATtGAaTTGGGgAGaTAtAagcTcAttgAaATaAaCCaATtGG?TT?GCgCCcACAGA	1384
SIVAGMB05	---G-C---T-----A---A-----G-----	702
SIVAGMB30	---T-----C-C-----G-T---G-T-A-T---	614
SIVAGMB53	---C-----AC---C-----	555
SIVAGMB14	---C-----C-G-----G-C-----A-G-T-T---	1452
SIVTAN1E	---G-C-----A---G-----C-T-C-----	896
SIVTAN17E	---T-----G-----AA---A---C-----A-T---	890
SIVTAN40E	---TC-T-AT-A---G-----G-A---T---A-C-G---	896
SIVTAN49E	---T-----C---G---G---C-A---A---C-G---A-C---P---	914
SABAEUS.CON	gaagcATtTGGGctAaTGAaTTgGGgaga?TAcAAAT?GTgGA?ATaAAaCCaATTGGcTTTGCtCCcAACaga	1509
SIVAGMD29	---G-T-----A-----	576
SIVAGMD30	---C-----T-T---G---A-----C---A-----	626
SIVAGMD45	---GTC---C---AG-T-----A-TAGC---T-----A---G-T---	594
SIVSAB2E	A-----C-----G---G-----	887
SIVSAB3E	A-----C-----A-A-G-----	881
SIVSAB4E	---G-T-----A-T-T---G-A-A-C---C---G-----A-	881
SIVSAB1C	A-----T-----A---A-C-G-----CA-	1544
SYRES.CON	AGGATGTTTGGAGAGCAGATCTGTTCAATTATAAAATTGATTCAAAATTAACCAATTGGCTTTGGCCCCACTGA	1460
SIVSYK	-----	1460
VERVET.CON	aGTaAGGaGaTACACgCG3agg?catgaaAGaCaAAA?GagPc...CCgTTcGTGCTAGGGTTC?T?GgcTTc	1535
SIVAGM3	---A-G-----T---C---AC---GC-----C-A---	1659
SIVAGMTYO	---C-T-----A---G-G---GA-----T-----T-A---	1638
SIVAGM155	---C-C---C-----C-A-G-G---AC-----C-A---	1644
SIVVER1E	---GCT...TC---G---GA-G-G---C-----T-G---	960
SIVVER2E	T---A-G-----C-----AC-----A---C-A---	969
SIVVER266E	---C-----C-----T-----GA-----T-G---	981
SIV2010E	---C-C-G---A-G-A---AA-G-T---C-----C-G-CT---	840
SIVVER385E	---G-A---C-----C-C---AC-A---T-G---	989
GRIVET.CON	tGTgAGAAGATACACTGGcCCC...aCaAgAgAaAAaAGGGTg...CCaTTcGTGCTaGGGTTccTAGGcTTc	1567
SIVAGM677	---A-----T-----	1596
SIVCRI3E	C-----C-----T-G---G-G-----G-----	957
SIVGRI2E	---A-A---A-C---T---T-T---A-T---	945
TANTALUS.CON	ggT?AGaaGATATgttCGGCCA...ACTCGAGaaAAGAGGGTG...CCCTTCGTGTtAGGGTtCCTAGGcTTc	1450
SIVAGMB30	A-A---AG-----	657
SIVAGMB14	---G-G---	1461
SIVTAN1E	---A-A-----G-----C-T---	963
SIVTAN17E	T-T---C-----C-----	957
SIVTAN40E	---C-----G-----	963
SIVTAN49E	A-G-A-G-----C---A---	981
SABAEUS.CON	?GTaAGAGATACACTGGcCCC...GAaAGgCAaAAaAGGGTG...CCATTGCTGCTaGGGTTcCTAGGcTTc	1575
SIVAGMD30	G-----A---A---G-----	669
SIVSAB2E	G-----	954
SIVSAB3E	A-----	948
SIVSAB4E	T-G---T---G-A-G-----G-----	948
SIVSAB1C	A-----A-----T---	1611
SYRES.CON	TCAACGTCGGTATGAGTTCCTAATACCAGGGAAAAAAGGGCAGCACCATTTGGCCTTTGGGTTCTGGGATTA	1533
SIVSYK	-----	1533

VERVET.COM	TTyGGaGcTcGctGGaACTGCAATGGGAGCagCGGCGAcagccCTGACGGTCCAGTctCaGcATT?CTTGCTG	1607
SIVAGM3	--A-----G-----G-----C-----G-----A-----	1732
SIVAGMTYO	-----G-----G-----G-----C-----G-----G-----	1711
SIVAGM155	-----G-----C-----G-----G-----C-----G-----	1717
SIVVER1E	-----A-----C-----T-----G-----G-----G-----	1033
SIVVER2E	-----G-----A-----A-----A-----A-----A-----	1042
SIVVER266E	-----G-----A-----A-----C-----G-----G-----	1054
SIV2010E	-----G-----A-----A-----C-----G-----G-----	906
SIVVER385E	-----G-----A-----A-----C-----G-----G-----	1042
GRIVET.COM	TTyGGaGcTcGCTGGaACTGCAATGGGCGCAGCGGACagCTGACGCTCCAGTCTCGGCATT?CTTGCTG	1640
SIVAGM677	-----G-----A-----A-----A-----A-----A-----	1669
SIVGRI3E	-----G-----C-----C-----C-----C-----A-----	1030
SIVGRI2E	--A--G--A-----T-----	1018
TANTALUS.COM	TTGGGagAGCTGGAGCTGCAATGGGTGCaACAGGgACaG?GPTGACGGTCCAGTctCAGCAATTACTTGctG	1522
SIVTAN1E	-----GGC-----C-----C-----A-----	1036
SIVTAN17E	-----G-----A-----A-----A-----	1030
SIVTAN40E	-----G-----C-----C-----C-----	1036
SIVTAN49E	-----AA--G--A-----	1054
SABAEUS.COM	TTGGGAGCAGCTGGAGCTGCAATGGGTGCaGCGGACAGCCCTGACGGTCCAGTCTCAGCagTTACT?GCTG	1647
SIVSAB2E	-----C-----	1027
SIVSAB3E	-----T-----	1021
SIVSAB4E	-----A-----T-----	1021
SIVSAB1C	-----C-----	1684
SYKES.COM	CTCTGGGCTGCGGGCACTGCAATGGGGGGCGCTGGGACAGCGCTAACACTCCAGTCTCAGACTTTGTGGGCTG	1606
SIVSYK	-----	1606
VERVET.COM	GGATAcTGCAGCAGCagAAGAAtcTGCTGGCGGCTGTGGagGCTCAACAGCagATGTTGAAGCTGACCATTtTG	1680
SIVAGM3	-----	1805
SIVAGMTYO	-----	1784
SIVAGM155	-----T-----GA-----	1790
SIVVER1E	-----T-----CT-----A-----A-----A-----A-----	1106
SIVVER2E	-----A-----T-----A-----A-----	1115
SIVVER266E	-----A-----	1127
SIVVER385E	-----	1115
GRIVET.COM	GGATATTCAGCAGCAGAGaGAActTGCTGGCGGCTGTGGagCAGCAaCAACAGcTGTGAAGCTGACCATTtTG	1713
SIVAGM677	-----A-----A-----T-----	1742
SIVGRI3E	-----T-----	1103
SIVGRI2E	-----G--TC-----G--A-----	1091
TANTALUS.COM	GGATaTTGCAGCAGCAGAGaGAATCTGCTGGCGGCTGTGGagCAGCAaCAACAGcTGTGAAGCTGACCATTtTG	1593
SIVTAN1E	-----G-----G-----	1109
SIVTAN17E	-----G-----A-----A-----	1103
SIVTAN40E	-----T-----A-----A-----	1109
SIVTAN49E	-----T-----A-----G-----	1127
SABAEUS.COM	GGATA7TGCAGCAGCAGAGAAAAATCTGCTTGGGCTGTAGAGCAGCAGcagCAAA7GCTAAAGCTCAC?ATtTG	1718
SIVSAB2E	-----C-----C-----	1097
SIVSAB3E	-----T-----C-----	1094
SIVSAB4E	-----T-----T-----C-----	1094
SIVSAB1C	-----C-----T-----	1757
SYKES.COM	GGATAGTGCAGCAGCAGCAAAAGTTGCTGGAGGCGAGTTGAGGCACAGCAACACCTCTTAGGCTTAACCTGCTG	1679
SIVSYK	-----	1679
VERVET.COM	GGGTGTgAAAAACCTCAATGCCCGCGCTCACAGctCTTGAGAAGTACCTaGaGGATcAGGCaCGgtTaAAcgCt	1753
SIVAGM3	-----C-----G-----G-----T-----	1878
SIVAGMTYO	-----T-----C-----AC-----T--C-----	1857
SIVAGM155	-----G--C-----	1863
SIVVER1E	-----	1116
SIVVER2E	-----	1125
SIVVER266E	-----	1137
SIVVER385E	-----	1125
GRIVET.COM	GGGTGTgAAAAACCTCAATGCCCGCGCTCACAGctCTTGAGAAGTACCTAGAGGATCAGGCACGGCTAAATTCA	1786
SIVAGM677	-----	1815
SIVGRI3E	-----	1113
SIVGRI2E	-----A-----	1101
TANTALUS.COM	GGGTGTGAAA	1603
SIVTAN1E	-----	1119
SIVTAN17E	-----	1113
SIVTAN40E	-----	1119
SIVTAN49E	-----	1137
SABAEUS.COM	GGGTGT?AaaAACCTCAATGCCCGCGCTCACTGCTCTTGAGAAGTACCTAGAGGATCAGGCACGGTTGAACATT	1790
SIVSAB2E	-----G-----	1107
SIVSAB3E	-----A-----	1104
SIVSAB4E	-----G--G-----	1104
SIVSAB1C	-----A-----	1830
SYKES.COM	GGGTGTCAAAAACCTTAAGGCCCGGCTCACAGctTTAGAGACGTATCTCAGAGACCAGGCCATCTTGTCAAAT	1752
SIVSYK	-----	1752

AGMENY

VERVET.COM	TGGGGGTGGCGaTGGAAaCAAGTaTGTcATACaACaGTaccgTGG	1805
SIVAGM3	-----G-----C-----C-----	1930
SIVAGMTYO	-----C-----GGA-----	1909
SIVAGM155	-----G-----C-----	1915
GRIVET.COM	TGGGGATGTGCGTGGAAACAAGTATGTTCACACCACAGTGGCATGG	1838
SIVAGM677	-----AAGTATA-----	1867
SABAEUS.COM	TGGGGATGTGCTTTCAAGCAGGTGTGTTCACACGACGGTCTTTGG	1842
SIVSAB1C	-----AAGTATA-----	1882
SYKES.COM	TGGGGATGGCGATTCAAGCAGATCTGCCATACAGCAGTGCAGTGGGAGAAAGCATGTGGAAACAACCTCCAAC	1825
SIVSYK	-----	1825
VERVET.COM	atAAT?ggAC?CC?GATGGaAtAATATGACTTGGTGGAGTGGGAAAGACAGATAgc?gatTTGGaagGcAA	1874
SIVAGM3	-----A-----C-----T-----C-----A-----	2003
SIVAGMTYO	CA-----C-----T-----G-----C-----A-----	1982
SIVAGM155	-----A-----A-----G-----C-----A-----	1985
GRIVET.COM	ATAAC...ACTCCCTAAGTGGGACAATATGACTTGGTGGAGTGGGAGAGACAAATTAATGCCCTTGAAGGCAA	1908
SIVAGM677	-----	1937
SABAEUS.COM	ATAAC...ACTCCCGACTGGGAAAACATGACGTGGCAAGAATGGGAGAGGCAAAATAGAAAAGTATGAAGCTAA	1912
SIVSAB1C	-----	1952
SYKES.COM	TCTGTCCCAAACACAGTGGGAAAAATATGACATGGCACAGATGGGAAACAAGAGGTAGACAAATTTGACTGACCA	1898
SIVSYK	-----	1898
VERVET.COM	CATaACaa?ACAATTaGag?AgGC?AGaGaACagGAAGAAaAGAAATtTgGATGC?TATCAaAAgTtaactagt	1943
SIVAGM3	-----C-----G-----A-----C-----G-----	2076
SIVAGMTYO	---T---GGC---T---A---T---C---A---C---A---C---G---A---G---A---G---A---G---	2055
SIVAGM155	-----A-----G-----AC-----A-----G-----A-----T-----GT-AGAC	2058
GRIVET.COM	CATAACTCAACTATTGGAAGAACACAAATCAGGAATCAAGAATCTOGATCTGTACCAGAAATTOGATGAT	1981
SIVAGM677	-----	2010
SABAEUS.COM	CATCAGTAGGATTTTAGAGCAAGCACATGAACAGGAACAAAGAAATTTGGATAGTTATCAAAATTOGTGTCT	1985
SIVSAB1C	-----	2025
SYKES.COM	TATTGATCGATTCTCAGAGAAACACAAGAACAACAGGAGAGAAATGTGCATGATCTAACAAAGTTACAGGAA	1971
SIVSYK	-----	1971
VERVET.COM	TGGTCagaTTTcTGGTCTGGTTOGATTTcTCAAAaTGGCTgaACATTtTAAaATaGGaTTTTTg?agTA?	2014
SIVAGM3	-----G-----C-----AT-----	2149
SIVAGMTYO	-----T-----G-----A-----T-----A-----G-----	2128
SIVAGM155	-----GAG-----T-----T-----G-----C-----C-----A-----	2131
GRIVET.COM	TGGTCAGGGTTCTGGTCATGGTCTCTCACTGTCAACTTGGTTAGGCTATGTTAAAAATAGGATTTTATGTATAG	2054
SIVAGM677	-----	2083
SABAEUS.COM	TGGTCAGACPTTTGGTCTGGTTTGACCTCACAaaATGGTPTGGATOGATGAAGATAGCCATCATGTTATAG	2058
SIVSAB1C	-----	2098
SYKES.COM	TGGGACTCGPTGTGGAGCTGGTPTGACTTGTGGAATGGTTCCTTTTATCTAAAGATAGGATTTTATGTAATAG	2044
SIVSYK	-----	2044
VERVET.COM	TAGG?atTaTAGGgTTAAGATTgCTTTTgACAGTATAT?otTGcATAGetAGGGTTAGGCAGGGtTaetCTCC	2085
SIVAGM3	-----T-----A-----T-----T-----T-----	2222
SIVAGMTYO	-----A-----A-----A-----GGA-----T-----TG-----A-----TGT---	2201
SIVAGM155	-----CG-----T-----A-----	2204
GRIVET.COM	TGATTATTCTAGGATTAAGATTGCAATGGGTATTATGGGGATGTATPCAGAAATATTAGGCAGGGATATAATCC	2127
SIVAGM677	-----	2156
SABAEUS.COM	CAGGCATCATAGTAGCTAGAGTTCTGCTAGTAATAATAGGAATTCCTAGGAAGTTTAGAAAAGGGTATGCACC	2131
SIVSAB1C	-----	2171
SYKES.COM	GAGCTTTAGTGTCTTTAAGGCTAGTTAGTTTTFAGTGTAGGAATCAFAAAGAAATTTGTTGGGAGGCTATGTTCC	2117
SIVSYK	-----	2117
VERVET.COM	TcTaTCTCCaCAGATCCATATCCAcccgttgG??AAGGGACaGCCAGACAACGCaGaagGcCAGgAGAAGGT	2155
SIVAGM3	-----T-----G-----	2292
SIVAGMTYO	-----AAGT---GGG-----G-----C-----C-----	2274
SIVAGM155	-----T-----T-----T-----G-----A-----	2274
GRIVET.COM	TCTC...CCCCAGATCCATATCCACAGT...TCAGCGGAACGCCAGACAACGGAGGAGGCAAGACAGAGGT	2194
SIVAGM677	-----	2223
SABAEUS.COM	TCTCTCTCTCCCTCCCTCTCTCTCAFPAA	2159
SIVSAB1C	-----	2199
SYKES.COM	TATTTTCAGAACCCCTACCCAGGCCCG...AAGCACCCAGGCAAAACCCGCAGACGAAGAAGGCCAGT	2184
SIVSYK	-----	2184
VERVET.COM	GGAgacAa??GCAaGa?CAaaTc?gaGccTTGG?AGAAAGA?TC?ggcacagcagagtggaAGAGCA??tggt	2219
SIVAGM3	-----GC-----A-----GC-----C-----A-----T-----AC-----	2365
SIVAGMTYO	-----CA-----G-----T-----TA-----T-----T-----C-----C-----GG---	2329
SIVAGM155	---AGA-CAG---A-G-----TAC---A-	2304
GRIVET.COM	GGAGAAAGCAGCAGCAGCAAAATGATAAGATTGCAGGAAGAGTCTTCAACACCTTCGAGGATCAACAACCTGGT	2267
SIVAGM677	-----	2296
SYKES.COM	GGAGACAGAGAAGGACTCAACGTCTCTACCTTCAGCAGAGAATCTTTGAGGCAATCTTTGGAAAGCAGGACAGC	2257
SIVSYK	-----	2257

VERVET.COM	G?A?GC??TTGACC??TGG??CGA??CA?CA??TGGCTATACAACAGTTGCCGTGACCCCTCTCT??TTCA	2272
SIVAGM3	-C-A--GA-----AAT---TGCT---TCAG--G--TC-----AG----	2438
SIVAGMTYO	-T-T--AG-----GCC---CTGA---GGCT--A--CT-----CA----	2402
GRIVET.COM	GGCTCAACTTCAAGAGCTGCAGCTTGAGAATAAGGACTTGGTGTACAAACATCTGCCGTGACCCCTCCTCATATT	2340
SIVAGM677	-----	2369
SYKES.COM	AGCTTTGGAGGACAGTC.....TGCAGCAGCTTCAGATCTCTGATTTAGACA	2303
SIVSYK	-----	2303
	<- rev nef cds ->	
VERVET.COM	XTTAGGA??GCTTT?CA?TAC?T?CAATATGGGCT?G??GAACTCAAA?CCG??GCACAAGAA??A?T???	2325
SIVAGM3	T-----GC-----C--G---A-A-----T-GG-----G---CA-----GC-G-TGTC	2511
SIVAGMTYO	G-----AA-----T--A---C-G-----C-CA-----A---GC-----AT-C-CCAA	2475
GRIVET.COM	CATCAGGACAGCAGTGGATACCTGCAGTATGGGCTCCAGCAACTCCAAGAGGCAGCAACAGGGCTTGGTCAA	2413
SIVAGM677	-----	2442
SYKES.COM	GCTTACTATCACCTGGGGATTTCATCAGCTATGGGTTCACAGAGCTCAAAATCGCAGCAGCTTCGCTCGGAAGG	2376
SIVSYK	-----	2376
VERVET.COM	?CT?TGGCA?GC?TTGCACAAAACGC??G??ACCAGATATGGCTTGCTTGCGAGATCCGCTTATAGG??ATC?	2386
SIVAGM3	G--T-----C--C-----GG-CT-----GCA---A	2584
SIVAGMTYO	A--C-----G--G-----AT-TC-----AAC---G	2548
GRIVET.COM	GCTCTGGCGAGGGCTGCCAGGGAAGGCTGGGGCAGACTGGGTGCTATTGTCCGATCCGCTTATCGGGCAGTCA	2486
SIVAGM677	-----	2515
SYKES.COM	GAAATACCGGATTTGGGTGGCGGCTATTGGCCAAGCAATATACGCCGCTACCAGAAGAGTTGTGAGGGCCGTTG	2449
SIVSYK	-----	2449
VERVET.COM	TCAAC??TCCAAGAGAGTGGCACAAGGCCTTGA?G?AATCCTTAAT	2429
SIVAGM3	-----TC-----A-G-----	2631
SIVAGMTYO	-----AG-----G-A-----	2592
GRIVET.COM	TCAACAGTCCAAGAAGAGTGGGCAAGGCCTTGAAAAAGTCCGCGG	2533
SIVAGM677	-----	2562
SYKES.COM	CAGCCTTGCCGCGGAGGCTTCGACAAGGCTTGGAGATCTACCTTAAC	2496
SIVSYK	-----	2496

SIVSAB1C

LOCUS SIVSAB1C 10036 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) complete genomic sequence, African green monkey isolate SIVagmSAB-1.

ACCESSION U04005

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PEMCs derived from a wild female asymptomatic sabaeus African green monkey captured in its natural habitat of Senegal.

REFERENCE 1 (bases 1 to 10036)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan-40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.

Virus sab-1 was derived from a wild female asymptomatic sabaeus African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.

Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaeus viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaeus strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaeus virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol.

NCBI gi: 466229

BASE COUNT 3293 a 1975 c 2517 g 2251 t

ORIGIN

```

1  tggaggggat ttattacagt gaaaggagaa agaagatctt ggcctcttat gctcttaattg
61  agtggggcat agtagatgga tggcagaatt atacagatgg gccagggtacc aggtacccaa
121 agtgcttttg atggtgcttt aaactgggtac ctgtggacct gtcagaggag gctaagaata
181 gtgagaacca ctgctgctg catcctgctc aagtggccta tgaggatgac gcttgaaaag
241 agaccctggt gtggaagttt gaccactac tagcagtgga ctatgttgct tggaggctgc
301 accagagaca ggtgccgtca gcacaaggct aaccacaaaa ccacatccta tggagttgtc
361 atggtgatga cattaagaac tgcctgactga gaGGGACTTT CCACGGTCCA GACTGGGCGG
                                NFkB ->      <--> Sp1  <- -> Sp1
421 TACTGGGAGT GGcttgcccc gtgctgctca aTATAAgcag ctgctttgag cctgtaaaac
    <--> Sp1  <-          signal ->
481 agtctctgct tggaggctgc cagattgagc ctgggtgttc tctggtaagt ctgaaccagc
                                ->      TAR I      <-          ->
541 ttgagcctgg gtgttcgctg ggaaggcaa gcctaagagc actctggtga ggtcttggtg
    TAR II      <-
601 gaccctcgcc tggcgactgg ccattgccag tagcagagac tccgcttgct tgccttgattg
661 cctcAATAAA gtcttgcttt agttacaag taagcaagt ggtgcccac ttattctca
    signal ->

```



```

721 ggacaacct ggttactaag gatccctgat agaaaaaatc cttagcagtg gcgccccgaac
      <- LTR end
781 agggacttga ctaaaagacgt cccggccgca gactagacca gcttgaggac tcccttgact
841 gagtaagacc ggacgggaaga aaggcggacc gtgcccggagt gcaggtaaga ggcgcctgag
901 gaccgcgagac accctcgggc tgtgtgtgtg tgagaagaga ggaaaaagag aagtcagaca
961 gggctgacaa ggctggctta gcagtggcag aaggctaggt cgagcgggaa aggaaaATGg
      gag cds start ->
1021 gtgcgagtaa ctcagtcctta agtggtagga aactagatgc attcgaatca gtcagggttac
1081 ggcaccaacgg aaagaagaag tataagttaa gacatctagt atgggcaagc aagggaactag
1141 atcgatttag cctttcagca aatctcttag agacaaagga aggggtcgtc aaaattctta
1201 gtgtactcct cccctagtt cctacagggt cagagaattt aattgctctc ttaatttat
1261 gctgcgtgtt agcttgtatt cagcagaaa taaaagtga agatacagag gaagcaaaag
1321 cgaagataaa agagggaagt ccagcagaaa tgacagaaag tgccacagcg acatctagt
1381 gccaanctaa ggaactgcag gcaaaaaaga aaaatgagcc aacagtgaac cctagtgggt
1441 gctcgcgaaa ttaccccatl gtcagtgtta ataataatg ggtccaccag ccattgtccc
1501 caaggacact taatgcttgg gtaaaagtga ttgaggaaaa gaagttcagt gcagaggtag
1561 taacctgttt tagcgcccta gcagaaggag ctatccctta tgatatcaat caaatgtctc
1621 atgcagttag agaacaccag ggggcactac aaatagtga agatgttacc aatgaagaag
1681 cggcagattg ggaccttaga ccccgccgc cgcaacagcc ccagctcag ggagtgttaa
1741 gggaccacac agggtcagac atagcgggca ctactagcac aatacaagag caaatagaat
1801 ggaccaccag agcacagaat gctgtcaatg taggaatat ctataaagga tggatcatcc
1861 tgggtctcca gaattgtgtc aaaatgtaca acccagtaaa catttttagc attaaacaag
1921 ggccataaga acctttcaag gattatgtag atagattcta caaggctcty agagcagagc
1981 agacagaccc tgctgtaaa aactggatga cacagtctct gttgatccag aatgccacc
2041 cagattgtta gcagtgctc aaagggttag gcatgaatcc aacctagaa gagatgtta
2101 ctgcatgcca aggcattgga ggagcacaac ataaggcacg tctaattggc gaagctatga
2161 cagcggcctt tcagcaacaa acagttggga acatctttgt gcagcaagg gcaagacca
2221 ggggtccctt aggggggaaga ggtagacctc tgaacccaaa tattaatgt tacaattgt
2281 gaaagcctgg tcacctggct agattttgta agccccctag gagacaaggc tgcgggaat
2341 gtggcagccc agaccatcag atgaaagatt gtcaaaaaca ggtTAAtttt ttaggggttt
      pol orf start -> (NH2 terminus
      unknown)
2401 gccctcgggg cagagggaaa cccaggaatt tccctctgac ctccatcaga ccaacagctc
2461 cccaatgga acgggattac agcaggccgg aggaaaattg gtatgcagac agacctcga
2521 ccagaggacc agggccagac gatccagcaa cagccctgtt aaagcagtat gctgttcagg
2581 ggaaacggca gaaacagcag tggcaaaacc actcgccaca acagagcccc tacgaggagg
2641 ottacagctc cctcaggtct ctctttggag aagaccaatG Aaaactgtgt acatagagg
      <- gag cds end
2701 acagaaagtc actgctctat tagacacagg agcagatgac tcagtaatac agggatatga
2761 attaggagac aattggaaac caagaataat tggaggataa ggaggatgta ttaattgtaa
2821 agcataccac aatcaggagg taaaaataga agataaaact tgtaaaagta ccatttttgt
2881 agggagaaac ccagtttaata ttataggaag aaatgtctta gcacaattag gacttactct
2941 aaatttgaca caaagagaga tagaacctat aaaagtccac ttaaagccag gacaagatgg
3001 gccaaaggata aggcaatggc ctttgtcaaa agaaaaataa gaggccttaa agggccattt
3061 tgaagactta gaaaaacaag gacatttggg aagaataggg ccagaaaaatc catataaac
3121 cccagtcctt gcaataagga agaaagataa aactcagtg agaatactca tggacttcag
3181 gcagtttaat aaaagtactc aagactttca ggaagtccag ttagggatcc ccccccagc
3241 aggccttacag caaaggggagc agattacagt gttggacata ggagatgctt atttttagct
3301 ccccttagat ccagacttcc agaaatatac agcattcacc atcccatcag tcaataatag
3361 ggagccaggc ataagatac agtataagggt gctcccacag ggttggaaag gctcaccac
3421 aatttttcag acaacagcca acaaaattct gcaggaaattt aggcaaaaga acccagatgt
3481 agatataat caatacatgg atgatatgtt aatagccagt gataggccaa aggcagaaca
3541 tctagtaatg gtacagcagt taagagacta tctagaaact tgggggttta aaaccttg
3601 aaagaagttt caaaaggatc caccatacct ctggatgggg tatgaattgt atccaaagaa
3661 atggcaactg caggagatca cctaccaga aagggaagaa tggacagtga atgatattca
3721 aaaatttagta ggaaggttaa attgggcaag tcaaatttac acagggataa agactaaaca
3781 tttgtgtaga ctaattagag gagctagacc tttgacagag atagtccaat ggacagaaga
3841 ggcagaacta gaattagaag aaaatagaca gatcttaaga caaaagcagc agggacaata
3901 ttatgacctt gctcttctt taagggttaa ggttcttaaa ctgggagatg gccaatggg

```

```

3961 atatacagata taccagccag aaaataaaat actgaagggtt ggcaaatatg caaaaatcaa
4021 aacagcacac accaatgaat taagaatggt agcaggcccta gtacagaaga taggaagga
4081 aagcatagtc atctggggac agatacctat tatggaactg coagtagaaa gagagctctg
4141 ggaacaatgg tgggtctgact atctggcaggt cacctggatc ccggaatggg aaatggtcag
4201 tactcctcaa ttgattagat tgtgggtacaa attagtaaaa gatcccatcc caggagaagc
4261 agtctattat gtagatgggg cagctaantag aaattccaaa gaaggaaaag caggatactt
4321 aacagataga ggggatcaaa aggtggtagc attagagaat actaccaacc agaagcaga
4381 gttagaggcc attctgttag ccttaagaga ctctggagc aaagtaata taataacaga
4441 ctcaaatat gctatgggta ttatagcagg ggaacctaca gaatcagata ataacatagt
4501 acaacaaatt atagaggaa taataaagaa ggaggcagtg tatatagcat gggtaacctg
4561 ccataaagga gtagggtggca atgaggaaat tgataaacta gttagtcaag gaataagaca
4621 agtactatct ctggacagaa ttgaagaagc acaagaagag catgataaat accatgccaa
4681 ttggaggagt atgcaacaag aatttggctt acctgctata gtagcaaaag aaatagtagc
4741 ggcattgctt aaatgccaga taaaagggga gtctgtacat gggcaagtag atgctagtcc
4801 aggagtgtgg caaatggatt gtacacacct agaaggaaaa atcatcatag tggctgttca
4861 tgttgctagt ggatttatag aagcagaagt aatccacag gaaacaggaa aagctacagc
4921 acacttcctg ctgaagttag ccagcagatg gcccatcaca cagctgcata cagataatgg
4981 taccaacttt actagccagc aagtagcagc tatttgcctg tggggaaaaa tagagcacac
5041 atttggggta cctacaacc cccaaagtca gggagttaga gaatctatga ataagcagtt
5101 aaaagaaatc ataggacaaa ttagagatga tgcagaaga ttggaacag cagtataat
5161 ggcagtgcac attcacaatt ttaaaagaaa aggggggatt ggggggtact ctgcacaga
5221 aagactaatc aatataatc atacagaact agaaccacaa acacttcaac aaaaatttc
5281 aaaaattcaa aattttcggg tctattacag ggaaggcaga gccctgtgtt ggaagggtcc
5341 ggctaagctc atctggaaag gagaaggagc agtagttata caagsacagg gtgaattgaa
5401 aacaatccca agaagaaaag ccaaaattat aaaagattAT Gaaaagcat tggatagtca
vif cds start ->
5461 ggccccctctg gaaggtaacg ggccgacagc aggagaggtg gactAGccta gtgaaatacc
<- pol cds end
5521 acatgcatgt gtccaaacag tgtgttcatt ggccgtatcc cctcatact aaataaagat
5581 ggaatttgga ctccatctca gaatgggtca taccctgaa ggatggagct ctaataaaag
5641 ttactaatta ttggcattta accccagaga aaggatggtt agaaacctat gcaacaggaa
5701 taggatacag taagggggaa tgggtttacag agcttgatcc ttggacagcg gaccatatta
5761 ttcattggtc ctatttcccc tgtttttacag atagagctgt ccaacaagcc atcaggggag
5821 aaaagtatct ttggtgtaaa caccaggtgg gacatcagcc aacagggcaa gtaccatcat
5881 tacagtatct agcacttagg gtatatacaa ATGgcctcag gaggggtggt cccaccagta
vpX cds start ->
5941 ggagggggtc ctcccaagga tcccccaag aatcccagag aagagatacc aggatggcta
6001 gaaacatggg atttgcccag agagccgttc gacgaatggc tccgagacat gttacaggac
6061 ctcaattcag agggccagtg ccacttccca aggaatctcc ttttcogtct ttggtggaat
6121 attgtggaag aaccagccat TGATcatggy caaaccaggc tggaaaggctg gtataaatat
<- vif cds end
6181 tgtagaattc tgcaaaaggc tttattttgt catatgaaag gtagatgttg taagccaaag
6241 actcaccctg cgtATGgacc aggagcagga ggcccgccc caggctctggg aggagcttca
tat cds start ->
6301 ggaggagctg catcgccgcg tccaggccctg TGATAaact tgcctctgca aagtttgcctg
<- vpX cds end
6361 ctttcattgt attctctgct tccataaaaa ggccctagcc attogctatt ATGtcctag
rev cds start ->
6421 gccagaaga gcttctaaga agattttcgca taatcaagtt tctctacaca actAGTaaagt
5' sj /\
<- tat cds
6481 ATGaaagcttc ttacagtact gctttggcta agtgggtgtt ggagcttagt gtggttagtc
env cds ->
6541 cagtatgtta cagtgtttta tggatccca gtaggaaaa atagttcagt ccaagccttt
6601 tgtaagaccc ccaatacgaa cctatgggcc tctactaatt gcatcccgga tgatgaacca
6661 gaaggaaaca tagcagaagt accataacct aetatcacag aaaaatttga tgcctggaaa
6721 aatcggaacc ctctagtagg ccaagcagag agtaacatac accttttgtt tgaagtagt
6781 ttaaagccat gtgttaagct ctacccaatg tgcataaaaa tgaactgcta tgccttagaa
6841 ggaggagcag caacaacaac atcaccatca acatcaacag caogcccgga ggtagttagt

```

```

6901 gtgggcttta atgactcagt gatagaacaa gagatggaaa aagaacaggc catgaactgt
6961 tccctttgcta tggcaggata tagaagggat gtaaagaaaa attactcaac agtatgggat
7021 gatcaagagg tggttttgtga agaggggaagg gaaaaatcaa acgcaacaca tactgtggga
7081 tgttacatga tccattgtaa tccctcagta atcaaaggag cttgtgataa aacttattgg
7141 gataccttta ggcttaggta ttgtgctcca gcagggttatg ctttgctaag atgtgcagat
7201 acagattatt caggacataa ggcattgtagg aatgtaacag tatcagcttg cactaggtta
7261 atcaatacta ctgtaagcac aggaatagga ataatggta gttatgtagc aatatagaaca
7321 gaaatctggc agaaaaatgg caattccaat gattctgtta ttattaggct caatagatat
7381 ttttaatttaa ctatcagatg tagaagacca ggaaataaga cagtattgcc tgaaccatc
7441 atggctggct tggctctcca tagccagaaa tataacacca gggtgaaaca agcttgggtgc
7501 tgggttggag gtaattggag aggagcctgg aaggaagtca aagaacaat agtgagacta
7561 ccaccaaaaa agtatagtgg cacaatatgc acaacaaga tttttctaca aagacaatgg
7621 ggagatccag aatctgaatt cttctttttt aattgtcaag gtgaattctt ctattgtaaa
7681 atggattggc tctcaacta tttgaataat aagtcagtag acccagacca taacaactgt
7741 gcaaaaaata atacaaagcc atgttggcag agaacctatg ttccctgcca cataaggcaa
7801 gtagttaatg attggtagac ctgtctctaa aagacatacg cccctcccac agaaggccac
7861 ttggagtgtg actcgacagc cacagcgtg tatgtagaat tgaactataa cagtaaaaaa
7921 aggaccaatg tgacgttatc acctcagatt gaaagcattt gggctaatag attggagat
7981 tacaataatg tggaaatcaa gccaatggga tttgctccaa ccaagtaag aagatacact
8041 ggccccgaaa gacaaaaaag ggtgccatc gtgctagggt tcttaggtt cttgggagca
8101 gctggagctg caatgggtgc agcggcgaca gccctgacgg tccagtctca gcagttactc
8161 gctgggatac tgcagcagca gaaaaatctg cttgcggtg tagagcagca gcagcaaatg
8221 ctaaagctca ctatttgggg tgtaaaaaac ctcaatgcc gcgtcactgc tcttgagaaa
8281 tacctagagg atcaggcaag gttgaacatt tggggatgtg ctttcaggca ggtgtgtcac
8341 acgacgggtc tttggaagta taataacact ccgactggg aaaaatgac gtggcaagaa
8401 tgggagaggc aaatgaaaaa gtatgaagct aacatcagta ggattttaga gcaagcacat
8461 gaacaggaac aaaagaattt ggatagttat caaaaattgg tgtcttggtc agacttttgg
8521 tcgtgggttg acctcacaaa atgggttggg tggatgaaga tagccatcat ggttatagca
8581 ggcatacatg tagctagagt tctgctagta ataataaggaa ttcttaggaa gtttagaaaa
8641 gggtagcac ctctctctc cctccctct tctcatTAAC AGatccatat ccacctggac
      /\ 3' sj
      <- env cds end

8701 aagggacagc cagacaaaga agaagagcaa ggcagcggtg ggctaaacag cgacagcagg
8761 tcatacactt ggcagagaga attcttgaga cactctgtgc acagattgat cacttggctc
8821 aggaatttga ccagttgggt ctcgacaatc ttcagcaacc tccatcgctg cctccaggac
8881 atccaacaga gaaccagaca gtaacagct catctTGhat ATGggtggca agagctcaaa
      <- rev cds end

      nef cds start ->
8941 gcagcagcag cgtcactcac tgtggctgtg gtccaagctg cgacaagcgc cagtgtattca
9001 gtacgacatg cttgcagatc cattgttagg gcagtcacg cacatccaag aagaatgggc
9061 caagagcttg agagatgggt taattaggca aggagactcc tctagaacag aagaaggtgt
9121 aaaaatgaaa catcaaggaa gacagccctc gtggtatgat gaggatgaag aagaagtagg
9181 ctttccagta aggccttgct taccactaag agcaatgaca tataagttgg caatagactt
9241 tggccacttt ttaaaagaaa aggggggact ggaagggtatt tattacagtg aaaggagaaa

      LTR start ->
9301 gaagatcttg gatctctatg ctcttaatga gtggggcata gtagatggat ggcagaatta
9361 tacagatggg ccaggtagca ggtacccaaa gtgctttgga tgggtgctta aactggtaac
9421 tgtggacctg tcagaggagg ctatgaatag tgagaaccac tgccctgtgc atctgtctca
9481 agtggcctat gaggatgacg cctggaaaga gacctgggtg tgggaagtttg acccaactct
9541 agcagtgagc tatgtttgct ggaggtgtca ccagagcag gtgcccgcag cacaaggcTA
9601 Accacaaaac cacatcctat ggagttgtca tgggtgatgac attaagaact gctgactgag
      <- nef cds end

9661 agggactttc cagggtggag actgggctgt actgggagtg gcttgcccg tgcgtctca
9721 tataagcagc tgcctttgag ctgtaaaaaca gtctctgtct ggaggctgcc agattgagcc
9781 tgggtgttct ctggtaagtc tgaaccagct tgagcctggg tgttcgctgg ggaaggcaag
9841 cetaagagca ctctgggtgag gtcttggtag acctcgctt ggcgactggc cattgccagt
9901 agcagagact ccgcttgcct gcttgattgc ctcaataaag tcttgcttta gtttacaagt
9961 aaagcaagtg tgtgccatt tattctcag gacaacctg gttactaagg atccctgata
10021 gaaaaaatcc ttgaca

```

SIV2010E

LOCUS SIV2010E 906 bp DNA VRL 24-AUG-1994
 DEFINITION Simian immunodeficiency virus (SIVbab) partial envelope sequence.
 yellow baboon sample SIVbab2010.
 ACCESSION U10897
 SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from
 uncultured leukocytes derived from an asymptomatic pregnant female
 yellow baboon from Tanzania.
 REFERENCE 1 (bases 1 to 906)
 AUTHORS Jin, M.J., Rogers, J., Allan, J.S., Desrosiers, R.C., Shaw, G.M.,
 Sharp, P.M. and Hahn, B.H.
 TITLE Infection of a yellow baboon with SIV from African green monkeys:
 Evidence for cross-species transmission in the wild
 JOURNAL J. Virol. (In press, 1994)
 COMMENT Two wild yellow baboons from Tanzania were identified in a previous
 study with strong SIV(agm) seroreactivity. They lived in a
 national park where yellow baboons and African green monkeys share
 the same habitat. In this study, molecular characterization was
 performed for virus derived from one of the SIVagm seroreactive
 yellow baboons (animal 2010) and from two vervet African green
 monkeys (animals ver266 and ver385). With this aim in mind,
 uncultured leukocyte DNA was PCR-amplified, cloned and sequenced
 over a 672 bp region of gag and a 906 bp region of env for animal
 2010 (only the env region was analyzed for isolates ver266 and
 ver385). Jin et al. note that the viral strain derived from
 the yellow baboon (animal 2010) clusters with SIV strains of the
 vervet subtype, including ver266 and ver385, and believe that this
 is the first evidence for simian-to-simian cross-species transmission
 of SIV in the wild.

BASE COUNT 308 a 163 c 320 g 215 t
 ORIGIN approximately 862 bp downstream from the start of env
 1 agtacaggac tattattgaa tgggagctat tccgagaaca aaacccagat ttggcaaaaa
 env cds ->
 61 catagagtaa acaatgattc ggtactgata ttacttaata aacattacaa tctgacagtc
 121 acctgtaaga gaccaggaaa taagacaatc ttgctgtgca ctatcatggc aggaatggtg
 -> v3 loop equivalent
 181 ttccattccc aaaagtataa taccagactg aggcgaagcct ggtgtcattt taatggcaat
 <-
 241 tggagaggag cctggggagga agtgaaaaga gagatagtaa agttaccaa agacagatat
 301 cagggaaaca acgatgcaca gaagatctat ttacaaagac agtttgggtga tccagaagca
 361 gcaaatattat ggttcaattg tcagggagaa ttcttctatt gtaaaatgga ttggttecta
 421 aattatataa ataatttgac agttgatgca aatcataaca ctaacagatg caataagaag
 481 gggaaaacaa gtggagcgcc aggtccttgt gcacaaagaa cttatgttgc ttgccacatc
 541 agatctgtcg ttaatgattg gtatacagtt agtaaatga catatgcccc accaagagag
 601 ggtcatttgg aatgctgctc caccgtgaca gggatgacag tagagctaaa ttataacaac
 661 cagagcagaa ctaatgtgac tttgagccca caaattgaaa gtatatgggc tgctgaattg
 721 ggcagatata aattggtaga gatcacacca attggatttg caccacaga agtcaggcgg
 781 tatacagggg gacaagaaag acaaaaaagg gttcccttcg tgctaggggt cctggctttc
 841 ttgggggcag caggcactgc aatgggagca gggcgacag cctgacggg ccagtctcag
 901 gatttg
 env cds ->

LOCUS SIV2010G 672 bp DNA VRL 24-AUG-1994
DEFINITION Simian immunodeficiency virus (SIVbab) partial gag sequence, yellow baboon sample SIVbab2010.
ACCESSION U10899
SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from uncultured leukocytes derived from an asymptomatic pregnant female yellow baboon from Tanzania.
REFERENCE 1 (bases 1 to 672)
AUTHORS Jin, M.J., Rogers, J., Phillips-Conroy, J.E., Allan, J.S., Desrosiers, R.C., Shaw, G.M., Sharp, P.M. and Mahn, B.H.
TITLE Infection of a yellow baboon with SIV from African green monkeys: evidence for cross-species transmission in the wild
JOURNAL J. Virol. (In press, 1994)
COMMENT Two wild yellow baboons from Tanzania were identified in a previous study with strong SIV(agm) seroreactivity. They lived in a national park where yellow baboons and African green monkeys share the same habitat. In this study, molecular characterization was performed for virus derived from one of the SIVagm seroreactive yellow baboons (animal 2010) and from two vervet African green monkeys (animals ver266 and ver385). With this aim in mind, uncultured leukocyte DNA was PCR-amplified, cloned and sequenced over a 672 bp region of gag and a 906 bp region of env for animal 2010 (only the env region was analyzed for isolates ver266 and ver385). Jin et al. note that the viral strain derived from the yellow baboon (animal 2010) clusters with SIV strains of the vervet subtype, including ver266 and ver385, and believe that this is the first evidence for simian-to-simian cross-species transmission of SIV in the wild.
BASE COUNT 245 a 141 c 161 g 125 t
ORIGIN approximately 304 bp downstream from the start of gag
1 gtagcaacag taaggcaatg ctgccaccta gtgaasaaag aaaaaagtgc agcagagaca
gag cds ->
61 tctagtggcc gccagaaaaa tgacaggtcc acagcgacat ctagtggaca gagtcagaat
121 tttcccgccac aacaacaggg aaatacatgg gtacatgtac cactctctcc acgcacgttg
181 aatgcgtggg taaaggcagt agaggagaaa aaatttggag cagaatatgt accaatgttc
241 caagcattgt cagaaggctg tactccatat gacatcaatc aaatgttgaa tgtccttggg
301 gatcatcagg gggcgctaca aatagtgaag gaaatcatta atgaagaggg agcccaatgg
361 gatatacaac acccgccccc agcaggggcc ttaccagcag gacagcttag agaccaaga
421 ggctcagacg tagcaggaac aactagtaca gtgcaagaac aattagaatg gatctacaca
481 gctaacccaa gagtagatgt aggggctatc tatagaaggt ggataatcct tgggctgcaa
541 aaatgtgtca aaatgtataa cccagtgtca atcctagaca tcaaacaggg accaaaagaa
601 ccctttaagg attatgtaga tagattttac aggaccatca gacagagaca ggctcagga
661 gaagtaaaac aa
gag cds ->

SIVGRI2E

LOCUS SIVGRI2E 1101 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmGRI-2.

ACCESSION U03995

SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from cocultured PBMCs derived from a grivet African green monkey from Ethiopia.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Jin,M.J., Hui,H., Robertson,D.L., Mueller,M.C., Barre-Sinoussi,F., Hirsch,V.M., Allan,J.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaues and tantalus viruses were derived from feral animals.

NCBI gi: 466209

BASE COUNT 375 a 182 c 268 g 276 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 actcttagat tgagatattg tgcaccagca ggctttgctt tggtgagatg cttagataag
env cds ->

61 aattataacg gctttgctcc tagttgtaag aatgtgacag ttactcactg cactagacta
121 ataaatacaa cagtaagcac aggtatagga cttaatggta gtagagctga gaatagaaca
181 gaaatatggc aagaaaaggg gaataagcaat aatacagtta ttattaagtt aaataagttt
241 tataacctta cagtaagatg tagaagacca ggaataaaaa cagtgttgcc agtgacaatc
-> v3 loop equivalent

301 atggccggat tgggtgttca tagtcagaag tataatacta gattaaaaa agcatgggtg
<-

361 cacttccaag gcgattggca tggagcctgg aaagaagtaa aacaagaagt taagaaagca
421 aaaaatttga caacagatga cataaaattg agaaggcaat ggggggatcc agaagctgca
481 aacttctggg ttagctgcca aggagaattc ttctattgta aaatggattg gtttattaat
541 tatataaata atgaacagc ggatgcagat gggataaat gcaacaacc gaaaggacca
601 ggaccaggtc ctgtgtgaca acgaacatat gttgcctgtc acataagaca ggtagtcaat
661 gattgggtata cctcactaa gaaaacatac cgtccgcaa gagaaggcca gttggagtgc
721 aattcttcag tcacagcact gtatgtggag ataaattaca ataatgagag tggccctatt
781 aatgtgacat tgtcacctca gataggaagt atttgggctt atgaattggg cgattacaaa
841 ttagtggaaa taacaccaat tggttttgct ccacggatg tgagaagata tacaggaccc
901 acaaaacaaa aaagggttcc tttgtgcta ggattcttag gtttcttagg ggcagctgga
961 actgcaatgg ggcagctgc gacagcgtg acggtccagt ctgggcattt gcttgctggg
1021 atattgcagc agcagaggaa tctgctggcg gctgtggagc agcagcaaca actgttgaag
1081 ctgaccattt ggggtgtaaa a
env cds ->

LOCUS SIVGRI2L 697 bp ss-RNA VRL 28-MAR-1994
 DEFINITION Simian immunodeficiency virus (SIVagm) LTR sequence, African green monkey isolate SIVagmGRI-2.
 ACCESSION U04007
 SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PEMCs derived from a grivet African green monkey from Ethiopia.
 REFERENCE 1 (bases 1 to 697)
 AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys
 JOURNAL EMBO J. 13, 2935-2947 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan-40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.
 NCBI gi: 466239
 BASE COUNT 172 a 162 c 197 g 166 t
 ORIGIN start of LTR
 1 aatagaagag aggcatacct aaatctgtat gctctcaatg aatggggcat tatagatgat
 LTR ->
 61 tggcaaggat acaccggggg cccaggaata agataccegc taacgtttgg cttctgcttt
 121 aagctagtgc cagtggactc tccgaggagg cagccaactg ttatagacat tgcttctgc
 181 accctgctca ggcagggggg agtgaagatc caagcggcca tggagaggtc ttggtctgga
 241 agtttgacct taagttaaca gtggaattta agccatggat gcatagagac atgcatgaac
 301 atgcttgata aacagcagga ggaagtgggt accgctagc tcggttacca tggagatgac
 361 attaagaact gctgacGGGA CTTCCagca aGGGACTTTC CAGGCGGTC CggGAGGTOT
 NFKb -> NFKb -> <--> Sph <- -> Sph
 421 GCGGGGAGT GCGtttacc tcagagctgc aTAAAagcag atgctcgtg gcttgtaact
 <--> Sph signal ->
 481 cagtctctta ctaggagacc agctagagcc tgggtgttcg ctggttagcc taaccggtg
 -> TAR I <-
 541 gccacggggg gtaaggactc cttggcttca tatagctcAA TAAAtctgcc cgttagttg
 signal ->
 601 ctatattgga gtcaagtgt cattgctacg tcgcatctca cgaggtgaac ctctcttact
 661 ggggtctcct gttaccagg caggagaaac tccagca
 <- LTR

SIVGRI3E

LOCUS SIVGRI3E 1113 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmGRI-3.

ACCESSION U03994

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a grivet African green monkey from Ethiopia.

REFERENCE 1 (bases 1 to 1113)

AUTHORS Jin,M.J., Hui,H., Robertson,D.L., Mueller,M.C., Barre-Sinoussi,F., Hirsch,V.M., Allan,J.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.

NCBI gi: 466207

BASE COUNT 395 a 188 c 271 g 259 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 acattaagag taagatactg tgcaccagca ggatatgctt cattaaaatg caatgataag
env cds ->

61 gattataatg gctttgctcc aaaatgtaag aatgtttcag tagtacactg tactagggtta
121 ataaatacta ctataaccac aggaataggg ctaaattggga gtagagtaga aaatagaacg
181 gagatatggc aagaaaaggg aaatagtaat gatacagtaa taataaagtt aaataagttt
241 tacaacttga cagtggagatg ccgcagacca ggcaacaaaa caatactgcc tgtaacaatc
-> v3 loop equivalent

301 atggcaggac gagtggttca ttctcagaaa tacaatacga aattaaaaa agcgtggtgt
<-

361 cactcccagg gaaattggct aggagccctg aaggaggtta gagaaaaaat aaaagaagaa
421 aagaatctca ctgaagtaaa aatagaaaac atacatttga gaaggcaatg gggagaccca
481 gaagcagcaa atttctggtt caactgtcac ggagaatttt totattgtaa gatggattgg
541 ttcattaatt atcttaataa tcagacagta gatgcagatg gaaaaagctg tacttgcac
601 aaaaagggtc aaggaccagg accatgtgcg aagaaaacat atgttgcttg ccatataaga
661 caagtagtaa acgactggtt cactctaacc aagaaaacat atgccccgcc aagagaaggg
721 catttggagt gtaattcgac agtcacggca ctatatgttg aaataaatta taacaataaa
781 tctggtccaa taatgtaac cctatcttct caagttagga gtatatgggc atatgagttg
841 ggagattaca aattagtaga aettacacca attgggtttg ctctacaga cgtgagaaga
901 tacactggcc cctcgagaga gaagaggggtg ccattcgtgc tggggttctt aggcctcttg
961 ggagctgctg gcactgcaat gggcgcagcg ggcacagccc tgacgggtcca gtctcggcac
1021 ttacttgctg ggatatggca gcagcagaag aacttgctgg tggctgtgga gcagcaacaa
1081 cagctgttga agctgaccat ttggggtgtg aaa
env cds ->

LOCUS SIVGRI3L 700 bp ss-RNA VRL 25-MAY-1994
 DEFINITION Simian immunodeficiency virus (SIV) LTR sequence, African green monkey isolate SIVagmGRI-3.
 ACCESSION U04006
 SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a grivet African green monkey from Ethiopia.
 REFERENCE 1 (bases 1 to 700)
 AUTHORS Jin,M.J., Hui,H., Robertson,D.L., Mueller,M.C., Barre-Sinoussi,F., Hirsch,V.M., Allan,J.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys
 JOURNAL EMBO J. 13, 2935-2947 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.
 NCBI gi: 466238
 BASE COUNT 177 a 155 c 204 g 164 t
 ORIGIN start of LTR
 1 aatagaaggg agaasatggt aaacctctac gccctgaata agtgggggaat aatagatgat
 LTR ->
 61 tggcaaggat acaccccagg cccagggtata agataccggc tgacatttgg cttctgcttc
 121 aagctagtac cggtaaccct ctctgaagag gcacagaact gtgatagaca ctgtctactg
 181 cactctgctc aggcggggagg agatggggat ccaagtggac atggggagac cttgggtctgg
 241 aaatttgacc ctaagtttagc agtggagttt aagccatgga tgcataaaga tatgcacggg
 301 atggaaaaga agtaaacagc aagaggaagt ggtttccgcc tagctcggtt accatggaga
 361 tgacattaag aactgctgac GGGACTTTCC aggctGGGGC TTCCAGGGC GGTGCggGAG
 NFKb -> NFKb -> <--> Spl <- -> Spl
 421 GCGGTACGGG GAGTGGTttt gccctcagag ctgcataaaa gcagatgctc gctggcttgt
 Spl <--> Spl <- signal ->
 481 aactcagttc cttactagga gaccagctag agcctgggtg ttgctgggtt agcctaaccg
 -> TAR I <-
 541 gttggccacc gggggtaagg actccttggc ttcatatagc tcaATAAAcc tgccctgctta
 signal ->
 601 gttgctatgt tggagtcaag tgcctattgc tgcgcggagt ctctgaggtg aacctctctt
 661 actgagttct cctgttacc aggcggggaga aactccagca
 <- LTR

SIVSAB2E

LOCUS SIVSAB2E 1107 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmSAB-2.

ACCESSION U03996

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild female asymptomatic sabaues African green monkey captured in its natural habitat of Senegal.

REFERENCE 1 (bases 1 to 1107)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.

Virus sab-2 was derived from a wild female asymptomatic sabaues African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.

Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaues viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaues strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaues virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol.

NCBI gi: 466211

BASE COUNT 362 a 206 c 265 g 274 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 acatttagat tgagatattg tgctccagct gggttatgctt tgctaagatg tgcagataca
env cds ->

61 gattatttcag gacataaggc atgtaagaat gtaacagtat cagcttgcac taggttaata
121 aataactacta taagcacagg gataggtatc aatggtagtt atgtagcaaa tagaacagaa
181 atctggcaga aaatggaaa tgccaatcat actgttatta ttaggcttaa tagacatttt
241 aatttaacta tcagatgtag aagaccagga aataagacag tattgctgtg aaccatcatg
-> v3 loop equivalent

301 gctggcttgg tcttccatag ccagaaatat aacaccaagt tgagacaagg ttggtgctgg
<-

361 tttggcggtg attggagagg agcctggcag gaagtcaaag aaaaaatagt ggctttacca
421 ccagataggt atagtggcac aaatgacaca aacaagatct ttctgcaag gcaatgggga
481 gatccggaag ctgaattctt cttttttaat tgtcaagggt aatttttcta ttgtaaaatg
541 gactggttcc tcaactattt gaataatcag tcagtagacc cagaccataa tccctgtgga
601 ggaaatgaca agaataacag gaaatgttgg cagagaacct atgttccctg ccacataagg
661 caaatagtaa atgattggta cacttgtct aaaaagacat acgcccctcc cagagaaggc
721 cactgggagt gtaactcgac agccacagcg ctgtatgtag aactgctota taacagttaa
781 aacaggacca atgtgacatt atcacctcac attgaaagca ttggggctaa tgaattggga
841 gactacaaat tgggtggagat aaaaccaatt ggatttgcct caacagaggt aagaagatac

```
901 actggccccg aaaggcaaaa aagggtgccg ttogtgctag ggttcctagg cttcttggga
961 gcagctggag ctgcaatggg tgcagggggg acagccctga cggccagtc tcagcagta
1021 ctogctggga tactgcagca gcagaaaaat ctgcttgccg ctgtagagca gcagcaaatg
1081 ctaaagctca ccatttgggg tgtgaaa
      env cds ->
```

SIVSAB3E

LOCUS SIVSAB3E 1104 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence. African green monkey isolate SIVagmSAB-3.

ACCESSION U03997

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild female asymptomatic sabaeus African green monkey captured in its natural habitat of Senegal.

REFERENCE 1 (bases 1 to 1104)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan-40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.

Virus sab-3 was derived from a wild female asymptomatic sabaeus African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.

Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaeus viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaeus strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaeus virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol.

NCBI gi: 466213

BASE COUNT 371 a 205 c 255 g 273 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 acccttaggc ttaggtattg tgctccagca ggttatgctt tgtaagatg tgcagataca
env cds ->

61 gattattcag ggcataagaa atgtaagaat gtaacagtat cagcttgca tagattaata
121 aatactactg taagcacagg gatagggtatc aatgggagtt atgtagcaaa tagaacagaa
181 atctggcaga aaaatggaas ttccaatgat tctgttatta tcaggcttaa tagacatttt
241 aatttaacta tcagatgtag gagaccagga aataagacag tgttgctgtg aaccatcatg
-> v3 loop equivalent

301 gctggcttgg tcttccatag ccagaatat aacaccaggt tgaacaagc ttggtgctgg
<-

361 tttggaggta attggagagg agcctggcaa gaagtc aaacaaatagt ggctctacct
421 ccagatagggt atcaagggcac aaaagacaca aataaaatct tctacaaaag acaatgggga
481 gatccagaag ctgaattctt cttttttaat tgtcaagggt aatttttcta ctgtaaaatg
541 gattgggtcc tcaactatct gaataatcag tcagtagacc cagatcataa tccctgtaat
601 aagactagcg acaaaaaatg ttggcagaga aactatgctc cctgccacat aaggcaagta
661 gtaaatgatt ggtacacott gtctaaaaag acatatgccc ctcccagaga agggcacttg
721 gagtgcaact caacagccac agcgtgtgat gtagaactga actataacag taaaaacagg
781 accaatgtga cattatcacc tcagattgaa agcatttggg ctaatgaatt gggagactac
841 aaattagtag agataaaacc aattggattt gctcaacag aagtaagaag atacactggc

```
901 cccgaaaggc aaaaaagggc gccattogtg ctaggggtcc taggcttctt gggagcagct
961 ggagctgcaa tgggtgcagc ggcgacagcc ctgacggtec agtctcagca gttacttgc
1021 gggatattgc agcagcagaa aaatctgctt gcggctgtag agcagcagca gcaaatgcta
1081 aagctcacca tttggggtgt aaaa
      env cds ->
```

SIVSAB3L18

LOCUS SIVSAB3L18 741 bp DNA VRL 28-MAR-1994

DEFINITION Simian immunodeficiency virus (SIVagm) LTR sequence, African green monkey isolate SIVagmSAB-3, clone 18.

ACCESSION U04012

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild female asymptomatic sabaues African green monkey captured in its natural habitat of Senegal.

REFERENCE 1 (bases 1 to 741)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, P., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan49). These isolates were derived from different African green monkey species inhabiting various geographical regions.

Virus sab-3 was derived from a wild female asymptomatic sabaues African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.

Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaues viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaues strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaues virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol. An alignment of this LTR sequence and another one of its clones SIVSAB3L19 (U04013) is shown below.

NCBI gi: 466244

BASE COUNT 184 a 165 c 209 g 183 t

ORIGIN start of LTR

1 gagaaagaag atcttggaac tctatgctct taatgaatgg ggcatagtag atggatggca
LTR ->

61 gaattataca gatggggccag gtaccaggta cccaaagtgc ttggatggt gctttaaact
121 ggtacctgtg gacctgtcag aggaggctaa gaatagagag gaccactgcc tgctgcaccc
181 tgetcaagtg gcctatgagg atgacctctg gaaggagacc ctggtatgga agtttgactc
241 aatattagca gtggactatg ctgcttggag actgcaccca gaacagggtgc cgacagtgtc
301 tgactgacca caaaaccaca tctatggag ttgtcatggt gatgacatta agaactgctg
361 actgagaGGG ACTTTCAGG GTGGAGACTG GCGGTAAGTG GGAGTGGCtt acccctgtct
NFKB -> <--> Sp1 <- -> Sp1 <- -> Sp1 <-

421 gctcaaTATA Agcagctgct ttgcgcctgt aaaacagtct ctgcttggag gctaccagat
signal ->

481 tgagcccggt tgttctctgg taagtctgaa ccagcttgag cctgggtgtt cgctggggaa
TAR I <- -> TAR II <- ->

541 ggcaagccta agagcactct ggtgaggtct cggtagatcc tcacctggcg actggccatt
601 gccagtagca gggactccgc ttgcttgctt gattgctcA ATAAAgctct gccttagttt
signal ->

661 acaagtaaag caagtgtgtg cccatttatt cctcaggaaa accctgggta ctaaggatcc
 721 ctggttagaaa aatccttagc a

<- LTR

CONSENSUS	GAGAAAGAAGATCTTGGATCTCTATGCTCTTAATGAATGGGGCATAGTAGATGGATGGCA	60
SIVSAB3L18	-----	60
SIVSAB3L19	-----	60
CONSENSUS	GAATTATACAGATGG?CCAGGTACCAGGTACCCAAAGTGCTTTGGATGGTGCTTTAAA?T	118
SIVSAB3L18	-----G-----C-	120
SIVSAB3L19	-----A-----T-	120
CONSENSUS	GGTACCTGTGGACCTGTCAGAGGAGGCTAAGAATAGAGAGGACCACTGCCTGCTGCACCC	178
SIVSAB3L18	-----	180
SIVSAB3L19	-----	180
CONSENSUS	TGCTCAAGTGGCCCTATGAGGATGACCCCTGGAAGGAGACCCCTGGT?TGGAAAGTTGAC?C	236
SIVSAB3L18	-----A-----T-	240
SIVSAB3L19	-----G-----C-	240
CONSENSUS	??TATTAGCAGTGGACTATG?TGCTTGGAGACTGCACCCAGAACAGGTGCCG?CAGTGTC	292
SIVSAB3L18	AA-----C-----A-----	300
SIVSAB3L19	TG-----T-----G-----	300
CONSENSUS	TGACTGACCACAAAACCATCTCTATGGAGTTGTCATGGTGATGACATTAAGAACTGCTG	352
SIVSAB3L18	-----	360
SIVSAB3L19	-----	360
CONSENSUS	ACTGAGAGGGACTTTCCAGGGTGGAGACTGGGCGGTACTGGGAGTGGCTTACCCCGTGCT	412
SIVSAB3L18	-----	420
SIVSAB3L19	-----	420
CONSENSUS	GCTCAATATAAGCAGCTGCTTTGGCGCTGTAAACAGTCTCTGCTTGGAGGCTACCAGAT	472
SIVSAB3L18	-----	480
SIVSAB3L19	-----	480
CONSENSUS	TGAGCC?GGGTGTTCTCTGGTAAGTCTGAACCAGCTTGAGCCTGGGTGTTGCTGGGGAA	531
SIVSAB3L18	-----C-----	540
SIVSAB3L19	-----T-----	540
CONSENSUS	GGCAAGCCTAAGAGCACTCTGGTGAGGTCTCGGTAGATCCTCACCTGGCGACTGGCCATT	591
SIVSAB3L18	-----	600
SIVSAB3L19	-----	600
CONSENSUS	GCCAGTAGCAGGGACTCCGCTTGCTTGCTTGATTGCCTCAATAAAGTCTTGCTTAGTTT	651
SIVSAB3L18	-----	660
SIVSAB3L19	-----	660
CONSENSUS	ACAAGTAAAGCAAGTGTGTGCCCATTTATTCTCTCAGGAA?ACCCCTGGTTACTAAGGATCC	710
SIVSAB3L18	-----A-----	720
SIVSAB3L19	-----G-----	720
CONSENSUS	CTG?TAGAAAAA?TCCTTAGCA	730
SIVSAB3L18	---T-----	741
SIVSAB3L19	---A-----A-----	742

SIVSAB4E

LOCUS SIVSAB4E 1104 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmSAB-4.

ACCESSION U03998

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild female asymptomatic sabaes African green monkey captured in its natural habitat of Senegal.

REFERENCE 1 (bases 1 to 1104)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan-40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.

Virus sab-4 was derived from a wild female asymptomatic sabaes African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.

Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaes viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaes strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaes virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol.

NCBI gi: 466215

BASE COUNT 368 a 194 c 262 g 280 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 accttttagat taagatatgt tgctcctgca ggctatggt tactaagatg tagagattca
env cds ->

61 gactacagag gacataagaa ttgtcaaaat gtatcagttt cagcttgca taggcttatt
121 aatactacag taagtacagg aatagggtatt aatggtagtt atgtagcaaa tagaacagaa
181 atatggcaag aaaatggaaa tagtaataat acagctataa taaggcttaa taggtattat
241 aatcttacc aagatgtag aaggccaggg aataagacag tattgctgt gaccatcatg
-> v3 loop equivalent

301 gctggcttgg tctttcacag ccaaaatat aacacaaggt tgagacaagc ctgggtgttg
->

361 tttgggggta agtggagagg tgcttggcaa gaagccaaag aagaaattgt gaaattatca
421 ccagaaaggt atagcgggac caaaaacaca tcacaaattt tcctacaaag acagtgggga
481 ggtccagaag cagaattctt ctttttttaac tgccaaggag aatttttcta ttgtaaaatg
541 gattgggtct taaattattt aaacaatcaa tcggttgacc ctaatcataa cacatgtact
601 ccaaaaagcg acaaaagatg ttggcagaga acttatgtgc cttgccacat tagacaagta
661 gtaaatgact ggtataccct gtccaaaag acatatgttc caccagggga aggccacttg
721 gagtgtact caacagcaac agctctttat gtggaattga actacaatag taagaacagg
781 accaatgtga cactctgcgc tcagattagg agtatctggg ctaatgaatt aggtgattat
841 aaattggtag aaatcaaacc cattgggttt gctccaacaa atgtgagaag atatactggc


```
901 cccgagagac agaaaagggg gccattcggtg ctgggggttcc taggettctt gggagcagct
961 ggagctgcaa tgggtgcagc ggcgacagcc ctgacgggtcc agtctcaga attacttgct
1021 gggatattgc agcagcagaa aaatctgctt ggggtgtag agcagcagca gcaaatgcta
1081 aagctcacta tctggggtgt gaag
env cds ->
```

SIVSAB4L10

LOCUS SIVSAB4L10 742 bp ss-RNA VRL 17-SEP-1994
 DEFINITION Simian immunodeficiency virus (SIVagm) LTR sequence, African green monkey isolate SIVagmSAB-4, clone 10.
 ACCESSION U04014
 SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PEMCs derived from a wild female asymptomatic sabaeus African green monkey captured in its natural habitat of Senegal.
 REFERENCE 1 (bases 1 to 742)
 AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Muller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Mosaic genome structure of simian immunodeficiency virus from west African green monkeys
 JOURNAL EMBO J. 13, 2935-2947 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.
 Virus sab-4 was derived from a wild female asymptomatic sabaeus African green monkey. She was caught in her natural habitat of Senegal and subsequently exported to the United States. From the point of capture and throughout the study, she was housed in an individual cage.
 Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaeus viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions (1). This feature was demonstrated for two viral sabaeus strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaeus virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol. An alignment of this LTR sequence and another one of its clones SIVSAB4L4 (U04015) is shown below.
 NCBI gi: 466246
 BASE COUNT 182 a 164 c 205 g 191 t
 ORIGIN start of LTR

```

1 aaggaageag atcttggatc tatatgcttt gaatgagtg ggcetagttag atggctggca
LTR ->
61 gaactacact gatggggccag gtacaaggta ccgaagcag tttgggttct gctttaagct
121 ggtacctgtg gctctatcag aagcagctca gaatgatgag catcactgcc tctgcaccc
181 tgctcaata gcttatgagg atgacccatg gaaggagact ctggtttggg agtttgatcc
241 aatactagca gtggactatg ttgcttggag actgcaacca gagcaggtgc cgtcagtgtc
301 tgctgacca caaaaccaca tectatagag ttgtcatggt gatgacatta agaactgctg
361 actgagaGGG ACTTCCAGG GTGGAGAcTG GCGGCTACTG GGAGTGGCtt accccgtgct
      NFkb ->      <--> Spl  <- -> Spl  <--> Spl  <-
421 gctcaatATA Agcagctgct ttctgcctgt aaaaacagtct ctgtttggag gctgccagat
      signal ->
481 tgagcctggg tgttctctgg taagtctgaa ccagcttgag cctgggtgtt cgtcggggaa
      TAR I      <-      ->      TAR II      <-
541 ggcaagcctt agagcacact ggtgaggtct tggtagatcc tcaactggtg actggccatt
601 gccagtagca gagactccgc ttgcttgctt gattgctcA ATAAAgcttt gccttagttt
      signal ->

```

661 acaagtaaag caagtgtgtg cccattcatt cctcaggaaa accctgggta ctaaggatcc
 721 ctgtgaagaa aaatccttag ca

<- LTR

CONSENSUS	AAGGAAGAAGATCTTGGATCTATATGCTTTGAATGAGTGGGGCATAGTAGATGGCTGGCA	60
SIVSAB4L10	-----	60
SIVSAB4L4	-----	60
CONSENSUS	GAACTACACTGATGGGCCAGGTACAAGGTACCCGAAGCAGTTTGGGTCTGCTTTAAGCT	120
SIVSAB4L10	-----	120
SIVSAB4L4	-----	120
CONSENSUS	GGTACCTGTGGCTCTATCAGAAGCAGCTCAGAATGATGAGCATCACTGCCTCCTGCACCC	180
SIVSAB4L10	-----	180
SIVSAB4L4	-----	180
CONSENSUS	TGCTCAAATAGCCTATGAGGATGACCCATGGAGGAGACTCTGGTTTGGAACTTTGATCC	240
SIVSAB4L10	-----	240
SIVSAB4L4	-----	240
CONSENSUS	AATACTAGCAGTGGACTATGTTGCTTGGAGACTGCACCCAGAGCAGGTGCCGTCAGTGTC	300
SIVSAB4L10	-----	300
SIVSAB4L4	-----	300
CONSENSUS	TGCCTGACCACAAAACCATCCTATAGAGT?GTCATGGTGATGACATTAAAGAACTGCTG	359
SIVSAB4L10	-----T-----	360
SIVSAB4L4	-----C-----	360
CONSENSUS	ACTGAGAGGGACTTTCCAGGGTGGAGACTGGGCGGTACTGGGAGTGGCTTACCCCGTGC?	418
SIVSAB4L10	-----T-----	420
SIVSAB4L4	-----C-----	420
CONSENSUS	GCTCAATATAAGCAGCTGCT?TTCCCGCTGTAAACAGTCTCTG?TTGGAGGCTGCCAGAT	476
SIVSAB4L10	-----T-----T-----	480
SIVSAB4L4	-----C-----C-----	480
CONSENSUS	TGAGCCTGGGTGTTCTCTGGTAAGTCTGAACCAGCTTGAGCCTGGGTGTTGCTGGGGAA	536
SIVSAB4L10	-----	540
SIVSAB4L4	-----	540
CONSENSUS	GGCAAGCCTTAGAGCACACTGGTGAGGTCTTGGTAGATCCTCAACTGGTGACTGGCCATT	596
SIVSAB4L10	-----	600
SIVSAB4L4	-----	600
CONSENSUS	GCCAGTAGCAGAGACTCCGCTTGCTTGCTTGATTGCCTCAATAAAGTCTTGCTTAGTTT	656
SIVSAB4L10	-----	660
SIVSAB4L4	-----	660
CONSENSUS	ACAAGTAAAGCAAGTGTGTGCCCATTCATTTCCTCAGGAAAACCTGGTTACTAAGGATCC	716
SIVSAB4L10	-----	720
SIVSAB4L4	-----	720
CONSENSUS	CTGTGAAGAAAAATCCTTAGCA	738
SIVSAB4L10	-----	742
SIVSAB4L4	-----	742

SIVSABD37P

LOCUS SIVSABD37P 651 bp ss-RNA VRL 17-SEP-1994
 DEFINITION Simian immunodeficiency virus (SIVagm) pol sequence, African green monkey isolate SIVagmSABD37, partial cds.
 ACCESSION U04018
 SOURCE Simian immunodeficiency virus, PCR-amplified DNA extracted from cocultured PBMCs derived from a wild sabaues African green monkey captured in its natural habitat of Senegal.
 REFERENCE 1 (bases 1 to 651)
 AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Muller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Mosaic genome structure of simian immunodeficiency virus from west African green monkeys
 JOURNAL EMBO J. 13, 2935-2947 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.
 Phylogenetic analysis suggests that each AGM species is infected with a distinct subtype of SIVagm. Interestingly, the sabaues viral genome seems to be a mosaic of many viruses infecting different primate species [1]. Virus sab-1 forms a distinct lineage in the 5' end of gag, clusters closest to the HIV2/SIVsm group in the 3' gag and 5' pol regions, and clusters closest to the other SIVagm lineages in the 3' pol, env and nef regions [1]. This feature was demonstrated for two viral sabaues strains, sab-1 and sab-D37 (sequenced only over the 5' end of the pol gene). Jin et al. conclude that an ancestor of the sabaues virus was a recombinant with two breakpoints: one roughly at codon 310 of gag and the other at 710 of pol.
 NCBI gi: 466250
 BASE COUNT 252 a 94 c 145 g 160 t

ORIGIN

```

1 tctgttattc aggggttaga attaggggat aattggaagc caagaattat aggaggaata
pol cds ->
61 ggaggctgta ttaatgttag ggcataatcat aatcaagaag ttgaataga aggaaaagga
121 tgtaaagcaa ctatttttagt gggagatact ccagtaaaca taatagggaag aaatgtacta
181 gctcaattag gagtaacttt aaacttagta gcaaaaacaa ttctccaat aaaagtgcag
241 ttaaagcctg gaaaagatgg tcccagggtta agacaatggc ctctatcaaa agaaaagatt
301 gaagccctaa gaacaatttg tgagcaatta gaaaagggaag gacatctgga aaggataggy
361 ccagaaaatc catacaacac tccagtcttt gcaattagaa agaaagataa atcacaatgg
421 aggatgttga tggacttttag ggaactcaat aaggcaactc aagactttca agaagtacaa
481 ttaggtatac cacacccggc agggttacag caaagagaac agataacagt cttagatata
541 gcagatgctt atttttcatg tcccttagat gaggattttc agaagtacac agcttttact
601 attcctagtg taaataacaa ggaaccagggt atgagatttc agtacaagggt g
pol cds ->

```

LOCUS SIVTAN17E 1113 bp ss-RNA VRL 25-MAY-1994
DEFINITION Simian immunodeficiency virus (SIVagn) partial envelope sequence, African green monkey isolate SIVagnTAN-17.
ACCESSION U04000
SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured FEMCs derived from a wild asymptomatic tantalus African green monkey captured in its natural habitat of Uganda.
REFERENCE 1 (bases 1 to 1113)
AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys
JOURNAL EMBO J. 13, 2935-2947 (1994)
COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions.
 Virus tan-17 was derived from a wild asymptomatic tantalus African green monkey. This monkey was caught in its natural habitat of Uganda in the same region as tan-49.
 NCBI gi: 466219
BASE COUNT 369 a 191 c 282 g 271 t
ORIGIN approximately 727 bp downstream from beginning of env cds
 1 accttgaggc ttagatattg tgcacctgca ggctttgcta ttttaaaatg caaggataca
 env cds ->
 61 aattatacag gctttgggat ttgtaggaat gtttcagtg taagctgcac agggctaagt
 121 aatactacag taagttcagc ttttggcttg aacggcagtc aagcggaaaa tagaacagag
 181 gtttggcaaa agcatggggt tagtaatgat tctgtaataa ttaagcttaa taagttttat
 241 aacttaacaa taaaatgtag gagaccaggt aataagacag tctaccagt aaccatcatg
 -> v3 loop equivalent
 301 gcgggtcttg tggtccactc tcaaaagtac aatacgaggc tcagacaggc ttggtgccac
 <-
 361 tttcaaggta aatggaaaga ggccctggaa gaggtgagag aaaccattgt gaaattaccc
 421 aaagaaaaat atcgaggaac aaatgataca aagaaaatat ggtaaggag acaatggggg
 481 gacccagagg ctgctaatat ttggttcaat tgccaaggag agttcttcta ttgtacacct
 541 gattggtttg taaattactt aaataatgac tcagcccgga ggagaaaaga ggatgtggaa
 601 gggaacgaat gtaaagatcc aaggaaatgt tttaaaagat cgtacgttcc ctgccacata
 661 aggtcgggtg tcaacgactg gtataccttg agtaagaaga cctatgcacc accgagggaa
 721 ggacacttgg aatgcacatc cacagtgcac tccatgatgg tgtcactaga ttacgtaaac
 781 aaaaacatga ctaatgtgac ctaaacagca aatttagaaa atatatgggc ttatgaattg
 841 gggagatata aaatcataga aatcaaacca attggatttg cgccaacaga tgttagacga
 901 tatgttgggc caactcgaga aaagaggggt cccttcgtgt taggggccct aggtttcttg
 961 ggacgagctg gagctgcaat ggggtgcaaca gcgacagagt tgacgggtca gtctcagcaa
 1021 ttacttgctg ggatattgca gcagcagaag aatctgctgg cggcagtaga gcagcagcag
 1081 caaatgctaa agctcaccat ttgggggtgt aaa
 env cds ->

SIVTANIE

LOCUS SIVTANIE 1119 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmTAN-1.

ACCESSION U03999

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild tantalus African green monkey captured in its natural habitat of Uganda.

REFERENCE 1 (bases 1 to 1119)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.

NCBI gi: 466217

BASE COUNT 382 a 187 c 280 g 270 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 accttaagat taagatactg tgcaccagca ggctttgcta tottaaaatg taaggatact
env cds ->

61 aattatacag gatttgggtgt ttgtagaat gtttcagtag ttagtgttac tggattgatg
121 aatactacag tgagctcagc atttggcata aatggcagtc aggcagaaaa cagaacagaa
181 atatggcaaa agcatggagt gagcaataat tctgtgataa taaaactaaa taaacattat
241 aagttgaaga tagtgtgtag aagaccagga aataagacag tcttaccagt aaccatcatg
-> v3 loop equivalent

301 gcaggtctcg tgttccactc acaacaatat aatacaaaat taagacaagc atgggtgcoat
<-

361 ttccagggtg attggaaaagg agcctggagg gaagtgagaa aaacaatagt ggagcttcca
421 aaagagaaat atcgagggac aaataacaca aggcagattt ggctaagtag acaatgggga
481 gatccagaag cagctaacat ttggctcaat tgccaaggag agtttttcta ttgtactcct
541 gattgggttg ttaattggct gaataatgag tctaatagtg gaagaaatgt agatgtagaa
601 ggtataaatt gcaccactgg aaaggatanaa cgctgctaca aaaggacata tgtccctcgc
661 catattaggt caattgtcaa tgactgggtac accctcagta agaaaaccta tgcaccacca
721 agggaaggac acttagaatg cacatcaaca gtgacatcta tgatgggtatc actggattat
781 aacagcaaag aaaggactaa tgtgacattg acagctaatt tggagaacat atgggcttat
841 gaattgggaa gatataagct catagaaatt gaaccaatcg gtttcgcgcc aacagagata
901 agaagatatg ttggggcaac tcgagaaaag aggggtgccct tctgtgtggg gttcctaggc
961 tttttggggg cagctggagc tgcaatgggt gcaacagcga cagcgttgac ggtccagtc
1021 cagcaattac ttgcagggat gttgcagcag cagaagaatc tgctggcggc agttgagcag
1081 cagcagcaga tgctaagct caccatttgg ggtgtgaaa
env cds ->

LOCUS SIVTAN40E 1119 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagn) partial envelope sequence, African green monkey isolate SIVagnTAN-40.

ACCESSION U04001

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild asymptomatic tantalus African green monkey captured in its natural habitat of Uganda.

REFERENCE 1 (bases 1 to 1119)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Virus tan-40 was derived from a wild asymptomatic tantalus African green monkey. This monkey was caught in its natural habitat of Uganda.

NCBI gi: 466221

BASE COUNT 383 a 194 c 268 g 274 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 acattaagat tacgatattg tgcaccagct ggctttgctc ttttaaatg cagggatact
env cds ->

61 aattatgcag gctttggtat ttgtaacaat gtttcagtag taagctgtac tggactaatg
121 aatactacaa taagttctgc ttttgggtta aacgggagcc aagcagaaaa tagaacagaa
181 atttggcaaa aacatggcaa tagtaatgat tcagtcatta taaagttgaa taaacattat
241 cagttaaaaa taaaatgtat aagacctggg aataagacgg tectaccagt aaccatcatg
-> v3 loop equivalent

301 gctggcttgg tattccactc acaaaagtac aatacaaagt taaggcaagc ttggtgtcac
<-

361 ttccaaggtg attggaaagg agcctggaag gaagtaaggg agacaatagt taaacttcca
421 aaggaaaaat accaagggac aaagaacaca acagagatat ggctaagaag acaatggggg
481 gatccagaag cagccaatat ttggttcaat tgtcaaggag aattcttcta ctgtacaccc
541 gattggatgg ttaattggct taacaatgag agtgattcag gaaaatcagt taacgtagag
601 ggaacactt gtaactcctga gaaagatagg ctatgtaaga agaggacata tgtagcctgc
661 cacatccggt cagtagttaa tgattggtac acactgagta aaaagaccta tgcaccgcca
721 agagaaggac acttagagtg tacatccaca gtgacctcaa taatagcgtc tcttgactac
781 aacaataaaa atatgactaa tgtgacctt acagcaacaa ttgaaaatct ttatgcatat
841 gagttagggaa ggtataaaact cattgaaatt aaacctattg gattcgcgcc gacagaggtc
901 agaagatatg ttggggccaac tcgagagaag aggggtgcct tcgtgttagg gttcctaggt
961 ttcttgggac gagctggagc tgcaatgggt gccacagcga cagcgttgac ggtccagtc
1021 cagcaattac ttgctgggat attgcagcag cagaagaatc tgctggcggc tgtagagcag
1081 cagcagcaaa tgctaaagct caccatttgg ggtgtgaaa
env cds ->

SIVTAN49E

LOCUS SIVTAN49E 1137 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmTAN-49.

ACCESSION U04002

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from a wild asymptomatic tantalus African green monkey captured in its natural habitat of Uganda.

REFERENCE 1 (bases 1 to 1137)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Mahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Virus tan-49 was derived from a wild asymptomatic tantalus African green monkey. This monkey was caught in its natural habitat of Uganda in the same region as tan-17.
NCBI gi: 466223

BASE COUNT 378 a 204 c 296 g 259 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 acactaaggc ttgatattg tgcaccagca ggctttgcta tcttaaaatg taaagatact
env cds ->

61 aattatacag gctatgggat atgtagaat gtatcagtag ttagctgcac aggcttaatg
121 aataccacag taagttcagc ttttggaact aatggcagtc aggcagaaaa tagaacagag
181 atctggcaaa agcataggat taacaatgat agtgtgataa ttaagcttaa taagtactac
241 aaattagggg taacatgtag gagaccaggg aataagacag tcttaccagt aaccatcatg
-> v3 loop equivalent

301 gcagggtctgg tgttccactc tcaaaaatac aatacgaagc ttagacaggc ctggtgtcac
<-

361 ttccatgggg actggaaagg agcctggaag gaagtggagg agaccattgt gaagctacca
421 agagacagat actccgggac aaataataca aatcaaatat ggctaaggag acaatgggga
481 gatccagagg cagccaatat ctgggttaac tgccaaggag agttcttcta ctgtacacca
541 gatttgggtt tcaattggct aaacaatgaa tctgcaggag gaaactacac agatgtagag
601 ggaaccctt gttccaatag cactaaaggg ggaatgacag gaagcagaaa gtgtttaaaa
661 agaacetatg ttggtgcca tatcaggtcg gttgtcaatg actggatatg cttgagtaaa
721 aagacctatg caccgccgag agagggacac ttggagtgca catccacggg gacctcaatg
781 atggtatcat tggattacaa ttcgaaaaat atgaccaatg tgaccctaac agcaaatttg
841 gaaaatatat gggcctatga gttggggaga tacaaactca tagaaatcaa gccaataggc
901 ttgcgccta cagaagtggag gagatatgtt gggccaactc gagaaaagag ggtgcccttc
961 gtgctaggat tcttaggttt ctggggacga gctggagctg caatgggtgc aacagaacgc
1021 gagttgacgg tccagtctca gcaattactt gctgggatat tgcagcagca gaagaatctg
1081 ctggcggcag ttgagcagca gcagcagatg ctaaagctca ccatttgggg tgtgaaa
env cds ->

LOCUS SIVVER1E 1116 bp ss-RNA VRL 25-MAY-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence. African green monkey isolate SIVagmGRI-2.

ACCESSION U04003

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PEMCs derived from a vervet African green monkey from East Africa.

REFERENCE 1 (bases 1 to 1116)

AUTHORS Jin, M.J., Hui, H., Robertson, D.L., Mueller, M.C., Barre-Sinoussi, F., Hirsch, V.M., Allan, J.S., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.

NCBI gi: 466225

BASE COUNT 375 a 193 c 278 g 270 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 gaattaagggc taagatactg tgcacctgca ggaatagctt tgctaaaatg caatgataga
env cds ->

61 aattatacag gctttgctcc aagatgtaga aatgttagtg tagtgcatg cacaggggtg
121 atgaatacaa cagtcagtac aggccttgcta ttaaatggta gttatacaga gaatagaaca
181 gaaatttggc aaaaacataa agtgagtaat aattcagtta ttattaggt taataagcat
241 tataatttaa caataaggtg tgtaagacca ggaataaaa cagtettacc agtgaccatc
-> v3 loop equivalent

301 atggcaggtc tgggtgtcca ctcccaaaaa tacaacacga aattaagaca agcatgggtg
<-

361 cacttcaaa gtgactggag aggagcctgg agagaaataa aacaaaaaat tgtagagcta
421 ccaaaagaaa ggtatcaggg gactaacgac acaaagcaaa tcttctaca aaggcagtg
481 ggagatccag aggcattctaa tttgtggttt aattgtcatg ggaattttt ctattgcaag
541 atggattgga tgctaaacta cattaacaac ctactgtag atgcgaatca taataattgt
601 aacagtacac agggaaaaag accagggcgg tgtgtccaaa ggacatatgt aggggtgcat
661 atcagatcag tcatcaatga ttggtacaaa ttggcaagaa agacttatgc acctccaga
721 gaaggtoatt tggagtgtac ctctacagtc acgggtatgg cagtagagct aaattatct
781 cataagaaca ggacaaatgt gactttgtca cccagatag aagccatag ggcacagaa
841 ttgggttaggt acaaattggt agaaattacg ccaattgggt ttgcaccaac cccagtaagg
901 agatatacgg ggccctcaag agaaaagagg gtgcccttcg tgctaggggt cttgggcttc
961 ttggggagctg caggcactgc aatggggagca gcggctacgg cgctgacggg ccagtctcag
1021 catttgcttg ctgggatatt gcagcagcag aagaacttgc tggcggcagt ggaggctcaa
1081 caacaaatgt tgaagctgac catatggggg gtgaaa
env cds ->

SIVVER266E

LOCUS SIVVER266E 1137 bp DNA VRL 24-AUG-1994
 DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVver266.
 ACCESSION U10896
 SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from cocultured leukocytes derived from a vervet African green monkey from Kenya.
 REFERENCE 1 (bases 1 to 1137)
 AUTHORS Jin, M.J., Rogers, J., Allan, J.S., Desrosiers, R.C., Shaw, G.M., Sharp, P.M. and Hahn, B.H.
 TITLE Infection of a yellow baboon with SIV from African green monkeys: Evidence for cross-species transmission in the wild
 JOURNAL J. Virol. (In press, 1994)
 COMMENT Two wild yellow baboons from Tanzania were identified in a previous study with strong SIV(agm) seroreactivity. They lived in a national park where yellow baboons and African green monkeys share the same habitat. In this study, molecular characterization was performed for virus derived from one of the SIVagm seroreactive yellow baboons (animal 2010) and from two vervet African green monkeys (animals ver266 and ver385). With this aim in mind, uncultured leukocyte DNA was PCR-amplified, cloned and sequenced over a 672 bp region of gag and a 906 bp region of env for animal 2010 (only the env region was analyzed for isolates ver266 and ver385). Jin et al. note that the viral strain derived from the yellow baboon (animal 2010) clusters with SIV strains of the vervet subtype, including ver266 and ver385, and believe that this was the first evidence for simian-to-simian cross-species transmission of SIV in the wild.

BASE COUNT 381 a 210 c 270 g 276 t

ORIGIN approximately 727 bp downstream from the start of env

1 gagttaaggc taagatattg tgcaccagca ggatttgccc tgctaaaatg taatgatgca

env cds ->

```

61 gattatgatg gttttaaac aaactgctcc aatgtatcag tggtagattg taccaacctg
121 atgaatacaa cagtgaactac aggggttgta ctgaatggaa gctattctga aaatcgaacc
181 caaatatggc agaacaatag gacgagcaat gactccgcac ttattttact taataagcat
241 tacaatttaa cagtgaactg taaaagacca ggaataaga cagtcctccc agtaacaata
      ->                                v3 loop equivalent
301 atggctggat tagtttttca ttctcagaag tataacttaa ggtaaggcca agcatgggtg
      <-
361 cacttcccaa gcaattggaa aggagcttgg aaggaaagtaa aggaagaaat agtaaacctg
421 cccaaagaaa gatacagagg caccaatgac ccaaaacgaa tcttcttcca aagacagtgg
481 ggggaccacg aaacagcgaa tttgtggttt aattgtcacg gagagttttt ctattgtaaa
541 atggattggt tcttaaatla tctgaataat ttaacggtag atgctgatca taatgaatgt
601 aagaatacaa gtggtacaaa aagtggtaac aagagggcac cagggccttg tgtacagaga
661 acttatgtgg cctgccatat tcatctgtc ataatatggt tggagacaat atcaaagaaa
721 acttatgttc caccaagaga gggtcacta gaatgcacat ccacggttac tggaaatgaca
781 gtggaactaa attatatccc caaaaataga accaatgtga cattaagtcc ccagatagaa
841 tctatctggg cagcagaatt ggaccgatat aaattggtag agatcacacc aattggcttt
901 gcaccacacg aagtaaggag atacacggga ggtcatgaaa gacaaaagag agtcccgctc
961 gtgctagggg tcttgggctt cttgggagct gctggaactg caatgggagc agcggcgaca
1021 gccctgacgg tccagtctca gcatttactt gctgggatac tgcagcagca gaagaattctg
1081 ctggcggtcg tggaaagctc acagcagatg ttgaagctga ccatttgggg tgtgaaa
      env cds ->

```

LOCUS SIVVER2E 1125 bp ss-RNA VRL 25-MAY-1994
 DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVagmVER-2.
 ACCESSION U04004
 SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from vervet African green monkeys from East Africa.
 REFERENCE 1 (bases 1 to 1125)
 AUTHORS Jin,M.J., Hui,H., Robertson,D.L., Mueller,M.C., Barre-Sinoussi,F., Hirsch,V.M., Allan,J.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.
 TITLE Mosaic Genome Structure of Simian Immunodeficiency Virus from West African Green Monkeys
 JOURNAL EMBO J. 13, 2935-2947 (1994)
 COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaues and tantalus viruses were derived from feral animals.
 NCBI gi: 466227

BASE COUNT 371 a 208 c 271 g 275 t

ORIGIN approximately 727 bp downstream from beginning of env cds

1 gagttaagat tgagatactg tgctccagca ggatatgctt tgttgaaatg taatgaccat
 env cds ->

61 gattataatg ggtacaaaca gaactgctct aatgtctcag tagtgactg tacaggggta
 121 atgaatacaa cagtaactac tggcttggtg ctaaattggc gctatcatga gaatcgaacc
 181 cagatctggc gaaagcatgg agtaagcaat gactcagtga taatattgat taataaggct
 241 tttaattctaa cagtcacttg cagaaggcca ggaataaaaa cagttttacc agtaacaata
 -> v3 loop equivalent

301 atggcaggat tagtggttca ttctcagaaa tataacacaa ggctgagaca agcttggtgt
 <-
 361 cacttecagg gtaattggag aggagcctgg gaggaatac ggcaggaaat aaccaaatta
 421 cctccacata aatacagggg gactaatgat acagagaaga tatttctgca aagacaattt
 481 ggagaccag aagcggcaaa cttatgggtt aattgccaaag gagaattctt ctattgtaag
 541 atggatttgt ttttaaatta tctaaataat cttacagtag atgctgacca taatagtgtc
 601 aatgggttcaa ggaagagccc tagggcaccg ggtccctgtg cccaaaggac ttatgttgct
 661 tgccatataa gatcagtcac aaatgattgg tacacaatat caaagaaaac atatgctcca
 721 ccaagagaag ggcatttaca atgcacctcc acgggttacc ggatgacagt agaattgaat
 781 tataatagtc gcaacagaac aaatgtaacg ttgtctcccc agatagaaaag tatctgggca
 841 gcagaattgg gcagatacaa attagtagaa attacaccaa ttggctttgc acccagagat
 901 gtaagaaggt atacgggagg ccatgaaaga caaaaacgag tcccgttcgt gcaagggttc
 961 ctaggcttct tgggggctgc tgggaactgc atgggagcaa cggcgacagc cctgacggtc
 1021 cagtctcagc atttacttgc tgggatactg cagcagcaaa agaatttgct ggcggctgtg
 1081 gaggtccaac agcaaatgtt gaagctgacc atttgggggtg tgaaa

env cds ->

SIVVER2L

LOCUS SIVVER2L 701 bp ss-RNA VRL 17-SEP-1994

DEFINITION Simian immunodeficiency virus (SIVagm) LTR sequence, African green monkey isolate SIVagmVER-2.

ACCESSION U04017

SOURCE Simian immunodeficiency virus PCR-amplified DNA extracted from cocultured PBMCs derived from vervet African green monkeys from East Africa.

REFERENCE 1 (bases 1 to 701)

AUTHORS Jin,M.J., Hui,H., Robertson,D.L., Muller,W.C., Barre-Sinoussi,F., Hirsch,V.M., Allan,J.S., Shaw,G.M., Sharp,P.M. and Hahn,B.H.

TITLE Mosaic genome structure of simian immunodeficiency virus from west African green monkeys

JOURNAL EMBO J. 13, 2935-2947 (1994)

COMMENT This sequence was obtained as part of a study designed to determine the breadth of genetic variation of SIV infecting African green monkeys. With this aim in mind, Jin et al. PCR-amplified, cloned and sequenced the LTR and partial envelope region from 12 isolates (ver-1, ver-2, gri-2, gri-3, sab-1, sab-2, sab-3, sab-4, tan-1, tan-17, tan40, and tan-49). These isolates were derived from different African green monkey species inhabiting various geographical regions. Most of the vervet and grivet viruses and all of the sabaeus and tantalus viruses were derived from feral animals.

NCBI gi: 466249

BASE COUNT 186 a 158 c 187 g 170 t

ORIGIN start of LTR

1 gaggaataaa atccttaact tgtacgcctt caatgagtgg ggaatcattg atgattggaa
LTR ->

61 tgcattggtct aaaggaccag gaatcagata cccaaagtgc ttgggtctct gctttaaact
121 agtgccagta gacttacatg aggaggcaaa acattgtgag agacactgct tggccatcc
181 agcacaagtg ggagaagatc cagatgggat caaccatgga gaggtcttgg tctggaagtt
241 tgatcccatg ttagcagtc aatacgaccc tacaagagaa tattttgaag acatgcattc
301 aacagttggt acaggaaact agcctaaccg cagcttgttg ttaagccgtt gcctaggaga
361 tgacattaag aactgctgac GGGACTTTCC agcaaGGGAC TTTCAGGGC GGGTCgAGGG
NFkB -> <- ->NFkB <--> Sp1 <- ->

421 CGGTACGGGG AGTGGCttta ccttcagagc tgcaTAAAAG cagatgctcg ctggtctgta
Sp1 <--> Sp1 <- signal ->

481 actccagtct ctacttagga gaccagcttg agcctgggtg ttgctggtt agcctaactg
-> TAR I <-

541 gttggccacc aggggtaagg actccttggc ttagaaagct AATAAAactt gctgcatta
signal ->

601 gagcttatct gactcaagt ccttcattga cgcctcactc gagggacggg tctgttctct
661 tactgggttc tctctcaacc caggcgagag aaactccagc a
-> LTR

LOCUS SIVVER385E 1125 bp DNA VRL 24-AUG-1994

DEFINITION Simian immunodeficiency virus (SIVagm) partial envelope sequence, African green monkey isolate SIVver385.

ACCESSION U10898

SOURCE Simian immunodeficiency virus from PCR-amplified DNA extracted from cocultured leukocytes derived from a vervet African green monkey from Kenya.

REFERENCE 1 (bases 1 to 1125)

AUTHORS Jin, M.J., Rogers, J., Phillips-Conroy, J.E., Allan, J.S., Desrosiers, R.C., Shaw, G.M., Sharp, P.M. and Hahn, B.H.

TITLE Infection of a yellow baboon with SIV from African green monkeys: evidence for cross-species transmission in the wild

JOURNAL J. Virol. (In press, 1994)

COMMENT Two wild yellow baboons from Tanzania were identified in a previous study with strong SIV(agm) seroreactivity. They lived in a national park where yellow baboons and African green monkeys share the same habitat. In this study, molecular characterization was performed for virus derived from one of the SIVagm seroreactive yellow baboons (animal 2010) and from two vervet African green monkeys (animals ver266 and ver385). With this aim in mind, uncultured leukocyte DNA was PCR-amplified, cloned and sequenced over a 672 bp region of gag and a 906 bp region of env for animal 2010 (only the env region was analyzed for isolates ver266 and ver385). Jin et al. note that the viral strain derived from the yellow baboon (animal 2010) clusters with SIV strains of the vervet subtype, including ver266 and ver385, and believe that this is the first evidence for simian-to-simian cross-species transmission of SIV in the wild.

BASE COUNT 376 a 210 c 272 g 267 t

ORIGIN approximately 727 bp downstream from the start of env

1 gagctaagat tgagatactg tgctccagca gggttcgccc tattaaaatg taatgactat
env cds ->

61 gactatgatg ggtttaaaca aaattgttcc aatgtatcag tagtgcatg tacagggcta
121 atgaatacaa cagtaactac aggggtgctg ttgaatggaa gctattctga gaatcgaacc
181 cagatctggc aaaagcatag agtaagtaat gattcagtgt taatacagat taataaacat
241 tacaacctca caattacttg caaaaggcca ggaacaaaa cagtcttacc tgtgacaata
-> v3 loop equivalent

301 atggcaggac tagtttttca ctctcaaaag tataacacga gattgagaca agcctgggtg
<-

361 cacttccaag gcaattggca tggagcttgg aaggaaataa aacaagaaat agttaagtta
421 cccaaagaca ggtacagggg gaccaacgac acagaacaaa tttatttgac aaggcagtg
481 ggagatccag aggcagcaaa tttgtggttc aattgtcaag gagagttctt ttattccaaa
541 atggattggt tcttgaacta ttttaataat aagacatgga atgcagacca taatcaatgt
601 gataatagta ctaagactgg aagggaacca ggacottgtg tacaaaggac ttatgttgc
661 tgccacatca gatctgtcat aaatgattgg tatacaatat caaagaaaac atacgcccac
721 ccaagagaag ggcatttaca atgcacatcc acggtaactg ggatgacagt ggagctaaat
781 tatatcaatg acaatagaac caatgtgaca ctgagtcccc agatagaaac catctgggca
841 ggggaattgg gcagatataa attggtagag atcactccaa ttggttttgc acccacagaa
901 gtgagaagat atacgggagg ccatgaaagc ccaaaacgaa tcccgctcgt gctaggggtc
961 ttgggtctct tgggagctgc tggaaactga atgggagcag cggcgacagc cctgacggtc
1021 cagctctcagc atttgcctgc tgggatactg cagcagcaga agaattctgt ggcggctgtg
1081 gaggtctaac agcagatggt gaagctgacc atttggggtg tgaaa
env cds ->

II

Amino Acid Alignments

The amino acid sequence alignments that follow were generated by the PIMA program ("pattern induced multiple alignment"; R.F.Smith and T.F.Smith, *PNAS* 87:118, 1990 and *Protein Engineering* 5:35, 1992). The alignments were also subject to some hand-editing performed with MASE (D.V. Faulkner and J. Jurka, *TIBS* 13:321, 1988). For information concerning both these UNIX tools, contact Dr. Randall Smith, Institute of Molecular Genetics, Baylor College of Medicine, Houston, Texas, 713-798-4735.

With few exceptions (most notably with HIV-1 *env*), only full-length protein sequences have been included in this compilation. Sequences not ready for publication at this time, but which are reflected in the consensus sequences and patterns will usually be identified. Eight sequence subtypes have been identified for HIV-1s (see Part I and Part III). Five sequence subtypes exist for HIV-2s. Sequence subtypes have been defined by "cladistic" criteria applied to nucleotide sequences; it remains to be seen whether the amino acid sequence subtypes inferred from that classification are valid, or whether a "phenetic" classification would be preferable.

The reference sequences for the alignments are mixed case consensus sequences in which upper case letters refer to amino acid residues which are conserved 100% and lower case letters represent amino acid residues conserved in at least 50% of the sequences. Again, consensus sequences have been generated for each of the defined subtypes (see Parts I and III). The user should keep in mind that these subtypes have been "cladistically" defined, not "phenetically" defined (the number of phenotypes remains to be discovered).

Within the sequences, "."s are used to indicate residues conserved with respect to the reference sequence and "-"s represent actual gaps. Blank spaces within the alignment indicate lack of sequence information over that region. Annotation of the *env* amino acid sequences utilizes "*" for conserved cysteine residues and "~" for potential N-linked glycosylation sites. A summary of *env* cysteine and N-linked glycosylation sites is found in Part III.

PART II Amino Acid Sequences and their Variation

Introduction	II-1
Contents(1994)	II-2
A. HIV-1 Alignments and Consensus Sequences	
GAG	II-A-2
POL	II-A-18
VIF	II-A-26
VPR	II-A-28
TAT	II-A-29
REV	II-A-31
VPU	II-A-32
ENV	II-A-34
NEF	II-A-64
B. HIV-2/SIV Alignments and Consensus Sequences	
GAG	II-B-1
POL	II-B-5
VIF	II-B-12
VPX	II-B-13
VPR	II-B-14
TAT	II-B-15
REV	II-B-16
ENV	II-B-18
NEF	II-B-31
C. SIVAGM and SIVSYK Alignments and Consensus Sequences	
GAG	II-C-1
POL	II-C-4
VIF	II-C-9
VPX	II-C-10
TAT	II-C-11
REV	II-C-12
ENV	II-C-13
NEF	II-C-21

This page intentionally left blank.

HIV1 GAG

CONSENSUS.A	MGARASvLsggkLDawekIRLRPgGkKkYrIKHlvwAsreLerFaLnPsILeT?egcqqineQLqsalkTy?e	71
HIVU455	-----K--S-----N-----K-T--G--A-----LG--P--Q--T-	73
HIVUG266	---V---E-----E---K-----K-----T---R-L---P--Q--A-	71
HIVMAL	-----G--G-----T---S-	73
HIVVI32	-----M-----D-----K-----E---T-	73
HIVVI57	-----T-----L---P--T-	73
HIVVI59	I-----E---X-----Q---X-K-----T---I---P--T-	73
HIVVI310	-----K-----Q-----T---I---P--T-	73
HIVVI354	-----I-----KR-----QI--I---K--K---G--K-A-----S-	73
HIVVI415	-----R--M-----D---G--A-----L---P--T-	73
HIVK7	I-----Q--RX-----T-R-----T-	73
HIVK29	-----M-----T-----I---P--T-	73
HIVK88	-----M-----D-----A-----T-	73
HIVK89	-----T-----K-----T-----P--T-	73
HIVK98	---T-----N-R-K--I-----G--T-----K---Q--T-	73
HIVK112	-----I-----K-----T-----T-	73
HIVK124	-----T-----S---T-	73
HIVCI4	-----N--I--G-----T-----A-----T-	73
HIVCI20	-----AG---L---AH-R'-S-	73
HIVCI32	-----R-----K-----S---LI--F--T-G--S-	73
HIVCI51	---I---S-----R-Q-----A---I-----S-	73
HIVCI59	-----G--AD---L---R--S-	73
HIVG141	-----D---Q---T-R-K--I-----D---S---K--LG--P--S-	73
HIVTN238	-----T-----R--KI-----G---V-----I---T--S-	73
HIVTN243	-----R--M-----G--A---I---T--S-	73
HIVTN240	-----R-----F--A---I---T--S-	73
HIVTN245	-----R--YM-----G--A--R--IR--T--S-	73
HIVLBV105	-----T-----M-----G--A---L---PSV--S-	73
HIVLBV23	-----R-----M-----G--A---G---T--Q--T-	73
HIVLBV2310	-----S-----G--A---L---R--S-	73
HIVIC144	-----GF--A---L---T--S-	73
HIVDJ258	-----S-----G--A---L---G--ESK	73
CONSENSUS.B	MGARASVLsgGeLDwEkIRLRPgGKkKkYrIKHlvwASReLERFAvNpGLLEtseGCRqIlgQLQpSLqTGsE	73
HIVSF2	-----K-----	73
HIVB2167	-----R--V-----AA-----	73
HIVB2190	-----K--K-----R-Q-----A-----	73
HIVB2200	-----K-----A-----K-I-----	73
HIVPH136	-----R-----G-----E---A-----	73
HIVPH153	-----K-----Q-----H-A--T-	73
HIVTB132	-----Q--K-----R-----	73
HIVLAI	-----	73
HIVHXB2R	-----	73
HIVBH102	-----	73
HIVPV22	-----	73
HIVMN	-----N-----V-----	73
HIVJH3	-----S-----	73
HIVJRC5F	-----R-----S-----K-----	73
HIVQYI	-----K-----Q-----I-----K-----	73
HIVNY5CG	-----K-----Q-R-----R-----	73
HIVNL43	-----K-----Q-----	73
HIVCDC4	-----Q-R-----K-----K-----	73
HIVHAN	-----K-----Q-----N-----	73
HIVCAM1	-----K-----	73
HIVRF	-----K--K-----R--R-----S--A-----A-----	73
HIVD31	-----R-----	73
HIVJRFL	-----K--K-----R-----S-----L-A-K---	73
HIVUG280	-----A-----K-----Q-R-----D-----A-----	73
HIVYU2	-----A-----K-----Q-R-----D-----	73
HIVBCSG3C	-----K--K-----L-----A-----	73

CONSENSUS.C	MGARASILRGgKLD7WEKIrLrPGGkKhYmIKHLVWASRELERFALnpgLLETseGCKqIikqLqPalQTGT	72
HIVUG268	-----V-----T-----K-----C-----M-----G-----S-----M-----D	73
HIVSM145	-----A-----R-----C-----M-----G-----S-----M-----D	73
HIVZAM10	-----K-----T-----C-----M-----G-----S-----M-----D	73
HIVZAM19	-----X-----A-----C-----M-----G-----S-----M-----D	73
HIVZAM20	-----K-----T-----C-----M-----G-----S-----M-----D	73
HIVDJ259	-----E-----A-----K-----R-----R-----L-----K-----S-----R-----R-----	73
HIVVT313	-----T-----R-----K-----K-----M-----D-----Q-----R-----K-----K-----	73
CONSENSUS.D	MGARASVLSGGKLD7WEKIrLrPGGgKKKY?LKHVWASRELERfaiNPGLEETseGCKQIigQLQpaitGse	71
HIVELI	-----K-----R-----Y-----T-----	73
HIVNDK	-----T-----R-----A-----LI-----T-----S-----	73
HIVZ226	-----A-----R-----L-----D-----R-----	73
HIVVI203	-----T-----S-----R-----E-----S-----T-----	73
HIVVI205	-----Q-----KS-----R-----K-----R-----E-----K-----	73
HIVG109	-----Q-----A-----R-----Q-----A-----E-----H-----ML-----S-----	73
HIVK31	-----E-----Q-----H-----K-----I-----P-----M-----N-----	73
HIVUG274	-----E-----N-----K-----I-----M-----LK-----	73
HIVUG270	-----E-----N-----K-----L-----I-----M-----S-----	73
HIVSB365	-----Q-----R-----K-----R-----	73
CONSENSUS.F	MGARASVLSGGKLDaWEKIrLrPGGgKKKYr?KHLVWASRELERFA??PgLLETseBGCrkIigQLQpSLQTGSE	70
HIVVI174	-----Q-----M-----ID-----P-----T-----	73
HIVVI169	-----R-----KM-----I-----LD-----	73
HIVBZ162	-----L-----IN-----Q-----	73
HIVVI325	-----T-----S-----L-----LN-----S-----T-----Q-----R-----	73
CONSENSUS.G	MGARASVLSGGKLD7WEKIRLrPGGKK?YR?KHLVWASRELE?FALNPGLEETTEGCGQI??QLQP?LQTGT	66
HIVVI191	-----A-----Q-----I-----K-----MR-----A-----	73
HIVLBV217	-----E-----R-----M-----R-----LQ-----S-----	73
CONSENSUS.H	MGARASVLSGG?LDAWEKIRLrPGGKK?YRLKHLVWASRELERFALNP?LL?T?EGCLQ?IEQLQPALKTGT	67
HIVVI525	-----R-----Q-----G-----E-----S-----I-----	73
HIVVI557	-----K-----K-----D-----D-----A-----L-----	73
CONSENSUS.O	MGATASVLTGSKLDAWE?IRL?PGSKK?YRLKHLVWASRELER?ACNP?LLETABG?E?LLQOLEPALKTGS?	64
HIVANT70	-----S-----Q-----K-----K-----F-----E-----N-----K-----D-----	73
HIVMVP51	-----R-----R-----R-----A-----Y-----G-----T-----Q-----E-----	73
STVCPZGAB	MGARASVLTGSKLDRWEKVRlRPGGRKRYmIKHLVWASRELERFACDPLMESKDECTKLLQOLEPALKTGS	73

HIV GAG

CONSENSUS A	ElkSLfNtvaatLyCvHqrIdvKDeKeeA	LdkLEel onkskqk	eqq	aaA???	Tg	120
HIVU455	R--Y---V-----	-N- M --N- R	--	- N	-	123
HIVUG266	R -Y- -V- - - -EI----	- - - M - - - - -	-	- N	-	121
HIVKAL	I -Y-----	- - - - -R--	-	---AQQAAAA	K	129
HIVVI32	I Y --V-F-- - - - - -	- - - - - K - - -	A -	---D	-	123
HIVVI57	---A -V-----E --D-	- - - - - K - - -	-	---D	-	123
HIVVI59	R --A -V- - - - -I- - -	- - - - - M - - - - -	-	- N	-	123
HIVVI310	R ---I-V- - - -E-----	- - - - - - - - -	K -	-E-D	-	122
HIVVI354	- - - - - K N - - - - -	- - - - - - - - -	-	-A	E	123
HIVVI415	I---Y- - - - - - - - -	- - - - - K - N - - -	-	- G	-	123
HIVK7	R -Y----- , -RQ-E-----	- - - - - - - - -	-	---D	-	123
HIVK29	R - - - - - - - - -E---A-	- - - - - L - - - - -	-	---A	-	123
HIVK88	R-----	- - - - - - - - -	-	- D	-	123
HIVK89	R----- -W---Q-N- - -	- - - - - - - - -N- -	-	---D	-	123
HIVK98	R--V-----	- - - - - - - - -	E--	---D	-	123
HIVK112	-R-----	- - - - - - - - -	-	- D	E	123
HIVK124	-IR --A-V-----E-----	- - - - - M - - - - -	-	---D	-	123
HIVCI4	-R--Y-----	- - - - - - - - -	-	---D	-	121
HIVCI20	- - - - -I---W- - - -I-----	- - - - - V- - - -	-	T--A	-	118
HIVCI32	-R--Y -I---W---R -EI-----	- - - - - - - - -	-	---AQAAAA	-	128
HIVCI51	--R - - - - -W-----I-----	- - - - - - - - -	K--	---A	-	121
HIVCI59	-- - - -AI-V-W-----I-----	--L--V-----	-	---A	-	123
HIVG141	- - - Y-----A--EI-----	- - - - - - - - -	-	---A	-	119
HIVTN238	- - - - -W-----E-----	- - - - - V ---Q--	-	---G	-	123
HIVTN243	- - - - -W-----E-----	- - - - - V ---Q--	-	---G	-	123
HIVTN240	- - - - -W-----E-----	- - - - - V ---Q--	-	---G	-	123
HIVTN245	- - - - -V--W-----E-----	- - - - - V ---Q--	-	---G	-	123
HIVLBV105	- R-----V--L-----E-R-----	I-----	-	---D	-	123
HIVLBV23	- - -H---V-----	-R--- -K---HQ---	-	- D	-	123
HIVLBV2310	- - - - -I---W-----	- - -L-----	-	---A	-	123
HIVIC144	-V-----I-----I-----IR-----	- - - - - - - - -	-	-V	-	123
HIVDJ258	---Y--I-V-W-----I-----	---AV-----	-	---A	-	123
CONSENSUS B	ElrSLyNtvaatLYCvHQRtevkDTKEA	LdkLEEE onkskKK	eqq	??aAad	??Tg	123
HIVSF2	- - - - -D-----	-E-----	-	AA --G	-	125
HIVB2167	- - - - -K-D-R-----	-E-V-- - - - -E-	-	A--	A-	124
HIVB2190	--K--F A--V-----Q-----	- - - - - I-----	T -	-	-	123
HIVB2200	- - - - -V---A--K-D-----	- - - - - D-----	-	-	-	123
HIVPH136	-K--H-----K-D-----	-E-----	-	-	-	123
HIVPH153	- - - - -N-----	-E-- - - - -N--	-	-	-	123
HIVTB132	- - - - -I-V-----K-----	-E-- - - - -	-	---N	-E	123
HIVLAI	- - - - -I-----	- - - - - - - - -	-	-	-	123
HIVXB2R	- - - - -I-----	- - - - - - - - -	-	-	-	123
HIVBH102	- - - - -I-----	- - - - - - - - -	-	-	-	123
HIVPV22	- - - - -I-----	- - - - - - - - -	-	-	-	123
HIVMN	--K-----K--I-----	-E-----	-	-	TGMR-	126
HIVJH3	--K--F-----	-E-----	-	-	-	123
HIVJRCSE	-T-----V-----I-----	-E-- - - - -T--M--	-	-	-	123
HIVQYI	-I-----V-----K-----	- - - - - - - - -	-	T--	-	123
HIVNY5CG	-R--F---V-----D-----	- - - - - - - - -	-	-	-	123
HIVNL43	- - - - -I-V-----D-----	- - - - - - - - -	-	-	-	123
HIVCDC4	- - - - -R-----	- - - - - - - - -	-	-	-	123
HIVKAN	- - - - -K-----	- - -V-- - - -N--	-	E--	A-	123
HIVCAM1	- - - - -K-D-----	-E-----	-	---G	-	123
HIVRF	--K---A-----N--R-----	- - - - - - - - -	-	-	-	123
HIVD31	---F-----	- - - - - - - - -	-P	-R-	A-	123
HIVJRFL	- - - - - - - - -	-E-- - - - -A--	E	-	-	121
HIVUG280	---P-----D-----	- - - - - - - - -	E--	---G	-	123
HIVYU2	- - - - -K-----	-E-----	-	-	-	123
HIVBCSG3C	--K-----I-V-----M-D-----	-E-- - - - -N--	-	-	-	123

HIV1 GAG

CONSENSUS.C	ELYSL?NTvAtLyCVH??IevrDTKEA..LDKieEE.QNK?QOK??TqQ...ak?AD.....?G	117
HIVGQ268	----Y-----KG-----,-----I-----,-----ET-----,-----K-	121
HIVSM145	----F-----EK-----,-----S-----,-----A-----,-----	123
HIVZAM18	--K--F--V--W--ED-T-----,--RL-----S--TK-----,--A-----,-----	125
HIVZAM19	--K--H--A--V-----XX-T-----,--R-----,--C-----,--E-----,-----	123
HIVZAM20	----Y-----AG-----,--E-----,--C-----,--T-----,-----	123
HIVDJ259	----Y-----AR--IQ-----,-----C-----,--K--,--ET-----,-----K-	124
HIVVI313	----I--H-----EK--I-----,-----K-----,--S-----,--AA-----,-----H-	124
CONSENSUS.D	elkSL?NTvATLYCVH?rI?vkdtKEA..leRMEEE.qnKSKkK.?aQQ???atAD.....tr	130
HIVELI	--R--Y-----KG-D-----,-----,-----,-----,--A-----,-----G	123
HIVMDK	--IR--Y-----,-----V-----,-----,-----T-----,-----,--A-----,-----	121
HIV2226	--R--F-----,-----,-----,-----N--K-----,-----,--A-----,-----AG	124
HIVVI203	----F-----Q-----,--D-L-----,-----,-----,-----,-----,-----	123
HIVVI205	--I--F-----K-----,--L-----,-----R-----,-----,--A-----,-----,--K	124
HIVG109	D--Y-----N-----,-----,-----,-----,--AQQ-----,-----,-----	126
HIVK31	--I--Y-----AG-K-TN-----,--D-I-----,--T-----,-----,--E-----,-----	123
HIVG274	--I--F-----K-T-----,--D-I-----,--T-----,-----,-----,-----,-----	123
HIVG270	--V--F-----K-A-----,--D-I-----,--T-----,-----,-----,-----,-----	123
HIVSE365	--R--Y--I-----K-----,--V-----,--R-----,--V-----,-----,-----,--SG	123
CONSENSUS.F	ELrSL?NTv?vLYfVHQrveE?kdtKEA..L?KLEEE.QNKSqQK...tqQ...?aAd.....KG	115
HIVVI174	----F--IV--Y-----I-----,--E-----,-----,-----,-----,-----,-----	122
HIVVI69	----Y--A-----K--V-----,--E-----,-----,--K-----,--V-----,-----,-----	123
HIVB2162	----Y--A-----V-----,--D-----,-----,--K-----,--A-----,-----,-----	123
HIVVI325	--K--F--VT--C----I--IR-----,--D-----,-----K-----,--A-----,--ET-N-----,-----	123
CONSENSUS.G	E?KSL?N??A?L?C?HQRI?VKDTKEA..LeEVERa.QknSQOK...tQq...????.....?kG	104
HIVVI191	-I---Y-TT-T-Y-V---EI---Q---,--Q-----,--K-----,--T-----,-----,--E-----,--E-	120
HIVLBV217	-L---F-AV-V-W-I---G-----,-----R-----,-----I-----,--A---D-----,-----,-----	121
HIVTAIG	-----,-----,--R-----,-----,--AAS-----,-----,--E--	30
CONSENSUS.H	ELQSLFNLLA?LYCVHQRID?KDTKEA..L?K?E?.QN?Q??..TQQ...AT?D.....K?	105
HIVVI525	-----V-----I-----,--S-IE-V---NR-QK-----,-----A-----,-----,--E	123
HIVVI557	-----T-----V-----,--E-LK-A---KS-NR-----,-----,--G-----,-----G	123
CONSENSUS.O	?L?SLNNAI?VLMCVHNR??I?DTQQA..IQKLKEVM.?SRKSA..?AA...KE?.....TS	104
HIVANT70	S-Q-----V-----YK-Q-----,-----G-----,--D-----,--D-----,-----,-----	122
HIVMVP51	D-K-----A-----FD-R-----,-----A-----,--E-----,--E-----,-----,-----	122
SIVCP2GAB	GLRSLPMTLAVLWCINSDITVEDTQKA..LEQLKR..HHG.....EQQ..SKTESNSG.....SR	122

HIV GAG

	p17 \ / p24	
CONSENSUS.A	nss?????...kvsqNYPIVQNaGgQm?hQ7ISPrTLnAwVKviEekaPepEVIpMfSaLSEGATpQdiNmM	182
HIVU455	S.....PV--A-----V-D-----	183
HIVUG266	--N.....T-R--TY-A-----V-----S-----	184
HIVMAL	--S.....I-AI-----	193
HIVVI32	--G.....I-S-----T-----	187
HIVVI57	--I.....AV-----	187
HIVVI59	--N.....IV-A--K--X--V-----E-----	187
HIVVI310	--H.....V-S-----R-----	186
HIVVI354	--V.....AM-----G-----N--I--	187
HIVVI415	--N.....I-AI-----	187
HIVK7	--R--I--N-----X-----L-----V-----	187
HIVK29	--T-----S-----	187
HIVK88	--I--T-----	187
HIVK89	--N-----IL-T-----	187
HIVK98	--I--N-----V-----	187
HIVK112	S-----I--N-----	187
HIVK124	--I-----S-----	187
HIVCI4	--R.....Q-----V-SI-----	185
HIVCI20	--T-----SM-----T-----	179
HIVCI32	--N-----L--V-A-----T-----T-----	189
HIVCI51	--T-----SI-----A--R-----	182
HIVCI59	--G-----T--PM-----T-----	184
HIVG141	SR-----ST-----V-AI-----	183
HIVTN238	S-----V-P-----V-K-G-N-----	187
HIVTN243	S-----V-P-----V--G-N-----	187
HIVTN240	S-----A--P-----V--G-N-----	187
HIVTN245	S-----A--P-----V--G-N-----	187
HIVLBV105	--NKGSSQ...N-----V--V--PV--D-----G-N-----	191
HIVLBV23	--G-----I--A-----A-----T-----	187
HIVLBV2310	--T--PM-----	184
HIVIC144	--T--PI-----	184
HIVDJ258	--V--PM-----T-----	184
CONSENSUS.B	nss.....qVSqNYPIVQnIqGQMvHQaiSPRTLNAWVKVvEBKAFSPPEVIMPfSaLSEGATpQdiNTM	187
HIVSF2	--V-----I-----	189
HIVBZ167	T-----M-----I-----	188
HIVBZ190	--N-----A-----	187
HIVBZ200	--K-----	187
HIVPH136	--K-----K-----L-----	187
HIVPH153	SC-----I-----	187
HIVTB132	--M-----	187
HIVLAI	H-----I-----	187
HIVHKB2R	K-N-----I-----	187
HIVBH102	H-----I-----	187
HIVPV22	H-----I-----	187
HIVMN	--I-E-----	190
HIVJH3	--K-----I-----	187
HIVJRCSE	--I-----	187
HIVQYI	--P-----A-----	187
HIVNY5CG	--P-----	187
HIVNL43	--N-----	187
HIVCDC4	--I-----A-----	187
HIVHAN	--RN-----	187
HIVCAN1	--P-----	187
HIVRF	--G-----	187
HIVD31	--N-----P-----	187
HIVJRF1	--M-----	185
HIVQG280	--I-----	187
HIVYU2	--I-----	187
HIVBCSG3C	--R-----I-----	187

p17 \ / p24		
CONSENSUS.C	KVSONYPIVQNLQGQMVHQaISPRTLNAWVKVIEEKaISpe?ipmftalsegatpqqdintm	177
HIVUG268	-L-----V-----	182
HIVSM145	-I-----	184
HIVZAM18	D-----I-----G-----I-----	186
HIVZAM19	-X-----	183
HIVZAM20	-F-----L-QRI-----	184
HIVDJ259	-F-----PL-----V-----	185
HIVVI313	-PM-----V-----	185
CONSENSUS.D	nsa.....QVSONYPIVQNLQGQMVHQaISPRTLNAWVKVIEEKaFapeviPMFsaLSEGATPQDLNcM	184
HIVELI	-N-----	187
HIVNDK	-----	184
HIVZ226	-N-----	188
HIVVI203	-L-----N-----	187
HIVVI205	-E-----T-----	188
HIVG109	-N-----L-----N-----A-----	190
HIVK31	-VM-----	187
HIVUG274	-----T-----	187
HIVUG270	-----A-----	187
HIVSE365	-N-----L-----D-----T-----	187
CONSENSUS.F	VSONYPIVQNLQGQMVHQaISPRTLNAWVKVIEEKaFSPEVIMFsaLSEGATPQDLNTM	175
HIVVI174	-----	182
HIVVI69	-----	183
HIVB2162	-S-----	183
HIVVI325	-M-----PL-----AD-----	183
CONSENSUS.G	nsS.....QVSONYPIVQnaQGQmvHQpISPRTLNAWVKVIEEKaFSPEVIMFsaLSEGATPQDLNcM	168
HIVVI191	-LT-----I-----N-----	184
HIVLBV217	DN-----	185
HIVTAIG	-XI-----PI-----A-----I-----	93
CONSENSUS.H	???.....K?SONYPIVQNAQGQ?VHQaISPRTLNAWVKVIEEKaFSPEVIMFsaLSEGATPQDLNAM	164
HIVVI525	KDK.....-V-----M-----	187
HIVVI557	NGN.....-I-----P-----	187
CONSENSUS.O	?R.....Q??QNYPIV?NAQGQMVHQaISPRTLNAWVKVIEEKaFNPEIIMFsaLSEGA??YDINTM	161
HIVANT70	.A.....-AG-----S-----IS-----	185
HIVMVP51	.P.....-TS-----T-----VP-----	185
STVCP2GAB	BGGASQGASASAGISQNYPIVQNAQGQMVHQaISPRTLNAWVKVIEEKaFSPEVIMFsaLSEGALPQDVNTM	195

HIV1 GAG

CONSENSUS .A	LNIVGQGHQAAMQMLKdtInEEAAewDR?HPVhAgPippgQmREPrGSDIAGtTStIqEqIgwT...sNPPiP	251
HIVU455	--V-----L-----V-----G-----	253
HIVUG266	--V-E-----L--Q-R-----V--VR-----	252
HIVMAL	-----D--V-----	263
HIVVI32	A-----G--L-----P-----	257
HIVVI57	-----L-----P-----	257
HIVVI59	--V-----X--L--R-----K-----RX--G-----	257
HIVVI310	-----V--Q--V-----V-----	256
HIVVI354	---K-----I--Q--S-I-----T-----G-----	257
HIVVI415	-----L-----T--A--G--N-----	257
HIVK7	-----X--L-----P--LQ-----	257
HIVK29	-----L-----P-----G-----	257
HIVK88	-----L-----P-----	257
HIVK89	-----L-----P--L-----	257
HIVK98	-----L-----P--LQ-----	257
HIVK112	-----L-----P-----	257
HIVK124	-----N-----L-----P-----	257
HIVCI14	-----G--V-----L-----V-----	255
HIVCI20	-----V-----V-----	249
HIVCI32	--T-----V-----A-----A--N-----	269
HIVCI51	-----T-----V-----T-----	252
HIVCI59	-----V--Q-----N-----	254
HIVG141	-----D--T--A-----	253
HIVTM238	-----E--V-----N-----	257
HIVTM243	-----E--V-----N-----	257
HIVTM240	-----E--V-----N-----	257
HIVTM245	-----E--V-----N-----	257
HIVLBV105	-----T--R-----I--V-----	261
HIVLBV23	-----V--Q--Y--L-----A--Q-----	257
HIVLBV2310	-----V-----N-----	254
HIVIC144	-----V-----N--S-----	253
HIVDJ258	-----V-----	254
CONSENSUS .B	LNTVGGHQAAMQMLKtINEEAAEwDRIHHPvhAGPiaPGQmREPRGSDIAGTTSTLQEQIGwT...nNPPiP	257
HIVSF2	-----V-----	259
HIVBZ167	-----D-----	258
HIVBZ190	-----	257
HIVBZ200	-----X-----	257
HIVPH136	-----V-----	257
HIVPH153	-----	257
HIVTB132	-----I--H-----	257
HIVLAI	-----V-----	257
HIVHXB2R	-----V-----	257
HIVBH102	-----V-----	257
HIVPV22	-----V-----	257
HIVMD	-----T-----	260
HIVJH3	-----AQ--T-----S-----	257
HIVJRCSE	-----	257
HIVQYI	-----	257
HIVNYSCG	-----H-----	257
HIVNL43	-----H-----	257
HIVCDC4	-----T-----	257
HIVHAN	-----	257
HIVCAM1	-----	257
HIVRP	-----	257
HIVD31	-----	257
HIVJRFL	-----	255
HIVUG280	-----	257
HIVYU2	-----	257
HIVBCSG3C	-----Q--V-----	257

HIV1 GAG

CONSENSUS.C	lntvgghqaanqmlkdtineaaaewdrllhpvhagpvapggmRePRGSDIAGTTSTLQEQIaWMT...?NPPiP	246
HIVUG268	-----N-----S-----	252
HIVSM145	-----Q-----D-----	254
HIV2AM18	-----N-----L-----N-----V-	256
HIV2AM19	-----N-----V-	193
HIV2AM20	-----I-----A-----V-	254
HIVDJ259	-----I-----N-----S-----	255
HIVVI313	-----Q-----I-----N-----	255
CONSENSUS.D	LMTvGGHQAAMQMLKETInaEAAEWDRlHPvHAGPIAPGQMREPRGSDIAGTTSTLQEQi?WMT...sNPPiP	253
HIVELI	-----A-----	257
HIVNOK	-----D-----V-----A-----	254
HIV2226	-----A-----	258
HIVVI203	-----G-----	257
HIVVI205	-----H-----I-----A-----	258
HIVG109	-----A-----A-----A-----	260
HIVK31	-----A-----G-----N-----	257
HIVUG274	-----V-----G-----	257
HIVUG270	-----I-----VG-----	257
HIVSE365	-----G-----N-----	257
CONSENSUS.F	LNTVGGHQAAMQMLKDTINEEAAEWDRlHPv?AGPIPPGQ?REPRGSDIAGTTSTLQEQIqWMT...sNPPvP	243
HIVVI174	-----Q-----I-----N-----	252
HIVVI69	-----H-----N-----M-----I-	253
HIVB2162	-----AQ-----I-----	253
HIVVI325	-----H-----M-----T-----	253
CONSENSUS.G	lntVGGHQAAMQMLKDTINEEAAEWDRlHPvQAGPIpPGQIRdPRGSDIAGTTSTLQEQIRWMT...SNPPiP	238
HIVVI191	-----E-----	254
HIVLBV217	-----Q-----	255
HIVTAIG	S-----S-----A-----	163
CONSENSUS.H	LN?VGGHQAAMQ?LkDTINEEAAEWDR?HPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIaWMT...?NP?iP	229
HIVVI525	--T-----I-----L-----G--A--	257
HIVVI557	--I-----M-----V-----S--P--	257
CONSENSUS.O	LNAIGGHQqALQVLKEVINEEA?EWDRTHPP??GPLPPGQIREPTGSDIAGTTSTQEQDI?WTT...R????iP	224
HIVANT70	-----V-----PV-----H-----FNQP--	256
HIVNVP51	-----A-----AM-----I-----GANS--	256
SIVCP2GAB	LNAVGGHQAAMQVLKEVINEEAAEWDRlHPvTHAGPIAPGQIREPRGSDIAGTTSTLQEQIGWTT...ANPPiP	265

HIV GAG

CONSENSUS.A	VGdIYkrwIiLGLNKIVRMYSFvSILDirQgPKEPPrdYVdrFfKtLRaeqAtQeVKnwMTETLLvQNANPDC	324
HIVU455	-----R-----D-----	326
HIVUG266	--E--R-----E--KE--K-----D-----	323
HIVMAL	-----	336
HIVVI32	-----K-----G-----	330
HIVVI57	-----K-----V-----S-D-----	330
HIVVI59	-----R-----V-R--K-----I-----	330
HIVVI310	--E-----A-----	329
HIVVI354	--K--KS-----E--N--K-----S-----	328
HIVVI415	-----T--VK-----I-----G-----	330
HIVK7	-----R-----A-----S-D--GX-----	330
HIVK29	-----K-----G-----I-----	330
HIVK88	-----VK-----G-----	330
HIVK89	-----K-----D-----	330
HIVK98	-----K-----A-----G-----	330
HIVK112	-----K-----G-----	330
HIVK124	-----K-----G-----	330
HIVCI4	--E--R-----C-----I-----	328
HIVCI20	--E--V-----	322
HIVCI32	--E-----Y-----S-----	322
HIVCI51	--E--K-----	325
HIVCI59	--E--V-----X-----	327
HIVG141	--E-----	326
HIVTN238	-----I-----Y-----	330
HIVTN243	-----I-----Y-----	330
HIVTN240	-----Y-----	330
HIVTN245	-----Y-----	330
HIVLBV105	-----K-----D-----D-----	334
HIVLBV23	---RK-----D-----	330
HIVLBV2310	--E--V-----V-----	327
HIVIC144	--E--V-----S-----	326
HIVDJ258	-----	327
CONSENSUS.B	VGeIYKRwIiLGLNKIVRMYSFtSILDirQGPKEpFRDYvDRFYKTLRAEQAsQeVKnwMTETLLvQnNPNDC	330
HIVSF2	-----D-----	332
HIVBZ167	--K-----I-----D-----	331
HIVBZ190	-----M-----K-----S-----	330
HIVBZ200	-----X-----	330
HIVPH136	-----M-----K-----	330
HIVPH153	-----D-----	330
HIVTB132	-----P-----	330
HIVLAI	-----	330
HIVHXB2R	-----	330
HIVBH102	-----	330
HIVFV22	-----L-----	330
HIVMN	-----S-----RT-----	333
HIVJH3	-----S-----	330
HIVJRCSE	-----V-----T-----	330
HIVQYI	-----V-----D-----	330
HIVNY5CG	-----	330
HIVNL43	-----	330
HIVCDC4	-----I-----	330
HIVHAN	-----K-----T-----	330
HIVCAM1	-----	330
HIVRF	-----I-----D-----P-----	330
HIVD31	-----T-----	330
HIVJRF1	-----	328
HIVUG280	-----K-----S-----	330
HIVYU2	-----	330
HIVBCSG3C	-----S-----S-----	330

HIV1 GAG

CONSENSUS.C	VGDIYKRWIIILGLNKIVRMYSFVSILDIKQGPKPEFRDYVDREFFKTLRAEQALQEVKNWMTDTLLVQMANPDC	319
HIVUG268	--E-----	325
HIVSM145	-----	327
HIVZAM18	-----E-----	329
HIVZAM19	-----R-----E-----	266
HIVZAM20	--E-----R-----	327
HIVDJ259	-----R-----S-----	328
HIVVI313	-----	328
CONSENSUS.D	VGeIYKRWIIILGLNKIVRMYSFVSILDIRQGPKEPFRDYVDREFFKTLRAEQAGQGVKNWMTDTLLVQMANPDC	326
HIVELI	-----V-----	330
HIVNDK	-----	327
HIVZ226	-----E-G-----	331
HIVVI203	-----T-E-----	330
HIVVI205	-KK-----K-----K-----	331
HIVG109	-----K--M-----	333
HIVK31	-----	330
HIVUG274	-----	330
HIVUG270	-----	330
HIVSB365	-----T-E-----	330
CONSENSUS.F	VGeIYKRWIIILGLNKIVRMYSFVSILDIRQGPKEPFRDYVDREFFKTLRAEQATQGEVKNWMTDTLLVQMANPDC	316
HIVVI174	-----	325
HIVVI69	--D-----E-----I-----	326
HIVBZ162	-----	326
HIVVI325	-----K--S-----N-----	326
CONSENSUS.G	VGeIYKRWIIILGLNKIVRMYSFVSILDIRQGPKEPFRDYVDREFFKTLRAEQATQ?VK?WMTDTLLVQMANPDC	309
HIVVI191	--D-----E-G-----	327
HIVLBV217	-----D-N-----	328
HIVTAIG	-----	185
CONSENSUS.H	VGDIYKRWIIILGLNKIVRMYSFVSILDI?QGPKPEFRDYVDREFFK?LRAEQATQ?VK?WMTDTLLVQMANPDC	298
HIVVI525	-----K-----T-----D-N-----	330
HIVVI557	-----R-----A-----E-G-----	330
CONSENSUS.O	VGDIYKRWIIILGLNKIVRMYSFVSILDI?QGPKPEFRDYVDREFFKTLRAEQATQGEVKNWMTDTLLVQN?NPDC	295
HIVANT70	-----K-----A-----	329
HIVMVP51	-----R-----S-----	329
SIVCPZQAB	VGDVYRRWIIILGLNKIVRMYSFVSILDIRQGPKEPFRDYVDREFFKTLRAEQASQEVKNWMTDTLLVQMANPDC	338

HIV GAG

	p24 \/ \/ 'p2'	\/ p7	
CONSENSUS.A	KsILraIg?gAtLeEMMTacQgVggPgHKArvLAEAmSqvq??n...??siMmQrGnf.?Ggkr?iKCPNCGk		385
HIVU455	.-----P-----QTS.....R-PR-----		391
HIVUG266	-----RPK-----R-R-S---Q-----QTS.....K-----		388
HIVMAL	-T-K--P-----S-----ATNST...AA-----K-----		403
HIVVI32	-----P-----H-QA-----I-R-GKKR-----		396
HIVVI57	-----T-----S-----QP-----R-RG-----		395
HIVVI59	.--KR-RVR-----K---I---QTS.....K-----		395
HIVVI310	.--K--T-----S---I-----A-HT-----I---K-----		394
HIVVI354	-T--G--A---K-----E-E-----HNTS.....K--K-----		393
HIVVI415	-T-----T-----HT-----K--R-----		395
HIVK7	---K---T-----S---K-----A-QT-----K---T-----		395
HIVK29	.--A-----S-----HT-----R-----		395
HIVK88	-----T-----HT-----R-----		395
HIVK89	-----T-----HT-----R-----		395
HIVK98	---K--T-----QA-----R-----		395
HIVK112	---G--T-----I-----HT-----R-----		395
HIVK124	-----A-----S-----ATNRI...NAAV-----K-P-KI-----		399
HIVCI4	-----T-----HT-----K--R-----		393
HIVCI20	-----P-----QS-----R--RT-----		387
HIVCI32	---K--P-----QS-----K--RKFV-----		398
HIVCI51	-----P-----S-----QS-----R--RT-----		390
HIVCI59	-----P-----QS-----G--RT-----		392
HIVGI41	-T-----P-----NA-----K--K-----		391
HIVTN238	---X--T-----S-----HA-QAS.....K--G-----R		395
HIVTN243	---X--T-----S-----HA-HA-----K-----		395
HIVTN240	---K--T-----S-----HA-HAT.....K-----R		395
HIVTN245	---K--S-----S-----HA-HA-----K-----R		395
HIVLBV105	-T-----P--S-----HP-----K---I-----		400
HIVLBV23	-T-----A-----HT-----V-----R-----		395
HIVLBV2310	-----PA-----QT-----R--RT-----		392
HIVIC144	-T-----P-----QS-----R--RT-----		391
HIVDJ258	-----P-----S-----QS-----D--R--RT-----		392
CONSENSUS.B	KTILKALGPaaTLEEMMTACQGVGGPgHKArvLAEAMSqv...tn...?setiMmQrGnF.rnqrKtVRCFNCGK		397
HIVSF2	-----P-M-----		399
HIVB2167	-----P--S-----K---I-----		399
HIVB2190	-----		397
HIVB2200	-----S-----X-----A-----		397
HIVPM136	-----S-----A-----I-----		397
HIVPM153	-----A-----M-L-----T-----		397
HIVTB132	-----K-K-----I-----		397
HIVLAI	-----I-----		397
HIVHXB2R	-----I-----		397
HIVBH102	-----T-----M-----		397
HIVPV22	-----T-----M-----		397
HIVMN	-----II-----		400
HIVJH3	-----T-----II-----		397
HIVJRCSP	-----P-----N-----		397
HIVQYI	-----V-V--K--K-----		396
HIVNY5CG	-----P--I-----K-----		397
HIVNL43	-----G-----P--I-K-----		397
HIVCDC4	-----R-G-----		397
HIVHAN	-----G-----A-----		397
HIVCAM1	-----		397
HIVRF	-----S---I-----L-K---D---I-----		397
HIVD31	-----V-----		397
HIVJRFL	-----P-----I-----		395
HIVUG280	-----PVN-----N-----I-----		397
HIVYU2	-----		397
XIVBCSG3C	-----G-----		397

	p24 \ / \ / 'p2' \ / p7	
CONSENSUS.C	KTILRALGPGAsLEENMTACQGVGGPpSHKARVLAEAMSQann?n.....imQrsNF.KGpkrivKCFNCGK	384
HIVUG268	-----G-----I-----I-----	391
HIVSM145	-----A-----I-----L-----G-----S-----	393
HIVZAM18	-----A-T-----T-----L-----S-----	395
HIVZAM19	-----A-T-----STT-----X-----RX-----	332
HIVZAM20	-----T-----T-SA-----V-----G-----T-----	393
HIVDJ259	-----T-----T-----K-----RRTI-----	394
HIVVI313	-----G-----I-----K-----I-----	394
CONSENSUS.D	ktILKALGPPATLeEMMTACQGVGGPpSHKARVLAEAMSQa?.tn..s?taiimQrgNF.KQprkikKCFNCGK	392
HIVELI	-----Q-----V-TA-----	398
HIVNDK	-----G-----V-----G-----A-V-----S-----	395
HIV2226	-----Q-----AA-V-----T-----	399
HIVVI203	-----G-----TNV-----KR-----	398
HIVVI205	-----Q-----K-----E-----S-----A-VI-----K-----	399
HIVG109	-----A-----S-----NA-----K-----VV-----R	401
HIVK31	-----G-----G-----N-G-----Q-----	397
HIVUG274	R-----G-----AN-----	398
HIVUG270	-----A-----S-----ANA-----A-----KRN-----	398
HIVSE365	-N-----G-----E-----S-----AA-----Q-----V-----	398
CONSENSUS.F	KTILKALGPGALLeEMMTACQGVGGPGHKARVLAEAMSQa..TN..??a?imMQksNF.KGQRriVKCFNCGK	381
HIVVI174	-----TA-A-----	393
HIVVI69	-----M-----T-----A-----T-----	392
HIVBZ162	-----T-L-----	392
HIVVI325	-----S-----I-----V-----TT-----RG-----K-I-----R	392
CONSENSUS.G	KTIL?ALGPGATLeEMMTACQGVGGP?HKARVLAEAMSQA..SG..?A?A?MMQK??F.KGP???IKCF?CGK	366
HIVVI191	-----R-----G-----V-A-M-----SG-----KRT-----N-----	395
HIVLBV217	-----K-----S-----T-T-I-----GN-----RKN-----D-----	396
CONSENSUS.H	K?IL?ALG?GASIEEMMTACQGVGGP?HKA?VLAEAMSQV..?TN..??A?A?MMQKGNF.KQQRRI?KCFNCGK	356
HIVVI525	-N--K--T-----S--R-----TH-AI-----V-----	398
HIVVI557	-T--R--Q-----G--K-----N-----VM-----I-----	397
CONSENSUS.O	KQILK?LGP?ATLeEMMVACQGVGGPTHKA??LAEANA?AQQDLKGGYTAVFMQROQN.P?RKG?IKCFNCGK	360
HIVANT70	-----S--G-----RV-----T-----I--T-----	401
HIVMVP51	-----A--E-----KI-----S-----N--P-----	401
SIVCPZGAB	KQILKALGPGATLeEMMTACQGVGGPpSHKARVLAEAMSMVQ.NQ..GRADVFPQKQGG.AGPKRKIKCFNCGK	407

HIV1 GAG

	pol cds ->	
	p7 \/ 'p1' \/ p6	
CONSENSUS.A	EGHIArNCrAPrKkGCwKCgkEGHQmKdCT.?e.rQANFlgkiwpsSKG.RPgNFPQSRp.....	441
HIVU455	-----K-----N-----	447
HIVUG266	-----K--K-----K-----I-----K-----R-----	443
HIVNAL	-----H-----L-----	459
HIVVI32	-----N-----N-----	453
HIVVI57	-----	451
HIVVI59	-----K-R-----	451
HIVVI310	-----N-----T-----N-L-----	450
HIVVI354	-----K-----I-N-----K-----N-----L-N-----	449
HIVVI415	-----S-----	451
HIVK7	-----X-R-----L-----RR-X-----	451
HIVK29	-----D-----L-----N-L-----	451
HIVK88	-----G-----N-L-----	451
HIVK89	-----E-----R-----N-L-----	451
HIVK98	-----R-----	451
HIVK112	-----R-----N-L-----	451
HIVK124	-----I-K-----R-----R-----H-----L-----	455
HIVCI4	-----	449
HIVCI20	-----K-----R-----	443
HIVCI32	-----I-K-----I-----	454
HIVCI51	-----K-----R-----	446
HIVCI59	-----K-----R-----G-----	448
HIVG141	-----N-----	447
HIVTN238	-----N-----	451
HIVTN243	-----R-----N-----	451
HIVTN240	-----N-----	451
HIVTN245	-----N-----	451
HIVLBV105	-----L-----T-----	456
HIVLBV23	-----	451
HIVLBV2310	-----K-----R-----	448
HIVIC144	-----K-----R-----T-----	447
HIVDJ258	-----K-----T-----	448
CONSENSUS.B	eGHIAkNCrAPrKkGCwKCgkEGHQmKdCt.?E?RQANFLGKIWPSHKG.RpgNFIQSRp????????..??	453
HIVSF2	-----R-R-----Y-----	455
HIVBZ167	-----S-----Y-----	455
HIVBZ190	-----S-----	453
HIVBZ200	-----R-----N-----N-----	453
HIVPH136	-----R-----	453
HIVPM153	-----R-K-----R-----	453
HIVTB132	-----R-----	453
HIVLA1	-----R-----Y-----EPTAPFFLQ..SR	464
HIVXB2R	-----T-R-----Y-----	453
HIVBH102	-----T-R-----Y-----EPTAPFFLQ..SR	464
HIVPV22	-----T-R-----Y-----EPTAPFFLQ..SR	464
HIVMN	-----R-----C-----R-P-----T-----	455
HIVJH3	-----L-R-----N-----S-----	453
HIVJRCSE	-----R-----E-----Y-----	453
HIVQYI	-----R-----N-----	452
HIVNY5CG	-----	453
HIVNL43	-----	453
HIVCDC4	-----R-K-----R-----	453
HIVHAN	-----R-----	453
HIVCAM1	V-----N-----	453
HIVRF	V-----N-G-----	455
HIVD31	-----R-----Y-----R-----	453
HIVJRFL	-----R-----Y-----	451
HIVGO280	V-----N.N-----Y-----	454
HIVYU2	-----	453
HIVBCSG3C	-----R-K-----R-----P-L-----	453

		pol cds ->				
		p7 \/	'p1'	\/ p5		
CONSENSUS.C	EGHIArncRAPRKKGCWKCG?EGHQMKDCT..E.RQANFLGKIWPs?KG.RPGNPLQsRPeptap???????				443	
HIVUG268	---L-----R-----,.,-----PN-----PA.....				454	
HIVSM145	-----K-----,.,-----H-----N-----PAESFRVR				462	
HIVZAM18	---K-----K-----,.,-----H-----R				457	
HIVZAM19	---KI-----X-----,.,-----H-----				388	
HIVZAM20	-----R-----,.,-----Q-----N-----PAENFR..				460	
HIVDJ259	---L-----R-----,.,-----N-----				450	
HIVVI313	---L-----K-----,.,-----Y-----N-----				450	
CONSENSUS.D	EGHIAKNCrAPrkkGCWKCGkEGHQMKDCT..e.RQANFLGKIWpshKG.rPGNFIQSRp.....				448	
HIVELI	-----R-----L-----,.,-----R-----				454	
HIVNDK	---T-----R-----,.,-----				451	
HIV2226	-----R-----L-----,.,-----				455	
HIVVI203	---L-----				454	
HIVVI205	-----K-----R-----I-----,.,-----L-----				455	
HIVG109	-----R-----,.,-----N-----				457	
HIVK31	---L-----R-----,.,-----RL---R---G-----				453	
HIVUG274	-----E-----,.,-----L-----				454	
HIVUG270	---L-----R-----I-----,.,-----				453	
HIVSE365	---R---K-----Q-----G-----S-----P-----				454	
CONSENSUS.F	eGHIAKNCRAPRKKGCWKCGrEGHQMKDCT..E.rQANFLGKIWPSnKG.rPGNFIQsRP.....				437	
HIVVI174	---R-----,.,-----				449	
HIVVI69	---V-----,.,-----				448	
HIVB2162	T-----G-----H-----N-----				448	
HIVVI325	---H-----K-----,.,-----K-----				448	
CONSENSUS.G	EGHLARNCRAPRK?GCWKCG?EGHQMKDCT..E.RQANFLG?IWPSHKG.RPGNFIQ?R?.....				417	
HIVVI191	-----K-----R-----,.,-----R-----I-T.				451	
HIVLBV217	-----R-----K-----,.,-----K-----N-P.				452	
CONSENSUS.H	EGH?ARNCRAPRKKGCWKCG?EGHQ?KDCT..E.RQANFLGKIW???KG.RPGNFIQSRP.....				406	
HIVVI525	---I-----R---M-----,.,-----PSS-----				454	
HIVVI557	---L-----K---L-----,.,-----SPN-----				453	
CONSENSUS.O	EGHIA?NCRAPRK?GCWKCGQEGHQMKDC?..NG?QANFLGKYWPPGGT.RPGNYVQ???				410	
HIVANT70	-----R-----K-----R---K-----RPA.				458	
HIVMVP51	---K-----R-----K---R-----KQV.				458	
SIVCFZGAB	EGHLARNCKAPRRKGCWRCGQEGHQMKDCT..G.RQVNFLLGKWPSRSG.RPGNFIQNRRT.....				463	

HIV1 GAG

	\ (minor)		\ (minor)		/ p6 terminus (80%)	
CONSENSUS A	EPtAPpAE?f?gmgeett s?	?pkqeekd	??ke??ppl?S1KSLFGNDpLSQ			483
HIVU455	- I - -KM P	A- L	R QT -V -			493
HIVUG266	IL R-K -M- -P	---K---	QT I			480
HIVMAL	---S- -F---K PS	Q - -	-LY -A-----Q---			505
HIVVI32	S -I- - A P	-	R-QSQ--V----			491
HIVVI57	L - -A- -P	- - -	RQGG--V- -			489
HIVVI59	- ----IL RIRK-M- -P	-----T	-SI-----			483
HIVVI310	-----S- ----D-- -S	Q-----	R-HP--V----			488
HIVVI354	S F-- -A PS	RP PRE	--RY --T----			486
HIVVI415	-----IV ---G-- -P	-----	R-QA--V----			486
HIVK7	-----I- RIE---A -	P-----	--QP -SV----			488
HIVK29	-----I- ---- -P	-----	R-QA--T----			489
HIVK88	-- - -IC ----A -P	----R-	R-QT--V----			489
HIVK89	- - ----IC ----A -P	S-----	R-Q ASI----			487
HIVK98	-----I- ---- -P	-----	R-QP --V----			488
HIVK112	----I- ----MA -P	-----	R-QT--SI----			489
HIVK124	-----S- -F---PS	Q--P--	--LY --A-V--			492
HIVC14	- - ----LL ---- -S	-----	R-QP -SI----			484
HIVC120	-----S- --R----S	----PR-	-GLY--A----			481
HIVC132	-----S- --E---A -S	----PG-	-GLY--A----			492
HIVC151	-----S- -L-----S	LQ--PR-	-GLY--A----			484
HIVC159	-----N- --E-Q-- -S	----PR-	-GLY--T----			486
HIVG141	--S-----S- -F---PS	Q-----	KD--LY --A----			486
HIVTN238	-----NW ---- -L	L-----	--HP--V----			489
HIVTN243	-----DN- R--T- -L	L-----	--DP--V-----			495
HIVTN240	-----NW ---- -L	-----	--HP--V----			489
HIVTN245	-----NW ---- -L	L-----	--HP--V----			489
HIVLBV105	-S-----S- -F---V- PS	-----	--LY--T----			494
HIVLBV23	-----MIF-----P	--K----	R-QPL-SV----			490
HIVLBV2310	-----S- ---- -S	----PR-	-GLY--T----			486
HIVIC144	-----S- -- --K- TS	S-Q--PR-	--LY--T----			485
HIVDU258	-----L- -S- ---- -P	----PRE	QGLY--T----			486
CONSENSUS B	?EPTAPpaeef rfgeetttPs???QKQEPiD	kelY?plasLrsLPGnDPsSq				500
HIVSF2	-----K----	-----	---T-----			502
HIVB2167	-----I- ----	----TK	---K-			493
HIVB2190	-----V	Q-K- -A--				491
HIVB2200	-----A- G--M -	P---QK-	EG -P-----			491
HIVPH136	-----R-----P	-----	-----K-			491
HIVPH153	-----R-----P	-----	---S--KL			491
HIVTB132	-----A- ----	----T-	---TA-K-----			497
HIVLAI	P-----S-V-----	-----	---T-----			512
HIVHXB2R	-----S-V-----P	-----	---T-----			500
HIVBH102	P-----S-V-----P	-----	---T-----			512
HIVPV22	P-----S-V-----P	-----	---T-----			512
HIVMN	-----S-V-----YQKQEK--T-	-D- -K-----L--				506
HIVJH3	-----R-	-----	-----P			500
HIVJRCSP	-----A- ---QKQE	-----	---T-----			504
HIVQYI	-----A- G-----P	-----	-G- -T-----			499
HIVNY5CG	-----	-----	-----S----			500
HIVNL43	-----	-----	-----S----			500
HIVCDC4	-----D-----	----R-	-----			500
HIVHAN	-----A-A--	-----	-----K--S----			499
HIVCAM1	-----K----	-----	-----			500
HIVRP	-----K-	-----	-----K-----			501
HIVD31	-----A--P	-----	-----			500
HIVJRFL	-----K- ----A--	-----	--M- -T-----			498
HIVUG280	-----V-----	----L-	---SS-----			492
HIVYU2	-----S--V	-----	-----S----			500
HIVBCSG3C	-----P	----RE-	--M- -----			500

HIV1 GAG

	\\ (minor)	\\ (minor)	/ p6 terminus (80%)
CONSENSUS.C	?ePtAppaEST.rF.Bett.Pa....pkQEpkd?????e???PLtsLkSLFGNDP		482
HIVUG268	-G-.-P-.....RE--I----		488
HIVSM145	P-----L-E-----		495
HIVZAM18	P-----N-----R-----		492
HIVZAM19	.X-X-XX-----N-----R-----		428
HIVZAM20	R-----SPTGE--A--		496
HIVDJ259	G-.-P-.-P...Q--L--KEPYK-...--A-R-		488
HIVVI313	S-----K-----		484
CONSENSUS.D	.EPTAPPAESF.GFGEit.Ps....qKQEqKd????kEly.PlasLkSLFGNDPlsQ		494
HIVELI	-T-----		500
HIVNDK	S--		497
HIVZ276	STA-----L-		501
HIVVI203	KA-....A--P-E-....-V----		490
HIVVI205	K-.....N-.-T--		492
HIVG109	N-.-P-.....KEKD-....-T--R-		498
HIVK31	M-		490
HIVUG274	K-.....PC-		491
HIVUG270	S-A--		490
HIVSE365	T-		491
CONSENSUS.F	.EPLAPPAESF.GFTEEIt.PS....PkQEqKd....eglypPLaSLKS		474
HIVVI174	G-.....G-----		487
HIVVI69	..-S-----R-.-S-----		486
HIVB2162	R-.-N-.....E-		486
HIVVI325	N-.-G-.....R-.-T-....KEQGS--T--		486
CONSENSUS.G	.EP?APPAESl.gfgEEI?..?S....PKQEPId....KELYpPL?SLKS		451
HIVVI191	..-T-----T-S-----V----		484
HIVLBV217	..-S-----A-P-.....-T----		487
CONSENSUS.H	.EPTAPPAESF.GFGEEMT.P?....PKQE??D....?E...?PL?SLKS		436
HIVVI525	S-----FR-....K-.-P-A----		490
HIVVI557	P-.....RK-....A-.-S--T----		489
CONSENSUS.O	.?PSAPPME.....E?VK.?Q....ENQ?QKG...??PELY.PFASLKSLEGTQDQ		443
HIVANT70	.N-----E-.-G-----E-....GPN-----		498
HIVMVP51	.S-----A-.-E-----S-....DQE-----		498
SIVCP2GAB	.EPTAPPIESY.GYQEEK.SQ....EKKEG....ESSLY.PPTSLKSLFGSDPSSQ		508

HIV1 POL

Sequence subtypes C and E are not currently represented in POL.

CONSENSUS.A	FFRE.MLAFQGGAREP.....SSE..QTRANSPTSRLWDOGGKDO..LP...CETGAER..QG..	49
HIVU455	-----L-----SSEQTRANSEPI-----P-----	49
CONSENSUS.B	FFRE.dLAFpQGGAREf??????????Se..QrTRANSptrrElqVWgrdnnsIS????EAGadR..QG..	50
HIVLAI	-----L-----SSEQTRANSEPI-----P-----	62
HIVHXB2R	-----L-----P-----S-----EAGA-----	50
HIVM81	-----L-----N-----EAGE---D-----	53
HIVJRCSP	-----L-----P-----S-----EAGA-----	54
HIVJRPL	-----L-----K-----S-----	50
HIVQYI	-----S-----R-----P-----	50
HIVSF2	-----L-----GE-----	50
HIVMYSCG	-----	50
HIVNL43	-----	50
HIVCAM1	---N---E-----E-----	50
HIVHAN	---E---K-----S---S-----	49
HIVD31	---L---A-----S-----	50
HIVRF	---N---L-----E-----	49
HIVYU2	---K---I---R---R-----	50
HIVBCSG3C	---S---G---P-----G-----	50
CONSENSUS.D	FFRE.dLAFpQGGAGEI.....Sse..QTRANSPTSRELrvWGrD.NPLS...eTGAER..QG..	49
HIVELI	---N---PK-----K-----	49
HIVZ226	-----F-----G-----	49
HIVNDK	-----F-----G-----	49
CONSENSUS.O	FFRe.7LAsggheARqI.....caE..ts?piSPt?????gGG??..??Segt?Esg?Ecg??r..	39
HIVMAL	---N---FPQGGK--EF.....PS---QTRAN---SRELrvW--DK.TL-----T-A---QG..	49
HIVANT70	---Q-I-----T---D-----G---T---PE---	46
HIVMVP5180	---V---V---N-----R---ES-G-SG---	46
SIVCP2GAB	FFRE.RLAFpQGGAREQL.....CAE..QNRINGPTDRELWVPGGRE.....EPGEERGREDO..	48
SIVCP2GAB2	-----	0

		\\ protease <- gag cds end	
CONSENSUS.A	.TDSFSFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEDINLPGRWPKPKIIGGIGGFIVRQYDQILIEI	121	
HIVU455		121	
CONSENSUS.B	.cvsfzfpQITLWQRPLVTVKIGGQLKEALLDTGADDTVleemhLPgrWPKPKNIIGGIGGFIVRQYDQILieI	122	
HIVLAI	N.....S.....	134	
HIVHB2R	N---V.....S.....	122	
HIVMN	.P-----I-----G---R-----T-G-	125	
HIVJRCSP	.I---N-----D-D-----P-D-	126	
HIVJRFL	.N---N-----D-D-----	122	
HIVQYI	NL-----I-----	122	
HIVSF2	N-----R-----K-----PV--	122	
HIVNY5CG		122	
HIVNL43	G-----	122	
HIVCAM1	P---	122	
HIVHAN	L-L-----V---S-----V---	121	
HIVD31	S-----	122	
HIVRP	I---V-----K-----	121	
HIVYU2	P---	122	
HIVBCSG3C	.N--L-----P---	122	
CONSENSUS.D	.TVSPnfPQITLWQRPLVTVKIGGQLKEALLDTGADDTVleemhLPGRWPKPKNIIGGIGGFIVRQYDQILIEI	121	
HIVELI	A-----P---	121	
HIVZ226	C-----	121	
HIVNDK	S-----I-----	121	
CONSENSUS.O	.avs?clPQIPLWdRPiVtarVGGhLcEaLLDTGADDTVLnnIqLeGkWKPKNIIGGIGGFIVKkeYdnvtvE1	110	
HIVMAL	.I--FSF---T--Q--V--V---Q-K-----EE-N-P-----RQ--QILI--	121	
HIVANT70	.L-V-----V-----	110	
HIVMP5180	--PI-----K-----R-T-----N---V	110	
SIVCPZGAB	.SISTNLPQITLWQRPLIPVKVEGQLCEALLDTGADDTVIERIQQLQGLWPKPKNIIGGIGGFIVKQFDNVHIEI	120	
SIVCPZGAB2		0	
		protease \\ p66, p51	
CONSENSUS.A	CGKKTIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPISTVPVKLKPMDGPKVKQWPLTEEKIKALTEICNE	194	
HIVU455		194	
CONSENSUS.B	CGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPISTVPVKLKPMDGPKVKQWPLTEEKIKALvEICTE	195	
HIVLAI		207	
HIVHB2R		195	
HIVMN	L-----I-----	198	
HIVJRCSP	V-----	199	
HIVJRFL		195	
HIVQYI	L-----V-I-----	195	
HIVSF2		195	
HIVNY5CG		195	
HIVNL43		195	
HIVCAM1		195	
HIVHAN	I-----	194	
HIVD31		195	
HIVRP		194	
HIVYU2		195	
HIVBCSG3C		195	
CONSENSUS.D	CG?KAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPISTVPVKLKPMDGPKVKQWPLTEEKIKALTEICTE	193	
HIVELI	--Q-----D	194	
HIVZ226	--H-----	194	
HIVNDK	--Y--M-----	194	
CONSENSUS.O	?GkevqGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPISTVPVKLKPMDGPKVKQWPLs?EKIEALTeICqe	101	
HIVMAL	C--KAI--I-----M--QI-----ET-----R-----TE--X--E--KD	194	
HIVANT70	E-R-----K-----	191	
HIVMP5180	Q-----L-----R-----	191	
SIVCPZGAB	EGRKVVGTVLVGPTPVNIIGRNLLTQIGCTLVFPISPISTVPVKLKPMDGPKVKQWPLSAERIKALTEICQE	193	
SIVCPZGAB2		0	

HIV POL

CONSENSUS A	MEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHTAGLKKKKSVTVLQVGD	267
HIVU455	-----	267
CONSENSUS B	MEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLQVGD	268
HIVLAI	-----	280
HIVHXB2R	-----	268
HIVMN	-----K--	271
HIVJRCSP	-----R--	272
HIVJRFL	-----K--R--	268
HIVQYI	-V-----	268
HIVSF2	-----	268
HIVNY5CG	-----Q-----	268
HIVNL43	-----Q-----	268
HIVCAN1	-----	268
HIVHAN	-----	267
HIVD31	-----	268
HIVRF	-----	267
HIVYU2	-----	268
HIVBCSG3C	-----	268
CONSENSUS D	MEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLQVGD	266
HIVELI	-----	267
HIVZ226	-----V-----	267
HIVNDK	-----	267
CONSENSUS O	MEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLQVGD	254
HIVMAL	--R--LK-----V-----N-----A--K-K-----	267
HIVANT70	-----G-----E-----	264
HIVMVP5180	-----R-----	264
SIVCP2GAB	MEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLQVGD	266
SIVCP2GAB2		0
CONSENSUS A	AYFSVPLDSEFRKYTAFTIPSIINNETPGVRYQYNVLPQGWKGSPIFQSSMTKILEPFRSQHPDIVIYQYMDQ	340
HIVU455	-----	340
CONSENSUS B	AYFSVPLDSEFRKYTAFTIPSIINNETPGVRYQYNVLPQGWKGSPIFQSSMTKILEPFRSQHPDIVIYQYMDQ	341
HIVLAI	-----E-----	353
HIVHXB2R	-----E-----	341
HIVMN	-----	344
HIVJRCSP	-----I-----	345
HIVJRFL	-----I-----	341
HIVQYI	-----	341
HIVSF2	-----	341
HIVNY5CG	-----C-----	341
HIVNL43	-----C-----	341
HIVCAN1	-----T-----	341
HIVHAN	-----	340
HIVD31	-----E-----V-----	341
HIVRF	-----E-----R-----K-----E-----	340
HIVYU2	-----HE-----T-----T-----L-----	341
HIVBCSG3C	R-----	341
CONSENSUS D	AYFSVPLDSEFRKYTAFTIPSIINNETPGVRYQYNVLPQGWKGSPIFQSSMTKILEPFRSQHPDIVIYQYMDQ	339
HIVELI	-----S-----	340
HIVZ226	-----K-----	340
HIVNDK	-----	340
CONSENSUS O	AYFSVPLDSEFRKYTAFTIPSIINNETPGVRYQYNVLPQGWKGSPIFQSSMTKILEPFRSQHPDIVIYQYMDQ	324
HIVMAL	-----V-----E-----I-----E--TK--IV-----	340
HIVANT70	-----RD--L--C-----	337
HIVMVP5180	-----V-----KS--V-----I--	337
SIVCP2GAB	AYFSVPLDSEFRKYTAFTIPSIINNETPGVRYQYNVLPQGWKGSPIFQSSMTKILEPFRSQHPDIVIYQYMDQ	339
SIVCP2GAB2		0

CONSENSUS .A	LYVGSdleIGQHRAKIEELRAHLLSNGFITPDKKHQKEPPFLWNGYELHPDKWTVQPIQLPEKDSWTVNDIQK	413
HIVU455	-----	413
CONSENSUS .B	LYVGSdleIGQHRCIEELRqHLLRWGFTTPDKKHQKEPPFLWNGYELHPDRWTVQPIvLPEKDSWTVNDIQK	414
HIVLAI	-----L-----	426
HIVXB2R	-----L-----	414
HIVMN	-----A-----R-----	417
HIVJRC9F	-----K-----	418
HIVJRFL	-----A-----	414
HIVQYI	-----M-----	414
HIVSF2	-----M-----	414
HIVNY5CG	-----	414
HIVNL43	-----	414
HIVCAM1	-----L-----M-----	414
HIVHAN	-----E-----K-----	413
HIVD31	-----T-----	414
HIVRF	-----I-----E-----K-----	413
HIVYU2	-----	414
HIVBCSG3C	-----	414
CONSENSUS .D	LYVGSdleIGQHRFKIEeLREHLLRWGFTTPDKKHQKEPPFLWNGYELHPDKWTVQsIkLPEKESWTVNDIQK	412
HIVELI	-----K-----R-----N-----	413
HIVZ226	-----	413
HIVNDR	-----P-N-----	413
CONSENSUS .O	LYVGSdlpl?eHRkriELLRSHlyqNGFTTPDKKHQKEPPFLWNGYELHPDKWTVQpIQLPdKevWTVNDIQK	396
HIVMAL	-----EIGQ--TK--E--LK-----S-----	413
HIVANT70	-----T-----S-----N-D-----	410
HIVMVP5180	-----A-----V-----	410
SIVCPZGAB	LYVGSdleIDQHRKKEELRQHLRWGFTTPDKKHQKEPPFLWNGYELHPDKWTVQPIQLPEKEVWTVNDIQK	412
SIVCPZGAB2	-----	0
CONSENSUS .A	LVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEALELAENREI...LKDPVHGVIYDPSKDLVAEI	483
HIVU455	-----	483
CONSENSUS .B	LVGKLNWASQIYAGIKVKQLCKLLRGtKALTEViPLTEEALELAENREI...LKDPVHGVIYDPSKDLIAEI	484
HIVLAI	-----P-----R-----	486
HIVXB2R	-----P-----R-----	484
HIVMN	-----V-----	487
HIVJRC9F	-----K-----V-----	488
HIVJRFL	-----L-----	484
HIVQYI	-----N-----V-----L-----	484
HIVSF2	-----E-----V-----	484
HIVNY5CG	-----R-----I-----V-----	484
HIVNL43	-----R-----V-----	484
HIVCAM1	-----V-----L-----	484
HIVHAN	-----P-----K-----C-----V-----	483
HIVD31	-----R-----A-----K-----	484
HIVRF	-----VQ-----K-----	483
HIVYU2	-----R-----	484
HIVBCSG3C	-----S-----VT-----	484
CONSENSUS .D	LVGKLNWASQIYpGIKVKQLCKLLRGTKALTEViPLTEEALELAENREI...LKDPVHGVIYDPSKDLIAEI	482
HIVELI	-----ER-----	483
HIVZ226	-----	483
HIVNDR	-----A-----K-----V-----L-----	483
CONSENSUS .O	LVGKLNWASQIYqGIrVkeLCKLLRGtKsLTevVPLs?EALELELAENRE?...LKDPVHGVIYqDpKDLwv?I	463
HIVMAL	-----P-----K-----Q-----L-----A-----DI-----TA-----A-----I-----D-----S-----IAE-----	483
HIVANT70	-----I-----R-----R-----R-----Q-----N-----	480
HIVMVP5180	-----K-----K-----S-----	480
SIVCPZGAB	LIGKLNWASQIYFOIKIKQLCKLLRGTKRLTDVVPLTPEEALELAENREI...VSTPVHGVIYDpKDLIAEI	482
SIVCPZGAB2	-----	0

HIV1 POL

CONSENSUS A	QKQGQDQWYQIYQEPFKNLKTGKYARKRSAHTNDVRQLTEVVQKVSTESIVIWGKIPKFRLEPIQKETWEAMW	556
HIVU455	-----T-----T-----T-----	556
CONSENSUS B	QKQGQDQWYQIYQEPFKNLKTGKYARKSGAHTNDVRQLTEAVQKISTESIVIWGKIPKFKLEPIQKETWEAMW	557
HIVLAI	-----T-----T-----T-----	569
HIVHXB2R	-----T-----T-----T-----	557
HIVMN	-----R-----T-----	560
HIVJRCSP	-----F-----T-----N-----I-----T-----	561
HIVJRPL	-----I-----T-----N-----I-----T-----	557
HIVQYI	-----TQ-----	557
HIVSF2	-----VS-----I-----	557
HIVNY5CG	-----K-----	557
HIVNL43	-----K-----	557
HIVCAM1	-----K-----T-----D-----	557
HIVHAN	-----E-----I-----G-----R-----	556
HIVD31	-----T-----	557
HIVRF	-----V-----	556
HIVYU2	-----T-----T-----	557
HIVBCSG3C	-----A-----T-----T-----C-----I-----R-----T-----	557
CONSENSUS D	QKQGQDQWYQIYQEPFKNLKTGKYARKSGAHTNDVRQLTEAVQKISTESIVIWGKIPKFRLEPIQKETWEAMW	555
HIVELI	-----R-----R-----	556
HIV22Z6	-----V-----	556
HIVNDK	-----D-----T-----T-----A-----K-----	556
CONSENSUS O	QKQGQDQWYQIYQEPFKNLKTGKYARKSGAHTNDVRQLTEAVQKISTESIVIWGKIPKFRLEPIQKETWEAMW	555
HIVMAL	-----Q-----QY-----I-----SA-----VK-----T-----A-----IA-----T-----IQK-----A-----	556
HIVANT70	-----GE-----T-----I-----I-----K-----	553
HIVMVP5180	-----H-----E-----V-----D-----A-----	553
SIVCPZGAB	QKQGNQDQWYQIYQEPFKNLKTGKYARKSGAHTNDVRQLTEAVQKISTESIVIWGKIPKFRLEPIQKETWEAMW	555
SIVCP2GAB2		0
p51 \ /		
CONSENSUS A	MEYWQATWIPWEPFVNTPLVLKLMYQLEKDPiAGAEtFYVDGAANRETKLGKAGYVTDGRGRQKVVSLETTNQ	629
HIVU455	-----	629
CONSENSUS B	MEYWQATWIPWEPFVNTPLVLKLMYQLEKDPiAGAEtFYVDGAANRETKLGKAGYVTDGRGRQKVVSLETTNQ	630
HIVLAI	-----S-----N-----T-----	642
HIVHXB2R	-----N-----T-----	630
HIVMN	-----TX-----V-----K-----N-----	633
HIVJRCSP	-----S-----	634
HIVJRPL	-----N-----	630
HIVQYI	-----D-----	630
HIVSF2	M-----IA-----	630
HIVNY5CG	-----I-----P-----	630
HIVNL43	-----I-----P-----	630
HIVCAM1	ID-----R-----P-----	630
HIVHAN	-----N-----	629
HIVD31	-----T-----	630
HIVRF	-----I-----	629
HIVYU2	-----I-----NK-----	630
HIVBCSG3C	-----E-----NK-----T-----	630
CONSENSUS D	MEYWQATWIPWEPFVNTPLVLKLMYQLEKDPiAGAEtFYVDGAANRETKLGKAGYVTDGRGRQKVVSLETTNQ	627
HIVELI	A-----L-----	629
HIV22Z6	V-----	629
HIVNDK	I-----	629
CONSENSUS O	MEYWQATWIPWEPFVNTPLVLKLMYQLEKDPiAGAEtFYVDGAANRETKLGKAGYVTDGRGRQKVVSLETTNQ	607
HIVMAL	T-----N-----V-----K-----DR R-----VVS T-----	629
HIVANT70	-D-----R-----S-----M-----Y-----D-----	626
HIVMVP5180	-----N-----N-----E-----	626
SIVCPZGAB	MEYWQATWIPWEPFVNTPLVLKLMYQLEKDPiAGAEtFYVDGAANRETKLGKAGYVTDGRGRQKVVSLETTNQ	628
SIVCP2GAB2		0

CONSENSUS .A	KTELHAINLALQDSGSEVNIVTDSQYALGIIQAQPDSESEIVNQIIEKLIKKEKVYLSWVPAHKGIGGNEQV	702
HIVU455	-----	702
CONSENSUS .B	KTELqAIhLALQDSGLEVNIVTDSQYALGIIQAQPDkSESELvsQIIEqLIKKEKVYLaWVPAHKGIGGNEQV	703
HIVLAI	-----N-----	715
HIVHXB2R	-----Y-----Q-----N-----	703
HIVMN	-----	706
HIVJRCSP	-----	707
HIVJREFL	-----	703
HIVQYI	-----	703
HIVSF2	-----	703
HIVNY5CG	-----	703
HIVNL43	-----	703
HIVCAM1	-----Y-----E-----	703
HIVHAN	-----	702
HIVD31	-----S-----I-----T-----	703
HIVRF	-----	702
HIVYU2	-----Y-----R-----	703
HIVBCSG3C	-----E-----	703
CONSENSUS .D	KTELQAINLALQDSGLEVNIVTDSQYALGIIQAQPDKSESELvsQIIEQLIKKEKVYLaWVPAHKGIGGNEQV	700
HIVELI	-----N-----	702
HIVZ226	-----	702
HIVNOK	-----	702
CONSENSUS .O	KaELma111aLQDSk*?VNIVTDSQYALG1Is*QPtqSeSpIVqQIIEeLtkK*?VYLtWVPAHKGIGGNEK1	678
HIVMAL	-T-Q-H-----GSE-----QA-DK-E-N-Q-IQ-DK-S-----QV	702
HIVAM770	-----T-----V-----Q-----	699
HIVMVP5180	-----V-I-----Q-----V-----D-----R-----	699
SIVCP2GAB	QAEKALLALQDSQGVNIVTDSQYVLGIITQSPDHSESELVNQIIEELIKKEKIYLSWVPAHKGIGGNEQV	701
SIVCP2GAB2		0
CONSENSUS .A	DKLVSSGIRKVLFLDGDIDKAQEDHEKYHCHNWRAMASDFNLPPVVAKEIVASCNKQQLKGEAMHGQVDCSPGIW	775
HIVU455	-----	775
CONSENSUS .B	DkLVsAg1rKVLFLGIDKAQeeHEKYHSNwrAMAsDFNLPPvVAKEIVASCDKQQLKGEAMHGQVDCSPGIW	776
HIVLAI	-----D-----	788
HIVHXB2R	-----D-----	776
HIVMN	-----D-----I-----	779
HIVJRCSP	-----D-----I-----	780
HIVJREFL	-----D-----K-----	776
HIVQYI	-----	776
HIVSF2	-----N-----	776
HIVNY5CG	-----	776
HIVNL43	-G-----	776
HIVCAM1	-----	776
HIVHAN	-----D-----N-----	775
HIVD31	-----D-----	776
HIVRF	-R-T-----D-----	775
HIVYU2	-----	776
HIVBCSG3C	-----V-----	776
CONSENSUS .D	DKLVSSGIRKVLFLDGDIDKAQEEHEKYHCHNWRAMASDFNLPPVVAKEIVASCDKQQLKGEAMHGQVDCSPGIW	773
HIVELI	-----	775
HIVZ226	-----	775
HIVNOK	-----	775
CONSENSUS .O	DKLVSkd1rrVLFLeGIDqAQEdHEKYHSNwrAlASdFgLPPIVAKEI1ASCPkCh1RGEA?HGQVDeSPe1W	750
HIVMAL	-----SG-K-----D-K-E-----M-N-----V-D-QL-----M-----G-----	775
HIVAM770	-----K-----E-----V-----I-----V-----	772
HIVMVP5180	-----T-----Y-----	772
SIVCP2GAB	DKLVsAGIRKVLFLDGDIDRAQEEHEKYHSNMKAMASDFNLPPIVAKEIVAHCDKQVVKGEAMHGQVDCSPGIW	774
SIVCP2GAB2		0

HIV1 POL

CONSENSUS A	QLDCTHLEGGKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPKVIHTDNGSNFTSAAVKAVCWAN	848
HIVU455	-----	848
CONSENSUS B	QLDCTHLEGGKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPKVIHTDNGSNFTScTVKAACWAG	849
HIVLAI	-----	861
HIVXB2R	-----GA--R-----	849
HIVMN	-----P-----T	852
HIVJRCSE	-----I-----T-----	853
HIVJREFL	-----I-----	849
HIVQYI	-----I-----I-----	849
HIVSF2	-----I-----	849
HIVNY5CG	-----V-----	849
HIVNL43	-----V-----	849
HIVCAM1	-----G--I-----	849
HIVHAN	-----I-----V-----V-----P-----	848
HIVD31	-----I-----	849
HIVRF	-----I-----I-----V-----	848
HIVYU2	-----T-----A-----	849
HIVBCSG3C	-----I-----	849
CONSENSUS D	QLDCTHLEGGKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPKVHTDNGSNFTSAaVKAACWAG	846
HIVELI	-----	848
HIVZ226	-----I-----I-----	848
HIVNDK	-----T-----	848
CONSENSUS O	Q ² DC ² THLEGGKIIIVAVHVASqfiEAEVIPAETGQETAYFILKLaRWPKVIHTDNGpNFTSaamKAACWan	822
HIVMAL	-L-----Y-----I--G--V--S--V-----	848
HIVANT70	-I--M-----TT-----	845
HIVMVP5180	-M-----D-----TG	845
SIVCP2GAB	QVDC ² THLEGGKVIIVAVHVASGYIEAEVIPAETGQETAYFLKLAGRWPKTIHTDNGENFTSAAVKAACWAD	847
SIVCP2GAB2	0	0
CONSENSUS A	IQQEFGIPYNPQSQGVVESMNKELKKIIQGVREQAEHLKTAVQMAVFIHNFRRKGIGGYSAGERIIDIIATD	921
HIVU455	-----	921
CONSENSUS B	IKQEFGIPYNPQSQGVVESMNkeLKKIIQGVREQAEHLKTAVQMAVFIHNFRRKGIGGYSAGERIvdIIATD	922
HIVLAI	-----	934
HIVXB2R	-----	922
HIVMN	-----I-----R-----G-----	925
HIVJRCSE	-----I-----	926
HIVJREFL	-----	922
HIVQYI	-----N-----	922
HIVSF2	-----N-----	922
HIVNY5CG	-----I-----	922
HIVNL43	-----I-----	922
HIVCAM1	-----N-----I-----	922
HIVHAN	-----	921
HIVD31	V-----	922
HIVRF	-----Q--Q-----	921
HIVYU2	-----	922
HIVBCSG3C	-----	922
CONSENSUS D	IKQEFGIPYNPQSQGVVESMNKELKKIIQGVREQAEHLKTAVQMAVFIHNFRRkgGIGGYSAGERIIDIIATD	919
HIVELI	-----RR-----	921
HIVZ226	-----	921
HIVNDK	-----	921
CONSENSUS O	IqheFGIPYNPQSQGVVEaMNKELKsIIQGVREQAEHLKTAVQMAVFIHNFRRKGIGGYtAGERIIDi ¹ Asq	895
HIVMAL	-KQ-----S-----K--G--E-----I-----S-----NI--TD	921
HIVANT70	-----R-----	918
HIVMVP5180	-----L-----	918
SIVCP2GAB	IKQEFGIPYNPQSQGVVESLNKELKKIIQGVREQAEHLKTAVQMAVFIHNFRRKGIGGYtAGERIIDIIATD	920
SIVCP2GAB2	-----M-----V--NL---	58

		vif cds ->	
CONSENSUS.A	IQTRELQKQISKIQNFRVYYRDSRDPINWKGPAKLLWKGEGAVVIQDMSDIKVVPRRKAKIIRDYGGKQMGDDC		994
HIVU455	-----		994
CONSENSUS.B	IQTRELQKQITKIQNFRVYYRDSrdPlWKGpAKLLWKGEGAVVIQDnsdIKVVPRRKaKiIRDYGGKmagDDC		995
HIVLAI	-----		1007
HIVHXB2R	-----N-----		995
HIVMN	-----N-----V-----T-----		998
HIVJRCSP	-----N--I-----V-----		999
HIVJRFL	-----E-----V-----		995
HIVQYI	-----E-----		995
HIVSF2	-----NK-----		995
HIVNY5CG	--I-----V-----		995
HIVNL43	-----V-----		995
HIVCAN1	-----		995
HIVHAN	-----E-F-----GS--		994
HIVD31	-----E-----		995
HIVRF	-----H-----		994
HIVYU2	-----		995
HIVBCSG3C	-----NK-----		995
CONSENSUS.D	IQTRELQKQIiKIQNFRVYYRDSRDPiWKGPAKLLWKGEGAVVIQDnSDIKVVPRRKVKIIRDYGGKQMGDDC		992
HIVELI	-----K-----		994
HIVZ236	-----T-----		994
HIVNDK	---R-----		994
CONSENSUS.O	IQTtELQKQILKi?nFRVYYRDSrDPIWKGPAqLLWKGEGAVVIQDkgDIKVVPRRKAKIIRDyGGKMGAGcDs		967
HIVMAL	---K-----T--Q-----N-----K-----NS-----D-C		994
HIVANT70	-----XHK-----E-----		991
HIVMVP5180	-----N-----		991
SIVCPZGAB	IQTSELQKQILKVQKFRVYYRDSRDPINWKGPA TLWKGEGAVVIQDQELKVVPRRKAKIIRDYGGKQMGDDC		993
SIVCPZGAB2	L--Q-----I-----G-----S--		93
CONSENSUS.A	MAGRQDED		1002
HIVU455	-----		1002
CONSENSUS.B	VASRQDED		1003
HIVLAI	-----		1015
HIVHXB2R	-----		1003
HIVMN	-----		1006
HIVJRCSP	-----		1007
HIVJRFL	-----		1003
HIVQYI	-----		1003
HIVSF2	-----		1003
HIVNY5CG	-----		1003
HIVNL43	-----		1003
HIVCAN1	-----		1003
HIVHAN	-----		1002
HIVD31	--G-----		1003
HIVRF	-----		1002
HIVYU2	--G-----		1003
HIVBCSG3C	-----		1003
CONSENSUS.D	VASRQDED		1000
HIVELI	-----		1002
HIVZ226	-----		1002
HIVNDK	-----		1002
CONSENSUS.O	MA?gQtEsES?EQP?EIP		982
HIVMAL	V-G--D-D		1002
HIVANT70	--S-----V--S--		1009
HIVMVP5180	--NR-----M--G--		1009
SIVCPZGAB	VASRQDED		1001

HIV1 VIF

<- pol cds		
CONSENSUS .A	MENRWQVMIVWQVDRMKIRTNWSLVKHHMYVSKKAQQWPFYRHHYESRHSRVSSSEVHIPLGE..ARLVVRTYWG	71
HIVU455		71
CONSENSUS .B	MENRWQVMIVWQVDRMRIRfCWksLVKHHMyis?KakgwifYrHHYEsthPrlsSEVHIPLGd..arLVITTYWG	70
HIVLAT	-----V-G-R-----P-----	71
HIVHXB2R	-----V-G-R-----P-----	71
HIVMN	---R---A---K---R---	71
HIVJRCSE	-----N-----G-----I-K-----N-V---Q-----	71
HIVJRFL	-----T-G---I-----	71
HIVNH52	-----V-----A-----R-----K-----	71
HIVQYI	-----V-K-----T-V---	71
HIVSF2	-----K-----V-----K-----	71
HIVNY5CG	-----N---R-----R---D-----N-K-----K-----	71
HIVNL43	-----N---R-----R---D-----N-K-----K-----	71
HIVCAM1	-----G---K-S-----E---V---	71
HIVHAN	---A---M---I-C-R---N-V-----N-V-----E---K-----	71
HIVD31	-----V-G-EK---K---N-----	71
HIVRF	-----R-----P---E-----	71
HIVYU2	-----A-----G-R-----P---K-----	71
HIVBCSG3C	-----I-----R-----K-----	71
CONSENSUS .D	MENRWQVMIVWQVDRMRI?TWKSLVKHhMYVSKKAnRWfYRHHYdSpHPKISSEVHIPLGE..ArLVvkTYWG	70
HIVELI	-----K-----E-----I-----	71
HIV2226	-----R-----S-----M-----	71
HIVNDK	-----N---Y-----H-----P---	71
CONSENSUS .O	MENRWQVlIVWQvDRqkvkawnSLVKyHkY?SkKa?NW?YRHHYESRnPKVSS?VyIPvg?..A??VVLTYWG	64
HIVMAL	-----M-----MRIRf-H-----N-M-V-----K-F-----H-----E-H---L-D---RL-R---	71
HIVANT70	-----R-R-TE-W-----R---S-----V---HV---	71
HIVMVP5180	-----I-----M---A---R-----A---AE---Di-----	71
SIVCP2GAB	MENRWQVMIVWQVDRMKIRTNWSLVKYHLYRSKICARGWfYRHHYDHPNPKVASEIHIFFRDY..SKLIIVTTYWA	72
CONSENSUS .A	LHTGEKDNHLGHGVSIENWLKRYSTQVDPDLADHLIHLHYFDCFSESAIRRAILQOQIVRPRCEYQAGHNKVGGS	144
HIVU455	-----	144
CONSENSUS .B	LHTgerdNHLGQGVsIEWrkkrySTQVDPDLADQLIHlyYFDCFSesAIRnaillG?ivSPrCEYqAGHnKvGs	142
HIVLAT	-----E-----D---K-L-H-----	144
HIVHXB2R	-----E-----D---K-L-H-----	144
HIVMN	-----H---H---D---K---NR---I---F-----P	144
HIVJRCSE	---M---TR-----H-----S---	144
HIVJRFL	-----V-----N-----H-----	144
HIVNH52	--SK-----V---K-R-----N-----H-----R-----I---	143
HIVQYI	-----E-----A-----G---T-----N-----P-----	144
HIVSF2	-----E-----A-----K-----G-----H-----K---YR-----	144
HIVNY5CG	-----H-----T---R-----	144
HIVNL43	-----H-----T---R-----	144
HIVCAM1	-----T-G-N-----K---V-RL-----	144
HIVHAN	-----A-----R-----N-----R-----S---	144
HIVD31	-----Q-H-----G-----R-G-S---R-----	144
HIVRF	-----R-----KPS-H-----	144
HIVYU2	-----K---YR-----	144
HIVBCSG3C	---H---V-----H-----	144
CONSENSUS .D	LHTGERewHLGQGVSIENWLKRRYSTQVDPDLADQLIHMYFDCFSesAIRKAILGhIVSprCEYQAGHNKVGGS	143
HIVELI	-----D-----	144
HIV2226	---D-----A-----H-----S---	144
HIVNDK	---K-----A-----S-----	144
CONSENSUS .O	LmpGERdeHLGHGVSIENWryK?YKTQIDPecAdrmIHLHYfCPTesAIRkAILGqrVltkCeY?agHsqVGC	135
HIVMAL	QT--K W-----Q-R-S--L- DL--QL---Y -D--S --- Q----H1-SPR D Q-- NK -S	144
HIVANT70	---K-----A V-----PT---	144
HIVMVP5180	---E-----Q-E-----L---	144
SIVCP2GAB	LSPGERAWHLGHGVSIQWRLGSYVTQVDPTADRLIKSQYFDCFAETAIRRAILGQLVAPRCEYKEGHRQVGS	145

vpr cds ->

CONSENSUS.A	LQYLAL.KAL...VTPTRAKPPLPSVKKLTEDRWNNKPKTKGHRGSRITLNRH	192
HIVU455	-----,-----,-----	192
CONSENSUS.B	LQYLAL.sAL...itpKkiKPPLPsV7KLTEDRWNNKPKTKGHRGShTmNGH	189
HIVLAI	-----,-----,-----T-----	192
HIVHXB2R	-----,-----,-----T-----	192
HIVMN	-----,T-----,-----K-----I-----	192
HIVJRCSE	-----,T-----,-----K-----K-----	192
HIVJRFL	-----,T-----,-----KS-----K-----	192
HIVNH52	-----,-----,-----T-----R-----K-----	191
HIVQYI	-----,-----,-----K-----T-----	192
HIVSF2	-----,-----,-----T-----K-----	192
HIVNY5QG	-----,-----,-----K-Q-----R-----	192
HIVNL43	-----,-----,-----K-Q-----R-----	192
HIVCAM1	-----,T-----,-----A-----R-----	192
HIVHAN	-----,-----,-----T-----I-T-----H-----	192
HIVD31	-----,-----,-----A-----R-----	192
HIVRF	-----,-----,-----T-----K-----	192
HIVYU2	-----,T-----,-----T-----K-----R-----	192
HIVBCSG3C	-----,-----,-----R-----T-----	192
CONSENSUS.D	LQYLAL.LAL...IAPKKIKPPLPSVRKLTEDRWNNKPKTKGHRGSHTMNGH	191
HIVELI	-----,-----,-----Q-----Q-R-----	192
HIVZ2Z6	-----,-----,-----V-----K-----	192
HIVNDK	-----,A-----,-----R-----	192
CONSENSUS.O	LQ?LAL.?AV...vk?k??kPPLPSVgkLTEDRWNNKp??irdqlgShahNGH	176
HIVMAL	--Y---,T-L---,IAP-KTR-----R-----QOTKGHR---T---	192
HIVANT70	--L---,R---,---ARSR-----HLR-----K-P-----	192
HIVMVF5180	--F---,K---,---V-RN-----R-----WK-----	192
SIVCPZGAB	LQFLAL.KAL...ISERRHRPPLPSVAKLTEDRWNNKHQRTKVHQENLTRNGH	193

HIV VPR

Sequence subtypes C and E are not currently represented in VPR.

		vif cds	
CONSENSUS.A	MEQAPEDQGPQREPYAEKALELLEELKHEAVRHFPRLWHLGQHIYNTYGDWEGVEAIIIRILQQLLFIFHR	73	
HIVU455	-----H-----N-----G-----	73	
CONSENSUS.B	MEQAPEDQGPQREPYNEWtLELLeELK?EAVRHFPRLWHL?LgQHIYETYGDTWagVEAIIIRILQQLLFiHFr	71	
HIVLAI	-----H-----N-----G-----	73	
HIVHXB2R	-----H-----N-----G-----Q	73	
HIVMN	-----O-A-----N-----G-----	73	
HIVSC	-----RK-S-----G-----	73	
HIVJRCSE	-----N-----S-Y-----	73	
HIVJRFL	-----N-----V-S-Y-----	73	
HIVNH52	-----N-----S-Y-----T-----	73	
HIVQYI	-----S-----S-E-Y-----E-----	73	
HIVSF2	-----R-----P-S-Y-----	73	
HIVNY5CG	-----S-----N-----	73	
HIVNL43	-----S-----N-----	73	
HIVCAN1	---T---Q---S---P-LA---P---L---S---	73	
HIVHAN	-----A-----S-----S-Y-----	73	
HIVD31	-----S-----S-----T---L---M---	73	
HIVRF	-----S-----L-S-----	73	
HIVYU2	-----H-----R-----P-G-----	73	
HIVBCSG3C	-----N-----G---Y-----LV-T---TL---	73	
CONSENSUS.D	MEQAPeDQGPQREPYNEWtLELLeELKSEAVRHFPRLWHLGQHIYETYGDTWagVEAIIIRILQQLLFiHFr	73	
HIVELI	---A---A-----V-----	73	
HIVZ226	-----Y-----L-----	73	
HIVNDK	-----	73	
CONSENSUS.O	MEQAPeDQGPQREPYNEWtLELLeELK?EAVRHFPRLWHLGQYIYETYGDTWagVMAIIIRILQQLLFiHFr	72	
HIVMAL	---AD---Q---H-----Q-----I---S---H-----E---S---I-P---	73	
HIVANT70	-----K---A---A-----V-----	73	
HIVMVP51	---L-----E-----Q-C-----	73	
SIVCPZGAB	MEQAPEDQGPQREPYQENALELLEELKNEAVRHFPRLWHLGQGFYDYGDTWVCVEAIIIRILQHLLFIHFR	73	
		tat cds ->	
CONSENSUS.A	IGCQHSRIGIIP...QRRG.RNGASRS	96	
HIVU455	-----	96	
CONSENSUS.B	igcPhSRIGItr...qRRa?RNGASRS	93	
HIVLAI	---R---V-Q---	96	
HIVHXB2R	MMVST	78	
HIVMN	---R---I-----	96	
HIVSC	---R---IQ---S---S---	96	
HIVJRCSE	---R-----	96	
HIVJRFL	---Q-----	96	
HIVNH52	---Q---M---Q---R---T---	96	
HIVQYI	---Q-----	96	
HIVSF2	---Q---IQ---R---	97	
HIVNY5CG	---Q---I-----	96	
HIVNL43	---R---V-----	96	
HIVCAN1	---R---IQ---S---	96	
HIVHAN	---R---P-----	96	
HIVD31	---Q-----	96	
HIVRF	---Q-----	96	
HIVYU2	---Q---IQ---R---	97	
HIVBCSG3C	---R---Q---R---	96	
CONSENSUS.D	IGCQHSRIGItr...QRRR.RNGSSRS	96	
HIVELI	-----I-----	96	
HIVZ226	-----	96	
HIVNDK	-----S-----	96	
CONSENSUS.O	IGCQHSRIGI?p???rgRgrRNGSSRS	95	
HIVMAL	-----TR---QR-A-----	96	
HIVANT70	-----N-----	97	
HIVMVP51	-----L-SNT-----	100	
SIVCPZGAB	LGCQHSRIGILP...QRRR.SNGSNRS	96	

	\\3'sj \\3'sj	
	rev cds ->	
CONSENSUS.A	M?PVDPnLEPwnHPGSqPTTaCskCYCK?CCwHCqLCFLnKGLGISYGrKKR..r?RRgtEQ.s?kDhQnp	64
HIVU455	-E-----K-----N---V-----K-----K---.KP---P---GS---TL	68
HIVSF170	-D-----P-----V---Y---PA---R-----Q-----N-----	68
HIV1UG273A	-D-----K---P-----V---Y---PV---K-----SQ-----N-----	68
HIV1UG275A	-D-----P-N---K---Y---TV-----KQ---R---CN-----	68
HIV1DJ263A	-E---S-----I-----R-----RQ-----	68
HIV1DJ264A	-E---S-----N---I-----R-----RQ-R---	68
HIV1DJ258A	-----K-----R-----RQ-----	49
CONSENSUS.B	MeFVDPrLEPwnHPGSQPKtactnCYCKKCCfHCqvCFttKgLgISYGRKKR..RQRzrapq.dsqtHQvs	68
HIVLAI	-----T-----A-----P---G-----	68
HIVHXB2R	-----I---A-----H---N---A---	68
HIVTH475A	-----T-----M---A-----H---N---A---	68
HIVNL43	-----M---A-----H---N---A---	68
HIVMN	-----T-----K-A-----E-----	68
HIVSC	-D-----A---S-----	68
HIVBAL1	-----K-----T-----I-----E-----	68
HIVJRCSE	-----S-----L-----P-----	68
HIVJRPL	-----S-----T-----L-----I-----A---	68
HIVNH5	-----P---T-----L-----	68
HIVQYI	-----SN---R---L---K-----K-----	68
HIVSF2	-----N-----R---N-----YA---R-----A---	69
HIVMY5CG	-----R-----R-----A-----E-----	68
HIVSF162	-----I-----P---EV---	68
HIVCAM1	-----R-----Q-----I-----T---S-K---A---	68
HIVCDC4	-----K-A-----H---N---A---	68
HIVSF33	-----N-----R-----I-----Q---D-S-N---D---	68
HIVRJS	-----E-----H---N---N---IP	69
HIVHAN	-----R-P-----L---R---S-----	68
HIVD31	-----E-----K-----E-----	68
HIVRF	-----N-----Y---L-----GP---G-----	68
HIVYU2	-----N-----R---N-----K-----P---S---	68
HIVBCSG3C	-----K-----L---M-----A---	68
CONSENSUS.C	m?FvDP?LEPwnHPGSQPKTaCtkCycKkCsYHC1VCPgtKGLGISYGRKKR..RqRRsaP?.sSDHqnp	65
HIV1DJ259A	SV I--R-----N-----P-----T---	67
HIV1DJ373A	-E---N-----S-----L---	68
HIV1SM145A	---C---NN---L-----Q-----	56
HIV1ZAM184	-D---K-----C---P---LN-----P---RS-Q-N-----	68
HIV1UG268A	-----P---FG-----Q---LN-----N---P-----L---	54
CONSENSUS.D	mdPvDPnLEPwnHPGSQPrTpCNkC?CKKCCYHCqvCFttKGLGISYGRKKR..RQRrrppq.g?QaHQvp	66
HIVELI	-----H-----P---LN-----G-----G-----	68
HIVZ226	-----I-----K-A---R-H-----S---G-T---D---	68
HIVNDK	-----S-----A---H-----K---D-----	68
HIVMAL	-----Y-----M-----N---D---	68
HIV1UG269A	-----S-K---G-----	26
HIV1SE365A	-----K---N-Y-----N-----T---N-T---S---	68
HIV1UG274A	SV I--R-----S---Y---L-----SS---D---	67
CONSENSUS.O	MDPVDPE?FPWHHPGS?PQ?PCNNCYCKRCCYHCYVCF??KGLGIS?GRKKR...?RPAAS?S?PD?KDP	57
HIVANT70	-----V-----Q---I-----VR-----Y-----G-----H---H---	66
HIVNVF5180	-----M-----K---T-----TK-----H-----R---A-Y---N---	67
CONSENSUS.CPZ	MDP?D?P?PW?HP?P?T?CNMCYCK?CCYHC?CF?TKKGLGISYGRK?R?P?R?T?P?S?NNQD?	45
SIVCFZGAB	---I-PDLE--K--GSQ-R-V-----A-----IY-----K-T.TR---APA-G-K---S---	69

HIV1 TAT

	exon1/exon	
CONSENSUS.A	ipKQplPqtqq??ptgpkESkKkVeSKteTDrf?	95
HIVU455	-----S-RV.SA-QE-----AK----A	101
HIVSF170	----SI----V,-----A-----D	101
HIV1UG273A	-----I-----I.S-----Q-----A	101
HIV1UG275A	----I----V.S--E-----T	101
HIV1DJ263A	V-----T-R-D.--D-----E-A-----QGD	101
HIV1DJ264A	VS-----T-R-N,-----E-----K--PCD	101
HIV1DJ258A	V-----T-G-NH--D-----E-A--K-QYA	83
CONSENSUS.B	LSKqP7sqprGD.PtgPkesKkkVBrEtEtDp7dQ	100
HIVLAI	-----T-----F-	100
HIVHKB2R	-----T-----F-	100
HIVTH475A	-----T-S--,-----F-	100
HIVNL43	-----T-S--,-----F-	100
HIVMN	-P---AP-P--,-----H-V-	101
HIVSC	-P---A-A--,-----V-	101
HIVBAL1	-----T-H--,-----E-----VH	101
HIVJRCSE	-P---S-Q--,-----DN	101
HIVJRF1	-----S-----D-	101
HIVMH5	-----S-----,DS-----VH	101
HIVOI	-----A-----,-----E-	101
HIVSF2	-----A-S--,-----T-----P-	101
HIVNY5CG	-----A-----,-----DH	101
HIVSF162	-P---A--Q--,-----VH	101
HIVCAN1	-----A-PQ--,-----G--H-G-	101
HIVCEC4	-----S-T--,-----P-E--A--L-	101
HIVSF33	-----S-----,-----E-----L-	101
HIVRJS	-P---A-----,-----VH	101
HIVHAN	-P---S-Q--,DS--K-----A-F-	101
HIVD31	-----A-----,-----T-----I-	101
HIVRF	-----T-----,-----E-----AV-	102
HIVYU2	-----T-L--,-----T-----VH	101
HIVBCSG3C	--R-T--,-----V-	101
CONSENSUS.C	IsKQpLpqtgrgd.pTGpEeSKKkVESKteDpFD	98
HIV1DJ259A	-----,-----A-R--	100
HIV1DJ373A	-----SR-Q--,-----KA--L-	101
HIV1SM145A	-----,-----S-----K----	89
HIV12AM184	-P--S--H-QRV.S-----E--A--R--	101
HIV1UG268A	-----P--,-----K-----	86
CONSENSUS.D	iPkQpSSQpRGd?PTGPKeQKKKVESkAetDpFDw	100
HIVELI	-----,-----E-----C	101
HIVZ226	-----,-----E-----	101
HIVNDK	--E-----S--,-----K-----E-----	101
HIVNAL	L-E-----H--H-----A-Q--	102
HIV1UG269A	-----,-----A-----	58
HIV1SE365A	V-----H--,-----Q--	101
HIV1UG274A	-----M,-----K-----A-----	100
CONSENSUS.O	VP7QS??T7RK.Q7RQEQE??VER??GP7G7P??SC??CTR??S?Q	85
HIVANT70	--K--PTI-K--,--E-----EE--KA--G-Y-BRKG--HC--T-E-	114
HIVMVP5180	--E--LSH-G--,--K-----KK--ET--S-Q-CHQD--NS--I-G-	115
CONSENSUS.CPZ	??K????SR??,??K??K??V??KT??G?C	56
SIVCPZGAB	IP-QPLSQ--GN.KEGSE--T-E-AS--EADQ D	101

Sequence subtypes C and E are not currently represented in REV.

	5' 3' sj	exon/exon	
CONSENSUS.A	MA7RSG??DE.?LLKA?RIIK?LYQSNP?P?P?G.SRQARKNRRRRWRARQQRQIDS?SERILS7CLGRPAEP		59
HIVU455	--R--NP--D--V--L--C--N--R--		70
HIVSF170	--G--DS--E--I--I--Y--K--K--		70
CONSENSUS.B	MAGragdsede.ellktVRLIKfLYQSNpPpPepG.TrqARRNRRRRWRARQQRqIraISewllstyLgRsaef		70
HIVLAI	-----D--A--R--		70
HIVHXB2R	-----IR--L--N--		70
HIVNL43	-----IR--L--N--		70
HIVMN	-----S--		70
HIVBRVA	-----V--R--A--Q--L--G--F--		60
HIVSC	-----E--K--		70
HIVJH3	KKRRQRR.RAAQA-----Q-----G--I--N--P--		67
HIVBAL1	-----IL--		70
HIVJRCSP	-----D--N--		70
HIVJRF	-----L--		70
HIVNH5	-----PT--R--RN--D--P--		70
HIVQYI	VPLQLPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		70
HIVSF2	-----R--L--		70
HIVNY5CG	-----Q--F--L--K--		70
HIVSF162	-----A--L--L--SK--		70
HIVCAM1	-----I--I--H--K--		70
HIVCDC4	-----G--R--TA--I--I--H--		70
HIVSF33	-----I--I--I--N--		70
HIVRJS	-----N--PT--		70
HIVHAN	-----R--F--P--		70
HIVD31	-----Q--Q--R--PEQ--		70
HIVRF	-----R--D--A--S--		70
HIVYU2	-----D--R--V--S--		70
HIVBCSG3C	-----L--D--N--		70
CONSENSUS.D	MAGRSGDsDE.dLLKa?RLIK?LYQSNPpPpPepG?TQARRNRRRRWRARQQRQIhaIgERILsTyLGRpeEP		68
HIVELI	-----V--F--		70
HIVZ226	-----R--V--I--		70
HIVNDK	-----N--I--F--		70
HIVMAL	-----R--I--I--MT--T--		71
CONSENSUS.O	MAGRSE?.D7.QLLQAIQILIKLYQSNP?P?P?G.SRNARKNRRRRWRARQQAQVD?LA?R?LATVVG?Q?N		59
HIVANT70	-----D--Q--S--R--		68
HIVMVP5180	-----E--Q--C--T--A--		69
STVCPZGAB	MAGRSEPDQDARLLQAVKIILYQSNPpPpPepG.TRKARRNRRRRWRARQQRQISETISGRVLATYLGRRPKP		71
CONSENSUS.A	VPLQLPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		100
HIVU455	-----I--R--SC--G--P--T--C--IS--V--N--		123
HIVSF170	-----L--H--DR--E--S--A--R--VP--I--K--		123
CONSENSUS.B	VplqLPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		116
HIVLAI	-----N--		116
HIVHXB2R	-----N--		116
HIVNL43	-----N--		116
HIVMN	-----Q--		115
HIVBRVA	-----N--		106
HIVSC	-----N--		116
HIVJH3	-----N--		113
HIVBAL1	-----D--S--R--		117
HIVJRCSP	-----L--N--N--E--T--		116
HIVJRF	-----N--		116
HIVNH5	-----N--		116
HIVQYI	-----N--N--A--		116
HIVSF2	-----N--A--		116
HIVNY5CG	-----N--		116
HIVSF162	-----N--K--		116
HIVCAM1	-----D--		116
HIVCDC4	-----D--		116
HIVSF33	-----D--		116
HIVRJS	-----N--		111
HIVHAN	-----P--		116
HIVD31	-----Q--S--		116
HIVRF	-----F--		116
HIVYU2	-----N--G--N--		116
HIVBCSG3C	-----N--G--N--		116
CONSENSUS.D	VPLQLPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		115
HIVELI	-----R--		118
HIVZ226	-----A--		118
HIVNDK	-----T--N--		116
HIVMAL	-----T--N--		117
CONSENSUS.O	N?VDLPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		78
HIVANT70	-----I--S--G--Q--SEANTVDPRADN\$CLQNLCSN-ILATRIAE		120
HIVMVP5180	-----L--N--A--R--PGTG--		103
STVCPZGAB	GDLELPPLERLIdcseDCGtSGTg.....?GVGsPqILvESpavLesGtKE		124

HIV VPU

Sequences of the A, C, and E subtypes are not currently available for the VPU

				env cds >	
CONSENSUS B	mqslq12aivaiivvaiaiaIvVWt	iv2ieyrKilRQR		>kidRl2idrireraeDsGNeSeGdges	61
HIVLAI	-PI-- --A--- I --S --I-----			- - --LI-- -----	59
HIVMN	P V-A -----G - S --F-----R---			- - - S - -----	62
HIVSC	P -LS----- --S --P-X-----			- - --X----- IRK	63
HIVBRVA	P -S----- --S -ALL--L-----			----- - -----	63
HIVJH3	---A ----V-----S --F--			-----	62
HIVBAL1	A S --G----- F - - - - -			-----T----- D	63
HIVCDC4	E-V --E----- F - - - -			---VN- ---E--N -	64
HIVQYI	1Y - L --VT----- --LL-----			-----D-----	63
HIVSF33	T L --I--VT----- I - -L-K-----			-----	62
HIVSF2	- L --S--V----- L - - - - -			X--- --K-----	63
HIVNY5	L X - - - - - X - - -P-----V---			-V--I-----	49
HIVNL43	-PIEV -----I--- --S --I-----			--- --LI----- -E	59
HIVCAM1	-LP--X-----V-----F-----R--K			QV-- --I-----	63
HIVJRCSP	--P--L-- --G---I--S --L-----			--- --K-----	63
HIVJRF1	---Y-L -----I-----S --L-----			--- --T-----	63
HIVSF162	-P- L -----V-----			--- --T-----	63
HIVHAM	-L---A -----V-----V-- - L---			-N-- --LT-----	63
HIVD31	---VA -----V-----F-----			--- --I-----	63
HIVRF	---E-L-----L-----G--I--T--K			X- ---D--E -	64
HIVBCSG3C	--PI A - -VG----- X - -			--- --K--SK-----	56
CONSENSUS D	WQPL2IIAIAALVVAIIIAIVVWT	IVFIEYRRIRKQR		KIDCI 1DRIRERAEDSGNESEGdREE	62
HIVELI	---G-----L-----			R--- L---T----- --K	63
HIV2226	---SQ-- --A-----			---I-----	63
HIVNDK	---V-----Y-----			--- --E--	63
CONSENSUS O	mh2221Aaiaalsail2Inv22W2fi2222lRkyl2Qr2qdr2e2iier	lrRlrEardDSdyESngseeqe			57
HIVMAL	IQPLVI--V--VVT-I-AIVV-T -VFIBI--IRR--			K-DRL ID---RAB--GN- E DT-E	63
HIVANT70	--HRD-----II---F---IL-G--			---E-KE--K-R-----	68
HIVMVP5180	-QEN---L---CL---LI-L-N --i--V--K---R-Q-----			--- --K-----E- Q-	68
SIVCPZGAB	MTLLVGL VLLLVGLIAMNICIMGYIRKGYRRYKRHR			LETEIERL NLILAERAEDSGNESNGEEZER	67
CONSENSUS B	LSAL222vem2GH2Apwd2ddl				77
HIVLAI	I -VENG--- --H---I---				81
HIVMN	--- --G- --D---VIN--				80
HIVSC	--- --D--G-I---				81
HIVBRVA	- ---G--D---I---				82
HIVJH3	-- --- --D---I---				80
HIVBAL1	--- --- --H---IN--				81
HIVCDC4	--- --- --H---NV---				82
HIVQYI	-- --- --H---V--M				81
HIVSF33	V- --- --D---VN--				80
HIVSF2	--- --- --L---V---				81
HIVNY5	--- --- --D---VN--				67
HIVNL43	V---VENG--- --H---I---				81
HIVCAM1	--- --M-- --N---VN--				81
HIVJRCSP	--- --R --L---IN--				81
HIVJRF1	--V- --R --L---IN--				81
HIVSF162	--- --R --L---V---				81
HIVHAM	--- --R --L---VN-Q				81
HIVD31	--- --H---V---				81
HIVRF	--- --R---V---				82
HIVBCSG3C	--- --- --H--L--I---				74
CONSENSUS D	LSKL VEM GHAPWD1DDL				80
HIVELI	---				81
HIV2226	---				81
HIVNDK	---				81
CONSENSUS O	vm2L v2shgfdn2m2Fe2l				72
HIVKAL	LSK- -EM -H-A-WDVDD-				81
HIVANT70	--D- -L-----P				85
HIVMVP5180	--E- IH---A--- --				85
SIVCPZGAB	LEQLIHNYNNHNFANFM FD L				88

This page intentionally left blank.

HIV ENV

← vpu cds
signal peptide/gp120

```

CONSENSUS.A      Mrvmgi.q?nyq?lwr????...Wgtmilg??iIcna??e??lWVtVyyGVFPVWkdaetTLfcASdAk      52
HIVSF1703        -----M-C-N-L-----ML--S--TS.K-----      58
HIVM455          -----R-PC-----LI-----Q.Q-----V-----      57
HIV2321          -K-K--G-W-MW-K-----L--LV--S-A-N-----      58
HIVRW020         -R--TSW-N-----MLV-YS-A-N--A-----G-----      56
HIVUG031fr       -R--R-W-H-L-----I--MI--S-RTD.K-----E-----      57
HIVMI            -----K-PL-----N-FWIM-----Q-----T-----      57
HIVRW009         ---T-LM--N-----ML--S--AN.N-----LG-----      56
HIVUG275A        -KA--T--M-C-MW-----ML--ST.A--D-----R-----      58
HIVUG273A        ---K--R-W-H-L-----I--LIT--ST.AK.N-----R-----      58
HIVV1191A        ---K--M-W-H-K-----L--LG-----N--F-----      57
HIVDJ264A        ---R--R-PLF-K-----MI-FWIM-----K-----R-----      57
HIVDJ263A        ---M--R-PL--K-----I-FWIM-----K-----RNT-----      57
HIVDJ258A        ---K--PL-----I-FIM-----K-EM-----R-K-----      56
HIVCARGAN        ---K--H-L-----F-LI--SCV.AD.Q-----N-D-----N-R      58
HIVCAR5AS        ---K-T--K--SP-----I--LI--S--T--K-----K-----      58
HIVCAR4054       ---KET--R-WHN--K-----RL--F-MLM-----K--N-----RE-N-----      58
HIVCAR286Afr     ---K-T--R-W-N-----L--F-MLM-----TN.N-----R-----      58
HIVCAR4023       ---KE..K--H--WGWK-----S-L--ILM-----DKDK-----R--N-----      60
HIVCAR423A       ---KE..K--H--WGWK-----S-L--ILM-----SD.K-----R--N-----      59
HIVKENYA         ---T--R-C-H-L-----MI--S--A-N-----N-----      58

```

```

CONSENSUS.B      ??mrvkgi.rknyqhlwr????????WgtmLiGmimicsa??e.?lWVtVyyGVFPVWkeatTLfCAsdAk      56
HIVJRCSF         -----K-----G-IL--T-----V--K-----T-----      58
HIVJRFL          -----S-Y-K-----G-L--I-----V--K-----T-----      58
HIVLA1           -----R--R-----I-----I-----T--K-----      58
HIVBRVA          -----K-----W-----G-M--I-----TD.K-----N-----      59
HIVJH3           -----R-----I-----I-----A-Q-----A-----      58
HIVBAL1          ---TE--S--W-----I--I--N--E--K-----R-----      58
HIVYU2           ---ATE--K-----G-L-----A-Q-----      58
HIVMN            ---R--W-G-----L-----T--K-----      58
HIVHXB2R         ---E..-----WGWR-----T--K-----      59
HIVLA1           ---E..-----WGWK-----I-----T--K-----      59
HIVNL43          ---E..-----WGWK-----I-----T--K-----      59
HIVMFA           ---E..-----WGWK-----I-----T--N-----      59
HIVCAM1          ---A-----C-R-----AD.K-----      58
HIVNY5CG         ---A-T-----A-Q-----      58
HIVADA           ---K-----K-----I-----V--N-----      58
HIVJFL           ---E-----C-----AG.K-----      58
HIVSIM184        ---K-E-----C-----V-Q-----      58
HIVD31           ---K-E-----G-L-----T--K-----      58
HIVSF162         ---E-----G-L-----V--K-----      58
HIVBCSG3C        ---E-----F-I--Y--AG.N-----      58
HIVOYI           ---TAR-T-----R-----A--N-----R-----      58
HIVSF33          ---ARET--C-----A--N-----D-----      58
HIVCDC4          MA--A--C-----AA.N-----      60
HIVSF2           ---K--T--R-----L-----T--K-----R-----      58
HIVSF2B13        ---K-ET-----IL-----K--K-----R-----      58
HIVHAN           ---K-ET-----R-----G-L--S-V-AG.N-----      58
HIVRP            ---MEM--C--K-----A--D-----E-----      58
HIVMAJ2          ---NR-C--I-----F--W--V--Q-----      58
HIVTB132         ---E-----C-----L--I--S--A--N-----      58
HIVBR020         X-EM--HW--W-K-----G-IL-----KD.N-----G-----      58
HIVTH014         X-T-----RQ-----I-----A--K-----      58
HIVUS711fr       X--T-----K-----A--N-----I-----      58
HIVUS712         X-TE--C--W-----I-----P-NM.A--ES-----      58
HIVUS715         X-T--C--W-K-----I--I--N--T--K-----N-----      58
HIVUS716         X-T--C--W-K-----M--I--NT--N-----      57
HIVHA593         XA-EM--HC-----R-I-----T--A--N-----      58
HIVHA594         X-T-----K-----I-----V--Q-----      58
HIVHA596         X-T-----K-----I-----V--Q-----      58
HIVHA599         X-----C-----W--L-----T--RF-----D-I-----      58
HIVHA651         XA-EM--HW--T-----I-----S--Q-----      57
HIVRJS          ---ARE--T-W-----T--N-----      58
HIVGUNfr         -----S-CR-----I-----A--K-----      29
HIVSC            -----S-CR-----I-----A--Q-----      58
HIVBR014         XA--C--WCW-----R--I-----T--K-----      59
HIVTH026         X-T-----AN.N-----G-----      58
HIVBR021         ---C-R-----I-----T--BED.K-----      59
HIVACH9          -----      18
HIVACPI          -----      18
HIVJM            -----      18
HIVMM            -----      18
HIVMA208         ---R--R-----WGTMLLGI-----AD.N-----      66
HIVP896          ---E--W-----G-IL-----AK--KT--I-----R-----      58

```

	<- vpu cds signal peptide/gp120		
CONSENSUS.C	mrVrgi.lzNwqqwWi.....WGILGFWMlmicnv??g.nLWVTVYYGVVWkeAktTLFCASDAK		57
HIVMA959	X---P-----F---S---W-----D-----		58
HIVMA960	X---S-----YS-L-----		58
HIVZAM20A	-----C-----G-V-D-----		58
HIVD1024	-----Y-G-----V-----		58
HIVBR025	-----E-K-----		28
HIVSM145A	-----SG-G-----R-----S-----		58
HIVZAM18A	---E---V-NY-VGD-----		59
HIVDJ259A	-K-M-T-Q--PR-V-----L--G-----D-NPP-----		56
HIVDJ373A	---M---QM-C-----G-----LM-G-M-----Q--NP-----		58
HIVSE364A	---T-M---C-P-----RL-Y--G-----		58
HIVUG268A	---M---Q-C-----I-----M-----T-----		58
CONSENSUS.D	mrvrei.erMyQhlw????????WGtmllGmlmtcs??as.??lWVTVYYGVVWkeAktTLFCASDAK		55
HIVJY1	---MG---RM-----K-----I-----I-----V---D-----		58
HIVMDK	---A-K---C-M--K-----I-----A-----D-----I-----		58
HIVMAL	---Q---NW-R-----M-----I---D-----		58
HIVELI	---A-Q---C-NW-K-----I-----I-----A--D-N-----		58
HIVZ226	---G---C-M--K-----I-----I-----M--D-N-----		58
HIVUG024	---N-Q---K-----I-----I-----GTE-QS-----		59
HIVUG021	X---N-R-C---G-----L--E---EY-T.TG.N-----		58
HIVUG038	---M---M---S-R-----L-----S---K-----N-----		57
HIVUG046	---MG---ERNQHSLSWR--I-----I-K--G-----		64
HIVUG266A	---KG-P---N-LR-----II---III--T--D-N-----D-E---G-----		58
HIVUG269Afr	-----R-----I-RV.TG.E-----		58
HIVUG274Afr	-----I-----Y-V---Q-----		58
HIVSE365A	---A--N-K---R-----V---K-----		58
HIVCAR4020	---G---R-----I---V---K-----E-----		58
CONSENSUS.E	mrVket.qpanwpnlWk????WGtlilGlviccsA.Sd.nLWVTvYYGVVWkeAktTLFCASDAK		58
HIVTH966	X---I-----		58
HIVTH975	X-----L-----		58
HIVTH022	X-----N-----K-----		58
HIVTH243	X-----G-----G-----		58
HIVTH006	X-----x---M-----		58
HIVTH011fr	---G---H-----R--E-----A--S-----		58
HIVCARELO	---K.YQH--RWGWR---SML--ILM-----A-X-----		59
HIVCAR4017	---RG--K-----M-----K-N-----		58
HIVCAR4039	---G---M-----		58
HIVCAR4071	---G---T--W-----N--N-----K--C-----		58
HIVCARMBA			
CONSENSUS.F	MRVRGM.QRNWQHLSK.....WGILLFLGIL?ICNA..E.nLWVTVYYGVVWkeAATTLFCASDAK		56
HIVBR7944	---EFG-----		28
HIVB2163Afr	-----L-----D-----		57
HIVB2126Afr	-----I-----		57
CONSENSUS.G	MRvkGI.qrNwqhlWk.....WGTLILGLVICCSA.sn.nLWVTVYYGVVWkeAdtTLFCASDAK		58
HIVV1525A	---R---T--N--N-----S-----E-----		58
HIVLBV217	---A---L-----I-----N--Q-----A--T-----		58
HIVCAR4067	-----T--S-----R-A-----R-----		58
CONSENSUS.O	MtvtMK.am?KrNrKlw.....?ylamALiiPCL??S??QLYATVYsGVVWkeA?PvLFCASDAN		54
HIVANT70	---I---E--K-----T-----T-----LR-----A-----		59
HIVRVPS180	---V-X-N--S-----S--I---L---SY-K-----E-A-----		60
HIVVAU	---AI---G-----G-----IWC-IL-----CN-----K--T-----		59
CONSENSUS.CPZ	M7?????.?????????.....?????????I?????T.??E.???VTVYsGVVWkeA?P?LFCASDAK		20
SIVCPZGAB	-KVMKK.KKRDWNSLS.....IITIIIT-LLTPECL-.S...LW---Y-----HD-D-V---SD-K		58
CONSENSUS.U	MrVm?i.qrNcq?lw?.....wgililgmi?ICNA.???nLWVTVYYGVVWkeAdtTLFCASDAK		51
HIVZAM184	---RG--PK-W-QW-I-----GFN-LM--SV.VG-----		58
HIVCAR4081	-K--ET-----PRS-R-----LVI-----SDN-----K-----		59
HIVK124A	---G---L-LG-----G--K---I---A-----R-----		57
HIVZ3fr	---KE---Y-H--K-----SLI---M--K--IE.KS-----K-----		58

HIV1 ENV

CONSENSUS.A	AydrE?hmvwaTHaCVPTDpNpQeI?Le NVTE?FnmwKnmMVeQmheDiisLWDqSLkPCvkLtPLCVTL	119
HIVSF1703	--ER-V-----D---Y---G-----T-----	128
HIVU455	--A-V-----D-V---E-----D-----	127
HIV2321	---K-----LS-G---K-D-----V-----	128
HIVRW020	---V---E-----H---D---R-----T-----	126
HIVUQ031fr	--K--M-----S--MH-K---R-----T-----	127
HIVD687	-----V-----H---K-----	120
HIVNT	-----V-----H---K-----	127
HIVRW009	---P-K-----T---H---E-----T-----P---	126
HIVUG275A	---M-----N-G---K-----	128
HIVUG273A	---E-M-----R---E-I-----T-----	128
HIVVI191A	--SE-K-----D-K-N-V---D---K---Q---	127
HIVDJ264A	---V-V-----H-K---K-----E---	127
HIVDJ263A	---A-V-----H---K-----	127
HIVDJ258A	---V-----Y-K---K-----	126
HIVCARGAM	---V-----N-V---E-----T---T---Q---	128
HIVCARSAS	---M-----MD-K---N-----T-----Q---	128
HIVCAR4054	--HS-K-----K---N-----E---	128
HIVCAR286Afr	--SA-NMSG-----P---N---R-S---ID---E---IE---	128
HIVCAR4023	--S-K-----T---D-R-P-K---N---S---Q---E---	130
HIVCAR423A	--S-K-----S---L---N---S-D---E---	129
HIVKENYA	--E-K-----H---E-----T-----	128
CONSENSUS.B	aydrEvHNVwathacvPTdPnPgEvvlv?NVTEnFnmwKnmMVeQMhEDIisLWDqSLKPCVKltPLCVtL	126
HIVJRCSF	-----D-----Q-V-N-----	128
HIVJRFL	-----H-----Q-----	128
HIVALA1	-----I-----	128
HIVBRVA	---f-----L-MG-----D-----	129
HIVJK3	-----K-----	128
HIVBAL1	-----E-K-----	128
HIVYU2	-----K-----	128
HIVM6J	---Q-----E-V-----	128
HIVMXB2R	-----V-----D-----S---	129
HIVLAI	-----V-----D-----S---	129
HIVNL43	---V-----D-----S---	129
HIVMFA	---C-----I-V---D---D-----N---	129
HIVCAM1	-----	128
HIVMY5CG	-----Q-----	128
HIVADA	-----K-----D-----	128
HIVJPL	-----K-----D-----	128
HIVSIMI84	---S---I---K-----	128
HIVD31	---K-----D-----D-----K-----	128
HIVSF162	-----I-----	128
HIVBCSG3C	---I-----I---D-----Q-----	128
HIVQYI	---A-----G-----D-----Q-----	128
HIVSF33	---G-----G-----D---V-----	128
HIVCDC4	---A-----N-----	130
HIVSF2	-----G-----Q-----	128
HIVSF2B13	-----G-----Q-----	128
HIVHAN	-----MG-----D---Q-----	128
HIVRF	---K---A---K-----L-----	128
HIVMAJ2	---S---A-----I-G-----	128
HIVTB132	-----S---A-V---I-T-I-----	128
HIVBR020	---KA-----I-----	128
HIVTH014	-----G-----	128
HIVUS711fr	-----E-----S-----	128
HIVUS712	---I-----	128
HIVUS715	-----D---E-----	128
HIVUS716	-----D-----	127
HIVHA593	---E-----L---Q-----	128
HIVHA594	---R-----D-----A---	128
HIVHA596	---R-----G-----A---	128
HIVHA599	---A-A-----K-----D---D-N-----TI---	128
HIVHA651	---H-----ME-K-----	127
HIVRJS	-----K-K-----	128
HIVGUNfr	-----N-----S-----E-----	99
HIVSC	---I-----G-----	128
HIVBR014	\$-----R-----GK-----D-----	130
HIVTH026	-----G-----E---Q-----	128
HIVBR021	---M-----D-----	129
HIVACH9	---A-----S---L-----	88
HIVACP1	T-----R---T-----D-----	88
HIVJM	-----H-I-P-I-----	88
HIVNM	---R-----G-----N-----	88
HIVMA208	---K-A-----LP---\$-----E---	136
HIVP896	-----G-----D-----E---	128
HIVBWB11A	-----	1

CONSENSUS.C	AYa?EwHNVWATHACVPTDPnPQEmvle.NVTenFNMWkNnmVQNMHeDiISLWDqSLKPCVKLTPLCVtL	126
HIVMA959	--R-----I-W-----I-----	128
HIVMA960	--Q-----V-G-----V-----	128
HIVD757	-----G-----R---N---V-----	20
HIVD747	-----Y---G-D---I---G-----V-----	63
HIVD760	-----V-----G-----V-----	64
HIVMOF	-----V-----	9
HIVZAM20A	--T-----S---A-----M-----	128
HIVD1044	-----	7
HIVD1024	--E-----N---V-----	128
HIVD744	-----RV---T-----	20
HIVD766	-----	4
HIVD808	-----P-----	21
HIVD868	-----P-----	18
HIVBR025	--DT-----E---E---Q-----G-----	98
HIVSM145A	--R-----S---L-G-----	128
HIVZAM18A	--R-----I-G-----R-----	129
HIVDJ259A	--KV-----S---LA-----G---Q-----	126
HIVDJ373A	--T-----S---L-M---G---Q-----EG-----	128
HIVSE364A	--R-G-----W---Y-----S-----Q-V-----	128
HIVUG268A	--T-----I---S-----Q-V-----	128
CONSENSUS.D	sY?EaHNIWATHACVPTDPnPrEiale.NVTenFNMWkNnmVQNMHeDiISLWDqSLKPCVKLTPLCVTL	123
HIVJY1	--EP-----M-----N-----	128
HIVMDK	A-KK-----Q-----	128
HIVMAL	--ET-V-----Q-----G-----	128
HIVELI	--ET-----Q-A-----	128
HIVZ226	--KT-----Q-----R-----	128
HIVUG024	A KA-----Q-----V-----	128
HIVUG021	--EA-----Q-V---I---D---P-----	128
HIVUG038	--IX-----	127
HIVUG046	--EP-----K-----D---V-----	134
HIVUG266A	A-DT-V---V-----K-----R---E-----	128
HIVUG269Afr	--KKG-----XK-----E-----	128
HIVUG274Afr	--IA-----D---R---D-----	128
HIVSE365A	--ET-K-----Q---A-----	128
HIVCAR4020	--KA-----Q-M-K---S-----	128
CONSENSUS.E	AKeTEvHNVWATHACVPTDPnPQeiyile.nvtenfNMWkNnmVQNMHeDiISLWDqSLKPCVKLTPLCVTL	128
HIVTH966	-----H---I-----	128
HIVTH975	-----H-----	128
HIVTH022	-----E-----	128
HIVTH243	-----N-----	128
HIVTH006	-----M-----H-----	128
HIVTH011fr	-----x-T---R---R-----	102
HIVCARELO	-----N---H-----	128
HIVCAR4017	--K-----M---A---I---ER-----	129
HIVCAR4039	--K-----M-----I-----	128
HIVCAR4071	--I-----F-----	128
HIVCARMBA	--V-M-----HTQ---V-M-----H-I---E-----	127
CONSENSUS.F	SYE?EVHNVWATHACVPTDPnPQEvVle.NVTENFNMWkNnmVQNMHeDiISLWDQSLKPCVKLTPLCVTL	126
HIVBRA7944	-----D---E-K-----	98
HIVB2163Afr	-----S-----	127
HIVB2126Afr	--K-----	127
CONSENSUS.G	sYSTEShNVWATHACVPTDPnPQEIle.NVTENFNMWkNnmVQNMHeDiISLWDQSLKPCVKLTPLCVTL	128
HIVV1525A	-----K-----E-----	128
HIVLBV217	A-----N-----	128
HIVCAR4067	--K-----V---E-----	128
CONSENSUS.O	LTSTEqHNIWAsQACVPTDPF?P?EypL?.NVTd?FNIWKNYmVQNMHeDiISLWDQSLKPCVqMTPLCVQH	120
HIVANT70	-----K-----T-Y---H---D---E---Q-----	129
HIVMP5180	-----N-H-F-G---N-D-----E---EK-----	130
HIVVAU	-----T-----S-N-E-K---GK-----I---D-----	129
CONSENSUS.CPZ	7?6TE?HN?WAT??CVF?DF?P?V?V?L?.N?P??PN?XN?MV??M?ED??L??QS?KPCVKLTPLCVTL	58
SIVCP2GAB	AH---A---I---QA---T---S-QE-F-P---VIES---MN---N---DQ-H---ITS-WD---L---L-VTL	128
CONSENSUS.U	aYETe?HNVWATHACVPTDPnPQEIvle.NVTENFNMWkNnmVQNMHeDiISLWDQSLKPCVKLTPLCVTL	120
HIVZAM184	--R-V-----D-D-----	128
HIVCAR4081	S-S-K-----L---X-----E-----	129
HIVK124A	--K-M-----K-----T-----	127
HIVZ3fr	--X-S-----S-S---L-G---K-----F-----	128

HIV1 ENV

CONSENSUS.A	7C????????n????????n????????n????????m?????.....??e..lkn	130
HIVSP1703	N-SH.....-ITTH.....-I-TMNP.....-I.....KY-E.....G.....	156
HIVU456	D-H.....-ITI.....-N-N.....-N.....NETD.GWR.....E..M--	153
HIV2321	S-H.....-I-IDK.....-M.....NVDE-..K.....E.....	152
HIVRW020	D-M.....ATAS.....-V-N.....-L.....SD-..N.....G.....	141
HIVUG031fr	N-T.....TVD.....-T-NV.....-IS.....KD-I.....G.....	150
HIVD687	N-S.....DVT.....-N-FS.....-VS.....YD-..K.....G.....	43
HIVNI	D-HNX.....-NSYS.....-SS.....-L.....SD-..N.....G.....	152
HIVRW009	E-M.....-ITNW.....-M-V.....-I.....DO-..K.....G.....	150
HIVUG275A	N-SDVN.....-VTVT.....-HSI.....-SSD-..K.....G.....	153
HIVUG273A	N-TNA.....-LTNA.....-KNKVNSAT-G.....IQS-E.....R.....	161
HIVV1191A	E-H.....-S-TPD.....-M.....TQSTII.....D-..TT-	150
HIVDJ264A	D-HN.....-V-SN.....-R.....ISRD-..Q.....G.....V--	150
HIVDJ263A	D-HNV.....-SS.....-SSTSI.....-R.....IDSK-..Q.....E.....	155
HIVDJ258A	E-H.....-VENS.....-TSSYR.....-IS.....SE-..R.....E.....	152
HIVCARGAN	E-H.....KP.....-T-DSTK.....-NSAE.....NMT..FAGI.....G.....-T-	157
HIVCARSAS	D-T.....-VA.....-DST.....-N.....LR-	143
HIVCAR4054	N-TDVKK.....-ST.....-.....LEQQ.....V..M--	147
HIVCAR286Afr	N-TDAKIDCTNGR-NSAGGOTCTGVQ-S-A.....EVGQ.....V..M--	167
HIVCAR4023	N-TDVK.....-ITV.....-S-S.TGNSA-TI.....NRT..VDSQ.....VG.....	164
HIVCAR423A	N-T.....-VTVT.....-SSN.....-V-KT.....NSSQSSKYNIDG.....KE.....	164
HIVKENYA	N-SNV.....-V-AT.....-SG.....THGGR.....GR..M--	151
CONSENSUS.B	nCtd?????????..?????????..??n?e????????n?????????mek?.....ge..lkn	141
HIVJRC5F	--K-V.....-A.....-TTSSS.....EGM--R.....G.....	153
HIVJREF	--K-V.....-A.....-TTNDS.....EGT--R.....G.....	153
HIVALA1	H-I-YL.....-A.....-TTSSS.....GTK--R.....G.....	154
HIVBRVA	--H-F.....-A.....-ATENS.....GKM--G.....M--	154
HIVJH3	--I-WG.....NDTSP.....-A.....-TTSSG.....GEK--G.....M--	159
HIVBA11	---LR.....NATNG.....-D.....-TTSSS.....RGMVGG.....M--	159
HIVYU2	---LR.....-A.....-TTSSS.....WET--G.....M--	154
HIVM6	---LR.....NTTNTN.....-S-AN.....NSNS.....EGTIKG.....M--	161
HIVHXB2R	K---LK.....-D-NT.....-SSSG.....RMI--G.....M--	156
HIVLAI	K---LG.....NATNT.....-SSNP.....-SSSG.....EMM--G.....M--	161
HIVNL43	K---LK.....-D-NT.....-SSSG.....RMI--G.....M--	156
HIVMFA	K---LK.....-D-NT.....-SSNG.....RMI--G.....M--	156
HIVCAM1	I---NV.....-M-RT.....-SSDWRD.....EGEKM.....M--	157
HIVNY5CG	---LT.....-A-YA.....-GSS.....B-R.....R--	151
HIVADA	---LR.....-V-NI.....-NSS.....EG-..R.....M--	152
HIVJFL	--INLK.....-V-NI.....-NS.....WKG--E.....M--	153
HIVSINI84	---NMKSA.....-S-SNT.....-SS.....LE--G.....M--	154
HIVD31	---LK.....-A-NTW.....-SSWTN.....TGE--G.....M--	156
HIVSF162	H---NLK.....-A.....-TKSSN.....WKE-DR.....M--	154
HIVBCSG3C	---ELK.....-A.....-TTSTNTPSGSWKKMER.....M--	160
HIVOIY	D---V.....NTTSSSLR.....-A.....-TTSSS.....WET--G.....L--	161
HIVSF33	---YLG.....-A-NT.....-NSSGG.....TV--G.....E--	155
HIVCDC4	---LNTN.....-T.....-TTLSIIIVWEORG.....M--	162
HIVSF2	---LG.....KA-NT.....-SSNWK.....EEI--G.....M--	154
HIVSF2613	---LG.....KA-T.....-SNWK.....EEI--G.....M--	153
HIVHAN	K---YN.....-A.....-SSMGR.....G-ID.....Q--	151
HIVRF	---ANL.....-G.....-VTSSSG.....GTM--N.....M--	156
HIVW6J2	--I-K.....-I-DWK.....-TTI.....IGG.....V--	151
HIVT8132	---NLR.....-D.....-.....S-..M--	142
HIVBR020	--I-LN.....NSTN.....NN-SS.....GVKTGID.....M--	156
HIVTH014	---LR.....-A-NT.....SST.....-G.....M--	150
HIVUS711fr	---KL.....I-A-YS.....-IS.....LEGEI.....D--	154
HIVUS712	---LS.....-A-NTN.....ATTTNSS.AGMM-DR.....M--	160
HIVUS715	---NLR.....-D-NTT.....R-AT..NTTSSETM--E.....M--	160
HIVUS716	K---LG.....-A-NTT.....-SSGRM--G.....M--	152
HIVHA593	E-H-VNMGATN.NGT.....T-V-ESGVNSDVTSMNVNMGATMEK.....M--	176
HIVHA594	---VE.....-S-NAT.....-SS.....GGIID.....M--	154
HIVHA596	---VK.....-S-NVT.....-SS.....GGII-Q.....M--	154
HIVHA599	H---AR.....-S-NVT.....IEV.....EKV--D--	153
HIVHA651	---AAE.....RNDTG.....SS-SS.....G-ID.....Q--	152
HIVRJS	---NMK.....-D-DT.....-TT..NSSS.GEK--E.....M--	157
HIVGONfr	---NLV.....-A-YS.....-SS.....WKK-DR.....Q--	124
HIVSC	---NLR.....NDTST.....-A.....-TTSSN.....RKK--G.....MT-	159
HIVBR014	---V.....NVINT.....-N.....TST-SS.....RLBE--G.....M--	161
HIVTH026	---LR.....-A.....-TSST.....-Q.....M--	150
HIVBR021	---NVNT.....NST.....-T.....STNAT..SSWEN--E.....M--	160
HIVACH9	---NL.....-S-NLG.....-NT.....DI--G.....M--	111
HIVACP1	---DLR.....-D.....-.....WEQ--G.....M--	108
HIVJM	--N-ASR.....NVTKR.....-D-NENATDS-NGTATSON.KTM--Y.....A--	131
HIVWM	---CVTS.....-C-NLR.....-VTNPKSSS.KEI--G.....M--	122
HIVMA208	---APIR.....TNT.....-T-MT.....-SS.....GEVK-R.....E--	166
HIVP896	---NL.....NITK.....-T.....-PTSSS.....WGM--G.....M--	157
HIVBWB11A	---NDLG.....DATNT.....-SS.....SEGM--I.....M--	28

HIV1 ENV

	peptide T region	
CONSENSUS.A	CsfnmTteirdkqkvyslFYrLDVvqi?????n?????n?????YrLinCnSaitQA	174
HIVSF1703	--Y-----K-----P-DK-----MG-----MN-----IRTQ-----	210
HIVU459	-----I-----KTD-----NS-----T-----	202
HIVZ321	--Y-----QR-I-----I-P-GG-----SS-NGDSSK-----	205
HIVRW020	-----I-----K-----Q-----K-----NE-----K-ET-----DK-----	191
HIVUG031fr	--Y-----I-----P-----ES-----SS-DMFGE-----	203
HIVD687	-----R-IPT-----D-----SS-----SE-----	90
HIVNI	-----I-----V-----K-MHA-----EN-----GS-----Q-----	201
HIVRW009	-----R-----I-----NS-----SN-----SS-----HQ-----	203
HIVUG275A	-----Q-----IHP-----P-----GSN-----RQ-----	202
HIVUG273A	-----TRR-----K-----K-----SSE-----	208
HIVV1191A	-TY-----V-----K-----NGI-----GTSN-----NS-----AQ-----	205
HIVDJ264A	-----Q-----SA-----ESS-----SS-----Q-----K-----	200
HIVDJ263A	-----SA-----ESSNSQ-----	204
HIVDJ258A	-----Q-----MSA-----SES-----NSQ-----	201
HIVCARGAN	-----K-----DEN-----CTR-----NS-----EL-----	211
HIVCAR5A5	--Y-----K-M-R-----H-----K-----I-P-NKINDASQ-ISEMI-----SSSGNITEQ-----S-----T-----	209
HIVCAR4054	-----I-----IN-R-K-EHA-----NS-----NS-ISN-----AN-----YND-----V-TAK-----	202
HIVCAR286Afr	-----I-----I-NR-K-E-A-----PTMM-----DD-STD-----YTHN-----A-T-K-----	222
HIVCAR4023	-----S-I-----R-K-E-A-----I-P-MN-----MN-NT-----NTND-----YND-M-----V-VK-----	221
HIVCAR423A	-----T-----II-----RKQE-A-----I-P-MN-----N-NTNE-----IA-TVK-----	214
HIVKE2NYA	--Y-I-----RR-----P-NEN-----SSNNE-----	202
CONSENSUS.B	csfnittsirdkvqkeyAlfyklDVvpid?????n?????n?????YrLisCnSvitQa	185
HIVJRCSP	-----K-----N-----NK-----NT-----K-----	202
HIVJRF6	-----E-----N-----N-----NT-----S-----D-----	201
HIVLA1	-----M-----S-----H-----N-----NT-----S-I-----	202
HIVBRVA	-----M-----H-----I-----NDKT-----T-----R-----	204
HIVJH3	-----N-----H-----H-----N-----NSTKDNKND-----ST-----R-----	216
HIVBAL1	-----N-----G-----IA-----NS-----NMR-----	208
HIVYU2	-----N-----N-----AS-----	200
HIVMN	-----M-----L-----I-S-----DS-----TS-----	209
HIVHX82R	-----S-----G-----F-----II-----DT-----TS-K-T-----	204
HIVLA1	-----S-----G-----F-----II-----DT-----TS-T-T-----	209
HIVNL43	-----S-----F-----I-----TS-----	202
HIVMFA	-----S-----N-----F-----IR-----TT-----	202
HIVCAM1	-----V-----N-----R-----KA-----TS-----T-H-----	205
HIVNY5CG	-----V-----I-----N-I-----R-I-----KD-----TS-----T-M-D-----	199
HIVAQA	-----K-D-----R-----ND-----TS-----N-T-----	200
HIVJFL	-----T-----KN-M-R-----KD-----TS-----T-N-T-----	201
HIVSIN184	-----V-----KN-----T-----H-----EDD-----TS-----T-N-----	202
HIVD31	-----H-----ND-----TS-----	204
HIVSF162	---KV-----N-M-----ND-----TS-----K-N-----	202
HIVBCSG3C	-----V-----LG-----K-A-----NDK-----ST-----K-----T-----	207
HIVQYI	-----T-----M-EQ-----L-----KN-----DTK-----P-H-----T-----	209
HIVSF33	-----G-----A-Y-----DD-----NT-----TS-----H-S-----T-----	205
HIVCD4	-----R-----E-----DNK-----NT-----K-----N-----	215
HIVSP2	-----I-----N-----RN-----NASPTT-----YT-----N-H-R-----	207
HIVSP2B13	---Y-----I-----N-----RN-----NASPTT-----YT-----N-H-R-----	206
HIVHAN	---KV-----N-----S-----T-----N-----NKTSMRD-----TT-----S-M-H-R-----	206
HIVRF	---QV-----R-----T-----K-----BKG-----NISPK-----NNTS-NT-----SYGN-T-H-S-----	217
HIVMAJ2	-----R-----H-----KGD-----N-----SS-----R-M-----	201
HIVT8132	-----K-----L-----I-----MND-----TS-----I-----	190
HIVBR020	---P-----VK-----EK-----N-----Q-G-----DN-----TS-----T-----	204
HIVTH014	-----KT-----K-D-----N-----D-TS-----N-----	197
HIVUS711fr	-----K-----K-----S-N-I-----G-----GD-TS-----MYK-----	202
HIVUS712	---V-A-----M-R-----FQ-----TS-----	206
HIVUS715	-----F-----E-----DT-----TS-----L-----	208
HIVUS716	-----M-----L-----I-----DN-----NS-----	200
HIVHA593	-----N-----M-----T-Q-----I-----E-----DQN-----RTNMTL-----N-----	228
HIVHA594	---V-----K-----TT-----H-----GNN-----SSKNGS-I-----T-----	206
HIVHA596	R---V-A-----K-----TT-----H-----SKMN-----SSKNGN-I-----T-----	207
HIVHA599	-----N-KG-I-----S-----L-----N-TG-----SSNSSR-I-----T-VK-----	206
HIVHA651	---V-----K-----T-QS-----R-----Q-----E-ND-----NSS-----N-----	202
HIVRJS	-----N-L-----G-----WD-TS-----T-N-----	205
HIVGUN1r	-----RK-N-----Q-----LNR-----I-----KTS-----T-X-----	170
HIVSC	-----S-----TS-----T-N-----	205
HIVBR014	---V-----N-M-----P-----KT-----R-N-H-----	207
HIVTH026	-----ST-----D-----S-----NT-----IS-----N-----	197
HIVBR021	-----KN-----K-D-----NR-----DN-----TS-M-N-----	208
HIVACH9	-----RGG-M-----L-----I-----NST-----SN-YT-----S-----	163
HIVACP1	---Y-----N-----N-M-----Q-----DR-----T-----DD-TS-----K-N-----	156
HIVJM	-----NMGG-MK-----NT-----V-NSH-----NN-----SN-DS-----I-T-----	184
HIVWM	---Y-----N-R-----T-----I-----NDKD-----TS-----	172
HIVMA208	-----N-K-RM-----A-T-----I-----G-----NDSTS-----NNNTS-NVT-----S-----D-----	224
HIVPB96	---Y-----N-----K-----NR-----ENT-----NT-----K-NHR-----	206
HIVBNB11A	-----R-----K-----K-----A-GS-----TS-----	76

	peptide T region	
CONSENSUS.C	CsFnaTTe?rDkqkvyaAlFYrIDivpl?????n?????n?e???.???YrLinChtSaitQA	182
HIVMA959	-----L--R-----K-----N-----EKD-----S-S-----GE-----S-TV-----	205
HIVMA960	-----I--H-----N-----ENEN-----N-----T-----	195
HIVD757	-----T--L--Q--A--K--V--DG-----D-----S-----T-----	89
HIVD747	-----T--L--Q--Q--K--V--TTD-----S-----T-----	134
HIVD760	-----P--I--RQ-----D-----MN-----ST-----	143
HIVNOF	-----YF--IG-----KP-V-----ER-----N-I-H-----T-----	77
HIVZAM20A	-----VIN--TS--H-----V-----NDENS-----XTNA-----ST-----KPTK-I-----T-----	217
HIVD1044	-----T-----I--R--H--K--D-----ESK-----D-----SRM-----	82
HIVD1024	-----I--R--T-----D-----SN-----K-----K-----	196
HIVD744	-----I--R--T--X-----D-----DRNER-----S-----KY-----	90
HIVD766	-----L--R--R-----T--L--N-----KKMS-----D-----Y-----	81
HIVD808	-----V--R--T-----FN-----KK-----S-----EY-----	97
HIVD868	-----I--R--T-----D-----KKK-----N-----SEY-----	94
HIVBR025	-----M--V--RE--H-----K-----ESS-----T-----GD--V-----	172
HIVSM145A	-----LS-----NEGE-----NSRK-----	205
HIVZAM18A	-----LK-----R-----K-----NEN-----N-SSE-----DS-----E-----	202
HIVDJ259A	-----V--L--R--A-----NE-----NSE-----	195
HIVDJ373A	-----I--L-----K-----K-----G-----GNEE-----S-----S-----	201
HIVSE364A	-----I--L--RRNE-----SGSK-----SS-----S-----E-----	204
HIVUG269A	-----V--I--TAM--K--V--S--EGNS-----T-----	209
CONSENSUS.D	CsFnitb??rDKKkqvyaLFYKLDvVqld?????n?e???.???YrLinChtSaitQA	180
HIVJY1	-----VIS--H-----R-----P--DDNS--A--T--NT-----YT-----N-----	216
HIVNDK	-----V-----RE-----I--P--NN-----RT-----ST-----N-----D--T-----	199
HIVMAL	-----PVGS--R--E--T--N--L--DSD-----SS-----V-----	209
HIVELI	-----V--VLK--Q-----R--I--P-----D-ST-----ST-----N-----	205
HIV2226	-----VV--T--H-----R--I--P--NO-----ST-----ST-----N-----	205
HIVUG024	-----V--EV-----E-----MND-----MT-----TS-----E-----	205
HIVUG021	-----EV-----A-----MND-----ST-----TS-----A--T-----	197
HIVUG038	-----S--IV--Q--EIN--I--R-----MT-----N--ND-----H--T-----	202
HIVUG046	-----V--KVI--EH-----M-----DN--T-----TN-----V-----	207
HIVUG266A	-----EI-----P--D-----N--ST-----TS-----T-----	203
HIVUG269Afr	-----Y--V--EI-----R-----MNSSES--I--T-----S-----	210
HIVUG274Afr	-----EL-----E-----D-----N--ST-----ST-----N--V--T-----	209
HIVSE365A	-----EI--RQN--S--R--L--P--GMSRK--ST-----YT-----S-----V--K--	212
HIVCAR4020	-----V--VI-----H-----I--P--EDEDWING--T--DNTK-----IT-----R-----S-----	236
CONSENSUS.E	CsFnmTtelrDkqkvyaAlFYKIDlvI?????n?e???.???YrLinChtSVIKQA	190
HIVTH966	-----I-----Q--EDK-----	206
HIVTH975	-----II-----OMEG-----T-----	201
HIVTH022	-----T-----I--Q-----R--I--QHNE-----	206
HIVTN243	-----I--Q--GDK-----SS-----	211
HIVTH006	-----I--QV-----	202
HIVTH011fr	-----I--V--K--E-----I--P--GD-----N--G--SEYSE-----R-----	184
HIVCARELO	-----G--H--S-----P--DGMIDG--N-----I--H-----	214
HIVCAR4017	-----V-----R--Y-----T--I--S--D-----SD-----M-----N-----	211
HIVCAR4039	-----I--K--TKEAY--P--H--P--KE-----SE-----V--LVTH-----	204
HIVCAR4071	-----I-----E--E--Y--RS--P--ERNSGE--N--G--S-----H-----	212
HIVCARMBA	-----S--I--VS--SMBA-----N-----	202
CONSENSUS.F	CsFnmTtelrDkqkvyaAlFYKIDlvIPIsn?????n?e???.???YrLinChtSTITQA	196
HIVBRA7944	-----V--M-----R--L--S-----G-----G-----	175
HIVB2163Afr	-----S--N-----K-----SS-----R-----H--D-----	205
HIVB2126Afr	-----S--N-----K-----SS-----R-----H--D-----	205
CONSENSUS.G	CsFnitTTE?rDKK?K?EYALFYrLDVVPIn?n?e???.???YrLinChtSTIKQA	188
HIVVI525A	-----V--L--Q--A--K-----DN--R-----D-----H-----	217
HIVL0V217	-----V--S--K-----G--E-----TTD-----A-----	214
HIVCAR4067	-----I--R--T-----EDQ-----T-----	204
CONSENSUS.O	C7FNVITTVlkDKKekKQALFYVvDL7k?N?ts???.N?T???.???YrLinChtSTIQA	176
HIVANT70	-----S-----I-----MEL--E--ST-----K--NSK-----T-----T-----	200
HIVMVP5180	-----E-----T-----Q-----S--V--DSNAV--G-----T--H-----I--K-----	204
HIVVAU	-----D-----Q-----T--V--I--A-----E-----R--I-----R-----	205
CONSENSUS.CPZ	C7FNT??T??rDKK?????FY??D??T???.N?e???.N?T???.???YrLinChtSTIQA	83
STVCPZGAB	-----S--VT--ELR--KQVYSL--VE--VNLG-----E-----N-----RIIN--A--A-----	202
CONSENSUS.U	CsfnitTEl?DK?ktvyaLFYKLDvV?n?n?e???.N?e???.???YrLinChtSaitQA	181
HIVZAM184	-----YRT--K--TLR--S-----Q--DF-----S--TM-----S-----S-----A-----	195
HIVCAR4081	-----IR--K--A--R--P--D--N-----N-----S-----K-----	205
HIVK124A	-----VR--RK--Q-----I--QM--E--QD-----S--GEYR-----	209
HIV23fr	-----KVP--K--TE--HT-----PL--VT-----I-----SST-----T-----	201

HIV1 ENV

	* * * * *	***	* * * * *	***		
CONSENSUS.A	CPKvsFEPIPIHYCaPAGFAILKc7dk?FnGtG.	pckNVStVQCTHG7IkPvvsTQLLLNGSLAe?ev?Ic				239
HIVSP1703	-----N-AB-----	-----R-I-----	GR-K-			279
HIVU455	-----K-PE--K---R-----	-----R-IR--				271
HIVZ321	-----R-EE-E-K---R-----	-----G-R--				274
HIVRW020	-----K-TE-----	-----R-I-----	EGIQ-			260
HIVUG031fr	---T-----R-NEDD-----	L-----X-----	KG-R-			273
HIVD687	-----KG-K-----	S.S-----	E-IV-			158
HIVNI	-----K-G-----	-----G-V--				270
HIVRW009	-----N-----K-K-----	-----E-II--				272
HIVUG275A	-----N-QE-----	I-N-----	GK-K-			271
HIVUG273A	-----N-E-----L-S-----	-----R-----	T-K-K			277
HIVV1191A	-----R-E-----R-S-----	-----EKIM--				274
HIVDJ264A	---T-----N-K-----	-----E-V--				269
HIVDJ263A	---T-----N-K-----L-----	-----E-V--				273
HIVDJ258A	---T-----N-K-----	-----A-----	G-V--			270
HIVCARGAN	---T-----T-----NE-K-----	-----R-S-----	T-----E-I--			280
HIVCAR5AS	-----K-TE-K-----	-----R-----	GKIM--			278
HIVCAR4054	-----N-TK-----	-----Y-----	E-M--			271
HIVCAR286Afr	---T-----X-----R-E-----	---X-----X-----	GG-I--			291
HIVCAR4023	---IT-----N-TK-S-----	-----EG-H--				290
HIVCAR423A	-----R-EK-----	-----N-----	G-K--			283
HIVKENYA	---T-----K-K-----	-----T-----	G-M--			271
CONSENSUS.B	CPKvsFePIPIHYCaPagfaillKcndkkFnGtG.	pckNVStVQCTHG.	frPvStqLLlNGSLaeceevir			254
HIVJRC6F	-----N-T--K--Q-K-----	-----K-----				271
HIVJRF1	---I-----T--K-----	-----K-----				270
HIVALA1	---Q-----K--K--T-K-----	-----K-----				271
HIVBRVA	---T-----N-----	-----K-----				273
HIVJH3	---I-----I-----	-----K-----				285
HIVBAL1	---I-----K--K-----	-----K-----				277
HIVYU2	-----I-----S-K--S-K-----	-----I-----				269
HIVMN	-----N-T-----	-----S-K-----				278
HIVHXB2R	-----N-T-----	-----K-----				273
HIVLAI	-----N-T-----	-----K-----				278
HIVNL43	-----N-T-----	-----D-----				271
HIVMFA	-----T-----K-----	-----K-----				271
HIVCAN1	-----T-----K-----	-----K-----				274
HIVNYS6G	-----T-----K-----	-----K-----				268
HIVADA	-----T-----K-----	-----K-----				269
HIVJFL	-----T-----K-----	-----K-----	H-----G-----			270
HIVSTH184	-----T-----M--K--R--K--Q-K-----	-----K-----	S--S-G--			271
HIVD31	-----T-----M--K--R--K--Q-K-----	-----K-----				273
HIVSF162	-----I-----N--T-----	-----K-----				271
HIVBCSG3C	---I-----H--T-----	-----K-----				276
HIVQY1	---I-----H--T-----	-----K-----				278
HIVSF33	-----T-T--L-----	-----K-----				274
HIVCDC4	-----T-T--L-----	-----K-----				284
HIVSF2	-----T-----N-T--K-----	-----I-----				276
HIVSF2B13	-----T-----N-T--K-----	-----K-----				275
HIVHAN	-----T-----R-----	-----K-----				275
HIVRF	-----T-----R-----	-----K-----				286
HIVWMJ2	-----T-----R--S-N--K-----	-----K-----				270
HIVTB132	---T-----T--Y--G-----	-----K-----	DI-----			259
HIVBR020	-----T-----Q--N-----	-----K-----				273
HIVTH014	---R-----F-T-----N-T--K-----	-----K-----				266
HIVUS711fr	-----F-----K-----	-----K-----	S-----G-----			271
HIVUS712	-----F-----K-----	-----K-----				275
HIVUS715	---I-----T-----RN-----	-----K-----				277
HIVUS716	-----T-----N-T-----	-----K-----				269
HIVHA593	-----T-----N-T-----	-----K-----				297
HIVHA594	-----T-----N-T-----	-----K-----				275
HIVHA596	-----T-----N-T-----	-----K-----				276
HIVHA599	-----T-----N-T-----	-----K-----				275
HIVHA651	-----T-----N-T-----	-----K-----				271
HIVRJS	-----T-----V--K--N-----	-----K-----				274
HIVGUNfr	-----T-----L-----	-----R-----	G-I--			239
HIVSC	-----RW--N-N-----	-----H-----	L-----			273
HIVSBA	-----T-----V--N-H-----	-----L-----				20
HIVSBB	-----T-----I-----	-----L-----				20
HIVSBC	-----T-----I-----	-----L-----				20
HIVBR014	-----T-----V--N-H-----	-----L-----				276
HIVTH026	---T-----T-----N-T-K--T-K-----	-----K-----				266
HIVBR021	-----T-----T-----K-----	-----K-----				277
HIVACH9	-----I-----R-EQ--K--T-----	-----K-----				232
HIVACP1	---I-----F-----N-T-S-K-----	-----K-----				225
HIVJM	-----T-----L-----T-----	-----K-----				253
HIVWM	-----Q-----V--M--N-T--S-----	-----R-----	DI-----			241
HIVNA208	-----T-----L-----T-----	-----K-----				293
HIVP896	---Q-----V--M--N-T--S-----	-----R-----	G-----			275
HIVWB11A	-----T-----L-----T-----	-----K-----				145

CONSENSUS.C	CPKvSFdPIPIHYCaPAGyAILKcNktFnGtG.PChnvsTvqcTNG.IkPVVStqLLLnGSLaE?eIiIR	250
HIVMA959	-----T-----D-----G-----	274
HIVMA960	-----D-----G-----	264
HIVD757	-----RT-----G-----	158
HIVD747	-----G-----I-----G-----	203
HIVD760	-----T-----S-----G-----	212
HIVN0F	-----N-----GD-----	146
HIVZAM20A	-----LQ-----FG-----T-----S-EG-----	286
HIVD1044	-----T-----T-----G-----	151
HIVD1024	-----T-----G-----	265
HIVD744	-----T-----T-----G-----	159
HIVD766	-----T-----T-----D-----R-----R-----	150
HIVD808	-----T-----T-----D-----Q-R-F-----R-----	166
HIVD868	-----AT-----F-I-----S-----D-EG-----	163
HIVBR025	-----N-----I-----M-H-----E-----	241
HIVSM145A	-----E-----RY-----G-M-----	274
HIVZAM18A	-----TL-----RD-N-----N-----E-----	271
HIVDJ259A	-----L-----RD-N-----N-----DD-----	264
HIVDJ373A	-----E-----RDTK-----K-----T-ED-----	270
HIVSE364A	-----L-----N-----G-----	273
HIVUG268A	-----TL-----N-----I-----I-E-----	278
CONSENSUS.D	CPKvSFdPIPIHYCaPAGfAILKc?dKkFNGtG.pCknVSTVQCTHG.IrPVVSTQLLLNGSLABEEIiIR	248
HIVJY1	-----K-----K-----	285
HIVNDK	-----IS-----F-----R-----S-----	268
HIVMAL	-----D-----N-----E-I-----K-----M-----	278
HIVELI	-----S-----R-----T-----V-----	274
HIV2226	-----S-----R-R-----T-----	274
HIVUG024	-----MN-M-----N-----	274
HIVUG021	-----IS-----N-----K-----	266
HIVUG038	-----I-----K-----T-A-----	271
HIVUG046	-----N-----S-----	276
HIVUG266A	-----N-----V-----	272
HIVUG269Afr	-----I-----T-----K-----K-----	279
HIVUG274Afr	-----Y-----M-----Q-----	278
HIVSE365A	-----N-----K-----K-----	281
HIVCAR4020	-----N-----K-N-----L-H-----K-----V-----	305
CONSENSUS.E	CPKvSFdPIPIHYCaPAGyAILKcNdKnFnGtG.pCknVSTVQCTHG.IKPVVSTQLLLNGSLABEEIiIR	259
HIVTH966	-----I-----T-----T-----	275
HIVTH975	-----I-----T-----T-----V-----	270
HIVTH022	-----I-----T-----T-----	275
HIVTH243	-----I-----T-----F-----	280
HIVTH006	-----I-----	271
HIVTH011fr	-----T-----	253
HIVCARELO	-----DF-----E-----	283
HIVCAR4017	-----N-----	280
HIVCAR4039	-----T-E-----F-----K-E-S-K-----T-----K-VV-----	273
HIVCAR4071	-----Q-----I-----K-----	281
HIVCARMBA	-----S-T-----	271
CONSENSUS.F	CPKvSWdPIPIHYCaPAGyAILKcNdKkFNGtG.pCKNVSTVQCTHG.IKPVVSTQLLLNGSLABEdIiIR	265
HIVBRA7944	-----R-E-----	244
HIVB2163Afr	-----G-----	242
HIVB2126Afr	-----	274
CONSENSUS.G	CPKvSFdPIPIHYCaPAGfAILKcRdKnFnGtG.pCKNVSTVQCTHG.IKPVVSTQLLLNGSLABEEI?IR	256
HIVVI525A	-----S-----E-T-----L-----K-I-----	286
HIVLBV217	-----D-----T-----K-----M-----	283
HIVCAR4067	-----Y-----R-----G-M-----	273
CONSENSUS.N	??V?IR	3
HIVCA13	EK-M--	6
HIVVI557	NH-I--	6
CONSENSUS.O	CPKvSFdPIPIHYCaPAGyAIFKCN?T?FNGtG.Ic?NItVVTCTHG.IkPTVSTQLLLNGSLsKgiRiM	242
HIVANT70	-----S-E-----T-R-----R-----M-----	269
HIVMVP5180	-----T-----D-D-----H-S-----RE-----	273
HIVVAU	-----C-----E-G-----K-V-----N-T-----	274
CONSENSUS.CPZ	C?K??PEPIPIHYCaPAG?AIL?C?D?DF?G?G.?C?NVS?VHC?TG.I?P?V?T?LL?NG?T?V?V?	126
SIVCPZGAB	-P-TS-----P-----K-N-K-----S-K-----K-T-----T-----K-V-T-Q-----I-----SLAEGNIT-R	271
CONSENSUS.U	CPKvSFdPIPIHYCaPAGfAILKcKdK?FNGtG.pCKNVSTVQCTHG.I?PVVSTQLLLNGSLABEEI?IR	247
HIVZAM184	-----T-----TE-----S-----S-----R-----G-M-----	264
HIVCAR4081	-----K-----K-----K-----E-I-----	274
HIVK124A	-----E-----K-----G-K-----	278
HIVZ3fr	-----N-----K-----R-----S-E-----	270

HIV ENV

	V3 neutralization loop										
***	***	***	***	***	***	***	***	***	***	***	
CONSENSUS A	SenitnNaktiivql??pv?inctrp	?nnctr	ks???	vr1???	gpgq??	afya	cgdi	igdirqa			290
HIVSF1703	-----NKT-E-----	N----	---	---	---	---	---	---	---	---	334
HIVU455	---F-----VN-K--S-	Y----	-N	I-RYSI-S-	--V	--K-	-----	-----	-----	-----	328
HIV2321	---F-D---I-----VK-N-T-M-	N----	---	IS-	--R	--P	-----	-----	-----	-----	329
HIVRW020	-----DKA-K-----	N----	-G	---	---	---	---	---	---	---	315
HIVUG031fr	- - -S-----K										284
HIVD687	-----R-----VK-R-----	G--N-N	-H-	----	----	----	R-R-	-----K-	-----	-----	213
HIVNI	-----AN-K-----	N----	-G	H-	----	----	----	-----	-----	-----	325
HIVRW009	-----H-----NET-L--S-	N----	---	H-	----	----	----	-----	-----	-----	327
HIVUG275A	-----N-L--TT-T-----	N----	T-	---	----	S----	----	-----K-	-----	-----	326
HIVUG273A	---SD---L---TT-K-----	G----	T-	---	----	----	----	-----	-----	-----	332
HIVV1191A	---F-D-G---VE-K-----	S----	-G	IH-	---R	---	-Q-	P - - - -	-----	-----	329
HIVDJ264A	-----N---VT-G-----	N----	---	---	----	T----	----	-----K-	-----	-----	324
HIVDJ263A	-----QN---VA-R-----	N----	---	---	----	T----	----	-----	-----	-----	328
HIVDJ258A	-----N---VT-R--S-----	G----	---	---	----	T----	----	-----	-----	-----	325
HIVCARGAN	-----VK-N-T-I--SRT--	-R	WH-	-S-	----	IDG-	T----	K-	-----	-----	335
HIVCARSAS	---PS-----NQS-E-----	N----	RR	MH-	---R	-I-	DA-	V-----	-----	-----	333
HIVCAR4054	---F-D---KN-R-----	G----	-G	IH-	---PGQ-I-	----	----	-----	-----	-----	328
HIVCAR286Afr	-----DS-G-----	G----	---	I-MHI	---R	S----	K-V	-----	-----	-----	348
HIVCAR4023	-----DS-I-----	N----	R	M-	----	T----	----	T - - - -	-----	-----	345
HIVCAR423A	---F-----DS-R-----	N-Y--	M	IPT	---GQVI-	---	-K-	-----K-	-----	-----	340
HIVKENYA	-----N---FAE-K-----	N----	-M	I-	----	----	----	-----N-	-----	-----	326
***	***	***	***	***	***	***	***	***	***	***	
CONSENSUS B	senitdnaktiivqlinesveinCtRp?nnctr	ks???	1?1??	gpgx	afyt	tg?1???	igdirqa				307
HIVJRCSE	-----K-----	S----	-H-	----	----	-E-	-----	-----	-----	-----	326
HIVJRFL	-D--N-----K-----	---	-H-	----	----	-E-	-----	-----	-----	-----	325
HIVALA1	-D-----K-----	IYRK	GR	-H-	----	-H-	RQ-	EN-----	-----	-----	326
HIVBRVA	-----N-V-----	---	-R	TM-	----	VY-	-Q-	-----R-	-----	-----	328
HIVJH3	-----K-P-V-----	SXT--	RR	-H-	----	----	KQ-	A--L--	-----	-----	340
HIVBAL1	-A--A--V-----	---	-H-	----	----	----	-E-	-----	-----	-----	332
HIVYU2	-----N-----V-----	---	-N	----	----	-L-	-E-	-----	-----	-----	324
HIVMM	-----H-----Q-----	-Y-K-	-R	-H-	----	----	KN-	-T----	-----	-----	333
HIVHXB2R	-V-----T-----	---	-R	R-QR	----	-V-	I-K	--NM--	-----	-----	329
HIVLAI	-A-----Q-----	---	-R-QR	----	-V-	I-K	--NM--	-----	-----	-----	334
HIVNL43	-A-----T-----	---	-R-QR	----	-V-	I-K	--NM--	-----	-----	-----	327
HIVMFA	-A-----T-----	---	-R-QR	----	-V-	I-K	--NM--	-----	-----	-----	327
HIVCAM1	-----N-----K-P-----	-L	---	-A-	----	TV-A	-DR-	-----	-----	-----	329
HIVNYSCG	-----N-V-----	---	-G	-A-	----	TL-A	REK-	-----	-----	-----	323
HIVADA	-S-----N-----K-----	---	-H-	----	----	----	-E-	-----	-----	-----	324
HIVJFL	---M-----K-IA-----	---	-TL	----	----	----	-D-	-----	-----	-----	325
HIVSIMB4	-----K-V-----	-SK-	T	LHM	---K-	---	-D-GGY	T-Y----	-----	-----	329
HIVD31	-D-----K-----	-Y-S-	-R	-R-	---	AR-	K-K-	-----	-----	-----	328
HIVSF162	-----K-A-----	---	-T-	---	---	---	-D-	-----	-----	-----	326
HIVBCSG3C	-A--S-----K-A-----	KK--	-R	-TT	----	VY-	-B-	V-----	-----	-----	331
HIVVOYI	-S--N--I-----K-----	---	NR	-S-	----	---	KQ-	-----	-----	-----	333
HIVSF33	-D--N--L--V-----	-R	RR	-TS	---K	VL-	-E-	-----K-	-----	-----	329
HIVCDC4	-----N-----V-----	-M--	-R	VTL	----	VH-	-E-	L-N----	-----	-----	339
HIVSF2	-D--N-----A-----	---	-Y	----	----	---	-R-	-----K-	-----	-----	331
HIVSF2B13	-D--N-----A-----	---	-G	-W-	----	-V-	-R-	-----K-	-----	-----	330
HIVHAN	-D--T--H-----	---	-G	-H-	----	-V-	-R-	V-----L-	-----	-----	330
HIVRF	-----V-----A-Q-----	---	-TK	----	VI-A	-Q-	----	K-	-----	-----	341
HIVMMJ2	-----H-----Y-V-R-	LS-	---	-R	RE-	---	-I-	-----	-----	-----	324
HIVTB132	-----K-P-----	-Y--	-X	-XM	----	XX-	-E-	-----R-	-----	-----	314
HIVBR020	---L-N-----KDP-D-----	---	-H-	----	---	-A-	-D-	-----	-----	-----	328
HIVTH014	-S-----RV-----	---	-HL	----	---	-W-	-Q-	-----	-----	-----	321
HIVUS711fr	-----N-----K-A-----	---	-HL	----	---	-A-	-E-	-----	-----	-----	326
HIVUS712	-A-LS---I-----M-I--	---	-H-	----	----	----	-D-	-----K-	-----	-----	330
HIVUS715	-A-LS---Q-----	---	-H-	----	----	----	-E-	-----	-----	-----	332
HIVUS716	-----N-----K-M-----	---	-H-	----	----	----	-Q-	-----	-----	-----	324
HIVHA593	-----N-----I--T-----	---	-S	-R	-S-	----	RA - K-	--N----	-----	-----	351
HIVHA594	-----N-----I-PK--V-----	-S-	-R	-S-	----	VW-	-Q-	-----K-	-----	-----	330
HIVHA596	-----N-----I-K--V-----	---	-R	-S-	----	VW-	-Q-	-----K-	-----	-----	331
HIVHA599	A--SN-----KA-L-----	---	R-VHSGH-	-G-	TLF-	-H-	V----	K-	-----	-----	332
HIVHA651	-----T-IQ-----	---	-N-	----	---	-W-A-	-N-	-----	-----	-----	325
HIVRJS	-----N-----IV-----Y-TKKI	RM	-H-	----	---	-V-	-G-	M-----	-----	-----	330
HIVGUNfr	-----N-----IV-----	---	-T-	-S-	----	HA	IEK-	--N----	-----	-----	294
HIVSC	-----K-A-----	---	-T	R-	-H-	----	-D-	-----	-----	-----	328
HIVSBA	-D--N-----K-A-----	---	-N-	----	---	----	-D-	-----	-----	-----	75
HIVSBB	-----N-----T-----	---	-S	-R	-S-	----	VA	ARE-	-----K-	-----	75
HIVSBC	-D--N-----Q-I-----	---	-S-	----	----	----	-E-	-----	-----	-----	75
HIVBR014	-----MN-----	N-IQ-	RI	-H-	----	HA	-G-	K-N--P	-----	-----	331
HIVTH026	YS--N--I-----	---	-PL	---	Q	-W-	-Q-	--N----	-----	-----	321
HIVBR021	-----I-----T-----	---	-HM	-W-	----	-A-	-E-	--N----	-----	-----	332
HIVACH9	-----N-R-----	---	-P-	----	----	----	-D-	-----	-----	-----	287
HIVACP1	-----N--I-----X-K--I--	---	-G	-G-	----	TV--	AEK-	-----	-----	-----	280
HIVJM	-D--N-----KT-V-----	---	-H-	A----	----	----	-D-	-----K-	-----	-----	308
HIVMM	-----N-----K-P-Q--I--	---	-G	-H-	----	-K-	----	-E-	--N----	-----	296
HIVMA208	-----E-M-I-----K-P-----	GM	---	PM	----	FA	-D-	-----K-	-----	-----	348
HIVP896	-----N-----V-----	---	RR	LS-	----	---	RKM	-----	-----	-----	330
HIVBWB11A	-----R--T-----I-----	---	-G	-NV	----	-L-	-D-	-----	-----	-----	200

		V3 neutralization loop	
CONSENSUS.C	SenltnNvktIivhln?aveivctRp.nnntr.ks...iRi...GPGQ...tiya.tgdI...IGDIRga		304
HIVMA959	-----E-----R-----V--NN--		329
HIVMA960	---AQ---F-E-----R-----NE-----N-E-		319
HIVD757	---A---Q-R-----		213
HIVD747	---Q---Y-----G-V-		258
HIVD760	---D---Q-V-----		267
HIVN0F	---I---V---K-D-----R---V---V---NA-		201
HIVZAM20A	---AQ---Q-K-I---A---G-----F---A-		341
HIVD1044	---P---A---Q-M-----		206
HIVD1024	---Q-----		320
HIVD744	---Q---E---S-----E-		214
HIVD766	---L---Q-----L-----R-		205
HIVD808	---YQ-----T-		221
HIVD868	---A-I---DQP---R-----R-		218
HIVBR025	---K---D---E-TN-----A---E-		296
HIVSM145A	---A---Q---R---Y-A-----V---N-		328
HIVZAM18A	---A---E---S-----A---G---N-		326
HIVDJ259A	---A-I---Q-E-A-N-----E-		319
HIVDJ373A	---A-I---Q-Q-N-----Q-		325
HIVSE364A	---A---E-I---A-----M-		328
HIVUG268A	---A-I---Q-K---N-A-----E-		333
CONSENSUS.D	SENltnNAktIivqlnesv?InCtRP?y?n?r?qr...?tpi...glgq...alyt?c?????..?g?i?ga		294
HIVJY1	-----V-----H-----E-----DNKIT--S-----R-I...K-D-R-		340
HIVMDK	-----V-----A-IV-----R-YT--S-----R---S---I-GKKKK...T-Y-G-		326
HIVMAL	---D-T-N---T-T-----GN-T--RG...IHF...-P---GI...V-D-RR-		332
HIVELI	---N-AH---K-T-A---Q-T-----S---RS...RSI-G-		328
HIV2226	---I---I---A-----R-I---S-----KT...RSI-G-		328
HIVUG024	---I---I---H---V-----N-I-----RR...IED-RR-		328
HIVUG021	---I---I---H---P-----DKVS-Y-----V-R---S---RI...K-D-R-		320
HIVUG038	---IS---T-S-A---FYAIE.RQ-R---Q---V---KK...I-R-G-		326
HIVUG046	---P---S---E-K-R---Y---KL...K-Y-RP-		330
HIVUG266A	---I---I---I---N-IRI-----R---P---TRR...K-IRG-		327
HIVUG269Afr	---I---A-T-T---TKT-H---AQ-R---WM---GI...T-D-R-		330
HIVUG274Afr	---I---T-D-A---E-I---T---QGRKK...I-K-G-		335
HIVSE365A	---F---M---N-K---V-H---TRV...K-D-R-		335
HIVCAR4020	---D---K---M-S---NNYK--G-----RV...I-D-RK-		359
CONSENSUS.E	SENltnNAktIivhLnksVeINCTRP.snntr.ts...?ri...GPGQ...VFYr.TGdi...IGDIRka		313
HIVTH966	-----Q-----TT-----T-N-		330
HIVTH975	-----V-----		325
HIVTH022	-----IT-----R-		330
HIVTN243	---D---G---X-P---ITX-----XR-		335
HIVTH006	---Y-----IT-----Q-		326
HIVTH011fr	---I---IA-----N-		308
HIVCARELO	---R---YK---A---K---S---T-		338
HIVCAR4017	---A---KKI---V---K---L-		335
HIVCAR4039	---I---Q-VTF-Q---I---V---K---S---T-		328
HIVCAR4071	---K---S---T-		336
HIVCARMBA	---A---P---V---G---E---R-		326
CONSENSUS.F	SenltnNAktIivqln?aveINCTRP.nnntr.ks...Iri...GpGr...?fya.tgdi...IGDIRka		317
HIVBR1944	---Q---A---H---E---Q-----HL---Q---A---T---E-		299
HIVBZ126Afr	---Q---A---HF-E-Q-----Y---YP...A-HT-A-K-		329
CONSENSUS.G	SENltnNAkvIIVQIN?sieLnCtRP.NNNTR.KS...I??...gpGq...afYA.TGdi...IGDIRQA		308
HIVV1525A	---N---ETV-----KF---T-R..VL---N-		341
HIVLBV217	---N---N---P-K-D-V-----HI...L---A-		338
HIVCAR4067	---L-N---I---K-V---I---SE		328
CONSENSUS.H	S?NITDNTKNIIVQLK?FV?INCTRP.NNNTR.?S??I?I...G?GQ...?F?A.?G?I...IGDIR?A		47
HIVCA13	---E-----N---V-----R-MR--G---R---T-H---I-A-----Q-		63
HIVV1557	---K-----S---P-----R---S---P---A-Y---T-D-----K-		61
CONSENSUS.O	gknI?dsgkNIIVTLN??i?mtCeRp.gn?tv.Qe...i?i...GPma...Wys.Mgl????n?k??s?r?A		287
HIVAM70	A-D-LEG-----STLN-----QI.DI...MR...-IGG...TAGNS--A-		325
HIVMP5180	---TE-A---TP-N---I-E---IAE---D---YT---R---R---T-KRSN-TSPR--V-		332
HIVVAU	---S---E---LI---TN-TIA---Q-I---K---MA...A-S---T-GDT-A-		330
CONSENSUS.CPZ	??K????V?????E?????C?RFGN?77R.??...?QI...GPGM...TFYN.?E??...?GDTR?A		151
SIVCP2GAB	VEN-SKMTD-WIVQLV-AVSLN-H---NT..GE..V---I-NV...V---S-		326
CONSENSUS.U	SenIT?NAktIIVqlnepV?InCtRP?nntr?ks...iRi...GPGQ...afYA.tgdi...igdirQA		299
HIVZAM184	---D---N---FT---Q-S---S---R---P---N-		318
HIVCAR4081	---D---I---H---S-E---N---V---T-		329
HIVK124A	---Q-L-N---VR-R---N-I---P---N-		333
HIVZ3fr	---N---T-K---GSDKKIRQ---K-V---K---G-ITG--		324

HIV1 ENV

CONSENSUS.A	hcnvar?ewh?tlq?V...a?qlr?..?f??Nkt...??iiF?n.as.GGDlEitthsFnCggEffy	336
HIVSF1703	Y----AD--K--G--N--KS..Y--S--AS--	388
HIVU455	-----RD--R-I-Q--E--KK..K--N--AS--I--	382
HIV2321	-----AT--D--SK--A--K..H--V-TS..TD--A--V--	385
HIVRW020	--H--A--D--RG--KK--E..H--K--I--I--	369
HIVUG031fr	XS--K--TQ--VKE--K..YW--T--	332
HIVD687	-----A--R--Q--EH--K..Y--K-N--AH..H--R--	267
HIVNI	-----KT--K--HQ--VT--KT..Y--K-T--A--PL--V--	379
HIVRW009	Y-T-NGTK--R--K--EK--SH..Y--E-I--T--K--	382
HIVUG275A	---I--A--E--K--VS--T..H--G--G--I--R--	379
HIVUG273A	-----S--KE--K--VK--T..HW--T--	385
HIVV1191A	---I--T--K--HQ--I--E..H--ITK-T--NS--	385
HIVDJ264A	-----SK--N--Q--I--K..HS--T--A--P--I--	377
HIVDJ263A	-----S--R--Q--T--K..H--A--I--	381
HIVDJ258A	-----SK--E--K--T--K..H--T--A--I--R--	378
HIVCARGAN	Y-E-NTT--Q--Q--VKK--T..YW--N-TK..P--	388
HIVCAR5AS	Y----AT--K--M--EE--K..Q--D--P--V--	386
HIVCAR4054	-----KA--DN--Q--KK--YE..V--VX-T..H--I--R--	381
HIVCAR286Afr	---TINKAK--Q--K--EK--G..I--N--T--H--V--V-L	402
HIVCAR4023	Y--KE--T--K--KK--GT..I--S--NS--V--R--	399
HIVCAR423A	Y--I--KA--KMM--RK--TK--ES..K--S--T-DK..P--I--IV--K--	393
HIVKENYA	-----A--T--K--VTK--E..Y--GN--K--A--	381
CONSENSUS.B	hCnlsrakWmptLkqi...v?klre??qf..g?Nkt...ivfnq.ss.ggdplvmhsfncgGeffy	360
HIVJRCFS	-----Q--E--N--TH--	380
HIVJRFL	-----D--I--E--H--	379
HIVALA1	-----ED--E--K--T--	380
HIVBRVA	---L--S--E--T--V--K--R--F--	382
HIVJH3	---N--R--A--G--K--V--R--	394
HIVBAL1	---L--D--NK--I--KH--T--	386
HIVYU2	---L--KTQ--E--E--AI--K--N--I--P--T--	379
HIVMN	-----D--R--S--K--K--	387
HIVHXB2R	-----AS--K--N--I--K--T--	384
HIVLAI	-----A--AS--N--I--K--T--	389
HIVNL43	-----A--AS--N--I--K--T--	382
HIVMFA	-----MS--AS--N--VI--K--T--	382
HIVCAM1	---L--ST--T--K--I--	383
HIVNY5CG	---L--D--T--K--R--	377
HIVADA	---L--T--N--AT--K--N--	379
HIVJFL	---L--VQ--I--K--T--	379
HIVSIMI84	-----D--K--K--R--K--	384
HIVD31	---G--DS--R--K--R--T--	382
HIVSF162	---GE--T--QA--K--	380
HIVBCSG3C	---L--D--R--I--E--	385
HIVQY1	---L--T-EK--E--AT--K--R--A-DR--	387
HIVSF33	Y----K--E--V--AT--K--R--	383
HIVCDC4	-----Q--Q--ATT--A--	393
HIVSF2	-----Q--E--K--N--R--	386
HIVSF2B13	---V--TQ--E--K--N--R--	385
HIVHAN	-----R--K--N--FR--IR--E--R--	386
HIVRF	---L--Q--V--T--D--TS--L--	395
HIVWMJ2	-----E--K--H--T--	378
HIVTB132	---ST--G--K--K--N--N--KE--I-W--X	368
HIVBR020	---L--Q--D--SK--T--E--K-QP..P--F--	382
HIVTH014	---L--ST--R--TE--	375
HIVUS711fr	---TLNKT--EA--M--G--I--X--EQ--	380
HIVUS712	-----AI--K--E--NTLM--	384
HIVUS715	---L--T--E--R--I--Y..E--L--	386
HIVUS716	---L--S--EN--TI--KP--L-L--	378
HIVHA593	---T--S--K--A--QP--	405
HIVHA594	---L--STQ--D--C--E--A--R--P--E--L--	384
HIVHA596	---L--STA--G--E--A--R--P--L--	385
HIVHA599	---L--I--Q--R--V--AK--L--D-R--I--K--T--T--	387
HIVHA651	---L--Q--E--KA--L--T--S--L--	379
HIVRJS	---L--E--ES--V--AT--K--T--	384
HIVGUNfr	-----KEN--T--K--T--	348
HIVSC	-----I--D--E--I--R--	382
HIVSBA	---I--T--R--GE--K--	129
HIVSBB	---IT--RL--S--Q--K--H--T-T--	130
HIVSEC	---L--N--A--K--E--K--N--	129
HIVBR014	Y-TLDK--T--K--SR--A--T--	385
HIVTH026	---L--ST--R--TE--Q--K--	375
HIVBR021	---L--IV--S--ET--KA--H--A--	385
HIVACH9	---L--A--T--	341
HIVACF1	---TE--V--Q--K--D--N--	335
HIVJM	---KEN--Y--V--AE--G--K--R--H--R--LIVE--I--	361
HIVMM	---VT--V--K--K--	350
HIVMA208	-----E--V--K--I--	402
HIVP896	-----Q--I--K--R--A--	384
HIVBWB11A	-----K--Q--T--V--K--L--	254

[illegible]

HIV1 ENV

CONSENSUS.A	cnccs?LPNstW??????...	N?C?????????	n?C?????????	ndtic?????lg	CrIkqIVnmWQR	369
HIVSF1703	-D-G-...-GG...	-S-DSIQES...	-S-...ES...	-I-...-P-...	-I-...	434
HIVU455	---G---I---	-GS-MS...	-DM-...	GP-G-...	-I-...	422
HIV2321	---G---G-L---	-G-...S...	-N-...WKIDTV---	-I-...-P-...	-I-...	429
HIVRW020	---G---ES---	-S-ES...	-N-...TS...	-T-...-I-...	-K-...	411
HIVUG031fr	---G---I-WE---	-G-...NNMP...	-S-...ET-G---	-P-...-K-...	-I-...	376
HIVD687	---N---D---	-S-...WD...	-N-...A-N---	-H-...-K-...	-K-...	307
HIVNI	---K---DNS---	-S-...A...	-H-...GS...	-P-...-I-...	-K-...	421
HIVRW009	---G---SKR---	-G-...WQS...	-G-...ELN---	-P-...-I-...	-C-...	424
HIVUG275A	-D-G-...-NG...	-TSKQE...	SSS-ES...	-GN-...-I-...	-I-...	423
HIVUG273A	---AG---NK---	-DSIKS...	-D-KS...	-I-...-I-...	-I-...	428
HIVV1191A	-S-G-...-E---	-G-QVA...	-D-GS...	-P-...-R-...	-R-...	427
HIVDJ264A	---E---DNSL---	-H-ES...	-H-E-...-W-	-P-...-K-...	-K-...	419
HIVDJ263A	---E---...---	-S-WD...	-SSNHIESNMTEGN-	-P-...-K-...	-K-...	427
HIVDJ258A	---S---...---	-SSSIHT...	-YSS-NDTD-N-	-P-...-K-...	-K-...	421
HIVCARGAN	---G---S-JONE-	-D-KSMKGL...	-KG-...-P-...	-R-...-R-...	-R-...	430
HIVCAR5AS	---G---K-WGP...	-S-TN...	-Q-IG...	-P-...-R-...	-R-...	429
HIVCAR4054	---TD---YNI---	-G-...G-M-Q...	-ET-G-...-IR-	-P-...-I-...	-I-...	424
HIVCAR286Afr	LQYIS---S-TDT-	-S-...-SSR...	-TG-...DIIT-	-P-...-I-...	-I-...	446
HIVCAR4023	-D-R-...-I-K-	-D-...G-I-GNITGWE-	-FNDISII-	-P-...-P-...	-P-...	446
HIVCAR423A	---K---K---	-G-...TEE-	-ESEDPIV-P-	-P-...-H-...	-H-...	435
HIVKENYA	---S-G---TMQES...	-S-ESN...	-P-...-I-...	-P-...-I-...	-I-...	421
CONSENSUS.B	Cn?tgLfinstW?????????	B?T?????????	n?C?????????	it??...	Lp.CrIkqIlnmWQe	391
HIVJRCSE	---S---...---	-D-...EKS...	-SG-...EGND-	-I-...-V-...	-V-...	421
HIVJRFL	---S---N---	-N-...EGS...	-N-...EGM...	-P-...-V-...	-V-...	420
HIVALA1	---S---...---	-S-EGPKGS...	-N-...EE-N-	-P-...-V-...	-V-...	422
HIVBRVA	---T---YR---	-T-...G...	-I-...EGNSP-	-P-...-V-...	-V-...	423
HIVJH3	---S---LS---	-S-W-NDTEGS...	-N-...GGND-	-P-...-V-...	-V-...	441
HIVBAL1	---S---...---	-V-...EES...	-N-...VERN-	-P-...-V-...	-V-...	427
HIVYU2	---S---...---	-D-...RKL...	-N-...GR-N-	-P-...-V-...	-V-...	416
HIVMN	---TSP---NG...	-N-...WN...	-T-...GSNNN-	-Q-...-K-...	-K-...	429
HIVHXB2R	---S---F---	-S-W-STECS...	-N-...EGSD-	-P-...-K-...	-K-...	429
HIVLAI	---S---F---	-S-W-STECS...	-N-...EGSD-	-P-...-F-...	-F-...	434
HIVNL43	---S---F---	-S-W-STECS...	-M-...EGSD-	-P-...-F-...	-F-...	427
HIVMFA	---S---F---	-S-W-STECS...	-N-...EGSD-	-P-...-F-...	-F-...	427
HIVCAM1	---T---T-LF-	-G-W-...D...	-EGLANTERN-	-P-...-R-...	-R-...	428
HIVNY5CG	---KT---LF-	-S-W-...D...	-ERS-DNNE-	-I-...-R-...	-R-...	421
HIVADA	---S---NF---	-G-W-...L-QSNGT-	EGND-	-P-...-R-...	-R-...	425
HIVJFL	---S---LS---	-G-W-...D...	-EESKDTREP-I-	-P-...-R-...	-R-...	424
HIVSIM184	---S-K---...---	-D-...STW...	-D-...TGSNKN-	-L-...-R-...	-R-...	427
HIVB31	---SA---...---	-D-...KES...	-N-...NG-	-P-...-R-...	-R-...	421
HIVSF162	---S---...---	-N-...IOP...	-N-...NG-	-P-...-R-...	-R-...	419
HIVBCSG3C	---KS---AG...	-N-W-...SSA-	ERSDDTCGN-	-P-...-V-...	-V-...	431
HIVVOYI	---TS---...---	-D-...TRA...	-S-...EV-	-P-...-V-...	-V-...	426
HIVSF33	---T-K---NE...	-S-W-...A...	-GND-	-P-...-V-...	-V-...	423
HIVCDC4	---S---A---	-V-...SNGTW...	SV-RKQKDTGDI-	-P-...-R-...	-R-...	440
HIVSF2	---T---N-RL-	-H-...EGT...	-KG-...ND-I-	-P-...-V-...	-V-...	426
HIVSF2B13	---TA---NP-	-N-...EGN...	-T-...DP-I-	-P-...-V-...	-V-...	424
HIVHAN	---S-K---...---	-N-...STAND-	-G-...ND-I-	-P-...-V-...	-V-...	426
HIVRF	---T---...---	-S-...EGS...	-N-...GGND-	-P-...-V-...	-V-...	436
HIVMMJ2	---S---...---	-G-D-...EKQDNK-	-S-...L-	-P-...-G-...	-G-...	419
HIVTB132	---T-K---...---	-S-...EGN...	DDG-EERN-	-P-...-V-...	-V-...	408
HIVBR020	---T---T---	-N-...EGTS...	-P-...GND-	-P-...-V-...	-V-...	424
HIVTH014	---T---...---	-D-...STWN-	-N-...GNG-	-P-...-V-...	-V-...	416
HIVUS711fr	---T---TTL-	-N-...GKS...	-S-...AGN-	-P-...-R-...	-R-...	422
HIVUS712	---S---...---	-S-...OLV...	-D-...GGN-	-Q-...-V-...	-V-...	423
HIVUS715	---S-K---...---	-G-...ES...	-N-...GDDP-V-	-P-...-V-...	-V-...	425
HIVUS716	---S---YN-	-C-...S-	-S-...EGNNTI-	-Q-...-Q-...	-Q-...	419
HIVHA593	---T---...---	-S-...EES...	-S-...EEG-	-P-...-V-...	-V-...	445
HIVHA594	---TSK---...---	-D-...KRS...	-ID-...G-	-P-...-V-...	-V-...	422
HIVHA596	---TSK---...---	-D-...KRS...	-SDIDGI-I-	-P-...-V-...	-V-...	424
HIVHA599	---TSK---EF-	-S-W-NGT...	EGLDETNAGTGGN-	-P-...-V-...	-R-...	436
HIVHA651	---T-K---...---	-S-...EES...	-KYNESN-LT-	-P-...-V-...	-V-...	420
HIVRJS	---S---MS...	-S-...G...	-D-...TGLN-	-P-...-R-...	-R-...	424
HIVGUNfr	---T---LV-	-G-W-...N-	-KRSND-	-Q-...-S-...	-S-...	390
HIVSC	---S---S---	-G-...EGS...	-N-...GGND-	-P-...-E-...	-E-...	423
HIVSBA	---T-K---TW-	-S-W-...G-	-G-...EGNNTN-	-P-...-V-...	-V-...	171
HIVSBB	---T-K---ND-	-SS-W-...V-	ERSKAEENDIL-	-Q-...-V-...	-V-...	176
HIVSBC	---S---YS-	-N-W-...G-	-G-...GGN-	-P-...-R-...	-R-...	169
HIVBR014	---T---A-ML-	-G-W-...TGS...	-M-...KN-	-P-...-R-...	-R-...	426
HIVTH026	---TS---NN-	-G-W-...G-	-G-...TGYS-	-P-...-V-...	-V-...	416
HIVBR021	---T---LSN-	-D-...V-H-	-V-...T-	-P-...-T-...	-T-...	423
HIVACH9	---S-K---...---	-V-...EET...	-SN-EGDSTIGND-	-V-...-V-...	-V-...	387
HIVACP1	---KTR---Y-SD-	-S-W-...G-	-D-...GGDN-	-P-...-V-...	-V-...	375
HIVJM	---DS---D-QLNKSDGNW-	-S-...SSIW-	-N-EG-NITGD-	-Q-...-V-...	-V-...	414
HIVMN	---A---I---	-G-...G-	-G-EGS-NITGDS-	-P-...-V-...	-V-...	392
HIVMA208	---T---TWN-	-G-...GESN-	-V-...GTEGNS-	-Q-...-F-L-K-	-K-...	449
HIVBP96	---TA---...---	-V-...GGT-	-G-...EGNDI-	-Q-...-K-...	-K-...	425
HIVBNB11A	---S---...---	-N-GEFNNIEG-	-N-...CKN-	-P-...-K-...	-K-...	298

CONSENSUS.C	CntS?LFnsty????p?.....ngt???.????.????.?IT.....lp.crikQiInmwqe	384
HIVMA959	---G---NF.....-D-EG.....NST.....LS.....	420
HIVMA960	---R-----N-----VS.....NDTE.....SNST.....	415
HIVD757	---N-----M-----KG.....NS.....NST.....IQ.....	306
HIVD747	---D---G---M-----KS.....NSS.....ST.....I.....	351
HIVD760	---G---G---M-----KS.....NL.....NST.....I-----V-L---	360
HIVN0P	---TD---SA.....-A-KF.....NES.....DPNAT.....	293
HIVZAM20A	---I---RSHLG.LP.....-E.....MNS.....PCN.....I-----N-----	436
HIVD1044	---N-----S-YML.....NGTT.....SNSTLN.....I.....	303
HIVD1024	---N-----M-----KSNST.....IL-S-----V-----	413
HIVD744	---N---A---M-----YMF.....NGTYMF.NSTSNAT.....I-----C---	316
HIVD766	---G---G---MH.....KSNSSSI.....I-----	298
HIVD808	---G---G---M-----KSNSSSI.....I-----	314
HIVD868	---N---G---M-----KSNSSST.....IL-----VV-----	310
HIVR025	---S-----T-----S-----	366
HIVSM145A	---G-----N-----N-----GSESNMN.....-Q-----	421
HIVZAM18A	---N-----K-----D-----NST.....YNPNT.....	421
HIVDJ259A	---D-----Q-----Q-----Q-----Q-----Q-----	405
HIVDJ373A	---A---DSITN.A-I.....NSST.....DAN.....-Q-----	418
HIVSE364A	---D---K---G-TMT.....ANS.....N-----IQ-----	414
HIVUG268A	---S-----SDMN.....S-----I-----	420
CONSENSUS.D	CntS?LFnstw????.....n?t??????.nst?..?????n?tit.....lp.CrIkQiInmwq?	377
HIVJY1	---R-----SSTWN.....-D-----LNSEQ-K.....	436
HIVNDK	---R-----Q-----GFM-G-V.....-V-L-R.....	419
HIVMAL	---K-----Q-----NGARLS.....ESTGS.....-K-----	429
HIVELI	---G-----ISA.MN.....-I-----ESMNST-TN.....-Q-----K-VAG	427
HIVZ226	---G-----DISKEWA.....ESDDK.P.....-Q-----G-----	426
HIVUG024	---K---HSVGE.....-N-LS.....I-N.....GV-TN.....-K---V-C-G	424
HIVUG021	---G---NSVWTS.....-S-I.....GA-G.....-G-----	413
HIVUG038	---Q---GS.....-D-S.....Y.....-N-K.....I---R---L-E	420
HIVUG046	---DA-R---R-Y.....-TSG.....GVN-S-K.....-G-----	423
HIVUG266A	---G---NSVWR.....-DYS.....GVNESV.....-K-----L-G	424
HIVUG269Afr	---D-R---DSY.....TSALNSTG.....ES.....EN.....-E-----	426
HIVUG274Afr	---R---NSGWENY.....IS.....GG-E-R.....-Q-----E	431
HIVSE365A	---G---XT.....KSG.....DN.....L-R	422
HIVCAR4020	---Q---HK.....-D-TGEMSEQ.....NRT.ES.....EN.....-R---R---R	458
CONSENSUS.E	CNTt?LFn?c????.....NeTm.....???.N?t.....???.?IT.....LP.CrIkQiInmwQq	395
HIVTH966	---R---N-CIG.....-I.....GGC.....-D-----I-----	423
HIVTH975	---K---S-WIG.....-G---BOR.....SS.....DTP-V.....	420
HIVTH022	---R---SN-CIK.....-AGC.....-G-----I-----	422
HIVTH243	---R---N-CIG.....-NGC.....-G-----	428
HIVTH006	---PK---G-MEGP.....-N-E.....GPN.....-T-----I-----V-C-	420
HIVTH011fr	---Q---N-CIG.....-KGC.....-D-----I-----	401
HIVCARELO	---N---S-WEI.....-K---KS.....-D-----NLSN.....-R---G---	433
HIVCAR4017	---Q---S-WRI.....-EKV.....-G-----NSNM.....-R-R-V---R	431
HIVCAR4039	---R---S-WGT.....-I.....ES.....-V-----NPRN.....-R-MV---R	423
HIVCAR4071	---R---IST.....-G---T.....-G-----R---V-K---E	425
HIVCARMB A	---K---S-WEM.....-K-I.....EEL.....-G-----R---V---	418
CONSENSUS.F	CNTskLFnd????.....Nnt??????.??????.LP?CrIkQIVmWQr	387
HIVB2126Afr	---G---TGS.....-G-----IT.....-L-----I-----E	414
CONSENSUS.G	CNTSgLFN7a????.....?N?c??????.n?????..?????i?c??..lp.CrIkQIVmWQr	379
HIVVI525A	---K---E---STWLF.....-S-GLVNSTW.....-DTR.....KD-T.....-Q---K---	439
HIVLBV217	---S---ILS.....-NNAPS.....-SS.....KN-T.....-Q-----K---	431
HIVCAR4067	---S-----S-KTTNDTKT.....-DT.....-T-----Q-----K---	426
CONSENSUS.H	CNT?KLFNSTW????.....N?????.N????..NGT.....????..L?.CRIKQI??mWQr	102
HIVCA13	---T-----TTNT.....-GSYVS.-NTE.....IT.....-Q-----VS---	161
HIVVI557	---S-----V.....-STS.....-DIS.....TIE.....P-----IN---	156
CONSENSUS.O	CNTs?mFNYtFsc??????.N7t??c??????.N7tn?.....?NgT??..ip.CrIkQIVrWwQr	371
HIVANT70	---AK-----G---T.....SVS-VSQG.....N-----L---K-----I---	426
HIVMVP5180	---G-----I.....-C-KSG-QEIKGS-E.....K-----K---L---K	436
HIVVAU	---NRL---H---KKNMTNKKI-C.....-IS-N.....S---QA.....D---	438
CONSENSUS.CPZ	C?????P??????.?IT.....NG??I.....??CRI?QIV??W??	187
STVCPZGAB	---NTSQI-TD.....N---II.....LP---R---SS-MR	413
CONSENSUS.U	CNTS?LFns?w??.....ns??????.n?c??????.?IT.....LP.C?IkQi??mWQ?	369
HIVZAM184	---N---TYNP.....-T.....ESNS.....DLNIT.....-R---IN---R	411
HIVCAR4081	---R---S-KG.....KYSWES.....-D-----K-R---VR---K	423
HIVK124A	---K---V-G.....-S.....-V-KV.NGTVKET.....-K---IN---E	431
HIV231r	---E---TCI.....-GT.....WD.K.-C-STEENCT.GN.....-R---VVRP-G	423

HIV1 ENV

CD4 -binding-		* ***		***	
CONSENSUS.A	vqg.AmyapPiQ.g?irc?snITGLlLTrDgG??...	n?????	???	net.FrPgGgdnrdnNrsE	416
HIVSF1703	---V-K-I---	D---	NSA---	T-N---	487
HIVU455	---V-E---	T---	NTK---	---K---	475
HIVZ321	---K-V-K-V---	I---	VG---	NTS---	482
HIVRW020	---P-V-E---	---	NST---	I-N---	463
HIVUG031fr	-E-P-I-D-	G---	DANS---	TNG-N---	432
HIVD687	---E-EL-E---	---	TSM---	TNET---	362
HIVNI	---I-D---	---	NST---	---	473
HIVRW009	T---V-S-V---	N---	NTT---	T---	477
HIVUG275A	A-K-I-K-V-V---	I---	GN---	SE---	476
HIVUG273A	---R-K-V-E---	---	NG---	SNT---	481
HIVVI191A	---M-E-E---	I---	VGN---	ST---	480
HIVDJ264A	---R-P-E-E---	---	SN---	ST---	472
HIVDJ263A	---L-S-E-K---	---	SN---	ST---	480
HIVDJ258A	---K-I-S-E-E---	I---	SN---	ST---	474
HIVCARGAN	A---T-I-E---	---	FN---	RT---	485
HIVCARSAS	---K-K-A-N-R---	---	G---	NRS---	482
HIVCAR4054	---EL-D---	---	V---	NT---	477
HIVCAR286Afr	---RV-E---	---	L---	ST---	499
HIVCAR4023	---S-P-E-S-N---	X---	DO---	ST---	499
HIVCAR423A	---T-E-K-E---	---	K---	TT---	487
HIVKENYA	A---I-K-V---	I---	DN---	SE---	474
* ***					
CONSENSUS.B	Vqk.AmyapPi?GqircsSnITGLlLTrDgG?????	n?k?	???	E1.FRpgGgdnrdnNrsE	439
HIVJRCSE	---K---	K---	ES---	EI---	474
HIVJREFL	---R---	INE---	G---	---	473
HIVALA1	---E---	---	SS---	NNT---	475
HIVBRVA	---R-K---	MN---	E---	TDT-N---	478
HIVJHJ	---E---	DNQ---	E---	T-N---	494
HIVBAL1	---R---	PED---	K---	V---	480
HIVYU2	---R---	KDT---	G---	---	469
HIVMN	---E---	KDT.DF---	D---	---	484
HIVHXB2R	---S---	NSN---	ES---	---	482
HIVLA1	---S---	NNN---	GS---	---	487
HIVNL43	---S---	NNN---	GS---	---	480
HIVMFA	---S---	KNT---	ES---	V---	480
HIVCAM1	---T-T-S---	RGE---	E---	T---	481
HIVMY5CG	---S---	DKE---	S---	T---	475
HIVADA	---R---	I---	T---	SS---	478
HIVJFL	---N---	KE---	GA---	N-K---	476
HIVSIM184	---E-H---	K---	ES---	ET-T---	481
HIVD31	---S---	KNKD---	E---	T-N---	475
HIVSF162	---R---	KEIS---	T---	---	473
HIVBCSG3C	---S---	T---	TSOT---	T---	485
HIVOYI	---S---	K---	K---	T---	480
HIVSF33	---E---	GDK---	S---	T---	477
HIVCDC4	---L-K-L---	GE---	Q---	T---	493
HIVSF2	---G-S---	T---	V---	NDT-V---	480
HIVSF2B13	---G-S---	I---	G---	NDT-V---	478
HIVHAN	---G-L---	I---	ND---	SS---	480
HIVRF	---S-K-I---	EDTT---	T---	L-N---	490
HIVVMJ2	---Q---	---	SS---	SRE-N---	472
HIVTB132	-X-L-D-G-	K-T-	NQ---	G---	463
HIVBR020	---SV-K-K-KS---	---	NE---	MNTT---	478
HIVTH014	---R---	R---	SE---	NKTT-T---	470
HIVUS711fr	---R-L---	S---	KN---	ET-T---	475
HIVUS712	---R---	IN---	KSE---	NGT---	478
HIVUS715	---R---	NS---	E---	NTT-N-K---	480
HIVUS716	---R-K---	T---	A---	N-T-N---	471
HIVHA593	---E---	MN---	K---	NGT---	500
HIVHA594	---S---	EN---	E---	T---	474
HIVHA596	---S---	EN---	G-E---	NNET-T---	481
HIVHA599	---AS-S-K---	NV---	ES---	KDN-T-K---	491
HIVHA651	---E---	NN---	D---	TET-T---	475
HIVRJS	---E---	I---	G-N---	DT-T---	478
HIVGUNfr	---Q---	---	GSE---	NT---	443
HIVSC	---K-VK---	ISK---	GSK---	NENT---	481
HIVSBA	---R---	I---	KNT---	E---	224
HIVSBB	---R-L---	NNNNNT---	NN---	T-A---	233
HIVSBC	---R---	I---	NNT---	V---	222
HIVBR014	---D-L-H---	---	T---	NRTE---	480
HIVTH026	---K-H---	---	NE---	NTT-K-K---	469
HIVBR021	---S-S---	---	IS---	ADN-T---	476
HIVACH9	---R-N-K---	N---	E-T---	DT-T---	441
HIVACP1	---N---	DE---	EIT---	NI-T-N---	430
HIVJM	---R---	NSNS---	D-N---	KT-T-N-K---	471
HIVNM	---A---	IN---	N-N---	NDT---	448
HIVMA208	---Q-N-S---	IN---	ASK---	DN-T---	504
HIVP896	---T---	---	S-E---	T-ET---	479
HIVBWB11A	---R---	SSS---	ES---	---	334

**CD4
-binding-**

CONSENSUS.C		VGR?amYAPPI?.gnlCCKSNITGLILLRDGG??????n?Tt???????ET.FrPGGGMrdMNRSE	431
HIVMA959	-E-	-L-G-PT....TK.....T..	473
HIVMA960	-A-	-L-HT...T.....NNT.	469
HIVD757	-E-	-E-V-TES...M-----N-	359
HIVD747	-A-	-I-V-IEL...D.....KT.....E-	406
HIVD760	-FA.R-	-V-EDT...D...-I-S-	413
HIVNOF	-TS-E-	-N-T-NA...T.....MYTT.-I-----N-K--N-	350
HIVZAM20A	-K-P-A-	-I-T-IDT...S-KITHETNGTAN-	500
HIVD1044	-E-	-T-E-PD...M-----I-----N-	355
HIVD1024	-E-	-V-TE...S-T-----N-N-	466
HIVD744	-E-	-V-RG.QES-T...-I-----N-	371
HIVD766	-E-	-V-TE.S-N-ET.....	353
HIVD808	-E-	-V-TE.S-M-----N-	367
HIVD868	-E-	-A-TE.S-H-----I-----N-	366
HIVSH145A	-Q-K-	-T-L-NDPK.RTEE.QNMT.-R-	480
HIVZAM18A	-Q-A-	-T-S-D-T.NT.....	475
HIVDJ253A	-K-I-	-V-T-T...-T-TEG.-I-	459
HIVDJ373A	-Q-Q-I-A-	-T-T...S-D.....	470
HIVSE364A	-I-	-A-T-TTN...N-----I-	467
HIVUG268A	-K-K-R-	-T-ETSET.-S-	475
		* * *	
CONSENSUS.D		vGk.AmYaPPIe.g7I7CsSnITGLILLRDGG?????PNs????????net.fRPGGGMrdMNRSE	425
HIVJY1	-L-K-T-	-VN-ST-----R--N-	489
HIVNDK	-L-K-	-AN...S...H-I-	472
HIVMAL	T--A-V-N-L-	-I-S-DNS...D--L-I-	484
HIVELI	.R.-I-	RN-L-IN-ST-----	479
HIVZ226	-Q-N-	-TN-S-	479
HIVUG024	-Q-L-R-	-T-Q-	476
HIVUG021	-T-Q-N-T-	-VK-N-Q-----N-	467
HIVUG038	-N-L-N-T-	-IGK-N-Q...N-D-	476
HIVUG046	-L-K-	-N-TRO-	476
HIVUG266A	-P-L-K-	-V-N-Q-	477
HIVUG269Afr	-I-R-	-IT-M-----N-	479
HIVUG274Afr	-I-Q-K-	-V-N-S.NETSSK-I-	490
HIVSE365A	-V-S-	-L-IH...-T.NET	477
HIVCAR4020	-Q-V-N-T-	-NTK...G...KNDI-I-	515
		* * *	
CONSENSUS.E		vGq.AmYaPPis.g7INCvSNITGLILLRDGG?????n?Tt?????net.fRPGGGMrdMNRSE	444
HIVTH966	A-L-	-R-A...N-TS...	476
HIVTH975	A--K-	-D...N-S-T...	474
HIVTH022	A--R-	-ANNA	474
HIVTN243	A--K-	-A-T-T...	481
HIVTH006	A--R-	-A-N-T...	473
HIVTH011fr	-R-	-D-IDKT...	454
HIVCARELO	--R-.SLV-	-P-PA.....M-	485
HIVCAR4017	-VV-	-I-E-HE	484
HIVCAR4039	-K-V-	-L-SNSGE...P--IMR-	476
HIVCAR4071	-D.EA-	-V-KIE...E-I--M-	479
HIVCARNBA	-N-V-V-	-ML-N-Q-----N-	472
		* * *	
CONSENSUS.F		VGr.AmYaPPIa.G7IqCnS7ITGILLTRDGG?????N?????TeT.lRPGGGMrdMNRSE	433
HIVB2126Afr	-A-E-N-T-N-	-QNS...QT...F--N-K-	467
		* * *	
CONSENSUS.G		vGQ.AmYaPPIA?GNlIC7SnITGLILTROGG.....?nntT?????net.fRPgGGGMrdMNRSE	430
HIVVI525A	-A--E-R-S-	-N-N-----G-	493
HIVLBV217	-K-	-NG...A-	484
HIVCAR4067	T--N--I-	-L-EN...RT-	481
		* * *	
CONSENSUS.H		?GQ.AmYaPPI?.GKITC?SnITGLILTSCHV.....NFSAS?.....?VT.fRPGGGMrdMNRSE	150
HIVCA13	A--N-V-	-R-E-	214
HIVVI557	V--K-E-	-S-N-	209
		* * *	
CONSENSUS.O		Ggs.gLYAPPI?.GNLIC7SnITGMILQ1O7PW.....NnsT?????n?t.lRP?GGGMrdDIWRte	418
HIVANT70	-Q--K--M-M-M-N-T-	-NN--V-F-I-	480
HIVMVP5180	-E-RI-	-P-M-Q...TGE...-V-----K-	489
HIVVAU	-P-V-R-	-T-XTHP...NST--G-----Q-	493
		* * *	
CONSENSUS.CP2		V7?.GIY?F2R.GN7?C?S?ITG?PL?????N?T?????T??PWT?	207
SIVCP2GAB	-GR---AP-I--IP-N-N--LL-TSDTPVT...	-NSG...NLT.FR-TGGNMK-I--S-	467
		* * *	
CONSENSUS.U		VG?.AmYaPPIe.GvIrCsSnITGl?LrDGGn?????Nstq?.....??ET.fRPGGGMrdMNRSE	419
HIVZAM184	-R-A-	-K-T-II-S-SG...NGT...	467
HIVCAR4081	-Q--P-	-E-I-N-K-NN...T-	477
HIVK124A	-R-L-N-	-L-E-N-S-N-	485
HIVZ3fr	-Q--T--L-	-L-Gx-A-N-N--K-	477

HIV ENV

		gp120/gp41	
CONSENSUS.A	LYkYKvVkieplgvaPtr.AkrRVV.....eREKRA?vg.IGavflgFLGaAGSTmGAaSiTltvQARqL		478
HIVSF1703	-----P-----	Q-----I-----I-----	549
HIVU455	-----I-----	-----I-----I-----	537
HIV2321	-----K-----	A-----I-M-P-----R-	544
HIVRW020	-----S-----	-----I-----I-----A-----	525
HIVUG031fr	-----	I-M-----I-----M-----	494
HIVD587	-----D-----	-----	424
HIVNI	-----	-----R-----	535
HIVRW009	-----	-----I-----	539
HIVUG275A	-----R-----	V-----	538
HIVUG273A	-----Q-----P-----	Q-----I-----	543
HIVV1191A	-----S-----	A-----L-----	542
HIVDJ264A	-----I-H-G-----	-----F-----	534
HIVDJ263A	-----	-----	542
HIVDJ258A	-----	-----F-----	536
HIVCARGAN	-----LT-TRIR-----	-----	502
HIVCARASAS	-----I-----	-----	500
HIVCAR4054	-----E-K-I-----	-----	495
HIVCAR286Aer	-----E-N-----	-----	517
HIVCAR4023	-----E-----	-----	517
HIVCAR423A	-----R-E-----	-----	505
HIVKENYA	-----K-----	I-----	536
CONSENSUS.B	LYKXKvVkieplGvAPTK.ahrRvv.....QrekRa..vg?iGanf1GfLGaAGSTMgaasmtLtVqaRqL		501
HIVJRCSP	-----	L-----R-----	536
HIVJREFL	-----	V-----L-----	535
HIVALA1	-----R-----	L-L-----R-----	537
HIVBRVA	-----	L-L-----L-----L-----	540
HIVJH3	-----L-----	V-----L-----	555
HIVBAL1	-----	V-----A-----L-----	542
HIVU2	-----	L-L-----I-----	531
HIVMN	-----T-----	A-L-----V-----L-----	545
HIVHXB2R	-----	L-----	544
HIVLA1	-----	L-----R-----	549
HIVNL43	-----	L-----CT-----	542
HIVMFA	-----T-----	L-----	542
HIVCAM1	-----	A-L-----VAL-----T-----	544
HIVNY5CG	-----	AL-L-----A-----T-----	538
HIVADA	-----	-----I-----I-----L-----	541
HIVJFL	-----	K-----T-----L-----	539
HIVSINI84	-----	L-L-----	544
HIVD31	-----	LL-V-----R-A-----	538
HIVSF162	-----	TL-----R-L-----	535
HIVBCSG3C	-----T-----	L-----	547
HIVQY1	-----R-----	ML-----R-----	543
HIVSF33	-----	V-----I-----K-----	540
HIVCDC4	-----	ML-----T-A-----	556
HIVSF2	-----I-----I-----	IV-----V-L-----	543
HIVSF2B13	-----I-----I-----	I-----R-----R-----	541
HIVHAN	-----	ML-----R-L-----	543
HIVRF	-----R-----R-----	T-----G-I-----H-----	553
HIVWMJ2	-----R-----	T-----G-L-----	535
HIVTB132	-----X-----	ASP-----X-PT-----P-----	525
HIVBR020	-----R-----I-----R-----	TL-----VA-----	541
HIVTH014	-----P-----	T-----T-----I-----	533
HIVUS711fx	-----I-----	V-----A-----L-----	537
HIVUS712	-----	F-----L-----	540
HIVUS715	-----R-----I-----R-----	A-L-V-----YA-----	542
HIVUS716	-----R-----	-----L-----L-----	533
HIVHA593	-----	IV-----L-----L-----	563
HIVHA594	-----R-----	AL-----	537
HIVHA596	-----K-----	VL-----	544
HIVHA599	-----S-----	VL-----L-----	554
HIVHA651	-----	TL-----L-----	538
HIVRJS	-----	V-----A-----	541
HIVGUMfr	-----I-----	T-V-----I-----	506
HIVSC	-----	T-----T-----L-----	544
HIVSBA	-----R-----I-----	-----I-----L-----	286
HIVSBB	-----IQ-----	VLF-----L-----	295
HIVSBC	-----	V-----I-----L-----	284
HIVBR014	-----R-----	R-----AL-----L-----	543
HIVTH026	-----R-----	T-----I-M-----	532
HIVBR021	-----R-----	T-----I-----PV-RI-----	539
HIVACH9	-----R-----	T-----	484
HIVACF1	-----R-----	I-T-----L-----	473
HIVJM	-----R-----	T-M-----L-----	513
HIVWM	-----R-----	L-----	490
HIVMA208	-----T-----M-----	-----	533
HIVP896	-----R-----I-----R-----T-----	V-----R-V-----L-----	541

		gp120/gp41	
CONSENSUS.C	LYKYKVVEiKPIGvaPT?.aKRRVV??...	ereKRA..vg?igAvfLGFLGaAGStmgAASitLTvqaRQl	492
HIVMA959	-----K-----	-----V-----	535
HIVMA960	-----E-----	-----	531
HIVD757	-----T-----	-----M-----	421
HIVD747	-----T-----	-----M-V-----	468
HIVD760	-----T-----	-----L-----	475
HIVNOF	-----I-G-S-----	-----L-----A-----	412
HIVZAM20A	-----I-G-P-----EK..QI-----	-----L-----V-----	564
HIVD1044	-----T-P-----	-----PL-----	414
HIVD1024	-----A-----	-----	523
HIVD744	-----A-----	-----G-----L-----	418
HIVD766	-----I-A-----	-----L-----	400
HIVD808	-----T-----	-----	424
HIVD868	-----V-T-----	-----M-----T-----	410
HIVSM145A	-----I-G-----	-----	542
HIVZAM18A	-----I-A-----E..T-----	-----A-----V-----	538
HIVDJ259A	-----V-I-K-----	-----G-----A-L-M-----	520
HIVDJ373A	-----L-K-----	-----A-LA-----	531
HIVSE364A	-----I-K-----	-----	529
HIVUG268A	-----V-----K-----	-----	537
CONSENSUS.D	LYKYKVvxiePIG?APT?.aKRRVV....	ereKRA..IG.LGa?FLGFLGaAGSTMGAesltLTvQARql	484
HIVJY1	-----I-R-----	-----V-----V-VA-G-----	551
HIVNDK	-----K-I-V-K-R-----	-----V-----V-----	534
HIVMAL	-----V-K-----	-----M-----	546
HIVEL1	-----Q-V-R-----	-----M-----R-V-----	541
HIV2226	-----K-V-R-----	-----M-----R-----	541
HIVUG024	-----IK-L-R-----	-----A-----V-----	538
HIVUG021	-----L-K-R-----	-----L-----	529
HIVUG038	-----I-M-S-----	-----L-----T-----	533
HIVUG046	-----L-E-----	-----M-----M-----	533
HIVUG266A	-----I-L-L-R-----	-----V-----L-----	539
HIVUG269Afr	-----L-T-----	-----L-----V-----	541
HIVUG274Afr	-----R-R-----	-----L-----T-----M-----V-----	552
HIVSE365A	-----L-V-T-----	-----K-----VM-----MA-----	539
HIVCAR4020	-----K-I-R-----	-----	533
CONSENSUS.E	LYKYKVQIEPIGiAPTR.AKRRVV....	EREKRA..VG.IGAMIFGFLGaAGSTMGAASITLTvQARQL	506
HIVTH966	-----	-----	538
HIVTH975	-----	-----	536
HIVTH022	-----	-----	536
HIVTH243	-----	-----	543
HIVTH006	-----	-----	535
HIVTH011fr	-----	-----P-----	511
HIVCARELO	-----	-----	503
HIVCAR4017	-----V-----	-----	502
HIVCAR4039	-----I-V-----	-----	494
HIVCAR4071	-----	-----	497
HIVCARMBA	-----	-----	490
CONSENSUS.F	LYKYKVVKIEPLGVAPTr.AKRRVV....	krakRA..?VG.igAvfLGFLGaAGSTMGAASITLTvQARQL	495
HIVB2126Afr	-----E-----	-----V-----L-----S-----	530
CONSENSUS.G	LYKYKVkiKIPLGvAPTr.ARRRVV....	?REKRA..I?.PGAVFLGFLGaAGSTMGAAS??LTvQVRQL	487
HIVVI525A	-----V-S-----	-----E-----G-L-----IT-----	555
HIVLBV217	-----K-----	-----G-----T-----LA-----	545
HIVCAR4067	-----I-----	-----	499
CONSENSUS.H	LYKYKVv??EPLG?APT?.A?RRVV....	?REKRA..VG.L?AVFLGFLGaAGSTMGAAS?LTvQARQL	204
HIVCA13	-----KI-V-K-K-----	-----G-----G-----I-----	276
HIVVI557	-----RV-L-R-R-----	-----E-----E-----V-----	271
CONSENSUS.O	LfkyKVvrvKPFsvAPtk.iARP?Igt?c?HREKRA..	VG.LgMLFLGvLSAAGSTMGAASATaLcVqhtl	482
HIVANT70	-----N-----R-----V-S-R-----	-----T-A-----	546
HIVMVF5180	-----YN-QI-----MS--I-NIH-P-----	-----R-SV-----	556
HIVVAU	-----T-----RS-----	-----A-A-I-----R-QH-----	559
CONSENSUS.CPZ	L??Y?V?V?I?PLSVAPT?..??R???Q?..?ROKR?..??G..?G??FLG?L?AAGSTMGAAS??LT?Q?R?L		245
SIUCPZGAB	-----YK-K-R-E-----K.AR-HIVAR-K.D-----A.AF-L-AL-----F-G-----AVT-V-A-Q-----		534
CONSENSUS.U	LYKYKVVKIEPLGVAPTr.aKRRVV....	EREKRA..VG.LGA?FLGFLGaAGSTMGAASmLTvQARQL	480
HIVZAM184	-----L-----K.PR-----	-----V-----I-----	529
HIVCAR4081	-----	-----	495
HIVK124A	-----L-S-----	-----I-M-----	547
HIV23fr	-----	-----M-L-----A-A-----	540

HIV1 ENV

CONSENSUS.A	1SGIVQQQnNLLrAieaQqhlLkLTVWGSIKQLQARvLA?ErYLrDQQLLGINGCSCGKliCtTnVPWNaSW.	547
HIVSF1703	-----V-----K-----	619
HIVU455	-----V-----Q-----T-----	607
HIVZ321	-----N-----I-----V-----K-----I-----P-----	614
HIVRW020	-----M-----V-----K-----	595
HIVUG031fr	-----R-----L-----	564
HIVD687	-----M-----L-----K-----T-----	494
HIVNI	-----K-----L-----	578
HIVRW009	-----K-----V-----	569
HIVUG275A	-----K-----V-----	608
HIVUG273A	-----L-----V-----T-----	613
HIVVT191A	-----H-----Q-----L-----P-----Y-----	612
HIVDJ264A	-----L-----T-----A-----	604
HIVDJ263A	-----L-----S-----	612
HIVDJ258A	-----L-----S-----T-----	606
HIVKENYA	-----V-----I-----	606
CONSENSUS.B	1sgIVQQQnNLLrAieaQqhlLqLTVWGSIKQLQARvLAvERYLkDgQLIgiWGCSCGKliCtFavPWNaSW.	571
HIVJRCSE	-----N-----M-----T-----	606
HIVJRF1	-----RM-----G-----	605
HIVALA1	-----R-----E-----T-----	607
HIVBRVA	-----M-----M-----E-----	610
HIVJH3	-----G-----I-----	625
HIVBAL1	-----R-----	612
HIVYU2	-----R-----T-----T-----	601
HIVBN	-----M-----T-----F-----	615
HIVHXB2R	-----I-----	614
HIVLAI	-----I-----	619
HIVNL43	-----D-----I-----	612
HIVNFA	-----I-----	612
HIVCAM1	-----I-----	614
HIVMYSCG	M-----K-----R-----T-----	608
HIVADA	-----L-----R-----	611
HIVJEL	-----Q-----T-----	609
HIVSIMI84	-----S-----V-----T-----	614
HIVD31	-----R-----	608
HIVSF162	-----R-----	605
HIVBCSG3C	-----T-----	617
HIVQY1	-----T-----	613
HIVSF33	-----K-----I-----R-----T-----T-----	610
HIVCDC4	-----K-----I-----F-----	626
HIVSF2	-----R-----	613
HIVSF2B13	-----R-----T-----	611
HIVHAN	-----R-----T-----T-----	613
HIVRF	-----R-----T-----	623
HIVMMJ2	-----D-----R-----T-----	605
HIVTB132	-----T-----	595
HIVBR020	-----N-----G-----T-----T-----	611
HIVTH014	-----R-----	603
HIVUS711fr	-----T-----K-----T-----T-----	607
HIVUS712	-----S-----N-----KT-----	610
HIVUS715	-----M-----R-----	612
HIVUS716	-----T-----T-----	603
HIVHA593	-----T-----T-----	633
HIVHA594	-----I-----T-----	607
HIVHA596	-----I-----M-----T-----	614
HIVHA599	-----S-----P-----	624
HIVHA651	-----R-----P-----S-----S-----	608
HIVRJS	-----T-----	611
HIVGUNfr	-----K-----I-----F-----R-----T-----T-----	576
HIVSC	-----R-----T-----T-----	614
HIVSEA	-----R-----	356
HIVSBB	-----L-----	365
HIVSBC	-----L-----R-----S-----	354
HIVBR014	-----K-----	573
HIVTH026	-----R-----	562
HIVBR021	R-C	542
HIVP896	-----M-----L-----R-----N-----S-----V-----	611

CONSENSUS.C	LSGIVQqqsnLLrAIEaqqHmLQLTVWGIRQLqRVLAlERYLkdQQLLG1WGC SGKLICTTAvpwnsSW.	562
HIVMA959	-F-----H-L-----A-----Q-----L-----T-----	605
HIVMA960	-----R-----A-----Q-----L-----T-----	601
HIVD757	-----L-----A-----R-----N-----	491
HIVD747	-----L-----I-----A-----E-----T-----	538
HIVD760	-----C-----M-----	545
HIVNOF	-----	458
HIVZAM20A	-----I-----A-----	634
HIVGM145A	-----RA-----E-----	612
HIVZAM18A	-----L-----M-----S-----	608
HIVDJ259A	-----N-----A-----	590
HIVDJ373A	-----K-----VH-----	601
HIVSE364A	-----	599
HIVDG268A	-----N-----Q-----	607
CONSENSUS.D	LSGIVQQQnNLLrAIEAQQHLLQLTVWGIRqIQARvLAVERYLkDQQLLG1WGC SGKbTCTTVPWNSSW.	553
HIVJY1	-----M-----S-----T-----	621
HIVNDK	M-----H-----R-----R-----N-----	604
HIVHAL	-----F-----Q-----R-----M-----F-----	616
HIVELY	M-----I-----N-----	611
HIVZ2Z6	-----I-----L-----T-----	611
HIVUG024	M-----S-----R-----P-----Q-----	608
HIVUG021	-----	559
HIVUG266A	-----M-----S-----S-----	609
HIVUG269Afr	-----S-----G-----R-----N-----	611
HIVUG274Afr	-----K-----I-----T-----	622
HIVSE365A	-----I-----R-----R-----N-----	609
CONSENSUS.E	LSGIVQQQsnLLRAIEAQQHLLQLTVWG1KQLQARVLAVERYLKDQKfLGLWG1CSGKLICTTAVPWNSTW.	576
HIVTH966	V-----	608
HIVTH975	-----M-----	606
HIVTH022	-----X-----X-----XX-----X-----LX-----X-----	613
HIVTM24J	-----	605
HIVTH006	-----	605
CONSENSUS.F	LSgIVQQQnNLLkAIEAQQHLLQLTVWGIRQLQARVLAVERYLQDQQLLG1WGC SGKLICTTNVPWNSSW.	565
HIVB2126Afr	-----E-----R-----	600
CONSENSUS.G	LSGIVQQQ?NLLRAIEAQQHLLQLTVWGIRQLQARVLAVER?L?DQQ?LG1WGC SGKLICTTNVPWNSSW.	553
HIVVI525A	-----S-----F-----K-----P-----	625
HIVLBV217	-----N-----Y-----Q-----L-----	615
CONSENSUS.H	LSGIVQQQsnLL?AI?AQQH?LKLIT	226
HIVCA13	-----R-----E-----L-----	301
HIVVI557	-----Q-----Q-----M-----	296
CONSENSUS.O	?KGIIVQQQDNLLRAIQAOQ?LLR1SvWGIRQLRARLJALET12QNQQ1Ln1WGCCKgkL1CYTSVYWN?tW?	549
HIVANT70	L-----Q-----X-----L-----S-----V-----R-----I-----	617
HIVNVP5100	L-----H-----Q-----R-----TS-----S-----	627
HIVVAU	I-----H-----P-----F-----NR-----K-----G-----	630
CONSENSUS.CPZ	??GIVQQQ?NLL?AIE?QQHLLQLS?HGVKQLQAR?LAVER?YL?DQQ?L?LWGC??K??C?TTVPWNSSW?	299
SIVCP2GAB	LS-----N-----K-----A-----I-----L-----R-----Q-----I-----G-----SG-AV-Y-----P-----	605
CONSENSUS.U	LSGIVQQQnNLLrAIEAQQHLLqLTVWGIRQLQARvLAVERYL?dQQLLG1WGC SGK1ICTTtVPWNSSW.	549
HIVZAM184	-----S-----K-----K-----A-----M-----Q-----	599
HIVK124A	M-----D-----S-----K-----H-----N-----	617
HIVZ3fr	-----ES-----L-----	610

HIV1 ENV

CONSENSUS.A	SNks?adlWdMTWlQWdKElsnYT?liY?LiEesqnQQEkNEqdLLALDkwanLWnWfIdIsn	607
HIVSF1703	RTQ-E-N-----D--N-----I-----E-----K	682
HIVU455	QE--N-----E--S--Q--Q-----L-----N--	669
HIVZ321	Q--K---E---V---QV--N-----T---I--R-----	677
HIVRW020	MNE-----Q--N-----S-----R	658
HIVUG031fr	Y-E-----N--EL--S--D-----	627
HIVD687	TY-----K--A---A-----TS--S---TK	557
HIVUG275A	Q-----N--EL--R--L-----G---	671
HIVUG273A	Q-E-----N--D---A-----S--S---	676
HIVVI191A	YD--N-----D--QL--G-----S--S---	675
HIVDJ264A	TYN--E-----D--N-----K-----ES---N-TK	667
HIVDJ263A	TYN-----V---D--N-----Q-----G-TK	675
HIVDJ258A	TFN-----D--T-----G--S--G-T-	669
HIVKENYA	Q-E--E-----D--N-L-----	669
CONSENSUS.B	SNksl??iw?nmTWmewarEldnYT?liYtLiesqnQQEkNEqdLLALDkwanLWnWf?Itn	629
HIVJRCSP	DS--N-----K--E--NT-----I-----G--K	669
HIVJREFL	DR--N-----SE-----D--K	668
HIVALA1	NQ--D-----S-----N---	670
HIVBRVA	SD--D-----N--S--D--I-----K-----N---	673
HIVJH3	EE--D-----S-----G-----T---	688
HIVBAL1	NX--D---I--D--N--SI--S-----D--K	675
HIVYU2	NE--D---K-----HI--S--Q-----A-----D--K	664
HIVMN	DD--N---Q-----S--L--K--F-----D---	678
HIVHXB2R	EQ--NHT---D--N--S--HS-----N---	677
HIVLA1	EQ--N-----D--N--S--HS-----N---	682
HIVNL43	EQ--N-----D--N--S--HS-----N---	675
HIVMFA	EQF--N-----D--N--S--HS--D-----N---	675
HIVCAN1	DK--N-----N-----KD---T-----D---	677
HIVNY5CG	DK--D-----G-----I-----D--K	671
HIVADA	T--DM--D-----E--G-----D--A-----D--S-	674
HIVJFL	DE--N---Q---S-----L-----D---	670
HIVSIN184	NQ--D-----G--R-----Q--D--K--T-----S--K	677
HIVD31	MDM--N-----S-----N--BN--S--D--S-	671
HIVSF162	DQ--N-----N-----D--SK	668
HIVBCSG3C	NQ--D---Q---E--D--N-----T---	680
HIVOV1	NE--D---Q---H-----G--S--S---	676
HIVSF33	DK--N-----S-----L-----S---	673
HIVCDC4	T--DQ--N-----D--H-----Q--Q-----T--SD--K	689
HIVSF2	ED--D---Q---NT-----L-----S---	676
HIVSP2B13	MED--D---Q--K---NT-----L-----N---	674
HIVHAN	T--DQ--N-----S-----Q-----S--YD--S-	676
HIVRF	NM--N---Q---GI--N--L-----N-----D--Q	686
HIVMKJ2	MNQ--D--L-----SI--S-----G-----D---	668
HIVTB132	DE--N-----N--G-----X-----LD-----D---	658
HIVBR020	DD--T-----K-----S-----R-----D---T--SK	674
HIVTH014	DK--N-----RE-----L-----D--K	666
HIVUS711fr	DK--G-----G-----L-----T---	670
HIVUS712	DQ--Q---Q---K--DV---G-----D--Q	673
HIVUS715	HDQ--Q---Q--K---S--N--V-----D---	675
HIVUS716	DQ--G---Q---G-----N---	666
HIVHA593	SE--D---Q---S-----G-----E---	696
HIVHA594	EK--N-----G--S-----D-----N--K	670
HIVHA596	EK--N-----G--S--G-----D-----N--K	677
HIVHA599	R--QY--N---I-----DI--S--K-----Q-----S--K	687
HIVHA651	EQ--N---L-----S--S--K-----N--E	671
HIVRJS	EE--D-----N--G-----Q-----A--T-----S--S-	674
HIVGUNfr	KQ--D--L-----GI--FN--A-----D-----G-----S-	639
HIVSC	DK--G-----S-----N-----D-----S--S-	677
HIVSBA	DQ--D---Q---E--S--N-----D-----S--S-	419
HIVSBB	ND--D---Q--D--NK--DS--Q-----D--K--E-----S--SK	428
HIVSBC	DQ--N---Q---E--S--N-----S-----S---	417
HIVP896	VDO--N-----DY--D--L--K--T-----K-----D---	674

CONSENSUS.C	sNksq?dIWdNMTWmqWdrEISNYtdtIYLLedSQQQEKNEkdLLALDsw?nLWnWfDItk	623
HIVMA959	S-----E-----I--K-----K-----N	668
HIVMA960	-RT-----K-----N-----N	664
HIVD757	-R--T-----K-----R-----K-----S--N	554
HIVD747	-R--T-----E-----R-----K-----S--N	601
HIVD760	-R--T-----N-----R-----K-----S--N	608
HIVZAM20A	T--TYEY--N-----EK--S--K-----T--Q--E--K--NI-----S-	697
HIVSM145A	-RT--EE--E-L-----K-----V-----GS-----	675
HIVZAM18A	-T-----E--I--R-----K-----N	671
HIVDJ259A	-R--KE-----I--E--N--GL--G-----Q-----K--Q-----S--	653
HIVDJ373A	EE-----M--E-----V--T--Q-----K--Q-----S--	664
HIVSE364A	-LE-----N--SI--S--E-----N-----N--N-----	662
HIVUG268A	-LG-----Q-----K--Q-----S-----	670
CONSENSUS.D	SNrSL?eiW?NMTWmqWdrEISNYTGLIYsLIBeSQ?QQEKNEqLLeLDKwASLWnWfPit?	612
HIVJY1	-K--E--N--I-----V-----N--I-----D--Q-----K	684
HIVNR	-D--Q-----I-----V-----I-----K-----K	667
HIVMAL	-DD--N-----Q--K--S--I--N-----I-----K-----SK	679
HIVELY	-N--Q-----T-----K-----Q	674
HIV2226	-ND--Q-----R-----T-----N--Q	674
HIVUG024	-K--DT--G-----S-----D-----I-----KD-----D--N	671
HIVUG266A	-K--E--N-----V-----Q-----N--Q	672
HIVUG269Afr	-E--H-----T-----Q-----N	674
HIVUG274Afr	-E--N-----T-----N-----Q-----V-----K	685
HIVSE365A	-K--T--Q-----T-----A-----D--N	672
CONSENSUS.E	SNrSfEETWnNMTWmqWdrEISNYTGLIYsLIBeSQ?QQEKNEqLLeLDKwASLWnWfDITn	639
HIVTH966	-S-----S-----	671
HIVTH975	-T-----D-----G-----L-----	669
HIVTH022	-K-----T-----K-----	669
HIVTH243	-X-----X-----S-----	676
HIVTH006	-L-----R-----K--D-----	668
CONSENSUS.F	SNKSQEEIWENMTWmqWdrEISNYTGLIYsLIBeSQQQEKNEqLLeLDKwASLWnWfDMSN	628
HIVB2126Afr		663
CONSENSUS.G	SNKS?N?IWdNMTWmqWdrEISNYTGLIYsLIBeSQQQEKNEqLLeLDKw?SLW?WF?I??	605
HIVV1525A	-Y--D-----M-----S--K-----I-----A--N--D--SS	687
HIVLBV217	-F--E-----I-----N--H-----L-----E--S--S--TQ	678
CONSENSUS.O	G.....??neSiWD?LIWQqWdggInNvSS?IYeeIQ?AQ?QQEKNEk?LLELDKwASLWnWfDITk	604
HIVAM70	-T-----E--R--S--I--T-----K--V--Q--K-----	677
HIVMP5180	RYNDG--N-----H-----I--D--A--D--V--A-----L--F--	690
HIVVAU	-D-----E-----F--K--E--E-----E-----	691
CONSENSUS.CP2	?????????????PIW?N?IWQ?WD?LV?NTGT?I??L??A??QQE?N??L?ELD?W?SLWnWfDIT?	333
SIVCP2GAB	G.....SN.STDD--G--L--Q--K--S--Y--K--FGL--EE--QS--K--ERD--L--Q--A-----K	668
CONSENSUS.U	SNKs1?e?WDrnTW?eWdrEISNYT??IY?LledSQ?QQEKNEqLL?LDKwASLWnWf?IT?	602
HIVZAM184	TQD--I--K--LQ--K-----E--Q-----N-----A-----N-----D--R	662
HIVK124A	-Q--T-----K--E--D--GL--S--I--K--T-----E--Q-----S--K	680
HIVZ3fr	-VTL--L--N-----QV--G-----K--S--K--E-----G--N	673

HIV1 ENV

	1/3' 5'	<- tat cds	
CONSENSUS.A	WLWYIriPimIVGGLIQLRIvFavLSiInRVRQGYSPISfQtltp?pr??pdrRperIeaeGG.eQdcdRSi		674
HIVSF1703	-----I-----S-----L-----T-----		751
HIVU455	-----L-VI-----T-----A-I-EG.LG--G-----GK-----		738
HIV2321	-----K-----HMQ-E-----G-----		746
HIVRW020	-----K-----V-----IR-N-KE,---LG--DG---		727
HIVUG031fr	-----K-----I-----H-N-G.L--G-----G-----		696
HIVD687	-----N-----HMQ-E-----G-----G-----V-----		626
HIVUG275A	-----L-----V-----H-N-G.L--G-----K-----		740
HIVUG273A	-----L-----L-F-V---K---V---H-N-D.L--A-----		745
HIVVI191A	-----PMMQ-E,---G--G-----G-----		744
HIVDJ264A	-----K-----L---TV-K-----HMQ-E,---G---X---V-----		736
HIVDJ263A	-----K-----T-----HMQ-G,---G---G-----G-----		744
HIVDJ258A	-----I-----T-----HMQ-EP-----G---K-----		739
HIVKENYA	-----K-----T-----H-A-G.L-----		738
CONSENSUS.B	wLWYIkiFImIvqGLvQLRIvFavLSiVNRVRQGYSPISfQT?lPaprg.pdrPegieaeGG.ardRdrSg		697
HIVJRCSE	-----I-----S-----L-----T-----		738
HIVJRFL	-----L-----T-----L-----		737
HIVALA1	-----S-----R---R-E,-----		739
HIVBRVA	-----I-----R---GR---,-----S-----		742
HIVJH3	-----R-----T-----R-----D-----		757
HIVBAL1	-----I-----S-----H-SS---,---G-----		744
HIVYU2	-----I-----V-----H---Q---,---D-----		733
HIVMN	-----L-----RP-V---,-----T-----		747
HIVHXB2R	-----L-----H---T---,-----I-----		746
HIVLAY	-----L-----H---T---,-----I-----		751
HIVNL43	-----L-----H---I---,-----I-----		744
HIVMFA	-----H---NRG---,-----V-----		744
HIVCAM1	-----I---I---TI-L-----RP-V---,---G---T---		746
HIVNYSCG	-----I---T---R---Q---,-----		740
HIVADA	-----I---T---H---,-----D---V-----		743
HIVJFL	-----S---A---R---T---,---G-----		739
HIVSIMI84	-----I---I---L---,-----		746
HIVB31	-----S---R---,-----D---N-----		740
HIVSF162	-----T---RP---,-----S-----		737
HIVBCSG3C	-----I---T---H---T---,-----		749
HIVQYI	-----R---I---R---TQ---,-----		745
HIVSF33	-----R---I---I---LI-Q---,---G---T-----		742
HIVCDC4	-----I---L---N---,---T---G---,---G---G---T-----		758
HIVSF2	-----R---V---,---D---V-----		745
HIVSF2B13	-----V---R---T---,---D---V-----		743
HIVHAN	-----I---L---T---,---RQ---E---,---G---V-----		746
HIVRF	-----R---K---H---,---G-----		755
HIVM472	-----I---T---H---T---,-----V-----		737
HIVTB132	X-----I---T---H---X---,---G---T-----		727
HIVBR020	-----A---I---K---V---R---Q---,---E---D-E-E---		743
HIVTH014	-----L---RP-Q---,---GG---Q---E-----		735
HIVUS711fr	-----I---R---,---Q---x-----		739
HIVUS712	-----S---R---A---,-----		742
HIVUS715	-----Y---I---H---R---,-----		744
HIVUS716	-----A---V---I---HH-L---,-----		735
HIVHA593	-----H---V---H---,---G---V-----		765
HIVHA594	-----I---L---Q---,---G---G---E-----		739
HIVHA596	-----L---Q---,---G---G---E-----		746
HIVHA599	-----L---I---I---A---L---,---L---G---,---GT---		756
HIVHA651	-----T---R---,---S-----		740
HIVRJS	-----R---T---,-----		743
HIVGUNfr	-----I---H---T---,---TG---,-----		708
HIVSC	-----T---R---SQ---,-----		746
HIVSBA	-----R---I-----		444
HIVSBB	-----		453
HIVSBC	-----I---I-----		442
HIVP896	-----RL-----I-----L---S---,---T-----		743

	\\3' sj	<- tat cds	
CONSENSUS.C	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTLTPNPRG?PDRIGRIEeGG.EQD?drSi		691
HIVMA959	-----K-----		737
HIVMA960	-----L-----K-S-		733
HIVD757	-----K-----G-----E--GG--K--		623
HIVD747	-----I-----C-----KA-----		646
HIVD760	-----		650
HIVZAM20A	-----R-----I--E-----R--		767
HIVSM145A	-----I-----V-----K--		744
HIVZAM18A	-----S-----R-K--		740
HIVDJ259A	-----P-G-----RG--		722
HIVDJ373A	-----I-----PRG-----KG--T		733
HIVSE364A	-----I-----R--		731
HIVUG368A	-----K-----T-----P-E-----R--V		739
CONSENSUS.D	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTLTPAPRG?PDRPEGIEEEGG.EqGR?RSI		680
HIVJY1	-----T-----D--		753
HIVNDK	-----L-----V-----V-----E--R--D--		736
HIVMAL	-----R--IV-----I-----L-----T--P-----G--		749
HIVELI	-----I-----T-----R--D--V		743
HIVZ226	-----R--D--		743
HIVUG024	-----RL--I-----T-----GET-----G--		740
HIVUG266A	-----L-----T-----G--		741
HIVUG269Afr	-----R-S-----F-----G--		743
HIVUG274Afr	-----R-----E-----E-----G--		754
HIVSE365A	-----V-E-----I-----D--		741
CONSENSUS.E	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTP?HbQRE.pDRPERIEEGGG.eQGRdRSV		707
HIVTH966	-----I-Q-----		740
HIVTH975	-----V-----E--		738
HIVTH022	-----I-P-----		738
HIVTN243	-----X-----L-----T-----L-----E--		745
HIVTH006	-----LT--G--A--		737
CONSENSUS.F	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSLQTHIPSRE.pDRPEGIEEGGG.EQGNERSV		697
HIVBZ126Afr	-----		732
CONSENSUS.G	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTL?HHQ?E.pDR?ERIEEGGG.E?DR?RS?		666
HIVVI525A	-----V-----R-----T--G--L-----H--T--T		756
HIVLBV217	-----I-----K-----A--R--P-----Q--D--V		747
CONSENSUS.O	WLWYIKIATIIIVGALIGVRVIMIVLMIVKNIRQGYQPLSLQIP??hQ?E.agTPGrTGegGG.eedEprwt		670
HIVANT70	-----V-----I-----NH--E--G--G--I		746
HIVHVP5180	-----I-----I-----VP-RQ--E--E--G--K--		759
HIVVAU	-----V-----IQQ-A--V--G--D--R--		760
CONSENSUS.CPZ	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTL?HbQ?E.pDR?ERIEEGGG.E?DR?RS?		369
SIVCPZGAB	-----L-A--G--I--INTVF-VV-R--S--L-TLIPV-RE.QGRIG--D-G--EQ--S-SV		737
CONSENSUS.U	WLWYIKIFIMIVGGLIGLRIVFAVLISIVNRVRQGYSPLSFQTLIP?PrG.pDRPgriEEeGG.EqgrdRSI		667
HIVZAM184	-----I-----I-II-V-----IPT-H-E--L-----G--		731
HIVK124A	-----V-----ML-L-----A-----EG--		749
HIV23fr	-----TVF-II-----T-----T-----EDMN--V		742

HIV1 ENV

CONSENSUS.A	rLVsGFLaLAWDDLRLCLFsfYHrLrDfiliIs.artVellghaslkglrig.....weglkYL.wM.Ll	735
HIVSF1703	-----RR--R-----	812
HIV0455	-----I-----N-----A-V--A--R-----	799
HIV2321	-----P-----C-----CA-----I--T-----I-R-----T--G--V	800
HIVRM020	-----S-----	788
HIVUG031fr	-----E-N-----	757
HIVD687	-----N-----	646
HIVUG275A	H-----Q-----	790
HIVUG273A	-----V-----K--R-----I-----	799
HIVV1191A	-----V-----I--T--RR-----A-----	798
HIVDJ264A	-----H-----V-----V-A-----GA-----G--	797
HIVDJ263A	-----V-----V-A-----A-Q--G--	805
HIVDJ258A	-----Q--V-----V-A-----A-Q--G--	800
HIVKENYA	---T-----TE-----R-----	799
CONSENSUS.B	rLVsGfLaliwDDLRLCLFsfYHrLrDfiliIs.artVellghaslkglrig.....?rrG.....WealkkYw.wn.LL	750
HIVJRCSP	Q--N-----V-----F-----T--T-----I-----	792
HIVJRFPL	---N-----V-----T--T-----V-----	791
HIVALA1	---D--L-----V-----T-----V-----	793
HIVBRVA	P--D-----F-V-----F-----T-----V-----	796
HIVJH3	Q--D--L--I--V-----T-----	811
HIVBAL1	P--N-----V-----F-----M-----LAG-----V-----	799
HIVYU2	P--D-----I--V-----T-----GV-----	787
HIVMN	---H-----I--V-----F-----H-----A-----V-----	800
HIVHXB2R	---N--S-----T-----	800
HIVLAI	---N--S-----T-----	805
HIVNL43	---N--S-----T-----	798
HIVMFA	---N--S-----T-----	798
HIVCAM1	---T-----	800
HIVNY5CG	P--M-----V-----C-----	794
HIVADA	---D-----F-----V-----	797
HIVJPL	Y--T-----E--W-----T-----	793
HIVSIM184	Q--T--PI--G-----F-----I-----	800
HIVD31	---K-----S-----L-----A-----V-----	794
HIVSF162	P--H--L-----I--A-----G-----	791
HIVBCSG3C	---N-----F-----I--V--T-----S-----	803
HIVVOYI	---D-----I-----V-----	799
HIVSF33	---N-----F-----T-----V-----	796
HIVCDC4	---H-----V-----V-----	812
HIVSF2	---D-----E-----R-----A--T--I-----H-----S-----	799
HIVSF2B13	---D-----E-----R-----A--V--T--I-----H-----V--G--S-----	797
HIVKAN	---S-----F-----R-----K--T-----V-----	799
HIVRF	GA--N--T-----WT--S-----V-----	809
HIVMMJ2	---H-----K-----	791
HIVTB132	---D-----IF--V--N-----V--T-----X-----	781
HIVBR020	---T--L--EI-----F-----R-----	797
HIVTH014	---S--L--T--E-----R-----	789
HIVUS711fr	---N--T--F--N-----I-----L-----	793
HIVUS712	Q--D-----I--V-----Y-----T-----	796
HIVUS715	Q--T-----I--V-----I-----	798
HIVUS716	Q--D-----TI--V-----T-----V-----	789
HIVHA593	---N-----I-----	819
HIVHA594	---I--N-----S-----	793
HIVHA596	---I--N-----A-----S-----	800
HIVHA599	---N-----I-----G-----	810
HIVHA651	P--AH-----I-----	794
HIVRJS	P--T-----	797
HIVCUNfr	---H-----F-----V--L-----	762
HIVSC	---D-----I--V-----	759
HIVP896	P--N-----F--V--N--L--L--N-----T-----R-----	797

CONSENSUS.C	RLvsGFLalaWDDlrslclfsYhrLRDfilia.ARaVELLgraslrgrlQrG.....WeaLKYL.Ga.Ly	752
HIVMA959	-----V-----K-----D-----N.-G	798
HIVMA960	---N---F-----Q---LL-V---V-----N---	794
HIVD757	---N---CGA-AS	642
HIVZAM20A	---N---V-----R---Q---T-----	828
HIVSM145A	---N---F-----V-----E-----N.-G	805
HIVZAM18A	---D---P-----A-----	801
HIVDJ259A	-----R---L-----T-----	776
HIVDJ373A	-----F-----L-V-V-----T-----	787
HIVSE364A	---SI-----V-----I-----A	785
HIVUG268A	--I-----R---LL-----	800
CONSENSUS.D	RLvnGfsaLiWDqLRnLCLFsyhrLRDlilia.aRIVellGR.....RG.....WeaLKYI.Wn.LL	734
HIVJY1	-----F-----T-----I-----S---	807
HIVNDK	---LF--F-----S-----	790
HIV44L	-----L-----T-----	803
HIVELI	--L-----S-----V-----DI-----	797
HIVZ326	-----	797
HIVUG024	--L--L-----R--H-----A-----I-----	794
HIVUG266A	--S--L--A---S-----L---T-----W---	795
HIVUG369Afr	--S-----T-----I-----	797
HIVUG274Afr	-----R---L---I-----K-I-----	808
HIVSE365A	-----G-----V-----S---	795
CONSENSUS.E	RLvsGFLalaWDDLRSLCLFsyhrLRDfilia.ArtVELLghSSLKGLRRG.....WESLKYL.GN.LL	768
HIVTH966	---T-----S-----R-----	801
HIVTH975	---I-E-----L-----L-----	799
HIVTH022	-----	799
HIVTN243	-----S---A-----	806
HIVTN006	-----R-----	783
CONSENSUS.F	RLVMGFLALVWDDLRNLCLFSYHRLRDlilia.ARIV.....NRGLRRG.....WEALKLL.GN.LT	751
HIVBZ126Afr	-----	786
CONSENSUS.G	R???GFL???WDDL?SLCLFSY???LRD?PLIV.?RTVELLGR???????G.....WE?LKYL.WN.LL	706
HIVV1525A	-SAN--PLV---N-----RL--LL--A-----R-----A-----	810
HIVLBV217	-LVS--SIA---R-----HQ--FI--T-----NSLKGLRL-----G-----	808
CONSENSUS.O	plsqQFL?ILYTDLRTIILWcyHLLSNLASgI.....qkII?yLqLQLWILGoktIeaCRlc?A.v?	727
HIVANT70	-S-----P-----V-S--R-----I-NV--I-A--T	807
HIVMVP5180	A--P--QQ-----I-----RR--D-----G--M	820
HIVVAU	-----H-----I-----E-----RH-----I--R-----FR--II	821
CONSENSUS.CPZ	PL?G?7L?WDDLRLNIGIWSYQSLTSLACNV.WRQLKTLGHLILHSLRL..RERLCLL.GG.II	423
SIVCP2GAB	R-VE-CLP-I-----	798
CONSENSUS.U	RLVnGFLalaWDDLRsLCLFSYHRLRDlilia.?r?Vellgr???????rG.....WEaLKYL.gN.Ll	719
HIVZAM184	---S-----C---A-T---HSSLKGLRL-----G--H--W---	792
HIVK124A	-----S--I---D-----T-I-----	803
HIVZ3fr	-----E---I-----V---KG-----V	792

HIV1 ENV

	<- rev cds	
CONSENSUS.A	lyWGRELK7SAInLLDtIAIavAgwtDRvIEigQrigRAilnIPrRIRQGIEraLI	790
HIVSF1703	-----N-----A-C-----F----	867
HIVU455	-----I---T---AV-V-----I-----T-----	855
HIVZ321	I---Q---N-----V-----D-----VV-A---P-----	856
HIVRW020	-----T-V-V-----R-----V---F-----G----	844
HIVUG031fr	-----T---F-----I-----F-----	813
HIVUG275A	A-----I---S-F---VVI-----L---LC---H-----F-----	846
HIVUG273A	-F-----I-----F-----F-----	855
HIVVI191A	-----I---T---AT-----R---I-----V-----A-K-Q	854
HIVDJ264A	T---Q---N-----N-----V---R---V-----	853
HIVDJ263A	-----Q---N-----L---N-----V-----	861
HIVDJ258A	-----Q---N-V-----V---N-----EV-----	856
HIVKENYA	-----Q---I---S---V-----V-----A-G-----H-----F-----	855

CONSENSUS.B	qYwqElknsavslinatAieVaEgTDRvievvgra?RailhiprRiRQGleralI	805
HIVJRC5F	-----I-----VY-----T-----	848
HIVJRF1	-----I---AL---TY-----T-----	847
HIVALA1	-----L---F-----F-----	849
HIVBRVA	-----G-R---A---F-----Q	852
HIVJH3	-----F---I-----LKIL---P-----T-----	867
HIVBAL1	-----V-----L---V-----	855
HIVYU2	---I-----V-----IL---F---V---V-----	843
HIVMN	-----S-----L---G---T-----	856
HIVHXB3R	-----G-C---R-----I-----	856
HIVLA1	-----G-C---R-----I-----	861
HIVNL43	-----N-----L-A-Y---R-----I-----	854
HIVMFA	-----G-Y---R-----I-----	854
HIVCAM1	-----R---PDTI-----C-----L-----	856
HIVNY5CG	---G---I-----V-----A-IC-G-----L-----	850
HIVADA	-----R-----IY-----T-----L-----	853
HIVJFL	-----I-----IG---V-----	849
HIVSIMB4	-----T---V-----IG-V-----P-----	856
HIVD31	-----W-----V-----	850
HIVSF162	---I-----FD-I-----I---A-IG-F-----F-----	847
HIVBCSG3C	-----N-V-----TY-----L-----	859
HIVYOY1	-----VI-----I---Y-F-N-----	855
HIVSF33	L-----V-----VG-----T-----F-----	852
HIVCDC4	-----V-V-----IY---F-----F-----	868
HIVSF2	---I-----W-----T-----A---Y---H-----L-----	855
HIVSF2B13	---I---R-----Y-----L-----	853
HIVHAN	-----F-TI-----C-----V-----	855
HIVRF	-----T-----I---A-IL-F-----	865
HIVMAJ2	---K-----G---I-----IC---I-----	847
HIVTB132	-----L---VY---N---T-----KGR	835
HIVBR020	-----S-----V-----I---IA-FF-OF-----Q	853
HIVTH014	---I-----V-----Y-----T-----	845
HIVUS711fr	-----H-----Y---I---T-----	849
HIVUS712	-----S-----L---IG---T-----	852
HIVUS715	-----SG-----L---G-----F-----	854
HIVUS716	-----S-----V-----L---W-----T-----	845
HIVHA593	-----N---D-----I---R---F-----T-----	875
HIVHA594	---I---N---S-----II---IG---V-T-----	849
HIVHA596	---I-----II---IG---V-T-----	856
HIVHA599	-----I---G-RGVC-----	866
HIVHA651	-----I---I---V---L-N-T-----	850
HIVRJS	---I-----	801
HIVP896	-----KI---C---RN---T-----	853

	<- rev cds	
CONSENSUS.C	QYWGIEIKKSAISLIdtiAIAVaEGTDRIie??Qri?Prai?nIPrRIRQGFEAALq	804
HIVMA959	L-----I-----IV--L-----L--SL---G---R---A-----L	854
HIVMA960	L-----I-----V-----L-----L-----L-----Y-----L	850
HIV2AM20A	-----L-----LV---C---R-----	875
HIVSM145A	-----T-----LA---C-GVR---T-----L	861
HIV2AM18A	-----V-A---T-----FI---C---R-----	848
HIVDJ259A	-----N-T---T-----KFI-GLM---C-----	832
HIVDJ373A	-----V-----T-----II---W---FC-----	843
HIVSE364A	-----I-----T-----LV---C---S-----	841
HIVUG268A	-----Q-----S-----VG-G-G---LH-----	856
CONSENSUS.D	QYWiqeLkNSaisL7dt?AIAVaEGTDRIieivqRa?PRAFL?iP?PriaRQGLERaLL	784
HIVJY1	---T---F---LNAT---I---LIR---F---V-H---R-V---	863
HIVNDK	---S---R---S---L---I---R---V---C---I-NV-R---L---	846
HIVMAL	---G---LN-T---C---FG---I-H---R---F---	859
HIVELI	---S---R---S---F-AI---I---C---V-N---R---S---	853
HIVZ226	---SR---S---L---I---R---C---V-H---T---L---	853
HIVUG034	---FI-T---V---I---L---IG---I-M---T---L---	850
HIVUG266A	---M-FN-T---A---V---I-N---T---F---	851
HIVUG269Afr	K-----F---I---V---VI---V---V-H---R---F---	853
HIVUG274Afr	---F---T---V---V-N---T---	864
HIVSE365A	---K-----L---I---I-DV---C---I-H---T---	851
CONSENSUS.E	LYWQELKISAIISLLNATAIAVAGWTDRIIEVAQGANRAILHIPPRIIRQGLERALL	824
HIVTH966	-----I-----	857
HIVTH975	V--S-----	855
HIVTH022	-----F-----I---R-----	855
HIVTN243	-----K-----	847
CONSENSUS.F	LYWQELKNSAISLLNTTAIVVTEGTDRIIAALQRIGRAILNIPRIIRQGFERALL	807
HIVBZ126Afr	-----	842
CONSENSUS.G	7YWGQELKNSAI?LL?T?AIA?A??TDRIIE??Q?ACRA?L?IPRIIRQGFERALL	749
HIVVI525A	Q-----H---M-T---V-EG---IV-I---V-H---F---	866
HIVLBV217	L-----N---D-V---T-NW---VA-R---I-N---L---	864
CONSENSUS.O	QYWLQELqSATSLLDT?AVaVANWTDgiIIGiQRIG?GILnIPRIIRQGLER?LI	780
HIVANT70	-----S---L-----A---T---R-----S---	863
HIVMP5180	-----K-----I---S-----L---Q-F-H---A---I-V	876
HIVVAU	-----T-----V-----ST---S---R-----L---	877
CONSENSUS.CP2	QYWGKELKISAIISLLDATAIAVAGTDRITDAFQVTLRIIRNIPRIIRQGLERALL	479
HIVCPZGAB	-----	854
CONSENSUS.U	LYWQELKNSAISLLdtiAivVAewTDRIIEIGQR?cRAILnIPrRIRQGFERALL	774
HIV2AM184	-----A-----A---R---T-----	848
HIVK124A	Q--TR-----V---G---S---K---LG---F-H-----	859
HIV23Fr	-----NAT-----G---V---I-----	848

HIV1 NEF

CONSENSUS A	MGGRWSKSKSRVENPEVRKRMRET	PAAAKGVGAVSQD	LDKYGAVTSSNTSSTNAS	54
HIVU455	-----	-----	-----	54
CONSENSUS B	MGGRWSKs's??GWPvRermrra??????????EpaAdgVQAvSRD	-----	IekHGaiTssNTaatNAd	50
HIVLAI	----S VV -T-----	-----A-----	-----A-----	54
HIVNL43	-----S-VI A-----	-----	-----N-A-----	54
HIVGLNEF3	-----R MG--A D Q--	-----A-----	-----	53
HIVGLNEF5	-----LG T-----	-----	-----A-----	52
HIVGLNEF6	-----LG T-----T	-----R A-----	-----A-----	52
HIVBRVA	-----MA -ST-----EPARERMRAEPRA	-----	-----	66
HIVMN	-----R VT T-----EPA	-----L-----A-----	-----L-----	56
HIVSC	-----R VV T-----KT	-----A-----	-----P-N-----	54
HIVBAL1	-----S IVV A--K--T	-----	-----N-----	54
HIVJRCSE	-----H VP--ST A-----EPATDRVR	QT-----V-----	-----	64
HIVJRFL	-----R VP--ST-----EPAADRVR	RT-----V-----	-----	64
HIVNH53	-----R -R VT--FT-----	-----	-----N-----	54
HIVQY1	-----C-MK--TI-- K -EL	QPP-----E-----A-----	-----	59
HIVSF2	-----R-MG--SAI-----EP	RA-----	-----	58
HIVSF162	-----R MS--SA-----R-EPA	-----	-----N-----	56
HIVCAM1	-----R-LG--SA-----Q-EP	RA-----E-----	-----	58
HIVHEI20BL	-----CT--SA-----Q-EPA	A-----X-----	X-----	56
HIVHEI28BL	-----S-IV--N-----	-----T-A-----	X-R-----N-----	54
HIVHEI39BL	-----R-RDD--TT-----Q-EPA	A-----A-----	-----R-----	58
HIVHEI3BL	-----xxMA--xx--DxxIXT	-----E-----A-----	X-----	53
HIVHEI4BL	-----S-MI--AI--D-----	-----V-----A-----	-D-----N-----N-----	54
HIVSF33	-----SKM--A-----K--	-----A-----	-----L-----	53
HIVSWB84	-----S-MS--AI-----RQ-	-----	-----	54
HIVHAN	-----C-----P-----KQ-EP	-----A-----	-K-----TN-A-----	54
HIVD31	-----S-VV--AI-----K--	-----E-----	-----P-----A-----	54
HIVRF	-----SKMG--A-----QK-	-----A-----	-----T-----N-A-----	54
HIVYU2	-----R-MA--T-----EPAAERMRA	-----	-----R-----	64
HIVBCSG3C	-----RMEG--HA-----	-----	-----N-----	54
CONSENSUS D	MGGRWSKSSIVGWPAIRERIRrT???	??dPAADGVGAVSRD	LEKHGAITSSNTasTNad	54
HIVELI	-----	N-----	-----	54
HIVNDK	-----L-----K-----	-----	-----DT-----	54
HIVZ6	-----R-----DPR	RT-----A-----	-----RD-----	59
CONSENSUS O	MGMA??K?KF?GW??VR?PMRR?	??????PCAPGVG??SRE	LA?RGGE?SSHTFQNNAA	41
HIVANT70	-----LR G -E--AA--E--T	RTFFESE-----QI----	-----A-----P-----	60
HIVMVP5180	-----WS-S--A--SE--D--S	SSDPQQ-----AV----	-----T-----S-----	59
SIVCP2GAB	MGTKWSKSSLVGVWPEVRRRIREA	PTAAEGVGEVSKD	LERHGAITSRNTPETHQT	54
CONSENSUS U	MG?KWSK???GWP??RERIR?T???	PP???GVGA?SQD???L?K?GA?SS??A?NN??	-----	33
HIVMAL	-----G--SSIV--KI-----R-	---TET---V---AVSQD D-C--AA--SP A--AS	-----	59
HIVZ321	-----N-----AV-----Q-PPA	---AAE---A-----	-A-H--IS--NT-T--PD	53
CONSENSUS A	CAWLEAQ	EEGDVGFFVVRPOVPLRPMTYKAADFSLPFLKEKGGGLDGLIHSQKROEILDLMVYHTQGFFPDWQ		125
HIVU455	-----	-----	-----	125
CONSENSUS B	caWLEAQ??EeEvGFFVrPOVPLRPmtYkaA?DISHfLkekGGLEglI?sqkrqdiLdlwyhtqgyfpdwq	-----	-----	119
HIVLAI	-----T-----V-----H-R-----I-----	-----	-----	125
HIVNL43	-----T-----V-----H-R-----I-----	-----	-----	125
HIVGLNEF3	-----T-----V-----VH-----	-----	-----	124
HIVGLNEF5	-----T-----V-----TY-----I-----	-----	-----	123
HIVGLNEF6	-----T-----V-----H-R-----I-----S-----	-----	-----	122
HIVBRVA	-----D-----K-----V-----H-Q-----	-----	-----	137
HIVMN	-----K-----L-----D-----Y-----	-----	-----	127
HIVSC	-----V-----Q-----TPREDKIS ICGSTTHKATSLIGR	-----	-----	125
HIVBAL1	-----D-----RS--I-----F-K-----H-----	-----	-----	125
HIVJRCSE	-----Y-----D-----I-----Y-----I-----	-----	-----	135
HIVJRFL	-----D-----G-V-----H-----	-----	-----	135
HIVNH53	V-----D-----G-L-----F-----E-----I-----	-----	-----	125
HIVQY1	-----D-----G-L-----Y-----	-----	-----	130
HIVSF2	-----D-----L-I-----W-R-----E-----I-----	-----	-----	129
HIVSF162	-----D-D-----L-----Y-----IH-----	-----	-----	127
HIVCAM1	-----D-----L-I-----I-Y-R-----I-----	-----	-----	129
HIVHEI20BL	-----D-----x-----G-L-----N-----VY-R-----E-----	-----	-----	127
HIVHEI28BL	-----E-----G-----x-----x-G-x-----R-----VY-x-----N-----	-----	-----	126
HIVHEI39BL	-----D-----SR-R-----D-M-Y-----E-----	-----	-----	129
HIVHEI3BL	-----x-x-----G-L-----I-Y-----	-----	-----	124
HIVHEI4BL	-----G-----V-----W-----	-----	-----	125
HIVSF33	-----D-----K-----L-----VY-----I-----	-----	-----	124
HIVSWB84	-----E-----RR-I-----R-----VH-----E-----	-----	-----	125
HIVHAN	-----E-----G-L-----Y-P-----E-----	-----	-----	125
HIVD31	-----K-----V-----VH-----	-----	-----	125
HIVRF	T-----ED-D-----F-----V-----D-VF-----	-----	-----	127
HIVYU2	-----H-----M-----K-----Q-----	-----	-----	135
HIVBCSG3C	-----I-----V-I-----F-----R-----T-----	-----	-----	125
CONSENSUS D	CAWLEAQE	ESeEVGFFVVRPOVPLRPMTYKAADFSLHFLKEKGGLEGLIWSKKROEILDLMVYHTQGFFPDWQ		126
HIVELI	-----D-----L-----	-----	-----	126
HIVNDK	-----	-----	-----	126
HIVZ6	-----L-----	-----	-----	131
CONSENSUS O	LAPL?SH	??E?VGFFVrPOVPLRPMTYKAADFSLHFLKEKGGGL?GLIYSHKRAEILDLMVY?TQGFFPDWQ		102
HIVANT70	-----QE-E-----A-----Y-G-----E-----V-M-----	-----	-----	131
HIVMVP5180	-----D-----KD-D-----R-----F-A-----D-----I-H-----	-----	-----	130
SIVCP2GAB	LAWLEEM	DNEEVGFFVVRPOVPLRPMTYKAADFSLHFLKEKGGLEGLVYSRRRQEILDLMVYHTQGFFPDWQ		125
CONSENSUS U	C??E??E	E?EeEVGFFVVRPOVPLRPMTYKAADFSLHFLKEKGGGLDOL??S?KROEILDLMVYHTQG?FPDWQ		92
HIVMAL	-----PP-----E-----Y-----H-----VW-P-----Y-----Q-----	-----	-----	128
HIVZ321	-----AWL-AQ-----S-----P-----F-----IY-K-----F-----H-----	-----	-----	125

CONSENSUS.A	NYTPGPGIRYPLTFGWCFKLVVPVDAEVE.EATGGENNLSLHPICQHGVDDEEKEVLMWKFDSSTLALKHRAVE	197
HIV455	-----	197
CONSENSUS.B	nytpGpG?ryPLtFGWCfKLVPepekvE.eanegannSLIHPmslHGmddpe?avL7WcFDrLAfhHwARE	188
HIVLAI	-----V-----Y-----D-----K-----T-----V-----R-----E-----V-----	197
HIVNL43	-----V-----D-----D-----K-----T-----V-----R-----E-----V-----	197
HIVGLNEF3	-----S-----V-----F-----D-----K-----T-----V-----R-----E-----V-----	196
HIVGLNEF5	KD-----V-----Y-----D-----K-----T-----V-----R-----E-----V-----	195
HIVGLNEF6	-----V-----Y-----D-----K-----T-----V-----R-----E-----V-----	194
HIVBRVA	-----V-----I-----I-----K-----C-----Q-----R-----Q-----	209
HIVMN	-----I-----I-----I-----K-----C-----Q-----WMTREKC	182
HIVSC	TTHQ-Q-SDI-C-----K-----I-----E-----R-----E-----N-----D	197
HIVBAL1	-----T-----P-----D-----C-----S-----K-----A-----K-----S-----V-----	197
HIVJRCSE	-----A-----V-----F-----D-----C-----Q-----K-----V-----K-----L-----V-----	207
HIVJRFL	-----I-----F-----C-----C-----Q-----I-----B-----K-----E-----K-----V-----	207
HIVNHS3	-----I-----G-----DE-----S-----A-----TE-----T-----R-----K-----N-----	191
HIVQY1	-----I-----C-----MD-----DQ-----I-----K-----V-----K-----R-----	202
HIVSF2	-----I-----I-----D-----OY-----A-----Q-----E-----A-----K-----V-----K-----	201
HIVSF162	-----I-----I-----D-----OY-----A-----Q-----K-----V-----K-----V-----	199
HIVCAM1	-----I-----I-----Q-----KR-----A-----Q-----K-----M-----K-----	201
HIVHEI20BL	-----I-----I-----D-----Q-----K-----C-----Q-----X-----K-----X-----M-----K-----X-----XX-----	199
HIVHEI28BL	-----V-----W-----A-----K-----K-----X-----R-----V-----K-----Y-----	198
HIVHEI39BL	-----T-----D-----D-----D-----K-----C-----Q-----K-----Q-----K-----Q-----V-----	201
HIVHEI3BL	-----V-----x-----E-----K-----R-----C-----x-----E-----T-----G-----R-----K-----x-----V-----	196
HIVHEI4BL	-----V-----F-----x-----x-----x-----I-----E-----N-----E-----x-----V-----	197
HIVSF33	-----V-----F-----E-----K-----V-----K-----H-----R-----	196
HIVSWB84	-----T-----W-----P-----E-----K-----V-----K-----S-----K-----	197
HIVHAN	-----V-----D-----S-----A-----TE-----T-----R-----K-----K-----H-----K-----	191
HIVD31	-----T-----P-----K-----Q-----R-----V-----K-----K-----	197
HIVRF	-----T-----W-----D-----T-----IC-----K-----V-----K-----V-----	199
HIVYU2	-----T-----W-----I-----A-----C-----Q-----R-----G-----E-----V-----	206
HIVBCSG3C	-----I-----D-----I-----Q-----R-----E-----	197
CONSENSUS.D	NYTPGPGIRYPLTFGWCFKLVVPVDAEVE.EaT2G2tNcLLHP?CQHGMDpERqVLKWRFNsRLAIEHKARE	197
HIVELI	-----Y-----D-----S-----I-----	198
HIVNDK	-----Q-----R-----D-----M-----Q-----M-----L-----	198
HIVZ6	-----R-----V-----T-----E-----	203
CONSENSUS.O	PYTGPpG?REPLTFGWLFKLVPS?EEAERLQNT?E?A7LLHPAC?HG?ED?H?ELL?W?FDRSLG?TH?A??	159
HIVANT70	N-----T-----E-----C-----R-----N-----A-----F-----T-----K-----M-----K-----N-----V-----MI-----	204
HIVNVP5180	C-----P-----A-----N-----D-----S-----N-----A-----A-----G-----K-----Q-----L-----I-----LQ-----	203
STVCP2GAB	NYTPGPGTRPPLCFGWCFKLVPLTEQVE.QANBGDMNCLLHPICQHGHEDEKKEVLVWRFDsRLALRHARE	197
CONSENSUS.U	NYTPGPGTR?PL?FGWCFKLV??P?EVE.EAN?GENNCLLHP?SQHGMD?D?EREVL7WKFDSsSLA?H?ARE	150
HIVNAL	-----I-----F-----T-----MS-----E-----E-----I-----E-----A-----K-----LR-----R-----	200
HIVZ321	-----T-----Y-----C-----VD-----R-----T-----N-----D-----D-----N-----RK-----L-----	197
CONSENSUS.A	LHPEFY.KD	205
HIV455	-----	205
CONSENSUS.B	lHpeYy.KdCLTSMLQGTFRWGISREARLGSTGEWRALRCCI	230
HIVLAI	-----P-----N-----	206
HIVNL43	-----P-----N-----	206
HIVGLNEF3	-----P-----N-----	205
HIVGLNEF5	K-----P-----N-----	204
HIVGLNEF6	-----P-----N-----	203
HIVBRVA	-----N-----	218
HIVSC	-----	239
HIVJRCSE	-----	216
HIVJRFL	-----	216
HIVNHS3	-----	200
HIVQY1	V-----	211
HIVSF2	-----	210
HIVSF162	-----	208
HIVCAM1	K-----F-----	210
HIVHEI20BL	x-----F-----	208
HIVHEI28BL	--x-----	207
HIVHEI39BL	-----	210
HIVHEI3BL	-----	203
HIVHEI4BL	-----	206
HIVSF33	-----	205
HIVSWB84	-----	206
HIVHAN	-----	200
HIVD31	-----N-----	205
HIVRF	K-----	208
HIVYU2	-----N-----	214
HIVBCSG3C	-----NR-----	206
CONSENSUS.D	mHPEFY.KdC	206
HIVELI	-----N-----	206
HIVNDK	L-----	207
HIVZ6	-----	212
CONSENSUS.O	?HPELF?KD	166
HIVANT70	T-----Q-----	213
HIVNVP5180	K-----P-----	211
STVCP2GAB	QHPEFY.KD	205
CONSENSUS.U	?HPEFY.KdC	157
HIVMAL	Q-----Y-----	209
HIVZ321	M-----F-----	205

This page intentionally left blank.

CONSENSUS.A	MGARNSVLRCCKADELEKIRLRPGGKKKYLKHHVMAANALDYFOLAESLLESKEGCGKILTVLDPVPTGSE	73
HIV2ROD	-----R-----K-----M-----	73
HIV2NIH2	-----K-----T-----IS-E-----	73
HIV2ISY	-----K-----T-----IS-E-----	73
HIV2ST	-----V-----K-----R-----	73
HIV2BEN	-----V-----K-----R-----	73
HIV2CAM2	-----V-----N-----R-----V-----K-----S-----	73
HIV2D194	-----S-----K-----	73
HIV2GH1	-----R-IS-----	36
HIV27924A	-----R-----	36
HIV260415K	-----I-----N-----	36
HIV2FA	-----	36
CONSENSUS.B	MGAR?SVLSGKKTDLEKVLRFPGGKK?Y?LKH??NAVNELDRFGLAE7LLSKEGCHKILTVLAPLVPTGSE	67
HIV2D205	---G---K-M---VW---S---Q---K---	73
HIV2UC1	---S---R-C---II---	73
HIV260667K	---E---N-G---A-E---	36
HIV2JA	---S---	36
HIV27312A	A---N---Q---M---	36
HIV2FT	---T---	36
HIV27810A	A---T-G---A---	36
HIV2ON	---R---	36
CONSENSUS.C	ANLDRFALAESLLETKEGCGKILTVLEPLVPTGSE	36
HIV22238	-----	36
CONSENSUS.D	MGARNSVLSGKKADELEKIRLRPGGKKKYLKHHVMAANLDRFGLAESLLENKEGCGKILsVLAPLVPTGSE	73
SIVMM251	-----	73
SIVMM32H	-----N-----	73
SIVMM1A11	-----	73
SIVMM142	---V---N---	73
SIVMM239	-----	73
SIVMNE	-----	73
SIVSMNH4	---E---V---N---D---	73
SIVSMNH9	---V---I---	73
SIVSMMPBJ	---R-Q---I---	73
SIVSTM	---S---V---S---IT-E---	73
SIVSTM22579	---S---V---S---T-E---	73
HIV2FO784	-----	36
CONSENSUS.E	ARELDRFGLAESLLENKEGCHRIITVLEPLMPTGSE	36
HIV2FA	-----	36
p16 \ / p28		
CONSENSUS.A	NLKSILNTVCVIVCLHAEKvKDTBAAK?ivq.RHLvASTyTaEKMPntSRPTAPpSgkggNIPVQq?gG?NY	142
HIV2ROD	-----G-Q-R-----S-----S-E-----Y-HV---	144
HIV2NIH2	-----G-Q-----E-----VA---	142
HIV2ISY	---Y-T---L-----R-G-----E-----I-----I---	144
HIV2ST	-----A-----QK-----K-T---S-----N-----VA---	144
HIV2BEN	-----L-----KLA-----R-Y---A---	144
HIV2CAM2	-----R-AL-----A-----D-----Y-SI---	144
HIV2D194	-----L-----KLA-----I-----A---	144
HIV2GH1	-----L-----KL-----G-----S-----R-R---T-G---	145
HIV27924A	-----V-----K-----Y-SI---	107
HIV260415K	-----A-----G-Q-----S-----E-----HV---	107
HIV2FA	---Q---L-----RKL-----S-----T-G---	108
CONSENSUS.B	NLKSILNTVCVIVCLHAEKvKDTBAAKIAQ.sHL?v...dtEKmPats?ptAppSG...rNYPVQqVAG.NY	131
HIV2D205	---I---P-----R-AA-----MK-----L---	139
HIV2UC1	-----R-AA-----R-----G-----I---	139
HIV260667K	-----G-----R-VG...N-----R---S---G-----I---	102
HIV2JA	---V-----A...N---T-----R---	102
HIV27312A	---W---K-----L-----V...-T-TADK-A-TS---	102
HIV2FT	---F---Q-----V...-A-V---KS-----	102
HIV27810A	---V-----AA...N-----K-----	102
HIV2ON	-----V...-A-T---R-----	102
CONSENSUS.C	NLKSILNTVCVIVCLHAEKvKDTBAAKIEEVE.RHLAAETKNAEKMPATSRPTAPSSGGGNGNYPVQqVAG.NY	107
HIV22238	-----	107
CONSENSUS.D	NLKSILNTVCVIVCLHAEKvKDTBAAKQIVq.RHLVvETGTae?MPkTSRPTAPpSgrGgNYPVQq?GG.NY	142
SIVMM251	-----T-----S-----I---	144
SIVMM32H	-----T-----S-----I---	144
SIVMM1A11	-----I-----S-----I---	144
SIVMM142	---M---T-----F-----I---	144
SIVMM239	---T---T-----S-----I---	144
SIVMNE	---T---T-----S-----V---	144
SIVSMNH4	-----DR-A-----V---	144
SIVSMNH9	---R-L-----DK-A-----V---	144
SIVSMMPBJ	---L-----DK-A-----K-----I---	144
SIVSTM	---F-----V-K-----NK-A-----V---	144
SIVSTM22579	---F-----V-K-----DK-A-----V---	144
HIV2FO784	---L---D-G-----DK-S-----D-R---V---	107
CONSENSUS.E	NLKSILFNAVCVVYCLHAAVVKVDTBAAKHVQqHVLVAGGATTEKLPQSRPTAPSSGGGNGNYPVQqIGN.NY	108
HIV2FA	-----	108

HIV2/SIV GAG

CONSENSUS .A	CHVPLSPRTLNAWVKLVBeKKFGAEVVPGFQALSEGCTPYDINQmLNCVGDHQAAMQIIREIINeeAAdWD7q	214
HIV2ROD	--I-----	217
HIV2NIH2	--I---G-----	215
HIV2ISY	V-L-----	217
HIV2ST	-----	217
HIV2BEN	V-----	217
HIV2CAM2	-----	217
HIV2D194	I-----	217
HIV2GH1	I-----D-----	218
HIV27924A	-----	180
HIV260415K	--I-----	180
HIV2FA	V-----I-D-----L-----	181
CONSENSUS .B	VH1PLSPRTLNAWVKLVBeKKFGAEVVPGFQALSEGCTPYDINQmLNCVGeHQAAMqiIREIINEEAAdNDQq	204
HIV2D205	-----	212
HIV2UC1	--M-----D-----	212
HIV260667K	S-----EN-----D-----IE-----E-----	175
HIV2JA	-----	175
HIV27312A	--V-----V-----	175
HIV2FT	--V-----I-----	175
HIV27810A	-----D-----	175
HIV2ON	--V-----I-----	175
HIV2EHOA	-----X-----	37
CONSENSUS .C	VHHPLSPRTLNAWVKLVBeKKFGAEVVPGFQALTEGCTPYDINQmLNCVGDHQAAMQIIREIINEEAAdNDEH	180
HIV2238	-----	180
CONSENSUS .D	VHLPLSPRTLNAWVKLiBeKKFGAEVVPGFQALSEGCTPyDINQmLNCVGeHQAAMQIireiINeEAAdNDiQ	215
SIVMM251	-----D-----	217
SIVMM32H	-----D-----	217
SIVMM1A11	-----D-----	217
SIVMM142	T-----S-----L-----D-----D-----	217
SIVMM239	-----D-----D-----	217
SIVMNE	T-----	217
SIVSMNH4	-----V-----	217
SIVSMNH9	T-----V-----X-----	217
SIVSMNPBJ	T-----	217
SIVSTM	-----V-----L-----K-----M-----	217
SIVSTM22579	-----V-----S-----	217
HIV2FO784	-----V-D-----D-----Q-----	180
CONSENSUS .E	VHSPLSPRTLNAWVKLVBeKKFGAEVVPGFQALSEGCTPYDINQmLNCVGeHQAAMQIIREIINEEAAdNDVQ	181
HIV2FA	-----	181
CONSENSUS .A	HP1P.GPLPAGQLRDPGSDIAGTTSTVeeQlqNmFRpqNpVPVGNiYRRWlQlGLQKCVRMYNPTNILDikQ	286
HIV2ROD	-----E-----	289
HIV2NIH2	-----E-----N-----	287
HIV2ISY	-----E-----Y-QE-----	289
HIV2ST	-----E-----S-----	289
HIV2BEN	-----D-----Y-----K-----	289
HIV2CAM2	-----A-----	289
HIV2D194	-----D-----Y-QP-----V-----	289
HIV2GH1	-----E-----Y-----V-----	290
HIV27924A	-----E-----A-----V-----	252
HIV260415K	-----E-----D-----	252
HIV2FA	-----D-----Y-A-----I-----V-----V-----V-----	253
CONSENSUS .B	HP2P.GPLPAGQLRDPGSDIAGTTSTVeeQlQmYRaqnPVPVGNiYRRWlQlGLQKCVRMYNPTNILDikQ	275
HIV2D205	--S---M-----	284
HIV2UC1	--I-----I-----	284
HIV260667K	--S---M-----E-----C-----I-----	247
HIV2JA	--V-----E-----E-----	247
HIV27312A	--V-----G-S-I-----V-----	247
HIV2FT	--I-----QP-----I-----V-----	247
HIV27810A	--S-----I-----	247
HIV2ON	--I-----QP-----I-----S-----V-----	247
HIV2EHOA	--S---M-----E-----F-----	109
CONSENSUS .C	HPVP.GPLPAGQLRDPGSDIAGTTSTVeeQlQmYRaqnPVPVGNiYRRWlQlGLQKCVRMYNPTNILDvKQ	252
HIV2238	-----	252
CONSENSUS .D	HpQp??Pq??GQLRePsgSDIAGTTSTVdEQlQmYRQqNPiPVGNiYrRWlQlGLQKCVRMYNPTnILDvKQ	284
SIVMM251	---A--Q-----S-----	288
SIVMM32H	---A--Q-----S-----	288
SIVMM1A11	---A--Q-----S-----	288
SIVMM142	---Q--A--Q-----E-----	288
SIVMM239	---A--Q-----S-----	288
SIVMNE	---Q--A--Q-----S-----	288
SIVSMNH4	---G-LPA-----R-----	289
SIVSMNH9	---x--G-LPA-----x-----x-----x-----	289
SIVSMNPBJ	---G-LPA-----R-----	289
SIVSTM	---PG-LPA-----SPE-----V-----	290
SIVSTM22579	---G-LPA-----E-----H-----V-----I-----	289
HIV2FO784	---G-LPA-----D-R-----E-----I-----	252
CONSENSUS .E	HPRG..QPpAQGLRDPGSDIAGTTSTPAEQIEWmYRNPNPiPVGDIYRRWlQlGLQKCVRMYNPTNILDvKQ	252
HIV2FA	-----	252

CONSENSUS .A	GPKEpFQSYVDRFYKSLRAEQTDpAVKNNMTQTLL?QnaNPDCKLVLKGLGmNPTLEEMLTaCQGvGGPgQKA	358
HIV2ROD	-----V-----	362
HIV2NIH2	-----I-----T-----	360
HIV2IGY	---S---A---V---S---I---	361
HIV2ST	-----V-----I-----	362
HIV2BEN	-----I-----	362
HIV2CAM2	---S---V-----	362
HIV2D194	---S---I-----S---	362
HIV2GH1	-----I-----	363
HIV27924A	---S---	260
HIV260415K	---S---	260
HIV2FA	---S---	261
CONSENSUS .B	GPKEpFQSYVDRFYKSLRAEQTDpAVKNNMTQTLLIQNANPDCKLVLKGLGmNPTLEEMLTaCQIGGGPGQKA	348
HIV2D205	-----	357
HIV2UC1	-----	357
HIV260667K	-----N-----	256
HIV2JA	-----	255
HIV27312A	-----	255
HIV2FT	-----	256
HIV27810A	-----	255
HIV2ON	-----	255
HIV2EMOA	-----	182
CONSENSUS .C	GPKEpFQSYVDRFYKSLRAEQTDpAVKNNMTQTLLIQNANPDCKLVLKGLGmNPTLEEMLTaCQIGGGPGQKA	260
HIV2238	-----	260
CONSENSUS .D	GPKEpFQSYVDRFYKSLRAEQTDpAVKNNMTQTLLIQNANPDCKLVLKGLGmNPTLEEMLTaCQIGGGPGQKA	356
SIVNM251	-----A-----V-----	361
SIVNM32H	-----A-----V-----	361
SIVNM1A11	-----A-----V-----	361
SIVNM142	-----A-----T-----	361
SIVNM239	-----A-----V-----	361
SIVMNE	-----M-----	361
SIVSMNH4	-----M-----	362
SIVSMNH9	---S---x---xI---x---x	362
SIVSMMPBJ	-----I-----	362
SIVSTM	---T---A---S---R---M---	363
SIVSTM22579	---T---A---P---M---	362
HIV2FO784	---T---	260
CONSENSUS .E	GPKEpFQSYVDRFYKSLRAEQTDpAVKNNMTQTLLIQNANPDCKLVLKGLGmNPTLEEMLTaCQIGGGPGQKA	260
HIV2PA	-----	260
CONSENSUS .A	RLMAEALKEam?PaPIPFaAAQ...r?aikcWNOGKEGHSArQcraPRRQGCWKGK?GHImanCP?RQAG	423
HIV2ROD	-----VIG-----K-F-----P-T-D-----	431
HIV2NIH2	-----V-A-----KTF-----WS-----S-V-D-----	429
HIV2IGY	-----R-----KR-----S-----D-----	430
HIV2ST	-----A-----RT-----A-K-E-----	431
HIV2BEN	-----G-S-----K-RV-----P-----E-----	431
HIV2CAM2	-----Q-P-----RT-----P-T-D-----	431
HIV2D194	-----LT-----R-R-----K-----S-----E-----	431
HIV2GH1	-----LT-P-----KV-R-----T-V-K-E-----	432
CONSENSUS .B	RLMAEALKEALT?PaPIPFaAAQqAGkrgTVTCWNCCK?GHTArQcraPRRQGCWKGK?GHINSKCPERQAG	420
HIV2D205	-----V-----Q-----T-----	430
HIV2UC1	-----V-----K-----	430
HIV2EMOA	-----STN-----FR--KR-----A-----R-----Q-----	254
CONSENSUS .D	RLMAEALKEalap?p1PFaAAQq?Gp?rkp1kCwNCgKeGHSArQcraPRRQGCWKGK?GHImanCP?RQAG	425
SIVNM251	-----V-I-----KR-----D-----D-----	433
SIVNM32H	-----V-I-----R-----D-----D-----	433
SIVNM1A11	-----V-I-----R-----D-----D-----	433
SIVNM142	-----A-I-----K-----D-----N-----	433
SIVNM239	-----V-I-----R-----D-----D-----	433
SIVMNE	-----Q-----K-----T-----Q-----D-----	433
SIVSMNH4	-----R-DQ-----V-K-Q-T-----K-----T-----E-----	434
SIVSMNH9	-----D-TQC-----x-V-K-Qx-I-x-x-x-----F-----x-----A-----E-----	434
SIVSMMPBJ	-----D-TQC-----V-K-Q-I-----A-----E-----	434
SIVSTM	-----VFO-D-----Q-----RTV-----A-T-K-KQ-----QQ-Q-----E-V-----	434
SIVSTM22579	-----EQ-G-----Q-----RTV-----T-K-K-----P-Q-----E-V-----	433

HIV2/SIV GAG

p8 \/ p1 \/ p6 \/ HIV2s only		
CONSENSUS.A FLGLGpWGKKpRNFPvqg?pqGL.....tPTAPP?DPavdLLEkYMqQG?rQRE.Q...R		471
HIV2ROD	-----V-----K-----	482
HIV2NIH2	---I-H---A---A-V-----L-----K-----	480
HIV2ISY	-----VFSS-----M-----RK-----	481
HIV2ST	-----I-----I-VE-----K-----	482
HIV2BEN	-----R-----T-A-----I---A---AE---R---RK-----	482
HIV2CAM2	---I-----V-----L-----K-----	482
HIV2D194	---M-R-Q---A-A-----I---I-----RK-----	482
HIV2GH1	---M-----A-P-----I---A-----R---RE-----	483
CONSENSUS.B FLGLGpWGKKpRNFPvqg?pqGL.....tPTAPP?DPavdLLEkYMqQG?rQRE.Q...R		489
HIV2D205	-----D-----K-----	499
HIV2UC1	-----F-----V-A---IV-----F---Q-I-----E---N---L-K...KOKEN	499
HIV2EHOA	-----F-----V-A---IV-----F---Q-I-----E---N---L-K...KOKEN	522
CONSENSUS.D FLGLGpWGKKpRNFPvqg?pqGL.....tPTAPP?DPavdLLEkYMqQG?rQRE.Q...R		475
SIVMM251	-----V-----K-----	484
SIVMM32H	-----V-----K-----	488
SIVMM1A11	-----V-----K-----	488
SIVMM142	-----V-----E-----H-----	484
SIVMM239	-----V-----M-----K-----	488
SIVMNE	---F-----M-----K-----	484
SIVSMH4	-----MP-----I-----KM-RK-----N	485
SIVSMH9	-----MP-----X-----KV-RR-----N	485
SIVSMPEJ	-----MP-----KM-RR-----N	485
SIVSTM	---F---H---IP-----A---RS---K-----	485
SIVSTM22579	---F---H---IP---TPTAPPMP-----V---A---RS---K-----	491
CONSENSUS.A eRPYKEVTEDLh1EqgetPhRe?CTEDLLHLNSLFGkDQ	/p6	510
HIV2ROD	-----Y-PP-----	522
HIV2NIH2	-----RF---A---C-----	519
HIV2ISY	Q-----F-Q---KQ---C-----N-----	520
HIV2ST	-----F-Q---KQ---C-----	521
HIV2BEN	-----R-----E-----	521
HIV2CAM2	Q-----A-C-----	521
HIV2D194	-----G-A-----	521
HIV2GH1	-----KA---A-----	522
CONSENSUS.B eRPYKEV.....TEDLLHLNSLFGEDQ		511
HIV2D205	-----TEDLLHLNSLFGEDQ	521
HIV2UC1	-----TEDLLHLNSLFGEDQ	521
HIV2EHOA	-----TEDLLHLNSLFGEDQ	544
CONSENSUS.D ekPYKEV.....TEDLLHLNSLFG?DQ		496
SIVMM251	-----G-----	506
SIVMM32H	-----G-----	510
SIVMM1A11	-----M-----G-----	510
SIVMM142	G-----G-----	506
SIVMM239	-----G-----	510
SIVMNE	R-----E-----	506
SIVSMH4	-R-----E-----	507
SIVSMH9	-R-----E-----	507
SIVSMPEJ	-R-----E-----	507
SIVSTM	KT-----E-----	507
SIVSTM22579	KT-----V-----	510

CONSENSUS.A	VLELWKGGLTGETVPSSTQNTGLLE7WQ?RTHHGK?P?KTGRFRRdgp?GKeapQ1PrgPsaGAdTNSTp??a	66
HIV2ROD	-----T-L-----	35
HIV2NIH2	--W--W-L--G-----P--N--IG-	35
HIV21SY	--W--AMTM-----KF--N--MG-	35
HIV2ST	--R--M--G-----M-R--G--V--M--F--C--NP-----DRP	55
HIV2BEN	-----V--A-----L-G-----V--T-----S--D--PS-----SGR-	55
HIV2CAM2	-----G--W-L--F-----T--N--IG-	35
HIV2D194	-----V--V-----L-G-----T--A-----S-----NR-	73
HIV2GH1	-----M--D--R--M-R-----SM-----S-----SR-	55
CONSENSUS.B	KTGNLEMMK?RTHHVKNMPKKTGgFFRVRLGKEASQ1PhdPsaGedtICTPDeP	54
HIV2D205	-----N-----	55
HIV2UC1	-----T-----W-----G-	55
HIV2EHOA	-----TA-----P-----F--XX--GTP--DSA--A--	52
CONSENSUS.D	VLELWBgTLCAMQaPktTgmLEMMWknGPcyQCMPrqTggFFR?WpmGKEAQfPHGp?aSGADaNCspRgp	71
SIVMM251	-----G-----P--S-----SS-----	73
SIVMM142	-----GR-----K-----P--L-----SS-----RT	73
SIVMM239	-----P--S-----SS-----	73
SIVMM32H	-----R-----P--S-----SS-----	73
SIVMNE	-----N-----L--T-----R--L-----N-----	73
SIVSMH4	-----K-----A-----D-----T-----S	36
SIVSMH9	-----K-----A-----D-----T-----x--S	39
SIVSMPEJ	-----K--S--A-----D-----T-----S	39
SIVSTM	R-E--RL--TT--SD--K-----L-----L--NT-----PS-R-	58
<- gag cds		
CONSENSUS.A	SaGptgeiyAarek...ae?AE?etiQr?DrqLcAp.Ragrdt?Qr?ddRGLAaPQFSLMKRPVVTAyiBqQ	129
HIV2ROD	-----S-----T-R--R--GS-----G--I--GATN-----	103
HIV2NIH2	-----K-----KG--R--V--GS-----F-----M--G-----H-----	102
HIV21SY	-----VH-----T-R--TK--S--A--S--R--T-----D-----	102
HIV2ST	-----R--R-VH-----R--R-A--S--P--A--ET--M-----HV-----	122
HIV2BEN	-----TV-----G--G--G--G--A--E--S--G-----D-----	122
HIV2CAM2	-----S-----G--T--G-----TR--GPH--G--N-----H-----	102
HIV2D194	-----V-----R--G--GG--G-----AP--G-----T-----F--D-----	140
HIV2GH1	-----SI--K--G--R--G--G--G--GR-----KS--S--G-----V-----	122
CONSENSUS.B	SrgHDTSGgDtICa...PCRSSaGDaekLHadgEtAeErEpRETLQG.GDRGFAAPQFSLMRRPVVkaCIEGQ	122
HIV2D205	-----E-----T-----	123
HIV2UC1	-----E-----R-----	123
HIV2EHOA	-----IR--C--S--T-----R--KE--TR--E--G--Q-----T-----	120
CONSENSUS.D	SCGSakeLHa?GQ?7777AERKQrealQGG...drGFAAPQFSLMRRPVVTAhiEgQ	119
SIVMM251	-----V--A-----	123
SIVMM142	-----L--A-----	123
SIVMM239	-----V--AAERK-----	127
SIVMM32H	-----A--AAERK-----	127
SIVMNE	-----K--V--T-----K-----	123
SIVSMH4	-----TE--ED--K--GE--T-----Y--E-----	86
SIVSMH9	-----TE--EG--K--GE--T-----NG--Y--E-----	89
SIVSMPEJ	-----TE--ED--K--GE--T-----NG--I--Y--E-----	89
SIVSTM	-----EK--A--E-----E--EDT-----	108
CONSENSUS.A	PVEVLLDTGaAdDSIVAGieLGaNYePKIVGGIGGFINTKeyKvEI?VLnKrVrATIMTGDTPINIFGRNilt	201
HIV2ROD	-----N-----E--K-----	176
HIV2NIH2	-----R--N-----E--G--K-----V-----	175
HIV21SY	-----D--R--K-----	175
HIV2ST	-----V-----R-----	195
HIV2BEN	-----D--T-----K-----	195
HIV2CAM2	-----E--G-----	175
HIV2D194	-----D--T-----K-----A-----	213
HIV2GH1	-----O--D--V-----I--I--K-----	195
CONSENSUS.B	sVEVLLDTGaDOSIVAGIELGSNYTPKIVGGIGGFINTKeyKvEIEVVGRVRATIMTGDTPINIFGRNiltN	195
HIV2D205	-----V-----	196
HIV2UC1	P-----	196
HIV2EHOA	-----N--N-----V-----	193
CONSENSUS.D	PVEVLLDTGADDSIVAGIELGphyTPKIVGGIGGFINTKeyKvKIEVLGRKIKGTIMTGDTPINIFGRNilt	192
SIVMM251	-----T-----E-----	196
SIVMM142	-----T-----E-----	196
SIVMM239	-----T-----E-----	200
SIVMM32H	-----T--K--L-----E-----	200
SIVMNE	-----A-----	196
SIVSMH4	-----N-----D--K--V-----	159
SIVSMH9	-----I--N-----D--K--V-----	162
SIVSMPEJ	-----N-----D--K--V-----	162
SIVSTM	-----LQ--V-----F--N--K-----	181
HIV2FORTC2	-----X-----	6

HIV2/SIV POL

CONSENSUS.A	aLGMSLNlqvAKiePiKlmLKPGKDGPrIrcWPLTKKIEALkeICEKMEKEGQLEEAAPTNPYNTPTFAIKK	274
HIV2ROD	-----V-----K-----	249
HIV2NIH2	-----K-----	248
HIV2ISY	-----V-VT-----KQ-----R-----R-----R-----	248
HIV2ST	-----K-----R-----	268
HIV2BEN	-----VT-----K-----	268
HIV2CAM2	-----R-----	248
HIV2D194	T-----LD-----VT-----K-----R-----	286
HIV2GH1	-----I-----VT-----R-----	268
HIV2S52907	MA-----V-----K-----	66
CONSENSUS.B	ELGMTLNFPVAKiEPVKV?LKPgKDGPKIRQWPLSKKKILALKEICEKMEKEGQLEEAAPTNPYNTPTFAIKK	267
HIV2D205	-----V-----E-----R-----	269
HIV2UC1	-----K-----	269
HIV2EHOA	S-----R-----Q-----E-----S-----	266
HIV27312AR	-----X-----S-----	24
CONSENSUS.D	algMSLNlPiAKvePvKvtLKPGKDGPKI?QWPLSKKKiVAlrEICEKMEKdGQLEEAAPTNPYNTPTFAIKK	264
SIVMM251	-----V-----K-----	269
SIVMM142	-----SP-----K-----	269
SIVMM239	-----F-----A-----K-----	273
SIVMM32H	-----A-----K-----	273
SIVMNE	-----E-----F-----R-----	269
SIVSMNH4	-M-----I-----E-----R-----I-----	232
SIVSMNH9	-M-----x-I-----R-----I-----x-----x-----	235
SIVSMMPBJ	-M-----I-----R-----I-----	235
SIVSTM	-----F-V-----I-K-----E-K-----	254
HIV2FORTC2	E-----F-V-----I-----R-----M-----R-----E-K-----RE-----	79
CONSENSUS.E	-----xKKGQLERAPTNPYNTPTFAIRK	24
HIV2PA	-----	24
CONSENSUS.A	KDKMKWRMLIDFRELNKVTQDFTEiQLGIHPAGLAKKcRItvLDVGDAYFSPILHEDFRqYTATFLPsvNNA	347
HIV2ROD	-----P-----	322
HIV2NIH2	-----	321
HIV2ISY	-----V-----Y-----	321
HIV2ST	-----K-----I-----	341
HIV2BEN	-----K-----SI-----A-----M-----	341
HIV2CAM2	-----	321
HIV2D194	-----R-----K-----	359
HIV2GH1	-----R-----K-----	341
HIV2S52907	-----P-----	139
CONSENSUS.B	KDKMKWRMLIDFRELNKVTQDFTEVqlgiPhpaglaeKcRITV?DVGDAYFSP?PLDP?FRQYTATFLPsvNNA	335
HIV2D205	-----D-----NWVF-----TRQV-----I-----I-----N-----V-----	341
HIV2UC1	R-----D-----L-----I-----N-----I-----	342
HIV2EHOA	-----E-----S-K-----L-----V-----D-----AV-----	339
HIV27312AR	-----E-----WF-----I-----V-----D-----I-----	97
CONSENSUS.D	KDKMKWRMLIDFRELNKVTQDFTEVQLGIHPPaGLAKR?RITVLDvGDAYFSP?PLDeeFRQYTATFLPsvNNA	336
SIVMM251	-----R-----K-----I-----	342
SIVMM142	-----R-----K-----I-----	342
SIVMM239	-----R-----K-----I-----	346
SIVMM32H	-----R-----K-----I-----	346
SIVMNE	-----I-----K-----	342
SIVSMNH4	-----R-----	305
SIVSMNH9	x-G-----x-----R-----	308
SIVSMMPBJ	-----R-----	308
SIVSTM	-----I-----R-----G-----	327
HIV2FORTC2	---S-----V-----R-----K-----T-----R-----	152
CONSENSUS.E	KDKMKWRMLIDFRELNKATQDFTEVQLGIHPAGLKEMEQTvLDIGDAYFSVPLDPEFRQYTATFTIPSVNNQ	97
HIV2PA	-----	97

CONSENSUS.A	EPGKRYIYKVLPQGWKGSPIFYQYEMRQVLEPFPRKANpDvIiIQYMDLILASDRtdLEHdkvVLQKELLNg	420
HIV2ROD	-----H-----K-----R-----	395
HIV2NIHZ	-----I-----E-----	394
HIV2ISY	-----V-----	394
HIV2ST	-----S-----I-L-----R-----	414
HIV2BEN	-----L-----G-----	414
HIV2CAM2	-----S-----L-----N-----	394
HIV2D194	---E---V-----PM---I-----L-----G-----	432
HIV2GH1	-----H-----L-----G-----	414
HIV2S52907	-----H-----K-----R-I-----	212
CONSENSUS.B	EPGKRYIYKVLPQGWKGSPIFYQYEMRQVLEPFPRKANpDV7IIQYMDLILASDRSDLEHDrVVSQKSLNng	405
HIV2D205	-----QS-C-----S-I-----I-----	414
HIV2UC1	-----S-I-----I-----	415
HIV2EHOA	-----L-----T-A-----N-T-----V-----N-----	412
HIV27312AR	-----R-N-T-----V-----K-----	170
CONSENSUS.D	EPGKRYIYKVLPQGWKGSPIFYQYEMRQVLEPFPRKANpDvTL?QYMDLILASDRtdLEHDrVVLQKELLNe	408
SIVMM251	-----H-----V-----	415
SIVMM142	-----H-----V-----	415
SIVMM239	-----H-----V-----S-----	419
SIVMM32H	-----H-----V-----	419
SIVMNE	-----H-----V-----V-----	415
SIVSMH4	-----I-----G-----	378
SIVSMH9	---KH-X-----I-----	381
SIVSMH9B	-----H-----I-----	381
SIVSTM	-----I-----R-----I-----N-----	400
HIV2PORTC2	-----I-----S-----I-----K-----D-----	225
CONSENSUS.E	EPGKRYIYKVLPQGWKGSPIFYQYEMRQVLEPFPRKANSDVLLIQYMDLLIGSNRSMTEHDrMVVLQKDNLNn	170
HIV2PA	-----	170
CONSENSUS.A	LGFSTPDEKFKQDPFqYMGYELWPTKWKLOKIQLPQKeIWTVNDIQKLVGVLMWAAQIYpGIKtkhLCrLir	493
HIV2ROD	-----H-----L-----	468
HIV2NIHZ	-----R-----V-----	467
HIV2ISY	-----H-----V-----K-----	467
HIV2ST	-----R-----V-----RN-----	487
HIV2BEN	---F---C-----L---D-----S-----	487
HIV2CAM2	-----R-----V-----	467
HIV2D194	---F-----K-----	505
HIV2GH1	---L-----L-----K-----	487
HIV2S52907	-----H-----L-----	285
CONSENSUS.B	MGPSTPEEKFKQDPFqYMGYELWPTKWKLOKIQLPQKeIWTVNDIQKLVGVLMWAAQLFPGIKTRHICKLir	478
HIV2D205	-----A-----	487
HIV2UC1	-----R-----	488
HIV2EHOA	-----I-----R-----	485
HIV27312AR	---Q-----I---R-----R---K-----	243
CONSENSUS.D	IGFScPEEKFKQDPFqYMGYELWPTKWKLOKIELPQRetWTVNDIQKLVGVLMWAAQIYpGIKTKhLCYLir	481
SIVMM251	-----	488
SIVMM142	---S-----	488
SIVMM239	-----	492
SIVMM32H	-----	492
SIVMNE	-----K-----	488
SIVSMH4	-----	451
SIVSMH9	-----N-----	454
SIVSMH9B	-----	454
SIVSTM	L-----DV-----K-----	473
HIV2PORTC2	L---L-----YH-----D-----	298
CONSENSUS.E	LGFSTPEDKFKQREPPLKQWGYLLYPKRWKLOKIELPEKERWTVNDIQKLVGVLMWAAQVYpGIKTKhLCrLir	243
HIV2PA	-----	243

HIV2/SIV POL

CONSENSUS.A	GKMTLTTEEVDWTELAEEAELEENKIIISQEQEGHYQEEK*LEATVgKdQdnQWLYKIHQeekiiLKvGRyAKIK	566
HIV2ROD	-----R-----E-----V-----G-----G-----IC-D-	541
HIV2NIH2	-----R-----K-----K-----V-----G-----G-----IC-D-	540
HIV2ISY	---P-----R-----V-----G-----V-----G-----V-----	540
HIV2ST	-----C-----GG-----V-----	560
HIV2BEN	-----Y-----I-S-GH-----V-----	560
HIV2CAM2	-----R-----V-----	540
HIV2D194	-----S-----E-----I-S-----A-----G-RV-----	578
HIV2GH1	-----Y-----I-N-----V-----	560
HIV2S52907	-----R-----D-----V-----	358
CONSENSUS.B	GKMTLTTEEVDWTELAEEAELEENKIIISQEQEGSYKegVPLEATVQKNLANQWYKIHQGnk?LKvGRyAKVK	550
HIV2D205	-----R-----V-----	560
HIV2UC1	-----RI-----	561
HIV2EHOA	-----F-----D-I-----	558
HIV27312AR	-----M-----G-----V-----	316
CONSENSUS.D	GKMTLTTEEVDWTEMAEAGYeENKIIISQEqEGCYyqEGKPLEATvIKsQDNQWSYKIHQSDKiLKvGRyAKIK	554
SIVMM251	-----S-----	561
SIVMM142	-----S-----	561
SIVMM239	-----	565
SIVMM32H	-----	565
SIVMNE	-----N-----	561
SIVSMNH4	-----I-----V-----V-----	524
SIVSMNH9	-----P-----P-----P-----	527
SIVSMNPBJ	-----A-----A-----Q-R-D-----V-D-----	527
SIVSTH	---A-----A-----Q-R-D-----V-D-----	546
HIV2FORTC2	---K-----S-----V-D-----Y-----	371
CONSENSUS.E	GKMTLTTEEVDWTELSAEALAEENKIIISQEQEGRYREBENLEATVLKNQDNQWSYKIHQGDRIILKvGRyPQIK	316
HIV2PA	-----	316
CONSENSUS.A	nthtngvrLLAQVQKIGKEALVIWQRIKPEHLPVERatWEQWMDNYNQVTWIPdWDFvSTPPLVRLAFNLVg	639
HIV2ROD	-----I-----I-----	614
HIV2NIH2	KYPYQR-----D-----	613
HIV2ISY	-----D-----I-----V-----K-----	613
HIV2ST	-----D-----I-----V-----K-----	633
HIV2BEN	-----E-----T-----	633
HIV2CAM2	H-----K-----V-----	612
HIV2D194	-----V-----D-----V-E-----T-----	651
HIV2GH1	-----E-----T-----	633
HIV2S52907	-----I-----I-----	431
CONSENSUS.B	NTHTMGVRLLAHVQKIGKEALVIWGEIPVfHLPVERETWQWMDYQWVTWIPeWDFvSTPPLRLAYNLVK	623
HIV2D205	-----	633
HIV2UC1	-----V-----	634
HIV2EHOA	-----M-----	631
HIV27312AR	-----I-----	332
CONSENSUS.D	NTHTMGVRLLAHVQKIGKEALVIWGEIPVfHLPVERETWQWMDYQWVTWIPeWDFvSTPPLVRLVFNLVK	627
SIVMM251	-----K-----	634
SIVMM142	-----K-----	634
SIVMM239	-----K-----	638
SIVMM32H	-----K-----	638
SIVMNE	-----K-----K-----	634
SIVSMNH4	-----V-----L-----E-----EI-----D-----V-----	597
SIVSMNH9	-----S-----V-----I-----R-----xEI-----V-----	600
SIVSMNPBJ	-----V-----R-----EI-----V-----	600
SIVSTH	-----I-----E-----A-----V-----	619
SIVSMNLB1	-----x-----x-----	25
HIV2FORTC2	-----V-----R-----S-----I-----S-----V-----	444
CONSENSUS.E	NTHTMGVRLLASVQK	332
HIV2PA	-----	332

CONSENSUS.A	dPipGaETFTDGSChRQsKegKagYvTDRGrDkVh?LeQTtNQQAeLEaFaMalcDSGPKvNIIVDSQYvmG	711
HIV2ROD	-----K-----K-----	587
HIV2NIH2	E-V-----I-----R-V-----A-----	586
HIV2ISY	-----I-----K-----RI-----V-----V-----	586
HIV2ST	---L-----K-R-----I-----RL-----V-----A-----	706
HIV2BEN	-----K-----V-----V-R-----A-----	706
HIV2CAM2	---T-----I-----A-----	685
HIV2D194	---T-----R-RV-----S-----A-----	724
HIV2GH1	-----R-----RV-R-----T-----	706
HIV2FA	-----	3
HIV2S132P1	-----x--	3
HIV2S52907	-----K-----K-----	504
CONSENSUS.B	DPLeqvETFTDGSChRasKegKagYvTDRGKDKVKVLEQTtNQQAeLEAFALALQDSgPQVNIIVDSQY7mG	695
HIV2D205	---R-----T-----T-E-----V-----	706
HIV2UC1	---K-----T-----V-----	707
HIV2EHOA	-----K-----P-----V-----	704
HIV27312AI	-----	3
HIV2JA	-----x--	3
HIV2ON	-----x--	3
CONSENSUS.C	-----xLG	3
HIV2POLB7	-----	3
CONSENSUS.D	ePi?g?ETfYvDgSChRqSkeGKagYvTDRGkdKvK?LeQTtNQQAeLEAFImAl?DSGPKaNIIVDSQYvmG	696
SIVMM251	D-E-E-Y-T-----K-----V-----T-----	707
SIVMM142	D-E-E-Y-----SK-----V-----T-----	707
SIVMM239	D-E-E-Y-T-----K-----V-----T-----	711
SIVMM32H	D-K-E-Y-T-----K-----V-----T-----	711
SIVMNE	D-K-E-Y-----K-----V-----A-----	707
SIVSMH4	---Q-A-----R-----V-----R-A-L-----YL-A-----	670
SIVSMH9	---Q-A-----R-----V-----RG-T-L-----YL-A-----	673
SIVSMHPEJ	---Q-A-----R-----V-----R-A-L-----YL-A-----	673
SIVSTM	---LE-T-----N-----A-----A-----V-----	692
SIVSMHLIB1	---Lx-V-----E-V-----R-----A-----K-----NQ-----E-----	98
HIV2POLC12	-----xI-	3
HIV2FORTC2	--LERT-----H-----V---DRE-T-A-Q-----L--Q-----	517
CONSENSUS.A	IvagQPtESEgrIVMqTIEeMIKKEaiYVAVPAHKGIGGNQeVDHLVSQGIQVLFLEKIEPAQEeHeKYhs	784
HIV2ROD	---S-----K-----	760
HIV2NIH2	-----N-----	759
HIV2ISY	---T---A---K-----I-----R-----G-----	759
HIV2ST	-----K-----	779
HIV2BEN	-----N-----V-----	779
HIV2CAM2	-----N-----	758
HIV2D194	-----N-----D-----V-----	797
HIV2GH1	---V-----N-----D-----V-----R-----	779
HIV2FA	---V-----N-----D-----R-----	76
HIV2S132P1	---S-----K-----D-----V-----I-----M-----	76
HIV2S52907	---S-----K-----	560
CONSENSUS.B	I?AgQPTESErIVnkIIEeMIKKEaiYVGVPAHKGIGGNQeVDHLVSQGIQVLFLEKIEPAQEeHeKYHq	767
HIV2D205	---I---A---A-----V-----	779
HIV2UC1	---V-----L---Q-----R-----	780
HIV2EHOA	---V---A---RE-----K-----I-----N-----	777
HIV27312AI	---V-----	76
HIV2JA	---I-----	76
HIV2ON	---I-----	76
CONSENSUS.C	IIAGQPTESErLVNRIIEELIRREAVYISWVPAHKGIGGNQeVDHLVSQGIQVLFLEKIEPAQEeHeRPHS	76
HIV2POLB7	-----	76
CONSENSUS.D	Ii?GqPTESErLVNQIIEeMIKKEaiYVAVPAHKGIGGNQeVDHLVSQGIQVLFLEKIEPAQEeHeKYHS	768
SIVMM251	---T---C---SE-----I-----D-----	780
SIVMM142	---T---C---TE-----I-----S-----	780
SIVMM239	---T---C---SE-----I-----D-----	784
SIVMM32H	---T---C---SE-----I-----D-----	784
SIVMNE	---T---C---TE-----I-----D-----	780
SIVSMH4	---A-----K-----	743
SIVSMH9	---VA-----K-----x-E-----E-----	746
SIVSMHPEJ	---VA-----	746
SIVSTM	---T-----K-----	765
SIVSMHLIB1	---VA-----NN-----T-----	171
HIV2POLC12	---VA-----K-----I-----D-----	76
HIV2FORTC2	---VA-----K-----I-----R-----D-----	590

HIV2/SIV POL

CONSENSUS .A	niKELtHKFgiPqLVARQIVNtCaqcQQKGEAIHQQVNAELG7WQMDCTHLeGKiIIVAVHVASGFIeAeVIP	856
HIV2ROD	-V---S-----N-----S-----T-----	833
HIV2NTHZ	-----S-----K-----HV-----T-----V-----	832
HIV2ISY	-V---A---L-N-----T-----	832
HIV2ST	-V---S---L-K-----T-----T-----	852
HIV2BEN	I-----L-----S-----I-V---Y-----	852
HIV2CAM2	-V---C---D-----Y-----V-T-----	831
HIV2D194	-I-----I-V-----	870
HIV2GH1	-M-----I-V-----	852
HIV2FA	-----I-V-----K---	149
HIV25132P1	-----A-----TK-----V-----G-----	149
CONSENSUS .B	NVKeLVHKFGipqLVakQIVNSCDKCCQKGEaIHGQVNaeLGTWQMDCTHLEGKvIIVAVHVASGFIeAeVIP	840
HIV2D205	-----I-----D-----I-----	852
HIV2UC1	-----L-----	853
HIV2ENOA	-----R-----I---S-----	850
HIV27312AI	-----	149
HIV2JA	-----HS-----	149
HIV2ON	-----I-----	149
CONSENSUS .C	NAKElTHKFGIPNLVAKQIVMTCHVCQQKGEAIHQQVNADLSTWQMDCTHLEGKvIIVAVHVASGFIeAeVIP	149
HIV2POLB7	-----	149
CONSENSUS .D	NVKeLVFXfglPrIVakQIVDTCKCHGKGEaIHGQVNaeLGTWQMDCTHLEGKiIIVAVHVASGFIeAeVIP	841
SIVMM251	-----I---R-----SD-----V-----	853
SIVMM142	-I-----SD-----V-----	853
SIVMM239	-----I---R-----A-SD-----	857
SIVMM32H	-----I---R-----SD-----	857
SIVMNE	-----SD-----	853
SIVSMNH4	-----	816
SIVSMNH9	---K---x-----L---x---xx-----G-----	819
SIVSMHBBJ	-----	819
SIVSTM	-----I-----	838
SIVSMMLIB1	-I-----Y-I-----	244
HIV2POLC12	-I-----K---R---I---R-----V-----	149
HIV2PORTC2	-I-----K---R---I---R-----VT-----	657
CONSENSUS .A	QeSgRQTALFLLLKLASRWPIThLHTDNGaNFtSQEVKMVAWwIGIEQsPGVPYNPQSQGVVEAMNHHLKNQIs	929
HIV2ROD	-----V---T-----D	906
HIV2NTHZ	-----V-----E	905
HIV2ISY	-----	925
HIV2ST	-----P-----V-----	925
HIV2BEN	-----V---T-----	904
HIV2CAM2	-----P-----	943
HIV2D194	-----S-----	925
HIV2GH1	-K-E-----S-----T-----	222
HIV2FA	-----S-----	222
HIV25132P1	-----	222
CONSENSUS .B	QETGRQTALFLLLKLASRWPIThLHTDNGaNFtSQdVKNVAMWwGIEqTFQVPYNPqSQGVVEAMNHHLKNQID	913
HIV2D205	-----PS-----	925
HIV2UC1	-----A---I-----	926
HIV2ENOA	-----A---I-----E-----	923
HIV27312AI	-----	222
HIV2JA	-----R-----	222
HIV2ON	-----R-----	222
CONSENSUS .C	QETGRQTALFLLLKLASRWPIThLHTDNGaNFtSQEVKMVAWwVQVEHTFQVPYNPQSQGVVEAMNHHLKNQID	222
HIV2POLB7	-----	222
CONSENSUS .D	QETGRQTALFLLLKLASRWPIThLHTDNGaNFtSQEVKMVAWwGIEqTFQVPYNPQSQGVVEAMNHHLKNQID	914
SIVMM251	-----G-----A-----H-----	926
SIVMM142	-----A-----H-----	926
SIVMM239	-----G---V-----A-----H-----	930
SIVMM32H	-----G---V-----A-----H-----	930
SIVMNE	-----G-----A-----H-----	926
SIVSMNH4	-----G-----	889
SIVSMNH9	x-----x---x---x---T---	892
SIVSMHBBJ	-----V-----T---	892
SIVSTM	-----V-----T---	911
SIVSMMLIB1	-----V-----N-----T---	317
HIV2POLC12	-----V-----	222

CONSENSUS.A	rIRBQANTiETIVLMavHcMNFkRGGIGDMTPaERlinMITtEQEIQPLQaKNSklknFrVYERBGRdQLWK	1002
HIV2ROD	-----I-----S-----D-----	979
HIV2NIH2	-----V-----I-----	978
HIV2ISY	-----M-----V-V-----N-Q	979
HIV2ST	-----V-----V-A-----Q-----	998
HIV2BEN	-----V-----S-----R-NF-Q-Y-----	998
HIV2CAM2	-----V-----S-----	977
HIV2D194	-----R-NF-K-Q-Y-----	1016
HIV2GH1	-----R-NF-Q-Y-----	998
HIV2FA	-----	236
HIV2S132P1	K-----	236
CONSENSUS.B	RiRDQAVSiETVVLMAcHcMNFkRGGIGDMTPaERlVnMITtEQEIQFIQaKNLKFQNFqVYYREGRDQLWK	986
HIV2D205	-L-----	998
HIV2UC1	-----A-----	999
HIV2EHOA	-----I-----L-T-----R-----	996
HIV27312AI	-----	236
HIV2JA	-----	236
HIV2ON	-----L-----	236
CONSENSUS.C	RIRBQANTMETIVL	236
HIV2POLB7	-----	236
CONSENSUS.D	rIReQANvETIVLMavHcMNFkRGGIGDMTPaERlVnMITtEQEIQFQSKNSKFKNFRVYREGRDQLWK	987
SIVMM251	-----I-----	999
SIVMM142	-----I-----	999
SIVMM239	-----I-----	1003
SIVMM32H	-----I-----	1003
SIVMNE	-----M-----L-----	999
SIVSMNH4	-----I-----	962
SIVSMNH9	-----I-----I-----I-----	965
SIVSMMPBJ	-----I-----	965
SIVSTM	-----D-T-V-----K-L-----	984
SIVSMMLIB1	K-D-----I-----T-----	390
HIV2POLC12	K-D-----	236
vif cds ->		
CONSENSUS.A	GPGLLWKGdGAViVKVGLdIKViPRRKAKIIRDYGGRQE7DSqshLEGARE?dGE?ACPCQVPEIQNKPRG	1072
HIV2ROD	-----E-L-----I-----M-----M-----	1036
HIV2NIH2	-----E-V-----K-----M-----M-----	1035
HIV2ISY	-----A-I-----K-----M-----M-----	1035
HIV2ST	-----A-I-----K-----M-----N-----V-----	1055
HIV2BEN	-----E-V-----L-SP-----M-----	1070
HIV2CAM2	-----J-----L-S-----N-V-----	1034
HIV2D194	-----A-V-----L-S-----V-----	1073
HIV2GH1	-----A-----L-----E-V-----	1055
CONSENSUS.B	GPGLLWKGdGAViIKVGTETIKViPRRKAKIIRHYGGGK*LDCCaDmEDTrQ?aremaqs?QVPEA	1050
HIV2D205	-----V-----G-----D-----	1058
HIV2UC1	-----L-----GT-----G-----	1064
HIV2EHOA	-----D-----N-----V-M-AREVAQSN	1056
CONSENSUS.D	GPGLLWKGdGAViIKVGTdIKVVPRRKAKIIRDYGGGKEVDSSSHMEDTgE.AREVA	1044
SIVMM251	-----	1056
SIVMM142	-----M-----	1056
SIVMM239	-----	1060
SIVMM32H	-----	1060
SIVMNE	-----V-----	1056
SIVSMNH4	-----E-----L-G-L-----	1019
SIVSMNH9	-----E-----L-G-L-----	1022
SIVSMMPBJ	-----E-----L-G-L-----	1022
SIVSTM	-----V-----G-L-R--G-----	1041
SIVSMMLIB1	-----	396
CONSENSUS.A	GALCSPFQGGMGHVDLQQGNIPTTRKKSSRNtGILEBNTKRKRMALLSCSKINLVYRKVLDRCPALCRHPNT	1144
HIV2BEN	-----	1142

HIV2/SIV VIF

<- pol cds			
CONSENSUS.A	NeegkrWlVVPtWRVPGR?ME?WHSLvKyLKyRTkDLQ?V?YVPHHKVGNAMWTCsRVIFPLkg?ShLEIQAY	68	
HIV2ROD	---D---K---K---K---C---N---	72	
HIV2NIH2	---I---R---K---C---EN---	72	
HIV2ISY	-DQ---A---K---I---Q---R---N---	72	
HIV2ST	---A---R---I---G---K---C---E---	72	
HIV2BEN	---DRM---K---A---E---R---Q---K---	72	
HIV2CAM2	---S---A---K---G---Q---C---R---D---R---	72	
HIV2D194	---N---R---K---E---R---E---E---	72	
HIV2GH1	---N---R---R---R---E---R---E---	72	
CONSENSUS.B	NEEEXNwlvVPTWRIPGR.LERWHSLiKYLKyRTkELQOVSYVPHHKVGNAMWTCsRVIFPLkga?LEVQGY	71	
HIV2UC1	--G---K---V---H---E---Y---	72	
HIV2D205	---D---G---I---NK---W---	72	
HIV2EHOA	---A---C---D---H---	72	
CONSENSUS.D	NEEEXNwlvVPTWRIPGR.LERWHSLiKYLKyRTkELQOVSYVPHHKVGNAMWTCsRVIFPLkga?LEVQGY	70	
SIVMM251	---A---V---F---	72	
SIVMM1A11	---A---V---F---	72	
SIVMM32H	---G---A---V---Y---	72	
SIVMM142	---V---A---	72	
SIVMM239	---A---V---F---	72	
SIVMNE	---A---V---K---Q---	72	
SIVSMNH4	---N---V---H---N---MA---RDET---	72	
SIVSMNH9	---N---V---G---K---H---N---A---RDEA---	72	
SIVSMNPBJ	---N---V---G---K---H---N---A---RDEA---	72	
SIVSTM	---V---G---H---N---E---S---A---GEA---	72	
CONSENSUS.A	WNLTPKQWLSS?SVR?TWYLEKFNTPDTPDQAD?LIHSTYFSCFTAGEVRRRAIRGEKILSCCNYPqAHk?QV	137	
HIV2ROD	---Y---I---V---P---R---RA---	145	
HIV2NIH2	---H---I---T---K---R---RS---	145	
HIV2ISY	---Y---M---S---G---T---S---SK---	145	
HIV2ST	---Y---L---S---Q---Y---	145	
HIV2BEN	---HA---L---I---A---	145	
HIV2CAM2	---YA---M---T---K---R---RS---	145	
HIV2D194	---K---L---S---A---	145	
HIV2GH1	---H---I---R---Y---I---V---	145	
CONSENSUS.B	WNLTPKQWLSS?SVR?TWYLEKFNTPDTPDQAD?LIHSTYFSCFTAGEVRRRAIRGEKILSCCNYPqAHk?QV	144	
HIV2UC1	---K---	145	
HIV2D205	---N---S---	145	
HIV2EHOA	---R---S---H---T---	145	
CONSENSUS.D	WNLTPKQWLSS?SVR?TWYLEKFNTPDTPDQAD?LIHSTYFSCFTAGEVRRRAIRGEKILSCCNYPqAHk?QV	141	
SIVMM251	---H---R---K---RY---	145	
SIVMM1A11	---H---V---K---N---R---Y---	145	
SIVMM32H	---H---R---K---RY---	145	
SIVMM142	---R---KD---E---R---R---H---	145	
SIVMM239	---H---K---N---R---Y---	145	
SIVMNE	---R---H---C---R---T---	145	
SIVSMNH4	---A---T---SE---K---K---N---	145	
SIVSMNH9	---x---A---T---G---SE---K---K---N---	145	
SIVSMNPBJ	---x---A---T---G---SE---K---K---N---	145	
SIVSTM	---E---T---S---C---Q---G---K---R---TK---N---	145	
vpx cds ->			
CONSENSUS.A	PSLQFLALRVVQGNQd.rpQrd?LtrKQWRRdYrGLRVRaQD?Rs?KQrgSEppapra?FPQvAkVLEILA	203	
HIV2ROD	---S---R---L---K---S---H---S---T---TY---E---	215	
HIV2NIH2	---S---GNSA---R---G---M---S---GY---S---PT---H---L---E---	215	
HIV2ISY	---K---N---N---L---G---H---QG---Y---	215	
HIV2ST	---G---N---N---G---H---Y---	215	
HIV2BEN	---Y---KG---A---HW---E---H---L---G---S---H---	215	
HIV2CAM2	---R---R---L---S---Y---S---G---Y---	215	
HIV2D194	---Y---G---KGAA---HW---Y---L---G---S---H---G---	215	
HIV2GH1	---Y---G---KGAA---HW---Y---L---S---H---	215	
CONSENSUS.B	PSLQFLALRVVQGNQd.rpQrd?LtrKQWRRdYrGLRVRaQD?Rs?KQrgSEppapra?FPQvAkVLEILA	215	
HIV2UC1	---I---N---H---	216	
HIV2D205	---N---G---M---D---IRTS---S---S---QG---	216	
HIV2EHOA	---T---G---M---D---IRTS---S---S---QG---	216	
CONSENSUS.D	PSLQFLALRVVQGNQd.rpQrd?LtrKQWRRdYrGLRVRaQD?Rs?KQrgSEppapra?FPQvAkVLEILA	209	
SIVMM251	---K---D---D---	214	
SIVMM1A11	---K---D---D---	214	
SIVMM32H	---R---Y---E---S---G---	214	
SIVMM142	---R---S---V---G---E---	214	
SIVMM239	---K---D---G---	214	
SIVMNE	---R---Y---R---S---VD---	214	
SIVSMNH4	---T---D---N---K---RN---GS---S---AE---	214	
SIVSMNH9	---T---x---x---N---x---L---T---x---RN---GS---ESFAx---	214	
SIVSMNPBJ	---T---x---x---N---x---L---T---x---RN---GS---ESFAx---	214	
SIVSTM	---K---E---R---TAR---G---GSI---V---T---G---H---P---S---E---TD---	214	

```

                                <- vif cds
CONSENSUS.A MtdPREtVPPGNSGEETigEAFpWLDRTvEaINReAVNHlPRELIQVWQrSWrYWHDeQGMS?SYTKVRYLC 72
HIV2ROD -----A--N-----E----- 73
HIV2NIH2 -----E--A-----H----- 73
HIV2ISY --N---I-----E--D-----R----- 73
HIV2ST -AG---I-----I----- 73
HIV2BEN -----R-----E--I--L-----A----- 73
HIV2CAM2 -----R-----E--A-----Q----- 73
HIV2D194 -A---R-----I--L-----A-----T----- 73
HIV2GH1 -----R-----I--L-----D-----P----- 73
CONSENSUS.B M.DPRERVPFGNSdEETvGEAF?WLErTitELNRVAVNHlPRELIQVWQrSWAYWREEQGMSiSYTKVRYLl 71
HIV2UC1 -,-,-----I---D-----C-----S----- 72
HIV2D205 -----,-----A----- 72
HIV2ENOA -----,-----G-----E---T-LEH-----K-----C----- 72
CONSENSUS.D MadPRErIPFGNSGEETigEAFpWLDRTvEaINReAVNHlPRELIQVWQrSWeYWHDeQGMS?SYvKYRYLC 72
SIVMM251 -----,-----Q----- 73
SIVMM1A11 -----,-----P----- 73
SIVMM32H -----,-----V-----Q----- 73
SIVMM142 -----,-----Q---T----- 73
SIVMM239 -----,-----P----- 73
SIVMNE -----,-----Q----- 73
SIVSMNH4 -----,-----R-----A-----R-----M---E--T----- 73
SIVSMNH9 -xx--x-----xx--x-D--H--x---A---x-----R-x-----M---V--T-x----- 73
SIVSMMPBJ -----,-----D--O-----A-----R-----M---V--T----- 73
SIVSTM -----,-----A-E-----H---D-----R-----P----- 73

CONSENSUS.A lmQkAvfiHikrgcTCLGgghGpgGw?r?GPPPPPPPGLV 110
HIV2ROD II---YM-VRK-----R-----P----- 112
HIV2NIH2 -I-M-MYM-A--DG-----M-QK-G.DQ----- 112
HIV2ISY -----M---K---R-E-----S----- 112
HIV2ST -----M---S-----S----- 112
HIV2BEN -----I-T-----W-EDM-RE-LEDQ----- 113
HIV2CAM2 -----,-----P----- 112
HIV2D194 I---Y---K---R---P----- 112
HIV2GH1 -----,-----S----- 112
CONSENSUS.B LMQKAMFvHytkGCrCLqEGHGPGGW.RSGPPPPPPPGLa 110
HIV2UC1 -----,----- 111
HIV2D205 -----,----- 111
HIV2ENOA -----I-FA---G--R----- 111
CONSENSUS.D LIQKALFmHcKkGCRCLGgghGAGGW.RpGPPPPPPPGLa 111
SIVMM251 -M-----,----- 112
SIVMM1A11 -----,----- 112
SIVMM32H -----,----- 112
SIVMM142 -----,----- 112
SIVMM239 -----,----- 112
SIVMNE -----,----- 112
SIVSMNH4 -----V-----E-----T----- 112
SIVSMNH9 -----,-----GE----- 112
SIVSMMPBJ -----,-----H-----GE----- 112
SIVSTM -----,-----S-R----- 112

```

HIV2/SIV VPR

< vpr cds		set cds	
CONSENSUS A	MCEApTefPpEdgTPpRepGdeWvleilBreIkeEALKHFDPrLLlaLQ?yIy?RHGDTLGGARELirvLQRAL		71
HIV2ROD	-A --L- V-----L-----I-----K---T-----K-----		73
HIV2NIH2	-----L-----R-----A-----E-----R-----T---R---T-----I-----		73
HIV2ISY	-----A-----T-----Y---T-----T-----Y---T-----		73
HIV2ST	---S---R-----T---X-----T---N---A-----G---I---		73
HIV2BEN	-----R-DL-SD---T-----R-----Y---HN-----KT---		73
HIV2CAM2	-----L-----D-----T---GH- A-----R-----		73
HIV2D194	---R---L-ST---T-K-----C---N---N-----		73
HIV2GH1	---R---L-GD---R---G-----N---HS---P-----I---		73
CONSENSUS B	MAEAapeIppenenPQREPWEewv?dvIEEIKQEALKHFDPrLLTALGNFIYsRHGDTLGGAGELIKLQRAL		72
HIV2UC1	---T---S-----E---M-----R-----		73
HIV2D205	-----IGEI-----		73
HIV2EHOA	---V-----DK-----Q---V-----N---N---E---L---		73
CONSENSUS D M e	?eRPPE?EgPqREPWGEWvEVLeElKeeALKHFDPrLLTALGNFIYsRHGDTLGGAGELIKLQRAL		68
SIVMH251	- - - - -N-----		69
SIVMH32H	- - - - -N-----		69
SIVMH1A11	- - - - -N---S---G-----G-----		68
SIVMH142	- - - - -N-----K-----		69
SIVMH239	- - - - -N-----		69
SIVMNE	- - - - -D-----		69
SIVSMNH4	- A---D-A-----Y---D-----		69
SIVSMNH9	- T---D-A-----x---x---Nx---x---Y---D-----G---		69
SIVSMNPBJ	- T---D-A-----I---N-----Y---D-----		69
SIVSTM	- TH---D-----I-Q---R-----S---Y---D-----K-----		69
CONSENSUS A	f?XIRAGCg?SRigQrRgnPlsA:PTPRgM?		100
HIV2ROD	-T-----H-----G-----N-Q		105
HIV2NIH2	-A-----H-----G-----H		105
HIV2ISY	-T-----H-----P-----N-Q		105
HIV2ST	LL-----R---P-----R		105
HIV2BEN	-V-----NR---		87
HIV2CAM2	-T-----NH-----G-----R-		104
HIV2D194	-V-I---DR--K---R-A-CP-A-----H		105
HIV2GH1	-V-L---NR--S---R-T-FF-A-----Y		105
CONSENSUS B	FLHFRAGCQHSRIQpGGGNPLSaiPPs		100
HIV2UC1	-----S-----T---P		101
HIV2D205	-----G-----		101
HIV2EHOA	-----G-----		101
CONSENSUS D	FmHFRgGC?HSRIQpGgGNPLstIPPsRgvL		99
SIVMH251	-----N-----		97
SIVMH32H	-----N-----		101
SIVMH1A11	-----I-----R---A---SM-		100
SIVMH142	-I---S---S-----SM-		101
SIVMH239	-----I-----A---SM-		101
SIVMNE	-----T-----S-----RM-		101
SIVSMNH4	-I---S-A-----SR-----		101
SIVSMNH9	-I-----R-----S-x-----x---		101
SIVSMNPBJ	-I-----R-----S-----		101
SIVSTM	-----R-----A---T---		101

		<- VDR	
CONSENSUS .A	NETPLKAPESL?syNEPES?TSeqdv??QelakqGeEiLsQLyrPLEaCtNsCyCK?CcyhCQICPlkKGLG	69	
HIV2ROD	-----K-C--F-R-----AT--R-----T-N-----R-----M-----N-----	73	
HIV2NIH2	-----E-C-----R-----AT--R-----K-----D-----Q-----	73	
HIV2ISY	-----G-----R-----AT--N-----T-N-K-F--G--F-----N-----	73	
HIV2ST	-----G--G-----C-----AAA-G-VSP-D--Y--Q-----D-K-----K-----M-----N-----	73	
HIV2BEN	-----KP-----C--R--TA-----L-A--H--P--K--R-SF-----S-----	73	
HIV2CAM2	-----E--G-C-----R--G--ATT-----L-----Q--E-D-----R-----	73	
HIV2D194	-----E-----B-----C--R--TA--R-----L-A--H-----Q-S-----	73	
HIV2GH1	-----H-----E-----C--G-TA-----L--H-----Q-SF-----	73	
CONSENSUS .B	MEIPLqEQESLkSSsepSSSTSEpVvNTQGLDNqGEEILsQLYRPLkACDntCYCKKCCYHCQLCPLKKOLG	73	
HIV2UC1	-----S-----E-----K-----	73	
HIV2D205	-----N-----	73	
HIV2EHOA	-----K-----N-----QH-----G-A-----R-----S-----S-----	73	
CONSENSUS .D	METPLrEQEnSLsSnerSScisEada?tpesAnlgeEILsQLYRPlEaCyNtCyCKKCCYHCQLCPlkKOLG	72	
SIVNM251	-----L-----T-----	73	
SIVNM32H	-----T-----	73	
SIVNM1A11	-----SGH--T--A-S--L-----	73	
SIVNM142	-----Y--A-AI-----	73	
SIVNM211	-----L-----T-----	73	
SIVNM239	-----S-----	73	
SIVMNE	-----K--G--T--A-P-L-----E-----E-----	73	
SIVSMH4	-----K-----CR-H--S--V-VP-----E-----K--R--H-----	73	
SIVSMH9	-----K--S--R-H--S--V-D--S-Ex-----xV--Kx--H--x-----	73	
SIVSMMPBJ	-----K--S--R-H--S--V-D--S-E-----K--R--H-----	73	
SIVSTM	-----K--S--R--S-P--T--V-A--GL--QE--W-----E-C-K-F-----L--VT-----	73	
rev cds ->		exon \ / exon	
CONSENSUS .A	Iwy?RkgRRRRrPKKtK?hessasD.KSISTRTy?SQptKkQKKFle?tvetd?glGr	120	
HIV2ROD	-C-E-----T-P-PPP-----D-----V-A-----T-P-----	130	
HIV2NIH2	---D-----A-P-----RN--E-----A-----L-----	130	
HIV2ISY	-C-D-----S--A--P-----N--TE-----QT-L-A-P-----	130	
HIV2ST	-E-----A-----M--E-----TAL--IG-P--	130	
HIV2BEN	-S-E-----R--TP-P-P--D-----E--S-A--V-TC--Q	130	
HIV2CAM2	-C-D-----A-A-----RN--A-----A-----P-----	130	
HIV2D194	---A-Q-----R--T-PPP-----D-----P-T--VSAC--H	130	
HIV2GH1	---A-S-----R--T-----D-----E--T-T-MV-TCS--	130	
CONSENSUS .B	IcYdrS..RKRSeKRAKtPassApd??SLSArTGDSQPPKKQKkeVETtg?TDIGPGRSNTSTSRFA	135	
HIV2UC1	-W--H--V--S--E--N-----TK--K-L-----	128	
HIV2D205	---R--A-----P--K-----T--P-----	128	
HIV2EHOA	---E--T--NRR-----RA-----N	139	
CONSENSUS .D	IcYEqsr?RRRcpKKaKantssASn?k?isnRtRbcQPKKakKetVE?AVatapGLgr	126	
SIVNM251	-----K-----L-P-----E-----K-----	130	
SIVNM32H	-----K-----NRP-----E-----K-----	131	
SIVNM1A11	-----K-----NRP-----E-----K-----	131	
SIVNM142	-S--K-HR-----EP-P--I-L-----A-----	130	
SIVNM211	-----K-----L-P-----E-----K-----	130	
SIVNM239	-----K-----P-----E-----K-----	130	
SIVMNE	-----R-----P--D-----EQ--A-----	130	
SIVSMH4	---H--T--PPP--SL-T-A-NR--E--TE--ADL--	128	
SIVSMH9	---Q--T--F--D--SLxx-A-N--E--xE--DL--x--	128	
SIVSMMPBJ	---Q--T--D--SL-R-A-N--E--AE--DL--	128	
SIVSTM	-T--R--VR--TVPI--NRSL-T-A-NS--EQ-KE--TE-EST--K	129	

HIV2/SIV REV

exon 1 / exon		
CONSENSUS.A	M7erAdeEg1qrkLRLLiRLHQT.....NPYPGpGTAgQRNRNRNRwz?rWzq??ALAdkiytEPDPPaDe	63
HIV2ROD	-N-----KQ---IL---S-----	67
HIV2NIH2	-T-----KQ---IL---S-----	67
HIV2ISY	-T-----VR-----L---R-----ER-K-IL---R-----P	67
HIV2ST	-N---E-E-R-----KQ---LV-----	67
HIV2BEN	-S-----G---L-----R-RQ-LRLV---N-LCAV---T--	67
HIV2CAN2	-T-----Q---IL---R-----P	67
HIV2D194	-RD---E-----I-----H-Q-----R-RQ-FRLV---T-LH-I---T--	67
HIV2GH1	-H-K-G-E-ER-----H-----Q-R-LRLV---L-----T--	67
CONSENSUS.B	Mcar..E?DLQKsLRLLHLLHQT?????nPYPOgPGTASQRNRNRNRWkrRGLQILALADRIhslsdsPrEe	64
HIV2UC1	--T---K-----T-----R-----	65
HIV2D205	---G---E-----F-TA-A--	65
HIV2EHOA	-N---R-----SEYGTTO-----Q-----P-P-----G	71
CONSENSUS.D	MaaH??EELRkLRLLiRLHQT?.....npYpGpGTANQRqrrRRWrgRWgQLLALADRIYBFPDPPTdT	65
SIVM251	---ER-----S-----R-----	67
SIVM32H	---ER-----T....D-----R-----	68
SIVM1A11	---ER-----I....D-----R-----	68
SIVM142	-R-TG---R-----S-S---K-----	67
SIVM1P11	---ER-----S-----R-----	67
SIVM239	--N-ER-----K---R-----	67
SIVMNE	---AE---R-----N-----N--	67
SIVSMH4	---T---R---F-----Q---R-----I---V--	65
SIVSMH9	---N---R---F-----xx---R---x---I-x---A--	65
SIVSMPEBJ	---N---R---D-----R-----I-----V--	65
SIVSTM	--DQ---QF---T....D---Q-----N-----N-----N-----AS-	66
CONSENSUS.A	PLdr?IQhLQsLTiqeLPDPPT?lps.....????nanQrLAET	97
HIV2ROD	---QT-----H-----	100
HIV2NIH2	---A-----D---N-----PEST-----A	107
HIV2ISY	---QT-Q-----T-----	96
HIV2ST	---EQT-----N-----SESID-S---I	107
HIV2BEN	---A---R-----D-----G---	103
HIV2CAN2	---T---E---D---HP-----A	100
HIV2D194	---A-Q---R-----D-----G---	103
HIV2GH1	---A-D-R---H-----D-----G---	103
CONSENSUS.B	PLdLaiQRLQnLlveOLPNEPtpstp???TaQA?tcIPVWDQLVPRSnPss?GcgrdSCE?r?dlmgssQes	130
HIV2UC1	---V---E-----F-----NE-E---H-KSP-E---KD	135
HIV2D205	--N-----PNLQSP-T---PG-V-----A-G-K-Y-N---C-R---G---	138
HIV2EHOA	---IK-----S---I-----Q-----RGE-V-P---	141
CONSENSUS.D	PLdLaiQQLQnLllesIPdPPTn?PeaLcdpte?srspQcTAEN	107
SIVM251	---T-----KG-----	107
SIVM32H	---T---G---N-----	108
SIVM1A11	---T-----KD-----	108
SIVM142	---I-----N---A	107
SIVM1P11	---T---KD-----	107
SIVM239	---T-----D-----	107
SIVMNE	---V-----I-I-H---NP-----	107
SIVSMH4	---G---EL---SA---P-N-VAK....	100
SIVSMH9	---G---XL-N-ASA-P-K-AA....	100
SIVSMPEBJ	---R---EL-N-ASA-P-K-AA....	100
SIVSTM	---V---G-S-QDL---P-L-KD-----	100
CONSENSUS.B	g?snHrdpQEdQTRt	144
HIV2UC1	SG-----	150
HIV2D205	-E-----N----	153
HIV2EHOA	-RRD-CNT-----G	156

This page intentionally left blank.

HIV2/SIV ENV

CONSENSUS.A	M?ggrnqlliaillLaSacliyc??qyVTVFYgiPaHrNASiPLfCATkNRDTWGTIQCIPDNDDYQEi?LNV	67
HIV2ROD	-M-----V--T-----V-T-K--T-----R-----T---	68
HIV2NIMZ	-KSK-----V---Y--H-K--F-----T-----T---	71
HIV2ISY	-S-KI--V-F--T-----T..K-----V-V-K-----P---	70
HIV2ST	-C-----FV-S--V-----V-V-----A---	70
HIV2BEN	-EP-----FVV--T---V--S-----K-----f---	71
HIV2CAM2	-ER-----RQ-----V--K-----P---	72
HIV2D194	-EP-----V---T-----K-----T---	71
HIV2GH1	-C-KSL-CV-S---Y-V--T-----V-V-----K-----P---	70
HIV2U07104	-.GA---T---T-----R-----S--G-----V-----A---	68
HIV2U07106	-TRR-A---S---T-----K-N-----T-----T---	71
HIV2U07107	-APQKT---S---T-----E-----A-----R-----A---	71
HIV2U07108	-ARE-T--I--T-----R---Q-----K-----A---	71
HIV2CBL21	-CS--L--V-S---I-----T-----V-----A---	70
HIV2CBL22	-S-KI--V---T---V---T-----V-----L---	70
HIV2CBL23	-NG-----V-----T-----TN-----T---	71
HIV2CBL24	-C-----VTP-----S---TK-----R-----A---	70
HIV2CAM1	-AYK-----T-----KKK-----P---	73
HIV2CAM2BR	-GC-----KQ-----V--K-----P---	72
HIV2CAM3	-TRKG-----Y---RQ-----T-----S---	72
HIV2CAM4	-AC-KS-----T-T-----A-----P---	71
HIV2CAM5	-TC-----T--Y-VH--K-----V-V-----A---	71
HIV2CAM6	-GRK-----I-T--I---K-----P---	71
HIV2ALI	-MSS-----VTP-----V--K-----V--K-----A---	71
CONSENSUS.B	MAH?sn?L?i?LLLIsiyGf?gk??nyVTVPFYGIPAWrNAtvPLfCATkNRDTWGTvQCLPDNgDYTEI?INI	65
HIV2UC1	---T--H-F-L-----V---L-HKK-----SV---	73
HIV2D205	---YF-SR-P-A---G-S--VC---Q-----I-----R---	71
HIV2EHOA	---VN-Y-LVT-----YM---F-----K--SI-----R-----D---Q---	71
CONSENSUS.D	MgCIGNQLLIAI111SvygiyC?..qYVTVPFYGvPAWrNATiPLFCaTKNRDTWGTTCCLPDNgDYSEIAInV	70
SIVMM251	-----T-----E-----	71
SIVMM32H	-----T-----E-----	71
SIVMM1A11	-----T..L-----	71
SIVMM142	-----I-----D-----	71
SIVMM239	-----T..L-----V-----	71
SIVMNE	-----F--A---I-----V-R-----D---I---	71
SIVSMNH4	-----L--V--LE-C-V--K-----D---I---	71
SIVSMNH9	-----LS-X-AS---V-----V-----R-----I---	71
SIVSMNPBJ	-----L--V--AS---V-----I-----V-----I---	71
SIVP209C19	-----L--V-ALK---V-----K-----R-----N-----V-I---	71
SIVSTM	-A-P-----ACLT--T-----I---	71

CONSENSUS.A	TEAFDAW?nNtvtTeQaiedVWnLFETSIRKPCVKLTPLCVaM?C?????????L?????????????..??	109
HIV2ROD	-----H-----K-SSTESSTGNN-----TS	123
HIV2NIH2	-----V-----N-TR.....NM-----	118
HIV2ISY	-----D-I-----V-----T-N-----NAS-----T	116
HIV2ST	-----V-----S-----R-----NS-----TA	117
HIV2BEN	-----V-----H-----N-SRVQG...NT-----TPNP...	125
HIV2CAM2	-----D-I-----K-----NIS-----TS	119
HIV2D194	-----D-----R-----M-----NI-----	116
HIV2GH1	-----D-----V-----S-----S-----NS-----T	116
HIV2U07104	-----T-----T...RKSSGNTS-----G	120
HIV2U07105	-----T-----S-KNSGNTSI-----	12
HIV2U07106	-----T-----N...SENSGNTS-----	122
HIV2U07107	-----T-----N-SKNYNMTSAGNITTTPTAESM.....	133
HIV2U07108	-----T-----N...SKDSGNTS-----	122
HIV2CBL21	-----D-----V-----S-----T-N...TTMQNIS-----	120
HIV2CBL22	-----D-----H-----N...TKTRRMNTI-----	121
HIV2CBL23	-----D-----R...NMTDARNT-----	122
HIV2CBL24	-----D-----V-----S-----K...NNT...AT-----TP	119
HIV2CAM1	-----K-----K...DGNITS-----	122
HIV2CAM2BR	-----D-I-----K...NISTSD-----	120
HIV2CAM3	-----S-----K...NNTQSSNTA-----	123
HIV2CAM4	-----I-----K...KRMTGNTT-----	122
HIV2CAM5	-----N-----D-V-----IK-E...KKLGNA7-----	122
HIV2CAM6	-----D-----V-----N-SRSRPS...NS-----GN	124
HIV2ALT	-----D-----V-----R-----I-K-SNISTESTTTS-----	124
CONSENSUS.B	TEAFDAW.GNTVT?QAvdDVWnLFETSIRKPCVKLTPLCVaM?CnktT?c???n??...	110
HIV2UC1	-----N-----E-----R-N-G...-TT-----	123
HIV2D205	-----Q-----R-----N-S-E-NPG-AS-----	124
HIV2EHOA	-----D-TK-----T-K...WSS-----	119
CONSENSUS.D	TESFDAM.?NTVTEQAIEDVW?LFETSIRKPCVKLTPLCItMrC.....NKsEtDrwGLTks?????????c	124
SIVMM251	-----E-----Q-----S-----S-----	128
SIVMM32H	-----E-----Q-----S-----K-----L-----	128
SIVMM1A11	-----E-----Q-----S-----S-----	128
SIVMM142	-----E-----Q-----S-----K-----S-----	128
SIVMM239	-----N-----Q-----S-----I-----	128
SIVMM5	-----E-----H-----K-----K-----S-----	128
SIVSMMH4	--A-----D-----N-----A-----T-----GNA...G-	129
SIVSMMH9	--A-----D-----N-----xx-----GxPAP-	130
SIVSMMPEJ	--A-----D-----N-----T-----GTFAPTT-	132
SIVP209C19	--A-----N-----N-----A-----T-----RG	126
SIVSMB670	-----T-----K-----GNA...I-	19
SIVSTM	--A-----D-----N-----A-----N-K-GKT...V	128

HIV2/SIV ENV

CONSENSUS.A	??????t?????????ine?spcir??nC5?G1?eBemv?CqfnMtGlerDkkk?YnetWys?DVVce??	154
HIV2ROD	KSTSTTT-TPTD...QEQE-S-DT--A-AD---G---TIN-----Q-----K-----T..	191
HIV2NIH2	TWTCRTD-Q....NITI--DT-H,A-AD--T--K---ID---S-----R-Q-T-A---K---D..	181
HIV2ISY	ESAVATTSPS....GPDm--DTD---QEN---R--D--E-----L---Q-S---K---S..	182
HIV2ST	KNTTSTP-TTTTA..NTT--G-N-S--TD--T--G---D-----L---R---S..	185
HIV2BEN	RTSSSTTSRPPTS..AASI--T-N--ENMT-A..-GY--MQ-E--K---Q--RR-KD--LE---D..	193
HIV2CAM2	DTTMIRT-TPSTA..KEAP-SDN-----TN---E--KI-K-H-----Q-----S---D..	187
HIV2D194	SGTTATPSPF....NITI-D-N-T--GDN--T--GK--V-E-E-----Q--RK--DA--R---DK..	182
HIV2GH1	NNTTTGTGSTT....GMSE--T--SY.SD--T--GK--I-N--Y-----Q-----K---S..	181
HIV2U07104	NATTSTP....NINETSI--T-----TN---R---S-----A--P-G-----S---T..	186
HIV2U07105	GS-TTTTPD..AAMV--T-Q--AN---R--L-S--D---T--P-R-----S---DA..	76
HIV2U07106	GS-TTPA...AISM--T-----TD---R--S-----S---T--P-R-----S---TT	186
HIV2U07107	N-S--AN---E--I-N--A--P--P-W-----R---N..	185
HIV2U07108	GN-PTST...ASV--T-S--HTDT--R--I-S--S---A--P-R-----S---AT	186
HIV2CBL21	TARPTTGISNTEHILK-N-S--STD-T-A--EQ--T-D-----Q-T---K---S..	186
HIV2CBL22	AQSIITTAkAVVINDTD...KFN--T--G--DI-N-E-D---V--QS-R---A---..	184
HIV2CBL23	..TPTTASPTTIKPVTE..S-N-S--AN---G--V-W-----Q-----K---..	188
HIV2CBL24	EGEAQATSTTAATATGANE--T--VSTN--T--GD--IS-E---Q---L---K---D..S	190
HIV2CAN1	GNT-TSTRARITSEE-K-N-S--SAD--P--K---M-----Q-----N---K..	188
HIV2CAM2BR	..TMMIRT-TPSTAKAPISDKFP--AN---IG--RI-R-H-----Q-----S---DN..	190
HIV2CAN3	ESPASMEINE--T--N--TD---K---S-----RR-T-S---T---A..	187
HIV2CAN4	AT-ASTLLNDTLL--S-T--GAN--T--E--LNE-----Q--V-R---K---DT..	188
HIV2CAN5	..LSNKT-TSSTANLAIDNT...--KNDT-P--ED-D--T-H---AL--P-R---T---D..	187
HIV2CAN6	VTTTIPS-TAATTTPGMV--T--K-AD---QR---E--S-A--A--P-Q---Q---..	194
HIV2ALI	PSPGSTLKLPL--SD--KAD--PR--GD---N-R---Q--P-Q---K---..	186
CONSENSUS.B	????Ttp?ctt?????lInetapCIkNDsC?GigIEa??dC?F?HTGLrRDE?KQYKDTWY??DLac???	159
HIV2UC1	TKPI--I--TKP..SEN--D-----T-P-----NTV--Y-N-----K-----EK---N..	191
HIV2D205	STTT-K-T--R...GLKTI--D-----T--L-E--IMQ-N-S---L-----SE-----	190
HIV2EHOA	ASKE--TSSA-LASSTOT--D-K--Q-----A-----MI--Q-K---K---S---KQ--V-EXG	191
CONSENSUS.D	tt?ctta?c?a?????k??vvNets?Ci??nnCt..GLEqEdmisCKFmWTGLKRDkkKEYNETWYS?DLvCEQ?	181
SIVMH251	--IT-A-P-S-PVS-E-IDM---S--AQ-----Q-----T-----T-----T-----	198
SIVMH32H	--AP--P-A-S...-IDM---S--THD---Q--G-----T-----T-----	195
SIVMH1A11	--T.A-PKAMSE...-INMI---S--VHD---Q--G-----A-----	195
SIVMH142	--AS-TT-TTAKSVETROI---P-VVHD---A-----	199
SIVMH239	--AS-TS-T-SA...-VDM---S--AQD---Q-----A-----	196
SIVMNE	--AP--IPTKAE...AIK---N-P--NHD---RR---A-----	196
SIVSMH4	--TAITT-ATPSV..AEN-I--SNP--KN-S-A-----G-----N-----R--I--S	199
SIVSMH9	--Q-STTPP...SPIIAK--DSDP--RS-----V-----R-----R-----N	199
SIVSMHBJ	Q-T--QASTTPTSPITAK--DSDP--KI---V-----R-----R-----N	204
SIVP209C19	AET--T-P-GTPSG-EK-I--GDP--RN-S-A-----G-----RR-----R--I--S	198
SIVSMH670	--A--APS-RTEKVTAN---S-VKN---V-----	76
SIVSTM	--V-PTAAA-AT...-PEL---S-VSM---E-SLVG-----R-----S--I--	196

CONSENSUS.A	??n?c?????????CyMnHCNTsvItESCDKHYWdamrFrYCaPPGfEallLrcnDtnYsGFpPnCsKVVvA	211
HIV2ROD	..N-S-...N...QTQ-----I-----Y-----A-----	254
HIV2NIH2	...N-...SS...QSK-----A-----	244
HIV2ISV	..D-S-...D...RKR-----V-----	245
HIV2ST	...D-...KK...EKT-----T-----	248
HIV2BEN	...T-...AGT--R---I-K-----K-T---	254
HIV2CAM2	...S-DQTN...ETT-----I---K---A---	253
HIV2DI194	..T-G-...GT--R---K-----K-----K---	243
HIV2GH1	...N-...KDG...KMR-----IK-----Y-----R---	245
HIV2U07104	..ET-N-...SN...QSR-----D-----	251
HIV2U07105	..PT-TSNQTSNITSQSR-C-----E---I---Y-----P---	147
HIV2U07106	GAT-TSN...QDR-----D-W---Y---YT-----	251
HIV2U07107	...NNPTQNR-----E---DI---T---Y---K---	247
HIV2U07108	GTNDPLN...QSR-----ND---S---Y---S---	251
HIV2CBL21	...STS...VTSKPMWIGK--S-----Y---K---T---	254
HIV2CBL22	...SG-...SNNTN.QSR-----KG---T-----	250
HIV2CBL23	...GNGTDT-----K-----A---	250
HIV2CBL24	TPG-S-...NQSR-----T-----	255
HIV2CAM1	ENG-S-T-...NOT-----T-----S---T-----I-	253
HIV2CAM2BR	...S-D-...QTTNET-----I---K---A---	255
HIV2CAM3	...D-N-...SSKSNITSKR-----I---Y-----	255
HIV2CAM4	...N-...NTNKTNDQGT-----T-----I---D---K---	256
HIV2CAM5	...S-...NOT--K---R-----T-A-----	248
HIV2CAM6	..KA-G-...SDCK-----R-----	258
HIV2ALI	...PFNTTPNQTR-----I---Y-----A---	251
CONSENSUS.B	??n?c.....?????CYrTCTNTsIQESCDKHYWDSLrFRYCaPPGfEallLRCNDtnYSGFHPnCsKVVv	216
HIV2UC1	..G-S-...STY--M---V-----Y-----K---	253
HIV2O205	...N-...RKVTSR-----F-----A---	254
HIV2EHOA	TRS-ES-...K--K-----K-----	253
CONSENSUS.D	??n?c.....?neSrCYMnHCNTSVIQESCDKHYWDalrFRYCaPPGYALLRCNDtnYsGFpPnCsKVVv	244
SIVMM251	...D-----T-----K---	262
SIVMM32H	...D-----I---T-----K---	259
SIVMM1A11	...G-----K---	259
SIVMM142	...G-----C--D---C-----	263
SIVMM239	...N-...G-----K---	260
SIVNNE	...E-----	260
SIVSMMH4	..A-ES-...-K--H-----S--L--A---	261
SIVSMMH9	SN--E-...D-K-----x-----A---T---	263
SIVSMM8BJ	..S-E-...D-K-----S---A---T---	266
SIVP209C19	AN--E-...-K--Q-----S---A-K---	262
SIVSTH	...V-...GE--R-----L-----T--A---	259

HIV2/SIV ENV

	* ***	***	***	***	***	*	***	-> "V3" loop	
CONSENSUS.A	atcTRMmEtQtSTWtGFnGtRaEnRTyiyWhgrdnRtiisLNkyyNLc>hCkrPGNktvvpitlMSGLvFHS?								282
HIV2ROD	S-----SL-----I-KQ-M---M---H								327
XIV2NTMZ	-----K-----NF---M-----L--P--FK---								316
HIV2ISY	S-----L-----IL-R-----RA---								317
HIV2ST	-----F---V-----								320
HIV2BSN	-S-----MR-----L-----								326
HIV2CAM2	S-----K-----H-SMY-R-----QR---								325
HIV2D194	-S-----X-----M-----RR---								315
HIV2GH1	S-----SI-----								317
HIV2U07104	-----SK-----H-SI-----								323
HIV2U07105	-----S-----H-SIY-----								319
HIV2U07106	-----C-A--K--F-SK-G-A--C-T-D-SIY-E-								323
HIV2U07107	--R-----SK-----TH-SIY-----								319
HIV2U07108	-----V-SK-----SI-----								323
HIV2CBL21	S-----I-----HR---								326
HIV2CBL22	S-----M-----H-I-----YK---								322
HIV2CBL23	-----F-----I-----R---								322
HIV2CBL24	T-----S-----I-----								327
HIV2CAN1	T-----S-S-----V-R-----								325
HIV2CAM2BR	S-----K-----SMY-R-----QR---								327
HIV2CAN3	T-----SL-----								327
HIV2CAN4	-S--K-----K-----TH-SV-----								328
HIV2CAN5	-S-----NIK--V--S-F--M-----								320
HIV2CAN6	T-----S-----L-----								330
HIV2ALI	-----M-----O--M-----I---								323
CONSENSUS.B	* ***	***	***	***	***	*	***		
SsectRMMEtQtSTWFGFNtTraENRTYIYW?KONRTIISLN?YYNLcthCkrPGNktVvPIrtvsOLIIFHS.									286
HIV2UC1	-----T--M--S-----K-----R-----I--TIM--N---								325
HIV2D205	-----S-----E-----T--SI-----								326
HIV2EROA	-LY-----G-----S-----M-----I---								324
CONSENSUS.D	* ***	***	***	***	***	*	***		
SSCTRMMEtQtSTWFGFNtTRAENRTYIYWhgndnRTIISLNkyyNLtckCRRPgNKTVLPVTIMSGLVPHS.									316
SIVMM251	-----								334
SIVMM32H	-----								331
SIVMM1A11	-----I-----Q-----								331
SIVMM142	-----R-----H-----A-----								335
SIVMM239	-----								332
SIVMNE	-----SK-----								332
SIVSMNH4	-----KS-----R--E-----								333
SIVSMNH9	-----x-S-----x--R-----								335
SIVSMMPBJ	-----S-----R-----								338
SIVP209C19	-----KS-----								334
SIVSTM	-----S-----								333

```

"V3" loop <-
*
CONSENSUS.A ?qpi?N?rPrQAwCwfkG?WkeAMqEVkqTlakHPRYkg.?tn?t???inf?apgkgSDpEvayMwtncRGE 343
HIV2ROD Y---K-----K-D-----E-----R---D-R..N-S-A----- 395
HIV2NIH2 .-VI-KK-----E-Q-----E-----RSRTEN-K-K--R-----T----- 385
HIV2ISY .-K-I-KK-----R-E-R-----V-----D-N..K--T---D----- 385
HIV2ST .-R-----E-----K-L-----D-E..K-R-I--ER----- 387
HIV2BEN .-T-----R-G-R-R-----VQ-----I-D-G..K--TK--A-----F----- 393
HIV2CAM2 .-R-I-K-----N-T-----E-----TK-I..D-T-K--ER-----T---S----- 393
HIV2D194 .R-VY-KK-G-----Q-N-I--R-----G---D-G..K--TK--I-----T----- 383
HIV2GH1 .-T-----K-R-----I-----D-K..N--TK--R----- 384
HIV2U07104 .-N-----K-----V-----E---K-E..D-H-A-----A----- 390
HIV2U07105 .-T-----S-L-Q-----N-G..K--T---S---A---I----- 286
HIV2U07106 .-N-----T-T-R-----V-----R---E-R..K--T---A----- 390
HIV2U07107 .-N-----R-----E-T-----R---D-G..K--A-----A----- 386
HIV2U07108 .-R-----T-K-----R-R-I-E-----R---N-G..N-K-V-----A----- 390
HIV2CBL21 .R-VI-RK-M-----E-K-R-----V-----R---K-E..N-T-AG-E----- 394
HIV2CBL22 .R-VL-TK-----Q-R-----N-E..S-T-T--R-----T----- 390
HIV2CBL23 .-VI-K-----E-G-----E-----E-K..N--T----- 390
HIV2CBL24 .-N-----K-----D-S..K--IG----- 394
HIV2CAM1 .-R-----N-T-----V-----R---D-G..N--T----- 392
HIV2CAM2BR .R-I-K-----N-T-----E-----KNIT..D-T-K--ER-----T---S----- 355
HIV2CAM3 .-T-----K-----E-----D-K..K--A--R----- 394
HIV2CAM4 .-M-----E-G-K-----LE-----ATNIN..N-RLAE----- 395
HIV2CAM5 .-K-----N-R-----E-VV-----R---EPNKT..V--I----- 386
HIV2CAM6 .-R-K-----Q-R-----E-K---K..F-V-N..N-T-E--R----- 397
HIV2ALI .-X-----E-R-----E-V-----D-N..Q--TK--R---A-EV----- 390
*
CONSENSUS.B .QPi..NkRPQAMCNFKGNWtEAI?EVKeTiikHPRYkg??tn?t???irl?eh??gSDPEvryMwtncRGE 346
HIV2UC1 .-L--T-----I--R-----N-E..R--VGPSA-----H----- 392
HIV2D205 .-R-----K--R-----GAK-I-SV.KL.VS--GK-----TT----- 395
HIV2EHOA .-R-----K-----Q-----KN---S--T---ISQ---A--ARS----- 391
*
CONSENSUS.D .QPi..NoRPQAMCNWfG?Wk?AIkeVktivKHPRYtG..Tmttd..kInltAPgGG.DPEVTFMMWtncRGE 380
SIVMM251 .-D-----K--D-----R----- 400
SIVMM32H .-D-----N--D---M-----R----- 397
SIVMM1A11 .-D-----R-K--D----- 397
SIVMM142 .-V-----R--N--E-----S---R----- 401
SIVMM239 .-D-----K--D----- 398
SIVMNE .-D-----R-E-N--E----- 398
SIVSMH4 .-E-S--K--Q--E-L-----D-R..-----A----- 399
SIVSMH9 .-E-E-R---xxL-----K---xxx..Q-K----- 401
SIVSMPBJ .-E-K--Q--E-L-----K-E..Q-K----- 404
SIVP209C19 .-R--N-SE-Q--E-L-----D-K..-----A----- 400
SIVSTM .-E-RG-----E-L-----D-A..--RIV----- 397
SIVSTM22579 ----- 8

```


HIV2/SIV ENV

CONSENSUS.A	f1YCNmTWFLmWVE...nkt????.....??hNYvpChIkQIINtWkVgknvYLPPReGELtCnStV	397
HIV2ROD	-----I-----HR--A-----R-----S-----	450
HIV2NIH2	S-----R-CQK.....QR--A--R-R-----R--L-----	443
HIV2ISY	-----GQ.....Q-----E-----S-E-----	442
HIV2ST	-----R-NQT.....Q-----Q-----	445
HIV2BEN	-----D-NQT.....RR-----A-E-----	448
HIV2CAM2	-F-----PNTT.....KR--A--R-----	451
HIV2Q194	-----NQT.....HG--A--R-----T-----	441
HIV2GH1	-----RPNQT.....Q--A--R-----Q-----	442
HIV2U07104	-----V-----I-----R-MNK.....TW-----H-----	449
HIV2U07105	-----DR-ANR.....TW-----	323
HIV2U07106	-----I-----R-VNR.....TW-----H-----	449
HIV2U07107	-----ANQT.....W-----H-----	445
HIV2U07108	-----I-----ENAT.....W-----H-----	449
HIV2CBL21	-----T-NQT.....Q-----R-----L-S-----	452
HIV2CBL22	-----GQT.....QR-----S-----	448
HIV2CBL23	-----I-----HR-----R-----	445
HIV2CBL24	-F-----R-NTT.....W-----S-----	452
HIV2CAM1	-F-----GE.....K-----	449
HIV2CAM2BR	-F-----P-NTT.....KR--A--R-----	453
HIV2CAM3	-----R-ENRT.....W-----R-----	453
HIV2CAM4	-----R-GGP.....Q-----R-----V-----	453
HIV2CAM5	-F-----GQE.....Q--A-----A--E-R-----I-----	444
HIV2CAM6	-----R-NQT.....QR-----R-----	455
HIV2ALI	-----GQE.....Q--A-----	448
CONSENSUS.B	F1YCNMcWFLmWVE...N?T???t.....?rNY?pChIrQI?NTWkVg?n?YLPPReGELSCNStV	398
HIV2UC1	-F-----R-G-----QK--VT--K--V-----KYV-----T-----S-----	450
HIV2D205	-----K-N-----A-----I-----K-I-----	453
HIV2GH2	-----N-----K-KS-----V-----I-----R-I-----L-----	58
HIV2EHOA	-----F-----R-G-----LK--AS--V-----I-R-V-----	448
CONSENSUS.D	FLYCKMnWFLmWVEdr?n?????????q?pkErhkrNYVpChIrQIINTWkVgknvYLPPReGDLTCnStV	441
SIVM251	-----DVTT.....R-----R-----	465
SIVM22H	-----DVTN.....R-----R-----	462
SIVM1A11	-----TAN.....K--Q-----I-----	462
SIVM142	-----SLTT.....K-----	466
SIVM239	-----TAN.....K--Q-----	463
SIVMNE	-----K-LTGT.....K-Q-----R-----	465
SIVSMH4	-----DQKG...RWKQ-NR--QQ-K-----	469
SIVSMH9	-----XIK-GS...RWTS-NQ--YOK-----X-----X-----	470
SIVSMH9	-----NIQ-GS...RWTS-NQ--QR-----	473
SIVP209C19	-----S-----DQNSN...RWKQ-KKP-QQ-----	470
SIVSTM	-----I-----RSTSENrdwnknk--QQ-----V-----Q-----	467
SIVSTM22579	-----RSTLEMKNWKNQ--QQK-----V-----R-----	78

	***	SU \ / TM ?	
CONSENSUS.A	TSIIANIId???nqt??nITFSaevaeLYrLeIGDYkLvEITpIGFAPt?ekkyss.ap?R?krgvfVLGfL		459
HIV2ROD	-----WQNN-----	K-----	519
HIV2NIH2	-----AGD-----A-----	SV-----	510
HIV2ISY	-----VDGD-R-----	A-----	511
HIV2ST	-----GGE-----	PV-----	513
HIV2BEN	-----IDK-R-HT-----	I-----	518
HIV2CAM2	-----ERD-----T-----D-----	SQ-----	520
HIV2D194	-----SDG-----	I-V-----	509
HIV2GH1	-----VNS-----	I-V-----	510
HIV2U07104	-----GG-----	SV-----	507
HIV2U07106	-----GG-----	S-N-----	508
HIV2U07107	-----GG-----	S-----D-----	504
HIV2U07108	-----GG-----	F-----P-----	509
HIV2CBL21	-----TYG-----D-----	SV-----	520
HIV2CBL22	-----L-----VNG-----	T-----	516
HIV2CBL23	-----ANG-----N-----D-----	A-----	513
HIV2CBL24	-----VKN-----	S-Q-----	520
HIV2CAM1	-----TDM-----G-ET-----	T-----	519
HIV2CAM2BR	-----ERO-----T-----D-----	SQ-----	522
HIV2CAM3	-----VEA-----	S-----	521
HIV2CAM4	-----L-----MR-DSP-----	I-V-----	521
HIV2CAM5	-----ETMCKT-----	D-----K-----	514
HIV2CAM6	-----N.VD-----S-----	G-----	523
HIV2ALI	-----L-----TCG-----	S-R-----	516
HIV2GB1222		-----L-----	8
CONSENSUS.B	TSIIANIId???sd?N?T??NI?vSAEVsELYrLEIGDYkLvEITpIGFAPt?kryss.vtPRNKRGV?VLGfL		460
HIV2UC1	-----DVYY-G-D-KT-----TM-----G-----	I-----	522
HIV2D205	-----N-----S-T-----S-----	DVR-----	521
HIV2GH2	-----I-----N-----AN-A-----S-----	P-----	126
HIV2EHOA	-----D.WI-K-L-----T-----	K-----	517
CONSENSUS.D	TSIIANIId.W???gMqT..nITMSAEVAELyRLLEIGDYkLvEITpIGfLAPT?VkrYTTggtSRnKRGVfVLGfL		508
SIVM251	-----TD-----S-----	D-----	535
SIVM32H	-----TD-----	D-----	532
SIVM1A11	-----TD-----	N-R-----	532
SIVM142	-----N-----TD-----S-----	N-----	536
SIVM239	-----ID-----	D-----	533
SIVM6	-----ID-----	N-----	535
SIVSMH4	-----E-----INS-E-----	I-----	539
SIVSMH9	-----E-----IN-K-----	F-----	540
SIVSMFPB3	-----E-----IN-E-----	AF-----	543
SIVP209C19	-----E-----INN-E-----	D-R-----	540
SIVSTM	-----I-----TNN-E-----	A-----	536
SIVSTM22579	-----I-----TNN-E-----	A-----	147

HIV2/SIV ENV

CONSENSUS.A	GFLa r aG s AMGaasLTl S SAQSRTLLaGIVQQQQQLldvVkrQQEmLRLTVWGtKnLQaRVTAIEkYLkdQAqI	532
HIV2ROD	-----V-----L-----Q---R-	592
HIV2NIH2	-----	583
HIV2ISY	---T---A-----FR-----A---R-	584
HIV2ST	---T---A-----	586
HIV2BEN	-----R-----H-	591
HIV2CAM2	---T---V---T-----L-----I-----	593
HIV2D194	-----G-----	582
HIV2GH1	-----	583
HIV2CBL21	---T---A-----	593
HIV2CBL22	-----	589
HIV2CBL23	-----	586
HIV2CBL24	---T---A-----R-	593
HIV2CAM1	---T---V---T-----L-----L-	592
HIV2CAM2BR	---T---V---T-----L-----	595
HIV2CAM3	-----T-	594
HIV2CAM4	---A---T---T-----E-----	594
HIV2CAM5	---R-----R-	587
HIV2CAM6	---T---A-----	596
HIV2ALI	-----T-A-----M-----R-	589
HIV2S132E2	-----	37
HIV27312A	-----	37
HIV2FA	-----	37
HIV2GB1222	-----V-----V-----T-----HV	81
HIV2CI171	-----	37
HIV2CI9	---I-----	37
HIV2IS053	-----I-----R-----	37
HIV2VI390	-----A-----	37
HIV2VI905	---T-----	37
CONSENSUS.B	GFLAMAGSANGatSLTL S SAQSRTLLaGIVQQQQQl?D V VKRQQELlRLTVWGtKNLQ t RVTAIEKYLKDQA?L	531
HIV2UC1	-L-----L-----L-	595
HIV2O205	-----PV-----A-----Q-	594
HIV2GH2	-----S-----L-----C-----Q-	199
HIV2EHOA	---T---A-----V-----A-----Q-	590
HIV2JA	-----L-	37
HIV2ON	-----L-	37
CONSENSUS.C	D V VKRQQEmLRLTVWGtKNLQ t RVTAIEKYLKDQA L	37
HIV22238	-----	37
CONSENSUS.D	GFLATAGSANGaASlTltAQSRtLLaGIVQQQQQLldvVkrQQELlRLTVWGtKNLQ t RVtAIEKYLKDQAqI	581
SIVMM251	-----	608
SIVMM32H	-----	605
SIVMM1A11	-----L-----	605
SIVMM142	-----V-----S-----	609
SIVMM239	-----	606
SIVMNE	-----	608
SIVSMHH4	-----V---S-----	612
SIVSMHH9	-----S-----xx	613
SIVSMHPIJ	-----S---V---S-----A-----	616
SIVP209C19	-----VS-----	613
SIVSTN	-----T-----	609
SIVSTM22579	-----A-----	220
HIV2F0784	-----	37
SMCI8	---T-----R-	37
SMCI2	---T-----R-	37

CONSENSUS .A	NSMGCAFRQVCHTtVPwvNdsLtPdnnMTWQeWe7qvryLeANIS7sLEQAQIQQEKNNMYELQKLNswDvfg	603
HIV2ROD	-----A-----D-----K-----K-----I-----	665
HIV2NIH2	-----S-----T-----QK-----Q-----T-----	656
HIV2ISY	-----T-----E-----HKI-F-----E-----	657
HIV2ST	-----T-----QRI-N-----E-----	659
HIV2BEN	-----S-----K-----K-----Q-----E-----IL-----	664
HIV2CAM2	-----A-E-----QK-----Q-----E-L-----N-----T-----	666
HIV2DI94	-----K-R-H-----Q-----	655
HIV2GH1	-----S-----K-----Q-----	656
HIV2CBL21	-----T-Q-----K-Q-I-----T-----	666
HIV2CBL22	-----R-D-----Q-I-----GQ-----	662
HIV2CBL23	-----T-E-----GKI-D-----QQ-----	659
HIV2CBL24	-----Q-----Q-H-----EL-----	666
HIV2CAM1	-----A-----QK-----Q-----R-----	665
HIV2CAM2BR	-----A-E-----QK-----Q-E-L-----N-----T-----	668
HIV2CAM3	-----QK-H-----EM-----I-----	667
HIV2CAM4	-----V-----N-----LRT-D-----QL-----S-----	667
HIV2CAM5	-----T-----HKI-D-----D-----S-----	660
HIV2CAM6	-----QK-----Q-----	669
HIV2ALI	-----N-K-D-----Q-----EQ-----T-----T-----	662
HIV25133E2	-----D-----ER-----Q-----	110
HIV27312A	-----D-----Q-K-I-D-----E-----	110
HIV2FA	-----Q-Q-I-----E-----	110
HIV2GB1222	-----Q-----ER-----	154
HIV2CI171	-----L-----K-----M-----QQ-----S-----	110
HIV2CI9	-----A-----E-----EI-----	110
HIV2ISO53	-----Y-----T-----Q-H-I-----EQ-----	110
HIV2VI390	-----K-----AK-F-----QQ-----	110
HIV2VI905	-----T-N-Q-----Q-Q-I-----A-----I-----	109
CONSENSUS .B	NSMGCAFRQVCHTtVPwvNdsLtPdnnMTWQeWe7qvryLeANIS7sLEQAQIQQEKNNMYELQKLN7MDvFg	602
HIV2UC1	-----T-----E-----R-----R-----S-----	668
HIV2D205	-----T-N-----H-----I-S-----	667
HIV2GH2	-----D-K-Q-----E-----S-----	272
HIV2EHOA	-----V-S-K-----R-R-----K-----Q-I-S-----	663
HIV2JA	-----H-N-Q-----QK-Y-----Q-----Q-----	110
HIV2ON	-----IE-----AK-Q-----Q-----Q-----S-----	110
CONSENSUS .C	NSMGCAFRQVCHTtVPwvNdsLtPdnnMTWQeWe7qvryLeANIS7sLEQAQIQQEKNNMYELQKLNswDvfg	110
HIV22238	-----	110
CONSENSUS .D	NSMGCAFRQVCHTtVPwvNdsLtPdnnMTWQeWe7qvryLeANIS7sLEQAQIQQEKNNMYELQKLNswDvfg	652
SIVMM251	-A-----A-----D-----E-----A-----	681
SIVMM32H	-A-----A-----D-----E-----A-----	678
SIVMM1A11	-A-----A-----K-E-----E-----A-----	679
SIVMM142	-A-----A-----E-----A-----	682
SIVMM239	-A-----A-----K-E-----E-----A-----	679
SIVMNE	-A-----AN-----N-E-----E-----A-----	681
SIVSMOH4	-----ET-V-N-----Q-----Q-----I-----	685
SIVSMOH9	-----X-----DT-----X-----KQ-N-----XX-----X-----XX-XX-----	686
SIVSMMPBJ	-----R-----DT-----N-----KQ-N-----QS-----T-----I-----	689
SIVP209C19	-----D-V-I-D-----Q-----Q-S-----Q-----F-----I-----	686
SIVSTM	-----D-V-----Q-----V-----	682
SIVSTM22579	-----D-V-----Q-----V-----	293
HIV2F0784	-----E-----ST-----Q-----Q-----N-----I-----	110
SMCI8	-----E-DT-Q-----I-IRD-----SEK-N-----	110
SMCI2	-A-----E-NT-Q-----I-IRD-----SES-----	110

HIV2/SIV ENV

		/ early stop	
CONSENSUS.A	NWfDltsWikYiqYGVyivvgiiaLRiviyvQmlsRlRkGYRPVFSspPgYiQQiHhkd?eQporeeTBeD		675
HIV2ROD	-----V-----L-I-AV-----RG-----N-----		738
HIV2NIH2	--L-F--VR-----V-----I-----Q-----		729
HIV2ISY	-----M-----V-----W-----D-----		730
HIV2ST	-----V-----V-----A-F-----R-----		732
HIV2BEN	-----V-----H-----A-----L-----P-----L-----RG-----N-G-----		737
HIV2CAM2	-----S-----V-----S-----P-----S-----		739
HIV2D194	-----G-----A-----I-----L-----L-----T-RG-----N-----		728
HIV2GH1	-----V-V-----A-----I-----L-----L-----T-RG-----N-G-----		729
HIV2CBL21	-----R-----I-----F-----P-A-L-V-----HRG-----PS-----		739
HIV2CBL22	-----F-I-VV-----V-----W-----D-G-----		735
HIV2CBL23	-----II-VV-----I-----L-----W-----		732
HIV2CBL24	-----A-----VII-----IV-----L-----RHW-----D-G-----		739
HIV2CAM1	-----V-----I-----L-----L-----		738
HIV2CAM2BR	-----S-----V-----P-----S-----		741
HIV2CAM3	-----R-----V-----A-----I-----F-----Q-----D-----		740
HIV2CAM4	-----R-----VV-----I-----I-----L-----T-K-----G-----		740
HIV2CAM5	-----R-----F-II-VV-----W-----NG-----G-----		733
HIV2CAM6	-----L-V-----C-A-V-----L-----N-G-----		742
HIV2ALI	-----L-A-V-----I-V-----Q-----T-C-----		735
HIV25132E2	-----A-V-----A-V-----A-I-----		150
HIV27312A	-----A-V-----V-VI-----IG-----R		150
HIV2FA	-----V-----V-----I-----R		150
HIV2GB1222	-----L-V-----H-V-----VV-----Q-----T-----		227
HIV2CT171	-----F-----VV-----R-----		149
HIV2CI9	-----L-II-V-V-----A-I-L-----		151
HIV2IS053	-----IA-V-----G-----R		151
HIV2VI390	-----A-V-----A-----F-----		151
HIV2VI905	-----G-----C-----L-TG-----R		150
CONSENSUS.B	NWfDFTSWmayirLGLYvVaGlivLRiviyi?QMLArLRKGYRPVFSspPgYiQQiHhkd?eQporeeTBeD		674
HIV2UC1	-----A-----V-----M-----H-----DE		741
HIV2D205	-----L-----IK-H-----I-----V-----V-----		740
HIV2GH2	-----L-----IK-H-----V-----V-----G-----		345
HIV2ENOA	-----I-I-IV-----A-----I-----		736
HIV2JA	-----K-----I-----K-----		150
HIV2ON	-----HL-----I-----M-----		150
CONSENSUS.C	NWfDFTSWVVKYIYLGFYIVAGVIVLRiAIYVVQMLMRLRG		151
HIV22238	-----		151
CONSENSUS.D	NWfDltsWikYiqYGVyivvgvi?LRiviy?vQmlLaklrgGYRPVFSspPgYiQQiHhkd?eQporeeTBeD		719
SIVMM251	-----A-----I-V-----L-----I-----T-T-DPA-----R-K-----		753
SIVMM32H	-----A-----L-----I-----T-----DPA-----R-K-----		751
SIVMM1A11	-----A-----L-----I-----T-----D-A-----R-K-----		750
SIVMM142	-----A-----I-I-----L-----I-----R-----T-T-DPA-----KK-----		754
SIVMM239	-----A-----L-----I-----T-----DPA-----R-K-R-----		752
SIVMNE	-----A-----R-----L-----I-----XT-----R-D-A-----T-----		754
SIVSMH4	-----R-----L-L-G-----V-----R-----A-V-----IP-HKG-EP-----E-----		758
SIVSMH9	-----x-----L-L-G-----V-----R-----AxVx-----IP-xTG-E-----E-----		759
SIVSMFBJ	-----R-----L-L-G-----V-----R-----A-V-----IP-TG-E-----E-----		762
SIVP209C19	-----R-----L-L-VG-----V-----R-----A-V-----IP-HKG-EP-----E-----		759
SIVSTM	-----VR-----L-I-LVM-----VA-----IN-L-----R-K-----CR-----IP-HKG-EQ-----T-EG		755
SIVSTM22579	-----A-V-----L-I-IVM-----VA-----V-----R-K-----H-IL-HKG-EQL-----T-EG		366
HIV2F0784	-----F-----V-----YI-F-V-----I-G-----V-LI-L-G-H-K		150
SMCI8	-----F-----V-----YI-F-V-----I-G-----V-LI-L-G-H-K		150
SMCI2	-----R-----F-----V-----YI-F-V-----I-G-----IV-LI-L-G-----K		150

```

                                nef start ->
                                <- rev cds
                                *
                                <- tat cds
CONSENSUS.A vg7n?Gdr?WFWpiayihflirllirllleglyniCRdLlSr7ff?tlqpi?g?l7???t77RDWlRlkaAyLQY 735
HIV2ROD G-S-G---Y-----Q-----R--S-----S-L---L-Y-N-----RT-F--- 804
HIV2NIH2 --S-G---S-----ISPI---F-S-QRAL-AI----- 802
HIV2ISY --NDV-S-S---E-----R--S-----...YL-L-P----- 792
HIV2ST --NSV--NW---R-----Q---NR-----S-Q---L-S--RRAL-AV---PNT--- 805
HIV2BEN --GDS-YDL---N-VQ---H--T---I-----KNSP-RRL-S-S---AI-----Q--- 806
HIV2CAM2 --G-G---S---Q-E---Q-----N---I-Q---L-N-----P-V-F--- 805
HIV2D194 A-DDS-FGL---LN--Q---N--T-----S--G---KNSP-RRL-S-S---AI----- 797
HIV2GH1 DRDD--YDL---N-----H--T---K---TNSP-HRL-S-N---AI----- 797
HIV2CBL21 --DSV--SS---I-----T-I-KS-Q---S-G-QRAL-AI---PG--- 811
HIV2CBL22 --D-A-SSS---R-----Q-V---W--SS---S-LI-A----- 796
HIV2CBL23 --N-V--SL---R-----HQ---N---ISL--R-VF-S-QRAL-AI---FD--- 804
HIV2CBL24 --D-V-TSL-----I-----SSLI---L-S-QRAL-AI---E--- 812
HIV2CAM1 --S-V---F-----R--S-----I-L---L-N-----AST-F--- 804
HIV2CAM2BR --G-G---S---Q-E---N---I-Q---L-N-----P-V-F--- 789
HIV2CAM3 --I-G---S---T-----R--S-----S-L---F-N-----RT-F--- 806
HIV2CAM4 --D-V--SL---L-----R--SS---S-P--RA---L-AI----- 806
HIV2CAM5 G-N-A--NS---Q-R---Q-----S--N--KS-L--L-----NT-F--- 795
HIV2CAM6 --N-G-Y-S-----R--SS---S-L---L-N-----RT-F--- 808
HIV2ALI --D-V---L---L---N-A---I---S---I-SI---F-S-QRAL-TI--- 808
HIV2GB1222 --S-V---F-----S---I---S--N---I-L---F-N-----P--RT-F--- 293
                                *
CONSENSUS.B GGndggYRSWFWQIEYIHFl?RQlRnLLlWLYngCrtBLlKtfg?Lqpi???lq.....DLRLpVAYLQY 734
HIV2UC1 --E-A-----A--I-----N---S-I--A-----SL--- 803
HIV2O205 --D-----L-----I-HQ-STN----- 805
HIV2GH2 --E-----L--P-----S-S-----RT---LQP-----L----- 407
HIV2BHOA --NE-----PI---D---S---S---T---V-----P---R--- 798
                                *
CONSENSUS.D GG7?GGnsSWFWQIEYIHFLIRQLIRLLTWLfsncRtllsR7yQlLqPi1Qr1S?tlqRiREVlR?eltYLOY 787
SIVMM251 --EG-----A-----A--R-V---T----- 826
SIVMM32H --EG-----A-----AA---T----- 824
SIVMM1A11 --EG-----V-----A--R---T----- 823
SIVMM143 --GS-----A-----F---A--R---L----- 827
SIVMM239 --EG-----V-----F---A---T----- 825
SIVMNE --GS-----A-----F--F-T-----T----- 827
SIVSMH4 --DR--SR-----S---DW-L-S---V--S--T---V---I-I-IA--- 831
SIVSMH9 --XR--R--x-----S---DW-L-xC---V--S--R---A---I-V-I--- 832
SIVSMMPBJ --QR--R-----S---DW-L-NC-T---V--S--R---A---I-VQIA--- 835
SIVP209C19 --DR--SR-----S---DW-L-S---V---F---V---I-I-I--- 832
SIVSTM --DR--IN-----T-----V-----YN-F-AC---I--T-H-TF---I-RI-----V-IGAA--- 828

```

HIV2/SIV ENV

CONSENSUS.A	GceWlQeAfqaFaratzeTlegawrPlwgPl7rIGRgILavPRRIRQGAEiaLL	786
HIV2ROD	-----A-----C-G-RV-E-----	858
HIV2NIH2	-----L-T-----G-D-RA-Q-----L-----	856
HIV2ISY	-----L-V-----TS-G-S-----A-G-----	846
HIV2ST	-G-----R-----G-TN--GF-T-GQ-----	859
HIV2BEN	-----T-----GW-EAAR-----L-----	860
HIV2CAM2	-----A--A-----C-DV--M-Q-----L-----	859
HIV2D194	-----R---TA---I-----G-CEAAQ-----	851
HIV2GH1	-G-----RT-----S--GG-CAAVQ-V-----	851
HIV2CBL21	-----L-----TSV--NFC-TMGQ-----I-----L-----	865
HIV2CBL22	-----R--KF-----D--G-G--E-----	850
HIV2CBL23	-----G-----T-D--A-Q-----	858
HIV2CBL24	-----L--LT-----N--A-Q-----L-----	866
HIV2CAM1	-----G--AK---TS-C-S--RT-G-----L-----	858
HIV2CAM2BR	-----A--A-----C-D--EV-Q-----V-----L-----	843
HIV2CAM3	-RQ-----TS-C-G--RT-DNF---SI-----	860
HIV2CAM4	-----TA-----N-EANG-L-----L-----	860
HIV2CAM5	-----TS-G-S--T-C-----L-----	849
HIV2CAM6	-----FA-----S-C-S--T-C-----I-----	862
HIV2ALJ	-----IA---TNT--D--AMQWV--R-----	862
HIV2GB1222	-----A-----C-G--T-G-----T-----	347
CONSENSUS.B	GisWFQEA??AAaRa??ETLASag?cIwaaLRRAAgaiIAIPRRIRQSiELaLL	783
HIV2UC1	-----IQ--T--AR---NT-RA--K---T-E-----	857
HIV2D205	-----LR---TG-----E-----R-----P-----	859
HIV2GH2	--D-----LR---VTR-----E-F--I--M--V-----R-----	461
HIV2EMOA	-----IQ---AG-----AR-S-GV-----E-----A-----	852
CONSENSUS.D	GwsyfqEAVqa?wrsaeLETLagAwgdIWetLgrvGRwIlaIpRrIRQGLELcLL	840
SIVMM251	-----H---G-----R-----R-G-----	880
SIVMM32H	-----VG-----R-G-----	878
SIVMM1A11	--N--H---VD-----A-R-G-----	876
SIVMM142	-----AQ-----E--A-Q-G-----	881
SIVMM239	-----H---V-----R-G-----	879
SIVMNE	-----VA-----	881
SIVSMNH4	--R-----W-KL-R---S---I-----G-----	885
SIVSMNH9	--R-L---AXXW-KFXR---S--R-----G---XX-----	886
SIVSMNPBJ	--R-L---A--W-KFVR---S--R-----G-----	889
SIVP209C19	-----W-KF-R---S--R-I-----G-----V-----A---	886
SIVSTM	-CIWI---A--A--A-G---S-GR-----R-G-----	882

		<- env cds	
CONSENSUS.A	MGASGSKKrs?ps?gl?ERLLrargg?cgg?nasgGeyasqgEgSgreQkapsCEGg?YqGGdEmNtPwrtP		66
HIV2ROD	---H-R-PR-Q---A-A---YW-E---RF---D---RQ---KD		73
HIV2NIH2	---K-LQ-Q---Q---ET---RC-E---G-L-H---N---R---V---		73
HIV2ISY	---R-R-Q---A---LNDE-E-G---FH---L---R---		73
HIV2ST	---E-R-R---QTP-EAS-HMDKL---L---G---RR---A---		73
HIV2BEN	---L-KH-R-R---DGY-QGRD---F-E---N---Q---EY-S---N-		73
HIV2CAN2	---R-LQ-Q---A-T---ECY---LE---SLR---D---N-L---R---		73
HIV2D194	---EH-Q-R---GYVKQR---S---Q---E---		73
HIV2GHI	---H-KH-QR-R---H-GYVQC---KG---Q-R---		73
CONSENSUS.B	MGSAOSKKqSkqQ?GLRERLLR??e?PyG?lSgerreqSarsPGgSDKdLNSPSCeQqna?gAE.....		58
HIV2UC1	---R---TQ-E---K---EQ-Q-K---R-PR---		64
HIV2D205	---R-ER-Q---K---VP-R---R---E---		64
HIV2EHOA	---P---ARRG-R-ES---Q-R-LQY---G---D---KTL---		64
CONSENSUS.D	MGgaISkrrsk?PgdlrqrLLrARGEtYgrLl?evadgyssqSiggldKGLsa1SCEgQKynqgq/MNTPWzNP		70
SIVMM251	---M---PA---K---G---S---G---R---		73
SIVMM11	---M---PT---K---G---S---G---R---		73
SIVMM32H	---R---SA---G---		73
SIVMM1A11	---T-M---RST---E---G---S---E---E---		73
SIVMM142	---K---PPR---N---FKG---S---E---		73
SIVMM239	---M---KPS---G---P---		73
SIVMNE	---T---SP---WE-L---S---S-L---F---K---		73
SIVSMH4	---KQY-RG-N-E-Q---WEGL-E---ASG---P---SE---		73
SIVSMH9	---VT---KQRXAG-N-xE-Q---WEGL-GE---xDASG---P---CE-F---		73
SIVSMHPEJ	---VT---KQRRRG-N-YE-Q---WEGL-GE---QDASG---P---CE-F---		73
SIVSTM	---ASC---KQR-QH-E-E---K---EGLGE-SGP-Q-AS---N-H---P-R-E-F---K---		73
		LTR ->	
CONSENSUS.A	AaegeK?ly?QONMDDvDedDDdlvgvpv.tprVPlR?mTyklAvDMSHfIKekgGLeGmFYS?rRhrilDiy		134
HIV2ROD	---R-N-R---Q-R-S---K---P-HR-I---L-TR---E-K-N---		145
HIV2NIH2	---R-E-K---L---Q-F---P-F---I---L-Q---L---		145
HIV2ISY	---T-K-ES-R---SD-S---A-RM---DL-D---Y-E---		145
HIV2ST	---G-S-K---E-R-R---L---LY-D-R-V---		144
HIV2BEN	---T-RQ-D-R---I---R-E---I---Q---R---		145
HIV2CAN2	---K-NA-R---I---A---T---L-E---		145
HIV2D194	---I-Q-NS-K---M---E---I---I---RE---LF---		145
HIV2GHI	---I-Q-K-K---I-S---A---R-D---RD---L---		144
CONSENSUS.B	GGGqQD?DesD.EDnE.VGv?v?rP?VPLRpMTyKlAvDMSHFIKekGgLEGiyYSERRHrILDY		118
HIV2UC1	---T-D---Y---NR---S---M-I---		127
HIV2D205	---A---D---AICKT-I---Q---M---		128
HIV2EHOA	---K-S-D---R---Q---F---L-E-F---K---		127
CONSENSUS.D	AaEreKLAYRkQNmDDiDeeDddLvGv?V.7prVPlR?mTyKLAIdMSHFIKekGgLEGiYYsERRHrILDmY		139
SIVMM251	---K---S-R-K---A---		145
SIVMM11	---K---S-R-K---A---		145
SIVMM32H	---V---P-M---T-S---V---		145
SIVMM1A11	---S-R-K---T-S---		145
SIVMM142	---K-P---I---IP-EA---T-S---I---		145
SIVMM239	---S---S-R-K---T-S---		144
SIVMNE	---G---I---N---P-R---IIS---V---E-K---		145
SIVSMH4	---T-A-G-Q---V-D---I-S-H---A---NE---		145
SIVSMH9	---T-GA-G-Q---V-N---CP-S---V-I---V-I-K-X-		143
SIVSMHPEJ	---T-A-D-Q---V-SA---CP-S---V-I---D-K-L---		143
SIVSTM	---A-SA-E-Q---V-N---A-H---E---L---S---E---		145

HIV2/SIV NEF

--max. HIV-1 nef similarity--		
CONSENSUS.A	LeKEEGIIpDWNQNYThGpGvRYPrfFGWLWKLVPvDvsQEgeD?E? ???t?cLVHpaQtSr?DD?HGETLvw?	198
HIV2ROD	-----A-----P-----T-----H-----KF--P-----E	214
HIV2NIH2	-D-----L-----T-P-----T-----L-M-S-V-F-P-----K	148
HIV2ISY	-----G-----K-----P--D-S-----H-----F-P-----R	214
HIV2ST	-----Y-----S-EL--A-ED-----AN-----H-E-----Q	213
HIV2BEN	-----A-----S-----T-TDTE-H-L-----H-M-----K	214
HIV2CAM2	-----T-----Y-----I--A-EV-----N-----Y-E-----R	212
HIV2D194	-----C-----A--D-----NY-T-----H-E-----L-R	214
HIV2GH1	-----C-----A--D-----NY-T-----H-E-----L-R	213
CONSENSUS.B	LENEEGIVSGWQNYThGPGiRYPK?FGWLWKLVPv??pae?r?eE...?ThCLvHPAQiSsWDDiHgETL?Wq	181
HIV2UC1	-----Y-----RT-----DI-E-E-GA-----S-----A-R	196
HIV2D205	F-----Y-----SV--AT-E-----E--M-----I--	198
HIV2EH0A	-----V-----F-----IMNI--PEO-----T-A--P-E--V--	196
CONSENSUS.D	LEkeeGIIpDWNQ?ytsGPGiRYPKtFGWLWKLVPvvsDEAQEde...?hyLmhPAQtsqWDDPwGEVlaWk	206
SIVMM251	-----D-----R-----Q-----K-----	214
SIVMM11	-----D-----R-----Q-----K-----	211
SIVMM32H	-----D-----R-----Q-----K-----	214
SIVMM1A11	-----D-----E-----P-----	214
SIVMM142	-----IH-----L-M-----I-----E--V-----	213
SIVMM239	-----D-----E-----	213
SIVMNE	-----M--N-----P-----G-----EN--L-----V--	214
SIVSMH4	-----N-----MHY-----D-----T-C-V-----Y-----	214
SIVSMH9	--x-----N--A-----MF-----x-----T-C-----	212
SIVSMMPBJ	-----N--A-----MF-----T-----	212
SIVSTM	-----V--N--A-----Q-----DM-N-----D...GT--V-----H-----V-----	215
CONSENSUS.A	FDpmlAy?Y?afilyPEEFghksGSlpeKewKAKLKARGIPfSE	239
HIV2ROD	---L---S-E---R-----E---R-----	256
HIV2NIH2	-----T-----N-----R-----	180
HIV2ISY	---HE-TT-----NE-DD-----	256
HIV2ST	---T---FS-E---R-----Y-----D---R-----	255
HIV2BEN	--S---N-K--T-----Y---	257
HIV2CAM2	--S---LK-E--TR-----D-----	260
HIV2D194	---S---R---H-----Y---	257
HIV2GH1	---T---D-K---H-----Y---	255
CONSENSUS.B	FDsLLAyDYVAPnRfPEEFgyQSGlPEKewKA?L?ARGIPT?	220
HIV2UC1	--P---H-----Y-----	225
HIV2D205	-----E---R-K-----D	240
HIV2EH0A	-----S-----M-----K-R-----E	238
CONSENSUS.D	FDftLaytYeAyvryPEEFgskglseesevrrRLtarGSLkmadkketsE	256
SIVMM251	-----A-----LEA-QACQ-K--EEG	247
SIVMM11	-----N---R--R	260
SIVMM32H	-----P-----N-----R	263
SIVMM1A11	-----N-----R	263
SIVMM142	-----I-----K--K--A--E--R---	262
SIVMM239	-----N-----R	262
SIVMNE	-----I-----R-----	263
SIVSMH4	---E---S-K-FIK-----K-----WLTRRKQA-	264
SIVSMH9	---x---x-K-F-EH-----Q---K---Q-----K---	261
SIVSMMPBJ	---K---N-K-F-EH-----Q---K---Q-----K---	261
SIVSTM	---L--H---P--H-----PK--E-----	264

VERVET.COM	MGaatSA.LnrRqLDqFEhiALRPnGKKKYQIKHliWAGKeMeRFGLEH?LEtEEGCKKiIEVlyPLEPTGS	71
SIVAGM3	-----K-----T-----R-----S-----	72
SIVAGMTYO	-----K-----R-----R-----	72
SIVAGM692	--SGS-- --SG-K-----L-----K-----L-----	72
SIVAGM155	-----E-----K-O-----K-----S-----	72
GRIVET.COM	MGOGHSA.LSGRSLDTFEKIRLRPNgKKKYQIKHliWAGKeMERFGLHEKLLetKEGQCKiIEVLTPLEPTGS	72
SIVAGM677	-----	72
TANTALUS.COM	MGaGHSa.LsgRNLDTFEKIRLrPMGKKKYQIKHliWAGKeMERFGLHdKLLetKEGQCKiIEVLaPLEPTGS	72
SIVAGM9	--V-----A-----R--K-----R-----E-----T--S---	72
SIVAGM17	-----H-----X-----E-----	72
SIVAGM27	--S--T-----R-----	72
SIVAGM40	-----P-----	72
SIVAGM49	-----T-----R-----	72
SABAEUS.COM	MGASNSV.LSGRKLDafESVRLRPNGKKKYKLRHLVWASKELDRFSLsANLLETKEGVVKILSVLLPLVPTGS	72
SIVAGMSAB	-----	72
SYKES.COM	MGAAGSAILTGRELDRYEKIRLRPNgKKRYLVRHLVWAKKELDRFGLSDQLMESKEGCEKILTIVLLPLEANGS	73
SIVSYK	-----	73
VERVET.COM	BGLKSLfNLvCVi?CvH?ekkVKDTEEEava?vrQcCHLV7Kek?A?.....????etssGq?K..nd	120
SIVAGM3	-----F--KD-E-----I-----E--RN-E.....RNTT--K--	132
SIVAGMTYO	-----Y-L-K-Q-----T--H--E--S-----T-----K--	128
SIVAGM692	-----T--IC-I-Q-A-----IRIK-Q--D-OEN-----AK-ID----	123
SIVAGM155	--M--Y-----L--Q-----L-I-----D--T-----VTPPG--Q--N	128
SIV2010G	-----T-----K--S-----A-----RQ----	28
GRIVET.COM	BGLKALFNLCCVINCIHAEQKVKDTEBAVVTVKQHYHLVDKNEKA.....AKKK...N	122
SIVAGM677	-----	122
TANTALUS.COM	BGLKSLfNLccVINCIHAEQKVKDTEBAV?iVhQhchLVEKEKta.....a?PsgGQkQ..ny	126
SIVAGM9	-----VY-----A--OS-----A--Q----	128
SIVAGM17	-----V-----T-----A--Q----	128
SIVAGM27	-----I-----A--Q----	128
SIVAGM40	-----T--V-----AT-R-----ST-----TE--S-----HS	128
SIVAGM49	-----Y-----V--R--A-----T-----	128
SABAEUS.COM	ENLIAlFNLCCVLACIHAEIKVKDTEAKAKVKEVPAENTESATATSSOOT.....KELQAKKK..NE	134
SIVAGMSAB	-----	134
SYKES.COM	ENLKSLFGIISVNAVHAKKEVEDTEQAKQKVKEACNWKDDPP..ATSGGQSENSSQMMASETSSGQKVVOQE	144
SIVSYK	-----	144
VERVET.COM	7g???ppG?SONfPAQDQ.gNaWVHVPLSPRTLNAWVKaveEKKFGASIVPMFQALSEGCTPYDINQMLNVLG	187
SIVAGM3	K-VTV--G-----I-----	204
SIVAGMTYO	K-IAR--C-----	200
SIVAGM692	.TTFT-S-R--Y-----N-V-----VI-----V-----	194
SIVAGM155	T-GTAT--G-----	200
SIV2010G	RSTATSS-Q-----T-----	100
GRIVET.COM	ETTAPFGGESRNYFVVNQ.NNAWVHQPLSPRTLNAWVKVEEKRWGAZVVPMEQALSEGCLSYDVNQMLNVIG	194
SIVAGM677	-----	194
TANTALUS.COM	n?aa?pgG?hGNYFVVQQ.NNQMVhLPISPRTLNAWVKtVEEKRFQASiVPMFQALSEGCLSYDINQMLNVIG	196
SIVAGM9	-L-----R-----Q-----V-----	200
SIVAGM17	-T-----N-----	200
SIVAGM27	-L-----RD-----Q-----YV-N-----C-----M-----YI-----V-----	200
SIVAGM40	RPD--SR-Q-----	200
SIVAGM49	-TT-----Q-----I-----	200
SABAEUS.COM	PTVTPSGGSr.NYPiVSV.NNQVHQPLSPRTLNAWVKVIEKKFSAZVVPMSALABGAIPYDINQMLNAVg	205
SIVAGMSAB	-----	205
SYKES.COM	KQKAATPPFRONYPLLRNPQNMiHTGVVVRTLTWVEAVNSKIFDASIVPLEQALTEGFIPYDLNGMLNAVg	217
SIVSYK	-----	217

AGM GAG

VERVET.COM	DHQQALQIVKEIINEEAAQNDI tHPpPAGPLPAGQLRDP tGSDIAGTTStvQEQLENIYTANPRvDVGAIYRR	260
SIVAGM3	-----A-----	277
SIVAGMTYO	-----V---L-----S-----	273
SIVAGM692	-----V-----I-----I-----	267
SIVAGM155	-----V-----G-----	273
SIV2010G	-----Q-----V-----	173
GRIVET.COM	DHQQALQILKEVINEEAAENDRTHAPPAGPLPAGQLRDP tGSDIAGTTSSIQEQIENTFNANPRIDVGAQYRK	267
SIVAGM677	-----	267
TANTALUS.COM	DHQQAMQIIKEVINEEAAQNDI tHPpPAGPLPAGQLRDP tGSDIAGTTStvQEQIEWTFNANPKVDVGRiYRg	269
SIVAGM9	-----N---A-----V-----I---T-----R-----	273
SIVAGM17	-----x-----S-----T-----	273
SIVAGM27	-----I-----SI-----S	273
SIVAGM40	-----R-----L---	273
SIVAGM49	-----I-----V-----	273
SABAEUS.COM	EHQQALQIVKDVINEEAAQNDI tHPpPAGPLPAGQLRDP tGSDIAGTTSTIQEQIEWTTAQNANVNVONIYRG	278
SIVAGMSAB	-----	278
SYKES.COM	DHQQAMQVIKDI tNEEAAQNDI tHPpPAGPLPAGQLRDP tGSDIAGTTSSIAEQIENWITRONNPVQVGEIYRR	290
SIVSYK	-----	290
VERVET.COM	WIIIGLQKCVKMYNPFvSLDI tQGPKepFKDYVDRFYkaIRAEQASgeVKQMMTtSLIIONANPDCKVILKGL	333
SIVAGM3	-----A-----	350
SIVAGMTYO	-----	346
SIVAGM692	-V-A-----R---TG-----R-S-S-----L-----QD--N---DT-----E-----	340
SIVAGM155	-----	346
SIV2010G	-----I---K-----RT-----	224
GRIVET.COM	WVILGLQKVVMYNPQKVLDIRQGPKepFKDYVDRFYKALRAEQAPQDVKNWMTQTLLIONANPDCKLILKGL	340
SIVAGM677	-----	340
TANTALUS.COM	WVILGLQKCVKMYNPFvSLDI tQGPKepFKDYVDRFYqALRAEQTPQDVKNWMTtTLLIqNANPDCKLVLKqL	342
SIVAGM9	-----	346
SIVAGM17	x---R-----I-----G-----R-----x-----H---R-----	346
SIVAGM27	-----A-----E-----E-----	346
SIVAGM40	-----I-----	346
SIVAGM49	-----T---Q-----	346
SABAEUS.COM	WIIIGLQKCVKMYNPFVNI tDIKQGPKEPFKYVDRFYKALRAEQTPAVKNWMTQSLIIONANPDCKTVLKGL	351
SIVAGMSAB	-----	351
SYKES.COM	WIIIGLQKCVQVYNPFVNI tDIKQGPKEIFKYVDRFYHCLRAEQADPSVKGWLTQTLIIONANPECRQILKAM	363
SIVSYK	-----	363
VERVET.COM	GMNPTLEEMLTACQGVGGF sYKAKVMAEMM... qnmQs q. NM. vQDg... G?rgRpRp?.. vkCYnCGKF	392
SIVAGM3	-----M-----Q-----	410
SIVAGMTYO	-----T--N-----PK--Q-----LR-----	405
SIVAGM692	-----Q---L-V-----NQ--GV-----A-----IG--G-GRP-----K-----	401
SIVAGM155	-----L-----G-----P-----	406
GRIVET.COM	GMNPTLEEMLTACQGVGGFPOHKAKLMVEMM... SN... GQ. NM. VQVG... PQKKGPRGP... LKCFNCGKF	398
SIVAGM677	-----	398
TANTALUS.COM	GMNPTLEEMLTACQGVGGFPOHKAKLMaEaM... QnmQGV. nm. vqg?... prgGRGRGRGP. pRCfKCGQI	403
SIVAGM9	-----I. T---AQG-----T-----	407
SIVAGM17	-----I-----Q-----I--H.... SK-----	408
SIVAGM27	-----V-----I-----I--P-----	408
SIVAGM40	-----I. T---AQG-----S---V---R---	407
SIVAGM49	--N-----V-V---Q---P.... S-----	408
SABAEUS.COM	GMNPTLEEMLTACQGI GGAQKARLMAEAMTAAPQQTvGNIFVQQGARPRGPLGGRGPLWPNIKCYNCGKP	424
SIVAGMSAB	-----	424
SYKES.COM	. VKPTLEEMLTACQGVGGFPLHKAKLMaEAM... VMAQSV. NM. VQGF... SKGRSM... IKCYNCGQI	420
SIVSYK	-----	420

VERVET.COM	GHMQRQCpgrk?kCLKCGKpGHlAkDCR.....GQVNFLGYG.NWMGAKPrNFp...aatlgaePsa	450
SIVAGM3	-----MR-----	469
SIVAGMTYO	-----T-----L-----	464
SIVAGM692	--V-KN-TQKGPVV-----F-R-----A-----K-----LEHRA-V---	459
SIVAGM155	-----I-----	465
GRIVET.COM	GHMQRDECKAPRQIKCFKCGKIGHMAKDCN.....GQANFLGYG.HWGGAKPRNFVQYRGDTVGLEPTA	461
SIVAGM677	-----	461
TANTALUS.COM	GHlQKDCprgGPiKCLKCGKpGHMAKDCRs.....GQANFLGpRMpEPwGSKPRNF...LEQgGaVPTA	463
SIVAGM9	-----N-----L-R---T-----D-----	466
SIVAGM17	--V-----R-----C-----R---A-X-----	468
SIVAGM27	-----Q-A-T-----R-----	468
SIVAGM40	-----K-----	466
SIVAGM49	--V-----	468
SABAEUS.COM	GHlARFCKAPRRQGCWKCSPDHQMKDCQK.....QVNFLGFG.PWGRGKPRNF...PLTSIRPTA	481
SIVAGMSAB	-----	481
SYKES.COM	GHMQRDECKKPLKAKCFNCGKTGHlARACRQPKRNQGPVQAQANFLQKGNQVSRPPANF.....PVRSEPSA	487
SIVSYK	-----	487
VERVET.COM	PPPPs.....n?tpYDpAkkLLqgYaekGkQlReq?r?.ppa?npdwEgYSLnSLFGEDQ	501
SIVAGM3	-----D-----RKK---V-----	521
SIVAGMTYO	-----GT-----K-N---M-----	519
SIVAGM692	---H-----PGA--E-TR--EK-TQE-A-Q-RKVEK..SSQAGREE-D---K-----	513
SIVAGM155	---N-----S-----M-N-N-N---N---N-----	520
GRIVET.COM	PPMET.....AYDPAKLLQQYAEKGQRLREEREQ.TRQKEKEVEDVSLSSLFGQDQ	511
SIVAGM677	-----	513
TANTALUS.COM	pp?paH.....gfptGpp?a?GAYDPAKRLLeQYAKKQdQLrkQKE.....KELEDYSISSLFGEDQ	517
SIVAGM9	--L-----ETSV--T.-R-----	522
SIVAGM17	--IA-----VA-----G-----R-----	518
SIVAGM27	KKM---V---L---A---R---E---K-----	524
SIVAGM40	--M-T-----Q-----R-----	522
SIVAGM49	-RL-T-----V---SOWR-----T---E-----	525
SABAEUS.COM	PPMERDYSRPEEMMYADRPPTRGPGPDOPATALLQYAVQGRKQKQWQNHSPQQSPLYEBATSLRSLFGEDQ	554
SIVAGMSAB	-----	554
SYKES.COM	PPLED....IEDGFNLTWSAQMSQQAQAKQNSPSKKPPTNREVLSPKE..SSGKEETKSLYPSLSSLFGEDQ	554
SIVSYK	-----	554

AGM POL

VERVET.COM	LMKIWPYAKTMSGTK??M??mWk?ra?SKRLQ.gTGgFFRVW	35
SIVAGM3	--KDE-LE--ET--F-----R--K-----	29
SIVAGMTYO	-----ENK-SKV--IGTF-----	42
SIVAGM155	---A--L-----	21
GRIVET.COM	VRQNNPYG.....KRLQEWTKGFFRVW	32
SIVAGM677	-----	22
TANTALUS.COM	SNAKHARDHNGSRTRKRRQKRTYPLQMN?????RL???RTn?vsQmWktRthgKrLQ?r?G?FFreN	58
SIVAGM9	-----TNRPHSE--PQG--KQ-----LS-G--D-A-K--K-	72
SIVAGM17	-----R-T-Q--K--	10
SIVAGM27	--PDWTYTK--SKS--E-P-V-E--A-----KWT-Q----	43
SIVAGM40	R---KM-----A-----V-A-K-----	30
SIVAGM49	---K-	5
SABAEUS.COM	FFRVW	5
SIVSAB1C	----	5
SYKES.COM	FFRER	5
SIVSYK	----	5
VERVET.COM	.pVDG?kT?KF...SRrySWgGt?.CApSTE???.....?iRPaKEAPAAiCRerEt?Eg..	77
SIVAGM3	---SE-Q-----AN.....S---C-----QG-AV----	76
SIVAGMTYO	.T---G--E---S--E--\$--RHN.....P-----T...	92
SIVAGM155	---D--K---SH---K---QLY.....TL-----V-----N-K..	71
GRIVET.COM	.PLGR\$ETKKF.CAIQRRHNSGTM.SPpMGN.....SLRSSKEAPPVCREGTAPER..	72
SIVAGM677	-----	72
TANTALUS.COM	a?tmgfqaAQF.FRARRSRSDGSp?a?AriphRS?s??gs.....lRpSQe?pgtvckkgy.....	106
SIVAGM9	D---IE-----I-C--DLR--H--T-----I---TTV-A-----	125
SIVAGM17	-R-LR-----DPR-W---A--CR-----S---T--A---RH.....	64
SIVAGM27	-NP---T-----T---KD-R-----A--C-----A--I-----	96
SIVAGM40	.D---IE-----TH-N-----S--T-----I---II---EE-----	83
SIVAGM49	-N-----I-N-WGS--VRLA-----A-R--HEER-----	60
SABAEUS.COM	.PLQRETOQEF.PSDLHQTNSSPwGTGLQQAGGKLVCRTSDQRTARRSSNSPVKAVCCSGETAETAVAKP	75
SIVSAB1C	-----	75
SYKES.COM	MOGLQEAPSQPLPSEERAICSPSGGHRMAMAH....MVSTDEP.....TSTSE.GTELSLEETPMQOGSA	65
SIVSYK	-----	65
VERVET.COM	?kek?ts?esrLdRGI.FFELPLWRRPIkTvyIeGvpikALLDTGADDTIiKe?DLQLsG...?NrPKIiGGI	141
SIVAGM3	T---T--S-----R-----A-----T-K-----	145
SIVAGMTYO	A--ES-GN--G-----N-----P-----	161
SIVAGM155	SEQ-PP-EQ---E---R-CI-G-TAV-----DT---R--S---V---	140
GRIVET.COM	GERTDKETEGERSGGC.FLELPLWRRPMXRVIIEGTFVQALLDTGADDTIiQEKDLHFPPHKPWRSKVVOGI	143
SIVAGM677	-----	143
TANTALUS.COM	.ptkeaegeg...tGGLFIEFPLWRRPI	130
SIVAGM9	.-AE-----	148
SIVAGM17	-----KR-----	81
SIVAGM27	.SV-----KR..A-----I-----	120
SIVAGM40	---EKT-R---I-----	106
SIVAGM49	.T---T-----	83
SABAEUS.COM	LAT.TEPLRGGLQ...LPQVSLWRRPMKTVYIEGQKVTALEDTGADDSVIQG?ELGDN...WKPRIIGGI	137
SIVSAB1C	-----I-----	138
SIVSABD37	-----L-----	20
SYKES.COM	LAEGEQ.WEGGDQESLSLSQLSLWRRPMYEVQVEGDLVQMLVDTGADDTIiREEDIQLHQP..WSEKIVGGL	134
SIVSYK	-----	134

VERVET.COM	GGGLNVKEY?drEV?1EDKILR.....GTiLiGaTPINTIGRN?LApAgakLVMSQLSe?IPiTPVvKlKEGA	204
SIVAGM3	-----S-----R-----S-----I-----Q-----	212
SIVAGMTYO	-----N-----KI-----L-----L-----VPR-----K-----V-----	228
SIVAGM155	-----DNV-----Q-----V-----F-----Q-----QT-----R-----	207
GRIVET.COM	GGGIHVKEYQGVQVQLEDKIIT.....QSILIGSTPINTIGRNILAQAGMKLVMSQLSSQIETKVLKEGK	210
SIVAGM677	-----	210
SABAEUS.COM	GGCINV?AYHMQEV?IE?K?CK.....ATIIVG?TPVMIIGRMVLAQLGVTLNL????I?..PIKV?LKPG?	190
SIVSAB1C	-----K-----K-----D-----T-----E-----TORE-E-----H-----Q-----	203
SIVSABD37	-----R-----E-----G-----G-----D-----VART-L-----Q-----K-----	85
SYKES.COM	GGNITVRQYRNIRFPVVKPSGKRXQVBCITLLVGPTPVNILGRNILTGLSVKLVMSQ..TALEPVKVSCLKPKK	204
SIVSYK	-----	204
VERVET.COM	RGP?1kQWPLSKEKI?ALQEIC?QLE?BGKiSivGGeNAYNTPvFCIKKKDKSQWRMLVDFRELNKATQDFF	272
SIVAGM3	---P-----D-----K-----KI-----	284
SIVAGMTYO	---CVR-----E-----S-----Q-----I-----	300
SIVAGM155	---R-----I-----KT-----E-----L-----D-----R-----	279
GRIVET.COM	DGPKLKQWPLSREKIEALTEICKQHEEBGKLSRIGGENFYNTPVFAIKKKDKTQWRMLVDFRELNKATQDFF	282
SIVAGM677	-----	282
SABAEUS.COM	DQPR?RQWPLSKEKIEAL?PICE?LEK?GHLEIRIGPENFYNTPVFAIRKKDK?QWR?LMDPR?LNK?TQDFFQ	253
SIVSAB1C	---I-----KA-----D-----Q-----T-----I-----Q-----S-----	275
SIVSABD37	---L-----RT-----Q-----E-----S-----M-----E-----A-----	157
SYKES.COM	ELPR1KQWPLSVEKLEALKAIVEDMLKAGQLEKASPTNPNYTPVTVIRKKDKKKWRMLIDFRKLINEATQDFF	276
SIVSYK	-----	276
VERVET.COM	EVQLGIPHPaG1?KM?gITv1DvQDAYYSIPLDPeFRKYTAFTIPtVMNqGPGIRYQFNCLPQGWKGSPTIF	342
SIVAGM3	-----S-PE--TE-----I-----S-----T-----	356
SIVAGMTYO	-----R-----R-----N-----	372
SIVAGM155	-----K-----K-----II-----E-----	351
GRIVET.COM	EVQLGIPHAPAGLQKKKQITVIDIGDAYYSIPLCKEPRKYTAFTIPSVNMTGPGIRYQFNCLPQGWKGSPTIF	354
SIVAGM677	-----	354
SABAEUS.COM	EVQLGIPHAPAGLQOREQITVLDI?DAYFSCPLD?DFQKYTAFTIPSVNN?EPG?R?QYKVLPGWKGSPTIF	320
SIVSAB1C	-----G-----P-----R-----I-----Y-----	347
SIVSABD37	-----A-----E-----K-----M-----F-----	217
SYKES.COM	EVQLGIPHPPGL.KRQKLTIIDLKDAYYSVPLDKERFPYTAFTVPSINNAPGERYQFTVLPQGWKGSPTIF	347
SIVSYK	-----	347
VERVET.COM	QNTAasILEEIKKkLk?LTIVQYMDLWVGSQE?e?LHD?LVeqLR?kLqawGLETPeKKvQkePpyENMGY	409
SIVAGM3	-----P-----D-----Y-----R-----M-----S-----D-----K-----	428
SIVAGMTYO	-----RN-PA-----N-H-----K-----T-----M-----	444
SIVAGM155	---SK-----Q-----SGPK--Q--QT--NR--E-----R--F-----	423
GRIVET.COM	QNTAANILEEIKRHTPQLEIVQYMDOLWLASDHDDETRHNQQVDIVRKNLLEKGLSTPDKKVQREPPWENMGY	426
SIVAGM677	-----	426
SABAEUS.COM	QTTANKILQEFQKMPDQVDIYQYMDOMLIASDRPKABHLVMSQQLRDYLETWGPKTPKKPKQDPPYLMWGY	392
SIVSAB1C	-----	419
SYKES.COM	QSTINQILOPPPKKYSDLTLIQYMDOLLIGTRSEKAHQEIVQOIIVTALLKVGFKVPKPKWQDQYFMQWLGY	419
SIVSYK	-----	419
VERVET.COM	KLWPHKW?1s1eLE?.KeeMTVNDiQKLVGKLNMAAQLYPGLrTknICKLir?GKKNLLe?V?WtpEABAE	475
SIVAGM3	---QI-----D-----R-----L-----T-----T-----E-----	498
SIVAGMTYO	---E--R-Q--E--D-----K-R-----TG-----L-A-----	514
SIVAGM155	---K-Q---K--Q---L-----L-----DV--E-----	493
GRIVET.COM	KLHPNKWTINKIELPPLBGEWTVNKKIQKVGVLMNASQIYPGIKTKHTCAML.R.GRKNLLEEIVWTEEAABE	497
SIVAGM677	-----	497
SABAEUS.COM	ELYPKKWQLQEITLPE.REEMTVNDIQKLVGKLNMAAQIYTGIKTKHLRLIR.GARPLTEIVQWTEEAABE	462
SIVSAB1C	-----	489
SYKES.COM	TLHPDKWQLQKIELPHIDDEITVNLQKLVGKLNMAAQIYSGIKTKELCKCIR.GTKQLTEVLTLEAAABE	490
SIVSYK	-----	490

AGM POL

VERVET.COM	YaEnkEILKTEQEGTYVYkPg?PiRAAVQKlegGQMSYQFKQB..GqVLKVGkyCK...QKnTNTNEIRVLWG	541
SIVAGM3	-----R-----	565
SIVAGMTYO	-----A-----I-----T---	581
SIVAGM155	-E-----A-EK-L-----GD-----KI-----FA-----A-----	560
GRIVET.COM	YKMNQGIQVQETQEGTYVYDPLKELIATVQKQCEQWTVQPTQE..GAVLKVGRYAK...QRETHNDLRTLAH	564
SIVAGM677	-----	564
SABAEUS.COM	LEENRQILRQKQQQYYPALPLRAKVLKLGQGWGYQIYQPE..NKILKVGKYAK...IKTAHTNELRMLAG	530
SIVSAB1C	-----	557
SYKES.COM	LEENRQILKEEQAGSYYPKKPLEAHITKLGSSQGWGYMIKQEQKGPPLITGRTAK...TFAAHSNDYQSLAQ	559
SIVSYK	-----	559
VERVET.COM	LVQKICKEALVING?LPVIELPIERavWEQNWADYQVSWIPEWdFVSePPLVKLWYTLTKEPIPKEDVYVY	612
SIVAGM3	---L---S---E---D-E---	637
SIVAGMTYO	-----I-----L-----	653
SIVAGM155	V---G---Q---TF---V---DT---V---P---G---	632
GRIVET.COM	LVQKICKEALTINGRLPRVQLPVQDKKTDMWQDYQVSWIPEWdFVSTPLLVKLWYSLVKEPIRGEDVYVY	636
SIVAGM677	-----	636
SABAEUS.COM	LVQKIGKE??VING??P?ELPVERE?WEQNWadyqvwswIPEW?fvStPpI??LWYcLcKePipgEdVYVY	592
SIVSAB1C	-----SI---QI-IM---L---S---T---EM---Q-IR---K-V-D---A---	629
SIVAGM266	-----G---L---T---V---VT---	40
SIVAGM385	--AL---EL-TL---V---D---WLK---LAK---	66
SYKES.COM	LLNKIGIQSLWYWGKVPFHLPVKREBEKMTDYQWATWVEVKFISTPPLVRMYNVLVPEPIPEAVTFYV	631
SIVSYK	-----	631
VERVET.COM	QGACNRnSkEGKAGYfLQYgKQrVe?LENTTNQQAeLTAIRMALEDSGPnVNIIVTDSQYAMGILTAQPTQSD	683
SIVAGM3	-----R-----K-----M-----	709
SIVAGMTYO	-----S-----T-----K-----	724
SIVAGM155	-----Q-----Q-----Q-----K-----	704
GRIVET.COM	DGAASKVTILGKAGYLSERGRSRIRELENTTNQQAELTAVKMALEDSGENVNIIVTDSQYVBNILTACPQESN	708
SIVAGM677	-----	708
SABAEUS.COM	DGAacNRNSrEGKAGYiTq?GkQrVekLENTTNQQAEL?AI??AL?DSG??VNI?TDSQYAMGIAGEPTESD	656
SIVSAB1C	---A---K---L-DR-D-K-VA---K---E-LL-R---Sk---I-----	701
SIVAGM266	-----Q-----T---KM---E---PR---V-----	100
SIVAGM385	-----Y-----	100
SYKES.COM	DGAANROSKTGNAGYVASDGTQRVQYLEQT?TNQQAELGLLNALQDSKDKVNIVVDSQYSYGILNTPNTTE	703
SIVSYK	-----	703
VERVET.COM	SPLVEQIIAlmVQK?qIYLQWVPahKGIGGNEEIDKLVSkg?RrILFlekIEEAQEEeHeRYHnMWZNLADTY	753
SIVAGM3	---i-----H-----D-----Q-M-K-----	781
SIVAGMTYO	-----I---Q-----I---V-----K-----K-----	796
SIVAGM155	-----Q---EA-----V---IGR---D---S-----P	776
GRIVET.COM	SPLVEQIIQALMKKRQVYLQWVPahKGIGGNTPEIDKLVSkgIRQILFLDRIEEAQDDHAKYHNMWRSMVQEF	780
SIVAGM677	-----	780
SABAEUS.COM	NNIVQQIIIEELIKKEAVYIAWVPahKGVGCGNEEIDKLVSQGIROVLFLDRIEEAQEBHDKYHAMWRSMQOEF	728
SIVSAB1C	-----	773
SYKES.COM	HPIVEQIIQEAIKKEAIYVTWVPahKGIGGNEAVDKLVSkgIRKILFLERIPQAQEDHERYHSMMEYLQOEF	775
SIVSYK	-----	775

VERVET.COM	GLPQIVAKEIVANCPKCQIKGEPVHGQVDASPGVWQMDCTHIEgKvVIVAVHVASGFIEAEVIFRETGKETA	825
SIVAGM3	-----I-----	853
SIVAGMTYO	-----T-----K-----	858
SIVAGM155	-----V-----I-----I-----I-----	848
GRIVET.COM	GLPNIVAKEIVAACPCKQIRGEPRKHQVVDASIEBQWQMDCTHLEGKVIIVAVHVASGFIEAEVIFRETGKETA	852
SIVAGM677	-----	852
SABAEUS.COM	GLPAIVAKEIVAACPCKQIKGESVNGQVDASPGVWQMDCTHLEGKIIIVAVHVASGFIEAEVIFRETGKETA	900
SIVSAB1C	-----	845
SIVAGM381	-----	13
SIVAGM384	-----	13
SIVAGM386	-----	13
SIVAGM385A	-----	13
SIVAGM385B	-----	13
SIVAGM385C	-----	13
SIVAGM385D	-----	13
SYKES.COM	HLPRQVAKAIIGQCPCQNRGEPKHGQVVDIYNWQMDCTHEEGKVICAVNTASGYIETKILKRETGDETA	847
SIVSYK	-----	847
VERVET.COM	KFLKXILSRWPITqLHTDNGPNFTSQEVAACWNGKIEHTTGVPYNPQSQGSIESMNLKQKEIIGKIRDDCQ	897
SIVAGM3	-----I-----I-----	925
SIVAGMTYO	-----IG-----K-----V-----	940
SIVAGM155	-----	920
GRIVET.COM	HFLKLLARWPVKHLHTDNGPNFTSQNVAAVCWNGNIEHTTGIPYNPQSQGSVESMNLKQKEIISQIRDDCE	924
SIVAGM677	-----	924
SABAEUS.COM	hFLKLASAWPITqLHTDNGPNFTSQQVAAICWNGKIEHT?GVFPYNPQSQGSVESMNLKQKEIIGQIRDDAE	871
SIVSAB1C	-----Q-----F-----	917
SIVAGM381	Q-----A-----P-----F-----E-----	85
SIVAGM384	-----Q-----F-----	85
SIVAGM386	-----H-----F-----	85
SIVAGM385A	-----Y-----E-----Y-----	84
SIVAGM385B	-----Y-----	85
SIVAGM385C	-----Y-----	85
SIVAGM385D	-----Y-----	85
SYKES.COM	LFLMQIASRWPIKQIHTDNGPNFVSDKFAACWNGGIEHTTGIPYNPQSQGIVESKNRYLKEAISQIRDDVT	919
SIVSYK	-----	919
VERVET.COM	YTECAVLMAChIHNFKRKGGIGGIT?AERLINmITTQLEIqhlQTKIQKILNFRVYYREGROPVWKGPaQLI	968
SIVAGM3	-----P-----L-----G-----	997
SIVAGMTYO	---A---IL---Q-S---I---	1012
SIVAGM155	-----A-----NT-----R-----	992
GRIVET.COM	RLETAVQMATHIHNFKRKGGIGGISSAERLVNMLTTQLELNTLQNDIQKILNFKVYYREGROPVWKGPaRLI	996
SIVAGM677	-----	996
SABAEUS.COM	RLETAVIMAVHIHNFKRKGGIGGYSAAERLINIHTLESTKTLQOKISKIQNFRVYYREGROPVWKGPaKLI	943
SIVSAB1C	-----I-----	989
SIVAGM381	-----	92
SIVAGM384	-----I-----	92
SIVAGM386	-----I-----	92
SIVAGM385A	-----	91
SIVAGM385B	-----	92
SIVAGM385C	-----	92
SIVAGM385D	-----	92
SYKES.COM	HLQTAVAMATFILNFKRKGIGGIGSPGERYINMLYTELQLO.QNTTSPKPSNFRVYYRQOKNE.WKGPALL	989
SIVSYK	-----	989

AGM POL

VERVET.CON	WKGE GA.VVlK7G?eLKvvPRRKAKI IKDYEPkr?Gde7?lEGagQsDnQMAGDS	1018
SIVAGM3	-----I-G-V---EY-----M---SN-----A--	1046
SIVAGMTVO	-----D-SD-----KQ-V-N-GDV--TR----	1061
SIVAGM155	-----E-E-----TL---TH-----H-----	1047
GRIVET.CON	WKGE GA.VVlKEGEDIKVVPRRKAKI IKDYGERKTMDSGSMGCVREANKQMEGDSDLQDQE	1057
SIVAGM677	-----	1057
SABAEUS.CON	WKGE GA.VVlQEQGELKTI PRRKAKI IKDYG..KALDSQAPLEGNNGRTAGEVD	993
SIVSAB1C	-----	1039
SYKES.CON	WKGE GAVVVQTEEGDIPAVPRRKAKI ITOHG..ERMDSGSHVENDEKTD	1036
SIVSYK	-----	1036

In this, and several of the AGM alignments that follow, only one consensus sequence was generated, as several of the species have only one representative. SIVSYK was not included in this consensus, although its similarity to the consensus sequence is shown.

CONSENSUS	H??eK?WvMvvtWkv?e?itldqgiVrYvwr??r?l?2W?Y?mHyqitWawytmsryvIP1???g?I?Vd?Ywh	56
SIVAGMTYO	-NPN-E-----PGDL-----Q-N-K,-N-Y-----IGKH-E-C--L--	73
SIVAGM155	-SQ--H---L---Q-EV-----N.K-N-K,-E-K-----PGS-E-H--I--	73
SIVAGM3	-NQ--E-----P-EL-----T-K-D,-K-R-----E---GQH-S-H--L--	73
SIVAGM677	-ER--Q-IV--V-R-S-RQ-SR-R--T-KI-.NKQ-P,-E-RH-W-VQ-QFW-Y-QFI---SKDDY-E-NI-HN	73
SIVSAB1C	'...-H-IV-PL--TGGQER-TSL-K-H-HVSKQCVH-R-TP-TK-R-N--SYQEW---KDGAL-K-TN--	73
SIVSYK	...-E-IVVP--RMTFRQ-DRL-H-IXTHRYKSKE-EKAT-KH---E-Q---YQWNT--VG.D-T-WITF-HN	72
CONSENSUS	LcPe?GWLstyavGiqY?s?????yrTEIDP?TADs?IH?hYF?CFtdRAiqqAlrG?rf?.fCqfpeGNk?tG	115
SIVAGMTYO	----Q-----V-NLESK,-----A---I--Q---N--KE-----H--V,-----S--	146
SIVAGM155	-A-KQ-----V-LVNDK,-----N---M--C---T-----N--I,-----Q---L--	146
SIVAGM3	----K-----E---L-NRDPW,-----A---L--T---T--E---RK--L-Q--T-----K--	146
SIVAGM677	----R---SHG--LS-YH..QRG,-K--V--G---RM--LY--N-----I--EKYT.W-T-K-----	142
SIVSAB1C	---K---E---T--G....SKGEWF-----W---HI--WS--P-----V---I--EKYL.W-RHQV--QP--	143
SIVSYK	-A--R---HMOGIR-----QMNQWN-D-T-AV--RL--NF--P---A--VN--V--ELLTSH-WT...-TD..	139
CONSENSUS	QVPSLQyLALLah.qNGIR?rs?r??,trr?rnlGs?Qgav???grMA?r???r?q??sq?Afw?rtp?psmeil	169
SIVAGMTYO	-----E--K-GK,---S---K-----Q--K-YVT-S-PGGEA---E---V-----	216
SIVAGM155	-----K--Q-GE---T---Q-----Q--YGR-M-QR--T---P---I-----	216
SIVAGM3	-----Q--Q-SK,-GQT--M-FE-----K-MAR-Y-SG--D---A-A-V-----	216
SIVAGM677	--Q---L---V-Y,T--I-K--K-TP,--MAG-----R---M,-----T-HAQGSKRR--K-L-NEHAN-----	212
SIVSAB1C	-----RVY,T---RVAPT...S--GSSQ--P-ESQRDT---RNMGFAQRAVRM-PRHV-GPQFRGPV	214
SIVSYK	-----QVYLKD-GGFLQSLPACA-NTNV-H-KKCR-DPKRDQCHCKGRTGSDR-IQ--YSSRNIV-L-SI	214
CONSENSUS	sgg???t??sh?gkg1???H	180
SIVAGMTYO	---RRK-MY--D---QIL	235
SIVAGM155	---RGE-GKT-S---I	232
SIVAGM3	---GRKESH--AR---	232
SIVAGM677	CR-GKE-	219
SIVSAB1C	PLPKESPFF-LVEVCGRTS-	234
SIVSYK	LKRRGRD	221

AGM VPX

CONSENSUS	MASG????????????rdPRE?rPG?iEiNDLsREPWDeWLrDWI?dIN?EA??MFgreiLfqvWnyccqEE?	53
SIVAGMTYO	-----A---EV-----K-----Q---Q---RL-----	59
SIVAGM155	-----E---G-----VEEI-N---KL-----Y-----	59
SIVAGM3	-----A---E-----E-I-Q---KM-----	59
SIVAGM677	-----PL--W---D-----Q---R---E--RR---MNM-IR---V---	59
SIVSAB1C	----GWLPPVGGDPPKDPKKN---EI--W--T---P---F-----Q---S--QC--P-N---RL-WNIV--P	75
CONSENSUS	gerhgtP??EraYKYrLvQKaIFVHfrcgCrrrqpfepYeerr?GqggGragrvppgl?AAPGL	114
SIVAGMTYO	-----MN-----D-----N-----E	119
SIVAGM155	---Q-R-IA-----N-----P-----D	119
SIVAGM3	--NR--ML-----K-----D-----	119
SIVAGM677	-R--N--WN-IG-----I---SM-----G--S-----N-----PFP-----A	118
SIVSAB1C	AID--QTRL-GW---C-IL-----MRGR-CKPKTHPA-GPGAG-PPP-LG-ASGGAAS-----	140

AGM TAT

CONSENSUS	MDkge?eq???hqdL?e????PLq?CtMkCfCKkCCyHCqlCFIqKgLGvtYhapR?R?rkki?s?n??	53
SIVAGMTYO	-----A--IVS----S-DYQK---T-K-----T-.-R-L-LA	68
SIVAGM155	-----ED-DVS---IKQYRK--ET-----F--R---I---F-T-.-A-ADRI	68
SIVAGM3	-----D--GAY---I-QLKA--KR-----Y--C-----I-.-APLDRF	68
SIVAGM677	---E-.-PHPLL-....HR--P-----Y-----E-----R--VS-K-.-TSTQD-QD	63
SIVSAB1C	--QEQEARPQVWE-E-Q-ELHR---A-D-T---V--F--I---HK-A--IR-YV--P-RAS-KI-H-QV	69
SIVSYK	M3STDQICQTQRVPPSFLEGTFL.EKG-PTP-.-N-----I--AR--K-AARS-SEDDSA	73
CONSENSUS	p???QSIStr?gRd7QTtqe?qakVet?A7?n????RKML??????GA?	82
SIVAGMTYO	-LQH-----KW---G---PTS-----T-GS-	100
SIVAGM155	-VPQ----I-.-S-----S-K---EQ-KA-LRIS----GDETRGPV--G	118
SIVAGM3	-EQK-----S---S---KG-----S-RTAPSLG----AQSGRAT--S	118
SIVAGM677	-IRQ-----V.Q-NG---E-GKTE--	88
SIVSAB1C	SLHN	73
SIVSYK	-TGTLEPRAG-TQANP--K...KKA---EKDSTSLPSAE--	110

AGM REV

CONSENSUS	MpLG?EErr?lkli??ly??npyp?eGTARQRRRARrKWr?qq?Qi??laeRil?c?????qlvqa?dqLv1	55
SIVAGMTYO	----P----FV---WL--ST-----SQ-----Q--D--RV-V--.....QE-,-Y-V-R-AD	67
SIVAGM155	----P----L----AF--RS----SV-----KNR-K--YA-----WG-R..QED-----I-----	72
SIVAGM3	----S----L----AF-NKN-----V-----QA-E-LRA-----WHSR..VEE-----I-----	72
SIVAGM677	-S--K--KQA-KI-KT--GS---QFS.....Q--K--Q--DKI-G-V-N-F..EDQ---AQLQE-Q-	72
SIVSAB1C	-S--Q--LLRRFR-IRFLYTNPV-PGQ-----Q--AK-RQ-VIH-----E-PVSQIDH-A-EF-----	75
SIVSYK	--DQGS-QLAVF-RMIAHLQE--GP..-PRQT---R--Q--QRRT-RLY-QQ--FEAIFGSRTAAE...-S...	70
CONSENSUS	d?ghl??Qq1PdpP??PTENQTANSSS	77
SIVAGMTYO	EA---AI-----HSA	84
SIVAGM155	-T---VT-----SQA	89
SIVAGM3	-Q---AI-----SSS	89
SIVAGM677	ENKD-VL-H----HIH	89
SIVSAB1C	-N...L--P-SL-PGH-----	99
SIVSYK	-QISD	77

```

VERVET.COM M??t1???qil?IGIGIvlstkqQWltVfYQvPvWKNSVQAFcMTPTTLWATTNCIPDDHDYT...EVPL 63
SIVAGM3 -KL--L.I---L---V--N-R---V-----S-----S----- 68
SIVAGMTYO -RY-IITL--IV-----I-----S----- 67
SIVAGM155 -TKF-GIFIV-G---GI----- 69

GRIVET.COM MGRLLIKILIIAIGISIGIG..NLYVTVFYGIPVWKNSVQAFcMTPTNTNMATTNCIPDDHDNT...EVPL 67
SIVAGM677 ----- 67

TANTALUS.COM QAFcMTPTNTNLWATTNCIPDDHDNYT...EVOL 29
SIVAGMB14 ----- 29

SABAEUS.COM MKLLTVLLWLSGCMSLWHL...VQYVTVFYGIPVWKNSVQAFcKTPNTNLWASTNCIPDDEPEGTIAEVPI 69
SIVSAB1C ----- 69

SYKES.COM MAAFRTYIVCLFSLISIGFMEKQYVTVFYGIPHWEDAYAPLFCCTSHKGGWATKNCVP.SADQI...EVRV 68
SIVSYK ----- 68

VERVET.COM .NIT.EPFEAWaDRNPLVAQAgSNIHLLFEQTLKPCVKLSPLCIKMcVELnsz???tct??sTTa?ctt??? 123
SIVAGM3 -----S-----EPT--PK----S--NITA 130
SIVAGMTYO -----G-----I---A-----H-----TRERA-TPT--PKS-G... 134
SIVAGM155 -----KG-ATS-PAT---G-K... 135

GRIVET.COM .NIT.EAFEA.W.D.NPLVKQAESNIHLLFEQTLKPCVKLSPLCIKMcVELNGTATTKATTATTWTTPC 134
SIVAGM677 ----- 134

TANTALUS.COM .NVS.EKFEAWKDRNPLVAQAESNIHLLFESTLKPCVKLTPNCIRMcCTRLPSPTPSSSSTTRTTTRNFCPGT 99
SIVAGMB14 ----- 99

SABAEUS.COM PNIT.EKFDAMKDRNPLVQGAESNIHLLFESTLKPCVKLSPMCIRMcCYRLEGGAAATTTSPSTSTARPEVVS 140
SIVSAB1C ----- 140

SYKES.COM .NITGEYPPAMNSSHMIROQILEDMSALFLOANRPCVKLAPNCIRMLCTLDNSPATSTPTTSPPTTPP... 135
SIVSYK ----- 135

VERVET.COM ???LPCV7nkt?7nLqSCN?tiIEKEmndepASNCTFAMAGYVRDQKKm?YSVVWMDAEI?CK????nn? 179
SIVAGM3 STTT---Q---STV-E---E---L-E-----K-----H---KG--S 206
SIVAGMTYO ---GPTSGE---AS--R--E-----Y-----Y--NKT--ST 200
SIVAGM155 ---R---DS---D-----A-----I-----F---RSTSH-G 201

GRIVET.COM .....QNCSTEQIEGEMAEEPASNCTFAIAGYQRDVKKN.YSMTWYDQELVCNNKTGSEK 189
SIVAGM677 ----- 189

TANTALUS.COM TN.....DECNSHLVTNSMEFENSSMCTFAMAGYMRDQKKT.YNSTWYDAELMCEPESKR.NS 155
SIVAGMB14 ----- 155

SABAEUS.COM VG.....FNDSVIEQEMEKEQAMNCSFAMAGYRRDVKKN.YSTVWDDQEVVCEEGREKSNA 195
SIVSAB1C ----- 195

SYKES.COM .....NET.MWGDNSTEPRENCSENLFGGPKOKKQQ.YRAFFYKODLM....KEEGNS 182
SIVSYK ----- 182

```

AGM ENV

VERVET.CON	?kE..CYMIHCNDSVIKEACDKTYWDeLRLRYCAPAGyALLKcND?dy?GfK?nCaNVSVVHCTgLaNTTVt	245
SIVAGM3	MR..-----F-----Y--A--T-----N-I-----	276
SIVAGMTYO	S--..-----Q-----E--N-Y-Q-----	270
SIVAGM155	T--..-----W--A--PE-----T-----	271
SIVVER1E	-----RN-T--APR-R-----S	46
SIVVER2E	-----H--N-Y-Q-----	46
SIVVER266E	-----F-----A--D--T-----N-----	46
SIV2010E	-----S	1
SIVVER385E	-----F-----Y--D--Q-----	46
GRIVET.CON	SKD..CYMIHCNDSVIKEACDKTYWDTLRLRYCAPAGyALLKcNDKdyGfAPKCKNVsVvHCTRLINTTic	259
SIVAGM677	---R-----	259
SIVGRI3E	-----	46
SIVGRI2E	---L-----F---R-L--N-----S---T-T-----VS	46
TANTALUS.CON	,RG..CYMIHCNDSVIKEACEKTYWDTLRLRYCAPAGfAiLKCKDTNYtGfG.iCrNVSVVSVCTGLMMTTvs	223
SIVAGMB14	,---AV-----T-----	223
SIVTAN1E	-----V-----	45
SIVTAN17E	-----	45
SIVTAN40E	-----L--R---A--N-----f-	45
SIVTAN49E	-----Y-----	45
SABAEUS.CON	THTVGCYMIHCNTSVIKEACDKTYWDTFLRLRYCAPAGYALLRCaDcDysGHK.aCKNVtVSACTRLINTTVS	266
SIVSAB2E	-----I-	45
SIVSAB3E	-----K-----	45
SIVSAB4E	-----R-S--R---N-Q--S-----	45
SIVSAB1C	-----R-----	266
SYKES.CON	SY...YYLLHCNTSVISAACEKQTFQPFPIQYCAPPGYSLLKcNDTNfEGDD.VCTNVTAVSCTQEFNTLAS	250
SIVSYK	-----	250

VERVET.COM	TGLLNGSYSENrTqIwqKHrvsNdsVII1?NKhyNLtvtCKRPGNKTVLPVTIMAGlVFHSQkyNtrLR..	314
SIVAGM3	-----V-F-----R-----	346
SIVAGHTYO	-----H-----N-NT-----F-----S-----R-----MX-----	340
SIVAGM155	-----G-----L-----	341
SIVVER1E	-----T-----E-----K-----N-----I-RI-----IR-V-----K-----	116
SIVVER2E	-----H-----R-----G-----I-----I-----AF-----R-----	116
SIVVER266E	-----T-----A-----L-----	116
SIV2010E	-----K-----N-----L-----I-----M-----	71
SIVVER385E	-----QI-----I-----	116
GRIVET.COM	TGIGLNGSR?ENRTEIWQekGNsNdTVIiKLNKfYNLTfVRCRRPGNKTVLPVTINAGlVFHSQkyNTrLK..	328
SIVAGM677	-----S-----KG-----D-----	329
SIVGRI3E	-----V-----I-----R-----K-----	116
SIVGRI2E	-----A-----N-----	116
TANTALUS.COM	SAPG1NGSQAENRTEIWQKHvsnDSViiKLNKhY?Lc1?CrRPGnKTvLPVTImAGlVFHSQkyNTKLR..	291
SIVAGMB05	-----Q-----V-----K-----N-----Q-----	52
SIVAGMB30	T-M-----	15
SIVAGMB53	IH-----	7
SIVAGMB14	-----N-----L-----M-----	293
SIVTAN1E	-----I-----N-----K-K-V-----Q-----	115
SIVTAN17E	-----V-----F-N-----K-----R-----	115
SIVTAN40E	-----N-----Q-K-K-I-----	115
SIVTAN49E	-----RIN-----Y-K-GVT-----	115
SABAEUS.COM	TGIGINGSYVANRTEIWQKNGNsd?VIIRLNr?fNLfIRCRrPGNKTVLPVTImAGlVFHSQkyNTrlr..	334
SIVAGMD29	--I-----A-----I-----V-----	30
SIVAGMD30	-----K-----	30
SIVAGMD37	N-----KF-----	11
SIVAGMD45	-----	28
SIVSAB2E	-----A-HT-----H-----K-----	115
SIVSAB3E	-----S-----H-----K-----	115
SIVSAB4E	-----E-----NTA-----YY-----	115
SIVSAB1C	-----S-----Y-----K-----	336
SYKES.COM	TWFQ1NGTYKAKDKVRPIKQDKNESVILVPEALRLQIICERPGNESIKNIQL.AAGYFLPVIQG.KLKTG	320
SIVSYK	-----	320

AGM ENV

VERVET.CON	...QAMCHPqgmwxGAWkEvkqeIvklPkdrYqGTndte?I?lqRq?GDPEaaNLMPNCqGEFFYcKMDWF	379
SIVAGM3	...-----N-----E-Y---LP-----	414
SIVAGMTYO	...-----E-----R---K-E---K---N-H-Y---W---S-----	408
SIVAGM155	...-----K-----QE-----E-----NK-F---F-----	409
SIVVER1E	...-----K-D---R-I---K-E---E-----KQ-F---W---S---H-----M	184
SIVVER2E	...-----E-IR---T---PHK-R---K-F---F-----	184
SIVVER266E	...-----PS-K---E---N---E-R---PKR-FF---W---T---N-----	184
SIV2010E	...-----N---E---R-----AQK-Y---F-----	139
SIVVER385E	...-----H---I-----R---Q-Y-T---W-----S---	184
GRIVET.CON	...QAMCHfQGdW?GAWKEVraev.Kk?K..NLTeV?IenIhLRRqWGDPEaANFWFnCqGEFFYCKMDWF	390
SIVAGM677	...-----K-----V-----S-----I-----S-----	394
SIVGRI3E	...-----S-N-L-----KI-EE---K-----H-----	181
SIVGRI2E	...-----H-----KQ---A---TDD-K-----S-----	178
TANTALUS.CON	...QAMCHPqGdWkGAWKEVreTivklPkexYqGTntkqIWLrRQWGDPEAAniWfNCqGEFFYCTPDWF	359
SIVAGMB05	...-----T---V-----	120
SIVAGMB30	...-----S-----Q---Q---A---P-----	83
SIVAGMB53	...-----S-K-E-E-K---A-----HS-----KQ-----	75
SIVAGMB14	...-----H---D-K-----S---L-----	361
SIVTAN1E	...-----R-K-E---R---R---S-----L-----	183
SIVTAN17E	...-----K-E-----R-D-K-----	183
SIVTAN40E	...-----K-TE-----M	183
SIVTAN49E	...-----H-----RDR-S---N-----	183
SABAEUS.CON	...QAMCWFGGnWrgAWqEvketivklpp?rYsGTndtskIFLQRQWGDPEaEfffnCQGEFFYCKMDWF	401
SIVAGMD29	...-----R-A-EL---EK-Q-----L-W-----	98
SIVAGMD30	...-----K-S-----KD-----S-----	98
SIVAGMD37	...-----N-----E---KEK-Q---H-EQ---S---I-----	79
SIVAGMD45	...-----T---K-R---AAN-NST-T---RNI-Q-----S-NLL-----	96
SIVSAB2E	...-----D---K-A-D-----N-----	183
SIVSAB3E	...-----A-D-Q-K-N-----	183
SIVSAB4E	...-----K---A-E---S-E---KN-Q---G-----	183
SIVSAB1C	...-----K-----R---KK-----N-----S-----	404
SYKES.CON	ROAKRAPCRVTGNWTEFFKQVHEQATKTWK.....NVINTTWRSQPGDLEVRTHWFCQCGEFPYC.....	381
SIVSYK		381

```

      ***          *          *          *
VERVET.COM  LNYIMNLT....vdadH?7N7C7n?7k7777raPGPCvQRTYVachIRSVindwytiskKTYAPPREGH 435
SIVAGM3     -----R-----P-----P-----G-----GKG.K-----A-----L-R----- 475
SIVAGMTYO   -----K-----W-----Y-----F-----SS-KKG.H-----Y-----S-----L----- 468
SIVAGM155   -----H-----KNNAGKG.-S----- 470
SIVVER1E    ---I-----N-----N-----S-QG...K-----G-----KLAR----- 243
SIVVER3E    -----S-----GSRTSP...-----A----- 245
SIVVER266E  -----E-K-TSGT-SGNK-----INLE----- 249
SIV2010E    ---I-----N-NT-R.-K-GKTSG...-----A-----V-----V-----L----- 202
SIVVER385E  -----K-----WN-----Q-D-S-----TG----- 245
      ***          *          *          *
GRIVET.COM  INYIMN?T....vDadG????C?..???kkgp..GPGPCvqrTYVACHIRQVNDWYtltKKtYaPPREGH 444
SIVAGM677   -----R-----E-----E-TNRT-D...-GK-----VS--V----- 454
SIVGRI3E    ---Q-----KS-----TSH---Q...-AKK----- 241
SIVGRI2E    ---I-----E-----NK-----NKP-----R-----Q 237
      ***          *          *          *
TANTALUS.COM vN?LNNesa?g???dVEG..N?C..???t??gltgs?rkClKRtYVachIRSVVNDWYtLsKkTYAPPREGH 415
SIVAGMB05    --Y-----N-TNL-----K...PDR-SG----- 187
SIVAGMB30    --W-----D-E-AQK-----D...PMT-MG-----Q----- 150
SIVAGMB53    --Y-----ED-TLL-----K...ANT-TPP-----T-----S-----I-----FH----- 143
SIVAGMB14    L-Y-----SE-SFT-----R...SNI-SG-----T-----GL-----LQ-R----- 429
SIVTAN1E     --W-----NS-RNV-----N...T-----KDKR-Y-----P-----I----- 244
SIVTAN17E    --Y-----D-RRRKE-----D...KDP--F--S--P----- 242
SIVTAN40E    --W-----DS-KSVN-----T...TPEKO-L-K----- 244
SIVTAN49E    --W-----G-MYT-----P...SNSTK-GMTGS-----G-----S----- 250
      ***          *          *          *
SABAEUS.COM  LMYLNMqs....VdPdH..N?C.....???k?ndkkCwqRtYvPCHirQvVNDWYtLsKkTYAPPREGH 454
SIVAGMD29    -----LT-----K-----NKG----- 153
SIVAGMD30    --F-----N-----T-Q-----I----- 154
SIVAGMD37    -----K-----G-----P-----S-N--R--N-----Q 135
SIVAGMD45    -----N-----Q-----KTN-----A-N-----A-----Q 152
SIVSAB2E     -----P-----GGNDK-NR-----I----- 241
SIVSAB3E     -----P-----N-TS-----N-A----- 239
SIVSAB4E     -----N-----T-----TPKS--R----- 239
SIVSAB1C     -----K-----N-----A-N-T-P----- 460
      ***          *          *          *
SYKES.COM    ....NVSKLFANITWGNASKMNYASNLR.....LSCAIRQIINDWRYVRKLIYLPPTAGH 432
SIVSYK       ..... 432

```

AGM ENV

VERVET.CON	LqCtSTVTGmtVELNyn?knRT.NVTLSPOIESIWAaELgRYKLVEITPIGFAPTeVRRYTGgherqKRV.P	504
SIVAGM3	-----S-----S-----T-----D-T-----	545
SIVAGMTYO	---R-----S-----	538
SIVAGM155	---NQ-----T-----Q-----	540
SIVVER1E	-E-----A-----IH-----A--S-----P-----P.S-E-----	312
SIVVER2E	-----SR-----D-----	315
SIVVER266E	-E-----IP-----D-----	319
SIV2010E	-E-R-----NQS-----Q-----	272
SIVVER385E	-----IND-----T-----SP--I--	315
* ***		
GRIVET.CON	LECMSavTALyVeInyNNkSGPINVTLSpQvRSIWAYELGDYKLVEITPIGFAPTDVRRYTGP.treKRV.P	514
SIVAGM677	-----A-D-----	524
SIVGRI3E	---T-----S-----S-----	311
SIVGRI2E	-----E-----I-----KQ-----	307
* ***		
TANTALUS.CON	LECtSTVTSmvSLDyNsKnRT.NVTLTANlENiWAYELGRyKlIeIkPIGFAPTeVRRYvGP.TREKRV.P	484
SIVAGMB05	-----K-----	234
SIVAGMB30	-----E-----	219
SIVAGMB53	-----I-----T-H-----N-N	185
SIVAGMB14	---R---L---D---	486
SIVTAN1E	-----ER-----E-----I-----	313
SIVTAN17E	-----VN-----I-----D-----	311
SIVTAN40E	-----LIA-----N-----I--LY-----	313
SIVTAN49E	-----	319
* ***		
SABAEUS.CON	LECnSTATALyVELnYNskNrT.NVTLSFqIrSIWanELGDyKLVEIKPIGFAPTeVRRYTGP.ERQKRV.P	523
SIVAGMD29	---T-----H-----	192
SIVAGMD30	-----M-----L-----L-----F-----K-----	223
SIVAGMD37	---H-----G-----	162
SIVAGMD45	---T-----H-----SY--S-----	198
SIVSAB2E	-----L-----H-E-----	310
SIVSAB3E	-----E-----	308
SIVSAB4E	-----N-----	308
SIVSAB1C	-----E-----K-----	529
* ***		
SYKES.CON	IKCTSMNTAVLTDIEYYPGSTL.NFTPTANVEDVWRADLFNYKLIQIKPIGFAPTDQRRYELPNTREKRAAP	503
SIVSYK	-----	503

VERVET.COM	FVIGFLGFLGAAGTAMGAaAaLTVQSQhLLAGILQOQKNLLAAVEaAQOQMLKLTWGVKNLNARVTALEKY	576
SIVAGM3	-----	617
SIVAGMTYO	-----SS-----R-----	610
SIVAGM155	-----G-----	612
SIVVER1E	-----	372
SIVVER2E	--Q-----T-----	375
SIVVER266E	-----	379
SIV2010E	-----A-----D-----	302
SIVVER385E	-----	375
GRIVET.COM	FVIGFLGFLGAAGTAMGAaAaLTVQSRHLLAGILQOQKNLLAAVEOQOQLLKLTIWGVKNLNARVTALEKY	586
SIVAGM677	-----T-----	596
SIVGRI3E	-----V-----	371
SIVGRI2E	-----R-----	367
TANTALUS.COM	FVLGELGFLGzAGAAMGATaT?LTVQSQQLLAGILQOQKNLLAAVEOQOQMLKLTWGVK	543
SIVTAN1E	-----A-----A-----M-----	373
SIVTAN17E	-----S-----S-----	371
SIVTAN40E	-----A-----	373
SIVTAN49E	-----E-E-----	379
SABAEUS.COM	FVLGFLGFLGAAGAAMGAaAaLTVQSQQLLAGILQOQKNLLAAVEOQOQMLKLTWGVKNLNARVTALEKY	595
SIVSAB2E	-----	369
SIVSAB3E	-----	368
SIVSAB4E	-----	368
SIVSAB1C	-----	601
SYKES.COM	LALGFLGLLSAAGTAMGGAaAaLTQSQTLLAGIVQOQKLLAAVEAQOHLGLTWGVKNLNARLTALEKY	575
SIVSYK	-----	575
* * * * *		
VERVET.COM	LeDQARLNawGCANWQVCHTTVPW.....?WnNrTbWnNMTWLEWERQI??LEgNIT?QLe?AReQEEK	636
SIVAGM3	-----Q-----SY-----T-----E-----A-----	682
SIVAGMTYO	-----S-----E-----P-T-----Q-----AD-----S-----G-----VK-----	675
SIVAGM155	-A-----T-----E-----K-----EG-----K-----Q-----	676
* * * * *		
GRIVET.COM	LEDQARLNswGCANWQVCHTTVPW.....KYNN.TPKWNNMTWLEWERQINALEGNITQLEEAQNQESK	650
SIVAGM677	-----	660
* * * * *		
SABAEUS.COM	LEDQARLNINGCAPRQVCHTTVLW.....FYNN.TPDWNNMTWQEWERQIENYEANISRILOAKEQSQK	659
SIVSAB1C	-----	665
* * * * *		
SYKES.COM	LRDQAILSNWGCAPKQICHTAVTWEKACGNNSNFCPKPQWNNMTWNRWBOEVENLTOHIDGLLEAQEQQER	647
SIVSYK	-----	647
* * * * *		
VERVET.COM	NLDAYQKLsawSGFWSWFOFSKWLNTLKIGFLV?GIIGLRLLYTVY?C1aRVRRQGYsPLSPQIHHPw?KG	704
SIVAGM3	-----D-L-----S-----	753
SIVAGMTYO	-----T-----M-----VIV-----G-----V-----V-----QVG-----	747
SIVAGM155	-----D-S-----A-I-V-----L-T-----	747
* * * * *		
GRIVET.COM	NLDLYQKLDDWSGFWSWFSLSLSTWLGYYKIGPLVIVYILGLRPAWVLWGCIRNIRQGYNPL.PQIHHS.SAE	720
SIVAGM677	-----	730
SABAEUS.COM	NLDSYQKLVSWSDFWSWFDLTKWPGWMMKIAIMVIAGIIVARVLLVYIGILRKFRRGYAPLSSLPSH	726
SIVSAB1C	-----	732
* * * * *		
SYKES.COM	NVHDELTKLQEWDSLWSWFDLSKNWFFYLLKIGFYVIGALVLLRLVSPSVGIKNNLEGGYVPILQNPQGR..KD	717
SIVSYK	-----	717

AGM ENV

VERVET.CON	qPDNA?ePgEGGd??k?kse?WQK?SgtaeWRS?wC??LT?W?????WLYNSCLTLL??LR?APQY?QYGL	756
SIVAGM3	-----EG-----KR-NS--P---E-----N--KR--N-CSISSI-----VH--S----I----	825
SIVAGMTYO	R---D-----NSRI-L-S- -D- -R.-MQ--A-LTRLNT-----IQ--K---L----	812
SIVAGM155	-----G--E---RTG-S--TH	768
GRIVET.CON	RPDNGGGQDRGGESSSSKLIRLQEESSTPSRINNWNLPKSCSLRIRTWCYNICLTLILFIRTAVGYLOVGL	792
SIVAGM677	-----	802
SYKES.CON	PGKPADEEBEGSGDREGLNVTFSRESLRQSLAQQLWRTV CSSFRSLIRQLTITWGFISYGF	780
SIVEYK	-----	780
VERVET.CON	?ELK??AQE????LA??AQNA??QIWLACRSAYR?I?NSPRVRQGLE?ILN	794
SIVAGM3	G---AA---AVVA--RL---GY-----A-I-----G---	877
SIVAGMTYO	A---TG---ILQT--GV---CH-----N-V-----E--	863
GRIVET.CON	QQLQEAAATGLAQALARAAREAWGRDGAIVRSAYRAVINSPRVRQGLEKVLG	844
SIVAGM677	-----	854
SYKES.CON	NELKIAAASLGREIRDWVAAIWQAIYAA?RRVVEAVAALPRRLRQGLEIYLN	832
SIVSYK	-----	832

CONSENSUS	MG??nSRpqhkk?l?lw?al?h?c??trygiladPLiGqSSt?QBECdKaLrksliRk?ng?mTpegrrlqegd?w	63
SIVAGMTYO	--SQ---A---YgK--Q---K-KV-----T---V-----Q--N--E-E-----T-	75
SIVAGM155	--LGS-----Q-TI-R---A-RH-----L-----G-----R--N-----D--Q-	75
SIVAGM3	--LQ-----Q-S--H---K-RA-----L-----KE---R--K---K---K-	75
SIVAGM677	--SS---R-QQGL-K--RG--RGKPGADWV--S-----V---G---K--W.G...K--D-----TF	72
SIVSAB1C	--GKS--Q-QRHS-W--SK-RQAPVIQ-DM-----L---HI---A-S--DG---QGSSR-E--VKMKHQGRQ	75
SIVSYK	--STS--S-QLRSEgKYAIGWRLFGKQ-TP-P-E-SRPLQPCRGQT--W-ST-TEPIDPHGPORDWGHSG-QKF	75
CONSENSUS	deWtsddEeE...VGFPVrPrvPLRqMTYKLAVDfshFLKsKGGLdGIYYSDRRnkILnLYALNEWGIIdDwna	133
SIVAGMTYO	E--.....	145
SIVAGM155	---E-D-.....I-----	145
SIVAGM3	---E-D-.....Q-----	145
SIVAGM677	---D-----Q-----S-----E--E-----Q-	142
SIVSAB1C	PS-YDE-----CL--A-----I-G-----E--E--K--D-----V-G-QN	146
SIVSYK	SPGDIVQD-GDTGL-----C-QT---TL-----I-L---I-N---Q-MN-CEK-DE--H--LQ--H---RI-	149
CONSENSUS	ws?GPGiRyPrCFGECPKLVPVdLhEEA?tcErHCLvHFAQ.....??edpDgi?hgE?LvwKFDpMLavqYdp	196
SIVAGMTYO	--K-----F-X-----Q-----MG-----S---I-----I---	214
SIVAGM155	--K-----A-----E-----LN-----N---I-A-----	214
SIVAGM3	--E-----E-----VR-----N--V-----	214
SIVAGM677	Y-P-----V-----RN-----M-----MG-----D--V-----K---E-R-	211
SIVSAB1C	YTD---T---K--W-----S---KNS-N--L-----VAYED-AWK...-T-----L---D-VA	213
SIVSYK	YTS---T---LI--MLWE---NEIEGCLEVE-HTL-L---SQGSSSNG--HVELQPPPGYTPGME-ARL-LER	223
CONSENSUS	?r?yftdmh???gkrk	207
SIVAGMTYO	N-E-----GLV----	229
SIVAGM155	S-E-----LYSTV-TGN	230
SIVAGM3	N-K-L---DL----	229
SIVAGM677	..DH-K---EH.A--	223
SIVSAB1C	W-LHPEQVPSAQ-	226
SIVSYK	QTGKPOELQSALS-NIS	240

III

Analysis

PART III Analysis

Contents	III-1
A. Gag Proteins (Robert J. Gorelick and Louis E. Henderson)	2
B. HIV-1 Tat: Structure and Function (Kuan-Teh Jeang)	11
C. HIV-1 and HIV-2 Sequence Subtypes (Gerald Myers)	23
HIV-1 Phylogenetic Trees	23
HIV-1 Subtype Summary	31
HIV-2 Phylogenetic Trees	32
D. HIV Synonymous and Nonsynonymous Substitution Frequencies (G. Myers)	34
HIV-1 and HIV-2 <i>gag</i> and <i>env</i> Substitution Frequencies	34
HIV-1 Intersubtype Comparisons	36
E. Global Variation in the HIV-1 V3 Region (Bette Korber)	37
Introduction	37
Subtype consensus sequences	39
Amino acids found in the V3 loop of each subtype(1994)	40
Analyses of variation at the immunogenic tip of the V3 loop	44
Country codes	47
Alignments of sequences by subtype	52
Brief descriptions and references for all sequences included in the study	66
Phenetic relationships of the tip of the V3 loop	84
F. Mutations in HIV-1 RT and Protease (John Mellors, Raymond Schinazi, Brendan Larder; Simon Wain-Hobson)	93

Gag Proteins

Robert J. Gorelick and Louis E. Henderson

*AIDS Vaccine Program, PRI/DynCorp, National Cancer Institute,
Frederick Cancer Research and Development Center, Frederick, Maryland 21702*

INTRODUCTION

All retroviruses utilize a *gag* gene to encode a precursor polypeptide (Gag precursor) that is initially incorporated in the budding particle and ultimately cleaved by a viral protease (encoded by a *pol* gene) into structural proteins of the mature virus collectively referred to as Gag proteins (See Figure 1 for biosynthesis of HIV-1 virion proteins). Gag precursors are self-associating proteins capable of assembling into virion-like particles without the requirement of other viral factors. For most members of the retrovirus family, including lentiviruses, assembly takes place on the cytoplasmic side of the cell membrane and nascent virion-like particles bud through the membrane into extracellular space. During normal viral replication the self-associating Gag precursors selectively bind viral RNA and package two plus strands of the genome inside the particle. They also form associations with viral Env proteins such that the budding particle is encapsulated within a lipid envelope containing viral Env proteins. All retroviruses express viral encoded enzymes from a *pol* gene as Gag-Pol fusion proteins and the self-associating properties of the Gag portion directs assembly of the Gag-Pol fusion protein into the budding particle. Thus, the Gag precursor orchestrates assembly of the nascent virion by forming the core structure and interacting with all other viral components (RNA, Pol proteins, envelope lipid and Env proteins) to facilitate their incorporation into the budding viral particle. The term "Gag precursor" was originally adopted to indicate that the major antigenic protein (capsid antigen, CA; [1]) or "Group antigen" (Gag) was synthesized as a precursor. The term does little to convey the important biological properties of the protein and a more descriptive name such as "Assemblin" might be more appropriate.

HIV-1 GAG PRECURSOR (ASSEMBLIN)

HIV-1 *gag* genes are generally about 500 codons in length. Figure 2 is a composite diagram indicating posttranslational modifications, proteolytic cleavages and sites for various interactions that are known for the HIV-1 Gag precursor (Assemblin, Pr55^{gag}). The upper panel in Figure 2 indicates the posttranslationally modified form of the protein that participates in the assembly of the immature virion particle. During or shortly after assembly, the viral protease (PR; product of the *pol* gene) is activated and the Gag precursor is cleaved (middle panel) into the proteins found in the mature virus (lower panel). The Gag precursor is thought to be organized into domains that roughly correspond to the final proteolytic cleavage products [matrix antigen (MA), capsid antigen (CA), nucleocapsid (NC), and p6] that make up the core of the mature virus.

During or shortly after translation, the initiator Met residue is removed leaving Gly-2 as the new N-terminal residue. Cellular myristoyltransferases catalyze the covalent attachment of a myristoyl group [$\text{CH}_3(\text{CH}_2)_{12}\text{COO}^-$] (Myr) to the amino group of Gly-2 via an amide bond and cellular kinases phosphorylate at least two sites on the protein in the CA domain (one serine and one threonine residue, but the exact sites have not been identified). The protein also binds two zinc ions to form zinc-fingers in the NC domain. As indicated in Figure 2, these posttranslational modifications carry through to the Gag proteins found in the mature virus. Thus MA is myristoylated at the N-terminal end, CA is phosphorylated at two sites and the NC protein contains two zinc fingers.

The myristoyl group and the first 31 N-terminal amino acid residues of the MA domain are necessary for targeting the Gag precursor to the cell membrane [2, 3]. The MA domain of the Gag precursor has also been implicated in binding and assembly of the *env* gene products into mature virus [4, 5]. Presumably the interactions with the *env* gene products involve the cytoplasmic domain of the transmembrane (TM) protein associating with the MA domain, however this has not been directly demonstrated.

The CA domain of the precursor appears to be involved in self-association or Gag/Gag interactions [6]. These interactions have been narrowed down to amino acid residues ~240-430 (see Figure 2, upper panel; [7, 8]). Located within the CA domain is a region called the major homology region (MHR) which is required for efficient viral replication and particle production [9]. The MHR includes residues 285-304 in the HIV-1 (IIIB) CA domain. Amino acid residues, 287/Q, 291/E, 296/Y, and 299/R (in the sequence of IIIB) are completely conserved in the MHR region of all primate lentiviruses.

A region between residues 178 and ~300 in the CA domain of the Gag precursor has been shown to bind Cyclophilin A [10]. Cyclophilin A is a peptidyl-prolyl cis-trans isomerase thought to play a role in immune regulation and has been found in preparations of purified HIV-1 (Henderson, unpublished). Other protein-protein interactions of the HIV-1 Gag precursor include interactions of the p6 domain with the viral encoded Vpr protein [11, 12].

Gag protein-protein interactions also play a significant role in packaging the viral enzymes encoded by the *pol* gene. The *pol* gene products are expressed as a Gag-Pol fusion protein which results from synthesis initiating as for the Gag precursor and progressing in the *gag* frame to a site just past the NC domain (Gag-Pol frameshift site, Figure 2) where a -1 ribosomal frameshift event occurs followed by synthesis continuing in the *pol* frame [13]. The resulting fusion protein is referred to as the Gag-Pol precursor (Pr160Gag-Pol) and contains the Gag sequences required for association with Gag precursors and incorporation into virion particles. The fusion protein is ultimately cleaved by the activated viral PR to the mature Pol products (PR; reverse transcriptase, RT; and integrase, IN) and the corresponding Gag proteins (MA, CA, p2, NC) and peptides derived from the *pol* reading frame preceding the codon for the N-terminal residue of PR. The "transframe" peptide isolated directly from whole virus [HIV-1 (MN)] has the sequence FLREDLAF [14], where the first two residues (FL) are encoded by the *gag* gene and correspond to the first two residues of the p1 peptide (lower panel, Figure 2) and the remainder of the transframe peptide is encoded by the *pol* gene.

In addition to the various protein-protein interactions involved in the assembly, the Gag precursor also binds and packages the viral genomic RNA. The NC domain of the Gag precursor is directly involved in the nucleic acid binding properties of the protein. The NC domain (Figure 2, upper panel) is rich in basic residues (Lys and Arg) and contains two copies of a zinc binding sequence [Cys(X)₂Cys(X)₄His(X)₄Cys] referred to as a retroviral CCHC zinc finger [see Figure 3; 15]. The basic residues confer a nucleic acid binding property to the domain and the CCHC zinc fingers are involved in the specificity for the viral RNA [16-18]. Apparently, during viral assembly, the NC domain of the Gag precursor specifically binds to a site near the 5' end of the genomic RNA referred to as the Psi site. If either of the two zinc fingers in the NC domain are disrupted by site directed mutagenesis, or the Psi sequence in the RNA is deleted, noninfectious particles are produced with reduced levels of packaged viral genome [16-19]. Mutants with the first finger replaced by the second finger also package reduced levels of viral RNA, but a mutant with the second finger replaced by the first (i.e., NC domain with two identical copies of the first finger) packages normal levels of viral RNA [20]. Thus, even though both zinc fingers are required for specific RNA packaging, they are not functionally equivalent in this regard.

MATURE GAG PROTEINS

After the Gag precursor assembles into the immature virion, the activated viral protease cleaves the protein into the mature Gag proteins found in the infectious virus (Figure 2). Based on *in vitro* analysis, the order of cleavage is as follows: p2/NC > MA/CA = p1/p6 > CA/p2 > NC/p1. The MA/CA and p1/p6 sites are cleaved approximately 12-fold slower and the CA/p2 cleavage proceeds about 400-fold slower than the initial p2/NC cleavage [21-23]. Following proteolytic cleavage (Figure 2) the mature Gag proteins rearrange to form the core structure of the mature infectious virus. The mature MA protein remains associated with the inner side of the lipid envelope while the CA protein forms the outer shell of the capsid structure which contains the NC protein bound to the genomic RNA. During the infection process, RT and IN together with the MA protein, NC

protein and the genomic RNA are taken into the cell as the "pre-integration complex" [24]. It has been suggested that proviral DNA synthesis takes place within the pre-integration complex. The mature MA protein plays a role in transport of the viral nucleoprotein complex into the nucleus for integration of provirus into the host genome [24, 25]. The MA protein detected in the nucleus is phosphorylated [26], whereas the MA protein in the virus is not. Thus, the phosphorylation of MA occurs during or after the viral infectious event. Within the MA protein there is a nuclear localization signal (NLS) [25] that is similar to other nuclear targeting sequences. The signal is located between residues 14-36 and 107-116 of the HIV-1 p17 MA protein (see Figure 2) and mutants lacking the signal accumulate proviral DNA in the cytoplasm. The NC protein may also play a role in the infection process. Mutants of the NC protein which still package RNA but have alterations within the zinc-finger structures have reduced infectivity [20] suggesting that a role is performed by the NC protein which may be at the level of the preintegration complex. The NC protein has additional functions in the virus life cycle. These include strong nucleic acid annealing and melting activities [27, 28]. It has been suggested that the NC protein is involved in annealing of the primer tRNA to the primer binding site of the viral genome. These activities may also have implications in strand transfer reactions that occur after the first strong stop synthesis of the proviral DNA, to complete synthesis of the provirus for subsequent integration. The NC protein has also been shown to catalyze the formation of the genomic RNA dimer found in virion particles [29]. A recent report showed that HIV-1 with a defective protease had an RNA dimer structure that was different from that found in wild-type HIV-1 [30]. This indicates that dimer formation is probably catalyzed by the mature NC protein and not by the Gag precursor.

References

- [1] J. Leis, D. Baltimore, J.M. Bishop, J. Coffin, E. Fleissner, S.P. Goff, S. Oroszlan, H. Robinson, A.M. Skalka, H.M. Temin, and V. Vogt (1988) Notes: Standardized and simplified nomenclature for proteins common to all retroviruses. *J. Virol.* 62:1808-1809.
- [2] W. Zhou, L.J. Parent, J.W. Wills, and M.D. Resh (1994) Identification of a Membrane-Binding Domain within the Amino-Terminal Region of Human Immunodeficiency Virus Type 1 Gag Protein Which Interacts with Acidic Phospholipids. *J. Virol.* 68:2556-2569.
- [3] M. Bryant and L. Ratner (1990) Myristoylation-dependent replication and assembly of human immunodeficiency virus type 1. *Proc. Natl Acad. Sci. USA* 87:523-527.
- [4] T. Dorfman, F. Mammano, W.A. Helms, and H.G. Göttinger (1994) Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. *J. Virol.* 68:1689-1696.
- [5] X. Yuan, X. Yu, T.-H. Lee, and M. Essex (1993) Mutations in the N-terminal region of human immunodeficiency virus type 1 matrix protein block intracellular transport of the Gag precursor. *J. Virol.* 67:6387-6394.
- [6] L.S. Ehrlich, B.E. Agresta, and C.A. Carter (1992) Assembly of recombinant human immunodeficiency virus type-1 capsid protein in vitro. *J. Virol.* 66:4874-4883.
- [7] J. Luban, K.B. Alin, K.L. Bossolt, T. Humaran, and S.P. Goff (1992) Genetic assay for multimerization of retroviral gag polyproteins. *J. Virol.* 66:5157-5160.
- [8] E.K. Franke, H.E.H. Yuan, K.L. Bossolt, S.P. Goff, and J. Luban (1994) Specificity and sequence requirements for interactions between various retroviral gag proteins. *J. Virol.* 68:5300-5305.
- [9] F. Mammano, A. Ohagen, S. Hoglund, and H.G. Göttinger (1994) Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.* 68:4927-4936.
- [10] J. Luban, K.L. Bossolt, E.K. Franke, G.V. Kalpana, and S.P. Goff (1993) Human immunodeficiency virus type 1 gag protein binds to cyclophilins A and B. *Cell* 73:1067-1078.
- [11] C. Lavallée, X. J. Yao, A. Ladha, H. Göttinger, W.A. Haseltine, and É. Cohen (1994) Requirement of the Pr55gag precursor for incorporation of the vpr product into human immunodeficiency

- virus type 1 viral particles. *J. Virol.* 68:1926-1934.
- [12] W. Paxton, R.I. Connor, and N.R. Landau (1993) Incorporation of vpr into human immunodeficiency virus type 1 virions: Requirements for the p6 region of gag and mutational analysis. *J. Virol.* 67:7229-7237.
 - [13] T. Jacks and H.E. Varmus (1985) Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* 230:1237-1242.
 - [14] L.E. Henderson, unpublished.
 - [15] J.M. Berg (1986) Potential metal-binding domains in nucleic acid binding proteins. *Science* 232:485-487.
 - [16] A. Aldovini and R.A. Young (1990) Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J. Virol.* 64:1920-1926.
 - [17] Gorelick, R.J., S.M. Nigida, Jr., J.W. Bess, Jr., L.O. Arthur, L.E. Henderson, and A. Rein (1990) Noninfectious human immunodeficiency virus type 1 mutants deficient in genomic RNA. *J. Virol.* 64:3207-3211.
 - [18] T. Dorfman, J. Luban, S.P. Goff, W.A. Haseltine, and H.G. Göttinger (1993) Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. *J. Virol.* 67:6159-6169.
 - [19] R. Carroll, J.-T. Lin, E.J. Dacquel, J.D. Mosca, D.S. Burke, and D.C. St. Louis (1994) A human immunodeficiency virus type 1 (HIV-1)-based retroviral vector system utilizing stable HIV-1 packaging cell lines. *J. Virol.* 68:6047-6051.
 - [20] Gorelick, R. J., Chabot, D. J., Rein, A., Henderson, L. E., and Arthur, L. O. (1993) The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J. Virol.* 67:4027-4036.
 - [21] Erickson-Viitanen, S., J. Manfredi, P. Viitanen, D.E. Tribe, R. Trich, C.A. Hutchison III, D.D. Loeb, and R. Swanstrom (1989) Cleavage of HIV-1 gag polyprotein synthesized in vitro: sequential cleavage by the viral protease. *AIDS Res. and Hum. Retroviruses* 5:577-591.
 - [22] Krausslich, H.G., R.H. Ingraham, M.T. Skoog, E. Wimmer, P.V. Pallai, and C.A. Carter (1989) Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. *Proc. Natl. Acad. Sci. USA* 86:807-811.
 - [23] Krausslich, H.G., G.H. Schneider, G. Zybarth, C.A. Carter, and E. Wimmer (1989) Processing of in vitro synthesized gag precursor proteins of human immunodeficiency virus (HIV) type 1 by HIV proteinase generated in *Escherichia coli*. *J. Virol.* 62:4393-4397.
 - [24] M.I. Bukrinsky, N. Sharova, T.L. McDonald, T. Pushkarskaya, W.G. Tarpley, and M. Stevenson (1993) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc. Natl. Acad. Sci. USA* 90:6125-6130.
 - [25] M.I. Bukrinsky, S. Haggerty, M.P. Dempsey, N. Sharova, A. Adzhubei, L. Spitz, P. Lewis, D. Goldfarb, M. Emerman, and M. Stevenson (1993) A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* 365:666-669.
 - [26] M. Stevenson, personal communication.
 - [27] F. Dib-Hajj, R. Kawa, and D.P. Giedroc (1993) Retroviral nucleocapsid proteins possess potent nucleic acid strand renaturation activity. *Protein Sci.* 2:231-243.
 - [28] R. Khan and D.P. Giedroc (1992) Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA. *J. Biol. Chem.* 267:6689-6695.
 - [29] J.-L. Darlix, C. Gabus, M.-T. Nugeyre, F. Clavel, and F. Barré-Sinoussi (1990) Cis elements and trans acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. *J. Mol. Biol.* 216:689-699.
 - [30] W. Fu, R.J. Gorelick, and A. Rein (1994) Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. *J. Virol.* 68:5013-5018.

III-6
DEC 94

Figure 1. Biosynthesis of HIV-1 Virion Proteins. Steps in the biosynthesis of proteins found in mature infectious HIV-1 virion particles are outlined in a schematic diagram. Proteins are indicated in color and designated according to convention. Naturally occurring proteolytic cleavage sites are indicated by red arrows and some posttranslational modifications are indicated (glycosylation sites in gp120 and gp41 are omitted).

Figure 2. Biosynthetic Processing of HIV-1 Gag Proteins. The biosynthetic processing of the HIV-1 Gag precursor (Assemblin) and its natural proteolytic cleavage products is indicated in the schematic diagram. The upper panel shows the posttranslationally modified Gag precursor that participates in assembly and budding of the immature viral particle. The protein is depicted as organized into domains that correspond to final cleavage products, matrix antigen (MA), capsid antigen (CA), nucleocapsid (NC) protein and p6 (lower panel) found in the mature virus. The Gag precursor and MA are myristoylated (Myr-) on the N-terminal amino group of glycine (Gly-2). The precursor and CA are phosphorylated at two sites (one serine and one threonine residues). The NC domain of the Gag precursor and the NC protein bind two zinc ions to form zinc-fingers. The MA protein is not phosphorylated in the virus but is phosphorylated during the infectious process (indicated in the lower panel). The middle panel indicates the sites for proteolytic attack by the viral protease during the normal maturation process. The lower panel indicates the mature Gag proteins and also the products of partial proteolysis of p1, p2, and p6 found in the infectious virus.

Amino Acid Sequence and Related Data

first array
linker
second array

⊕ ⊕ ⊕ ⊕ ⊕
⊕ ⊕ ⊕ ⊕ ⊕
⊕ ⊕ ⊕ ⊕ ⊕

MQRGNERNORKIIKFCNCGKEG
HIAKNCRAPRKRG
CKGKEGHOMKDCTERQAN

Total Residues.....55
Basic Residues.....15
Acid Residues..... 4
Net Charge..... +11
IEP.....10.77

Molecular Weight.....6451.5
280nm Molar Absorption.6050

NC Complex With Two Zinc Ions

First Zinc Finger

Second Zinc Finger

Figure 3. HIV-1_{MN} Nucleocapsid (NC) Protein. The upper panel shows the amino acid sequence of the NC protein and indicates the distribution of charged residues. Amino acid side chains that coordinate with zinc ions are indicated as well as aromatic residues that are required for specific RNA binding. The panel also lists fundamental molecular data for the protein. The lower panel is a cartoon to indicate the zinc-finger structures formed by binding metal ions.

HIV-1 Gag Consensus Sequences

CONSENSUS.A	mGARAsvLsggkLDawekIrLRPgGkKkYrIKHlwAsreLerFaLnPslLeT?egcqqimeQlgsalkTg?e	71
CONSENSUS.B	-----e-r-----k-----i-----v-g---ts---R-lg---ps-q---s-	73
CONSENSUS.C	-----i-r-----7-----r-----h-Mi-----pg---s---k-ikq--P-Q--T-	72
CONSENSUS.D	-----?-----g---?---i-----f-l-G---s---k-ig--P-igt-s-	71
CONSENSUS.F	-----?-----?-----??-g---s---rk-Ig--ps-Q--S-	70
CONSENSUS.G	-----?-----?-----?-----?-----G---T---??-P?--Q--T-	66
CONSENSUS.H	-----?-----?-----?-----?-----?-----L-?I---P---T-	67
CONSENSUS.O	---7---T-S-----?---?---S---?-----?---C---?---A---?E?LLQ--EP-----S?	64
CONSENSUS.A	ElkSLfNtvalIyCVHqridvkdEKeA..ldkiEei.qnkskqk..tqq...aaa???????.Tg	120
CONSENSUS.B	--r--y-----e-----E-----k--a--.?--ad.....???	123
CONSENSUS.C	--r--?-----??-e-r-----e-E---?Q---??-ak?-D-----?-	117
CONSENSUS.D	e---7-----e-e-d---..e-m-E---k--?a---??-t-D.....tr	120
CONSENSUS.F	--r--?---?v--f---vE?---..7-L--E---q---?--d.....K-	115
CONSENSUS.G	-?---?--?--?--?--?---k---eEV-Ka--kn-Q---???	104
CONSENSUS.H	--Q---LL-?-?-----?---?--?--?---??Q??...-T?D.....K?	105
CONSENSUS.O	?-?-W-AI?V-W--N-??I--QQ---IQ-LK-VN.?SR-SA..?AA...KE-----S	104
p17 \ / p24		
CONSENSUS.A	nes?????..kvSqNYPIVQNaQgQm?hQ?lSPrTlNawVKvIEkaFspEVIpmFsaLSEgATpQdLmM	182
CONSENSUS.B	-----q-----nlq--V-ai-----v-----s-----T-	187
CONSENSUS.C	-----y-----L---v-ai-----f--e?ip-ft-lsegat-q-lntm	177
CONSENSUS.D	-----Q-----L---V-ai-----k-----i-----t-	184
CONSENSUS.F	-----l---V-ai-----ae-----T-	175
CONSENSUS.G	-----Q-----n---v-pis-----V-----t-	168
CONSENSUS.H	???.....?-----?v-AI-----V-----A-	164
CONSENSUS.O	?R.....Q??-?-----V-AI-----AV-----N-I---M-----?Y-I-T-	161
CONSENSUS.A	LNivGQHQAAMQMLKdtINeEAAewDR?HPVhagPippgQmREPrGSDIAGtTStlqEqigwmf...snPPiP	251
CONSENSUS.B	--T-----e-----l---v---a-----a-----n-----	257
CONSENSUS.C	Intv-ghqaamqmk--in-aaa--drhlpv-a-pva--q--e-----a-----?	246
CONSENSUS.D	--Tv---q-----E--n-----l---v---A-----?-----	253
CONSENSUS.F	--T-----L---v?-----?-----q---...v-	243
CONSENSUS.G	l-T-----l-----f--pQ-----I-d-----R-----	238
CONSENSUS.H	--?-----?-----?-----?-----A-----?--?--	229
CONSENSUS.O	--AI---G-L-V--EV-----?---T--P??--L---I---T-----Q---?--T--..R????--	224
CONSENSUS.A	VgdIYkrwIiLGLNKIVRMYSPIVSiLDiRqPKREPrdYvdrFFKcLRAeqAcQeVKnwMTeTLlVQNaNPDC	324
CONSENSUS.B	--e-----l-----t-----p---v---Y-----s---m---l-q-a---	330
CONSENSUS.C	-----k-----D-----d-----	319
CONSENSUS.D	-ge-----l-----r-----e-----Y-----s-d-----	326
CONSENSUS.F	--e-----p-----q---D-----	316
CONSENSUS.G	--e-----?-----?-----D-----	309
CONSENSUS.H	-----?-----?-----?-----D-----	298
CONSENSUS.O	----RK--V-----M-K-----?-----Y-----?-----	295
p24 \ / 'p2' \ / p7		
CONSENSUS.A	KsILraLg?gAtLeEMMTacQgVggPgHKArvLAeAmSqvq?n...???iHmQrGnf.?Gqkr?IKCFNCGk	385
CONSENSUS.B	-T--K---Pa-----t-----..t---?sat-----rn-rKtv-----	397
CONSENSUS.C	-T-----P--s-----s-----aon--...m---s---K-p--iv-----	384
CONSENSUS.D	kt--K---P?-----s-----a?..t---s-ta-m--g---K-prki-----	392
CONSENSUS.F	-T--K-l-P-----a---a..T---?a-m-ks--K--R-iv-----	381
CONSENSUS.G	-T--?---P-----?-----A..SG..?A-A?---K??..K-P??-----?	366
CONSENSUS.H	-?--?---SI-----?-----?-----T---?A--?---K--R-I?-----	356
CONSENSUS.O	-Q--K?--P?-----V-----T---??-----A?A-QDLKGGYTAVF---QN.P?R-G-----	360

Gag Proteins

```

                                pol cds ->
                                p7 \ /   'p1'   \ / p6
CONSENSUS A EGHIArNCrAPrKkGcWkCgkEGHQMkDCT ?e rQANFlgkiwpSSKG RpgNFPQsRp 441
CONSENSUS B e--i-k-----k-----k-----t --?-----h-- -p---l-----? ? ? ? ? ? ? ? 453
CONSENSUS C ---i-n-----?-----?-----s?--- -L---spTap? ? ? ? ? ? ? ? 443
CONSENSUS D ---i-k-----k-----k-----h--r---l--- 448
CONSENSUS F e--i-kn-----f-----n-- r---l--- 437
CONSENSUS G -----?-----?-----H-- -L-?-- 417
CONSENSUS H ---?-----?-----?-----L--- 406
CONSENSUS O ---I-?-----?-----Q-----? NG?-----Y--PGGT ----YV-?? 410

```

```

                                \ / (minor)   \ / (minor)   / p6 terminus (80%)
CONSENSUS A EPrAPPAR?f?gmgeezt s? ?pkqeqkd ??ke??pp1?SlKSLPGNDpLSQ 483
CONSENSUS B ?-----ees- rf---t-tPa??-q---pi- --LY?-as-rs---n--s-q 500
CONSENSUS C ?e---p---S- rF ---t- Pa ----p-??-??-ts-k----- 482
CONSENSUS D -----eS- -F-----Ps q-----??-ly --as-k-----is- 494
CONSENSUS F -----s- -F?----PS -----egly---a--- 474
CONSENSUS G -?-----gl -f-----?S ----Pr- --LY----- 451
CONSENSUS H -----S- -F---M- P- ----??- ?- ?----- 436
CONSENSUS O ?-S---M- -?VK ?Q EN-?-G --?-LY -FA-----T-Q 443

```

HIV-2/SIV Gag Consensus Sequences

```

CONSENSUS A HGAINEVLRGKKADELEK1RLRPgGKKkYrLKH1VMAANEldrFGLaESLLESKEGCKiItVLdPlVPTGSE 73
CONSENSUS B ----?-S---T-----V-----?-?-?-v---d-----?-e---h---v-a-v----- 67
CONSENSUS C -----A-----T-----E----- 36
CONSENSUS D --a-n--s--k-----m--v-----en-----s--a----- 73
CONSENSUS E -----R-----N---HR-----E--M----- 36

```

```

                                p16 \ / p28
CONSENSUS A NLkSLfNTvCV1WC1HAEEKVKDTEeAk?lvq RHLVAETgTaEKMpntSRPTAPpSgkggWfPVQq?gG?NY 142
CONSENSUS B -----t---y-L-ee-----K-A- s--?v dt--m-a-s?pt-p-- r-Y---VA- -- 131
CONSENSUS C -----Y--T-----L--Q--H---RNEV-E --A---KN-----A-----S--GR--Y-- VA- -- 107
CONSENSUS D ----y--c-----h---Q-----v---e?-k-----r--Y----- 142
CONSENSUS E -----A---VY-L-AV-----KH--QH---GGK-T--L-PQ-----G--Y---I-N -- 108

```

```

CONSENSUS A tHvPLSPrtLNAWVKLvBeKKFGAEVVPFGFQALSEGCTPYDINQmLNCVGDHOAAmQIIREIINeeAAAdWd?g 214
CONSENSUS B v-l-----ae-----c-----e---mqi-----Q- 204
CONSENSUS C V-K-----T-----L-----EH 180
CONSENSUS D v L-----l-----p-----t-y-----e-----re----- l- 215
CONSENSUS E V-S-----E-----V- 181

```

```

CONSENSUS A HPIp GPLPAGQLRdPRGSDIAGTTSTveEQIqWmfrpgNpVpVGNiYRRW1QIGLQKCVrMYNPTN1LD1kQ 286
CONSENSUS B --?- -l-----q---y-a-n-----l-----p-----q 275
CONSENSUS C --V-----Y-A-----L-----V- 252
CONSENSUS D -pQp??-q??-e-s-----tvd---y-Q---l-----r-----L-----tn---v-- 284
CONSENSUS E --RG QP--QG --S-----PA--E -Y-NP I--D-----L-----V-- 252

```

```

CONSENSUS A GPKEpFQSYVDRFYKSLRABQTDpAVKMMTQTLL?QnanPDCKLVKGLGmNPTLEMLTaCQGVGGpQKA 358
CONSENSUS B -----s-----I-----I--- 348
CONSENSUS C ---- 260
CONSENSUS D -----t-av-- --q-l f--- --g?----- g --- a 356
CONSENSUS E - -S--- 260

```

```

                                p28 \ /   p2   \ / p8   pol cds >
CONSENSUS A RLMAEALKEam?PaPIPFAAAOQ r?aikcWNCQKEGHSArQcrAPRRQGCWKCCK?GH1ManCP?RQAG 423
CONSENSUS B -----LT--p1-----a-qkAGk-gTVT- ---? T-----kq ---SK -E---- 420
CONSENSUS D -----e-lap?pl--a-a q?Qp? kp--w g-e -s---- a---q -----kmg-v--K --- a 425

```

```

                                p8 \ /   p1   \ / p6   \ / HIV2s only
CONSENSUS A FLG1GpWGKkpRNFPvq?pqGL tPTAPP?DPavdLLEKYMqQG?rQRE Q R 471
CONSENSUS B -----mt-v --vtPSAPPmnpAeGmTPrGa- S- -A- --EN-Ks--qm rqqre? ? ? ? s 489
CONSENSUS D -----r- M---h- ? ? ? ? ? ? ? ? -p p-ed ----kn--ql kq-- ? ? ? ? sr 475

CONSENSUS A eRPYKEVTED1LhiEqgetPhRe?tTEDLHLNLSLFGkDQ 510
CONSENSUS B -- - --E- 511
CONSENSUS D k - v -- i- ? 496

```


HIV-1 Tat: Structure and Function

Kuan-Teh Jeang

Chief, Molecular Virology Section

LMM, NIAID, National Institutes of Health, Bethesda, MD 20892-0460

INTRODUCTION

Most viruses encode functions used to regulate genome transcription. Examples include the SV40 T-antigen, the adenovirus E1a protein, and the herpes virus immediate-early proteins. For the human immunodeficiency viruses (HIV), Tat functions similarly, though not identically, to those above. Over the past decade, we have learned much about Tat, both in structure and in function. While a goal of this Los Alamos database is to completely compile relevant tat sequences, the intent of this brief accompanying overview is not similarly archival, but rather to "add flavor" to raw data. It is written for the purpose of apprising, in a short format, the readers on some of the current thoughts about Tat. For more in-depth discussions, complete subject reviews can be found elsewhere (e.g. Felber and Pavlakis, 1993; Jeang and Gatignol, 1994).

TAT FUNCTION

Tat is a small nuclear protein of 86 to 101 amino acids in size (depending on the viral strain) which is encoded from two separate exons (see Section I). Analyses of "full-length" Tat have been performed commonly using the 86 amino acid version. However, it should be noted that while many laboratory strains (e.g. HXB2 and NL4-3) have the smaller Tat (86 aa) most HIV-1s have the 101 aa protein (see compendium Part II).

A. Transcription. Despite intensive efforts, the mechanism of Tat action remains incompletely understood. It is, however, accepted that Tat is required for optimal HIV viability (Fisher et al., 1986; Dayton et al., 1986). Tat's role in critically directing transcription (Peterlin et al., 1986; Rice and Mathews, 1988; Laspias and Mathews, 1989) from the HIV LTR is one indispensable function suggested for this protein. However, increasingly there is evidence that Tat has other important effects on the virus and on the host cell (see Concluding Perspectives, below).

In activating transcription from the LTR, Tat differs from other prototypic viral transcription trans-activators in requiring a bipartite responsive element consisting of DNA and RNA. To our knowledge, Tat is the first characterized eukaryotic transcription factor that binds to a nascent leader RNA, TAR (Berkhout et al., 1989; Dingwall et al., 1989; Cordingley et al., 1990; Roy et al., 1990; Calnan et al., 1991), and then influences formative events at the TATAA-enhancer-promoter (Berkhout et al., 1990; Selby and Peterlin, 1990; Southgate et al., 1990; reviewed in Jeang et al., 1991; see Fig. 1A). TAR RNA has extensive secondary structure containing a stem, a bulge, and a loop (Muesing et al., 1987; Berkhout and Jeang, 1989; Roy et al., 1990c; see Fig. 1B). Earlier, studies have indicated that the specific UCU sequence of the bulge is critical for binding by Tat (Dingwall et al., 1989; Roy et al., 1990a; Calnan et al., 1991; Cordingley et al., 1990). In comparison, the structure of the stem, but not its specific sequence, was proposed to be important for function. However, more recent findings reveal that both for Tat binding and for transcription there are sequence specific requirements for the immediate stem nucleotide pairs that flank the bulge (Weeks and Crothers, 1991; Berkhout and Jeang, 1991; Churcher et al., 1993). The loop of TAR RNA is an important binding site for cellular factor(s) that cooperate with Tat in the activation of the LTR (Sheline et al., 1991; Wu et al., 1991). A more extensive discussion of the role of cellular factors that bind TAR RNA (Gatignol et al., 1989; Gaynor et al., 1989; Gatignol et al., 1991) is presented in the compendium Part IV.

While it is clear that Tat binds TAR RNA and interacts with enhancer-promoter-binding factors (Berkhout and Jeang, 1992), the mechanism by which these physical events influence transcription

are not wholly evident. A number of models that attempt to explain the transcription function of Tat have been proposed. These include i) anti-terminating (Kao et al., 1987) stalled RNA polymerase II (RNAP II), ii) increasing processivity/elongation of transcribing RNAPII complex (Laspi et al., 1989; Marciniak et al., 1990), and iii) facilitating initiation of RNAPII complexes at the promoter (Laspi et al., 1989; Jeang and Berkhout, 1992; Jeang et al., 1993a; reviewed in Cullen, 1993). Currently, there are conflicting findings that support each of the models. It is possible that Tat, like basal transcription factor TFIIF (Boratowski, 1994), functions simultaneously as an initiation and as an elongation factor. Alternatively, conflicting attributions could arise from differences in the rate-limiting events (Jeang et al., 1993a) for the various experimental systems used to define Tat action.

A couple of observations indicate that Tat is not a typical transcriptional trans-activator. First, unlike other activators (Tjian and Maniatis, 1994), a modular activation domain has been difficult to define for HIV-1 Tat (see below, Activity domains of Tat). Second, in different settings, neither Tat nor Tat-chimera, in contrast to modular activators, is able to directly activate transcription from a minimal TATA-promoter (Southgate and Green, 1991; Berkhout and Jeang, 1992; Kamine and Chinnadurai, 1992). Instead for physiological activity, Tat requires the presence of an upstream enhancer-binding factor such as Sp1 (Harrieh et al., 1989; Kamine et al., 1991; Southgate and Green, 1991; Zimmermann et al., 1991; Jeang and Berkhout, 1992; Jeang et al., 1993b). Thus it appears that Tat is more aptly classified with transcriptional co-activators (Dymlach et al., 1991) in bridging conformationally a stereospecific architecture (Tjian and Maniatis, 1994) between Sp1 (Jones and Tjian, 1985; Jones et al., 1986) and TATA-binding protein (TBP; Jeang et al., 1993b; Huang et al., 1993; Kashanchi et al., 1994).

B. Translation. There is evidence that Tat also functions in regulating translation (Rosen et al., 1986; Cullen, 1987). Indeed there are findings that TAR RNA can inhibit translation (Parkin et al., 1988; SenGupta and Silverman, 1989) of HIV-1 mRNAs, most likely through activation of double-stranded RNA-dependent protein kinase and 2-5A synthetase (SenGupta and Silverman, 1989; Edery et al., 1989). Addition of Tat was found to reverse this translational inhibition (SenGupta et al., 1990; Braddock et al., 1990).

ACTIVITY DOMAINS IN TAT

Tat is synthesized from an mRNA joined from two coding exons. The first exon codes for amino acids 1-72, while (in most strains of HIV-1) the second exon codes for amino acids 73-101 (see Fig. 2). In most functional assays, the first 72 amino acids of Tat fully effect transcriptional trans-activation of the LTR. In fact, a truncated 58 amino acid form of Tat is virtually completely active in co-transfection assays (Seigel et al., 1986; Garcia et al., 1988; Kuppuswamy et al., 1989).

A. First coding exon. The combined results from many laboratories have permitted an arbitrary demarcation of "domains" in Tat (Kuppuswamy et al., 1989). For instance, the N-terminus of Tat (domain 1; Fig. 2) has 13 amino acids with amphipathic characteristics. Mutations that alter the acidic composition of this region were felt originally to affect trans-activation (Rappaport et al., 1989); however, results from a later study conflicted with this interpretation (Tiley et al., 1990).

Amino acids 22 to 37 (domain 2, Fig. 2) contain seven cysteines and are highly conserved between different isolates of HIV-1s, group O as well as group M. Individual mutation in six of the seven cysteines abolish Tat function (see Table I). Although originally proposed as a metal-chelating dimerization domain (Frankel et al., 1988), this region was recently shown to be used for intramolecular disulfide bond formation in monomeric Tat proteins found within cells (Koken et al., 1994). Currently, it is believed that Tat is active functionally as a monomer rather than a dimer (Rice and Chan, 1991; Koken et al., 1994).

Domain 3 (amino acids 40 to 48) contains a RKGLGI motif that is conserved between HIV-1, HIV-2 and SIV Tat. This region, in conjunction with the amino terminus and the cysteine domain, has been suggested to circumscribe the minimal activation domain of HIV-1 Tat (Carroll et al., 1991; Derse et al., 1991). Domain 4 (amino acids 49-72) contains a basic RKKRRQRRR motif. These amino acids confer TAR RNA-binding properties to Tat (Dingwall et al., 1989; Roy et al., 1990;

Weeks et al., 1990; Chang and Jeang, 1992) and are important for nuclear localization of the protein (Ruben et al., 1989; Hauber et al., 1989). However, recent studies suggest that this short basic domain is insufficient in determining the entire specificity of Tat-TAR binding since amino acids outside of the domain also contribute to this interaction (Churcher et al., 1993; Luo et al., 1993).

Table I summarizes 75 point mutations in the first coding exon of Tat collated from the work of eight laboratories (Garcia et al., 1988; Sadaie et al., 1989; Kuppuswamy et al., 1989; Ruben et al., 1989; Hauber et al., 1989; Meyerhans et al., 1989; Rice and Carlotti, 1990a; Rice and Carlotti, 1990b; Siderovski et al., 1992). As alluded to above, since the first 58 amino acids of Tat recapitulate well the trans-activation function of the whole protein, it is not surprising that most of the mutations are clustered within amino acids 1 to 58. In many cases, individual amino acids have been changed to more than one counterpart. This heightens the validity of the resulting phenotype. Examination of these mutants reveals that the region spanning amino acids 1–21 is remarkably tolerant of changes. In contrast, changes in amino acids 22 through 40 were generally deleterious for trans-activation. Finally, although the basic domain (amino acids 49–57) as a unit is necessary for Tat function, individual amino acid changes do not significantly affect activity.

B. Second coding exon. Less information is available about the second coding exon of Tat. It is clear that in routine transfection assays of HIV-1 Tat, absence of the second exon does not alter greatly Tat activity. However, findings from HIV-2 and SIV Tat are quite clear in demonstrating that this exon contributes towards optimal trans-activation (Viglianti and Mullin, 1988; Tong-Starksen et al., 1993). Recently, there have been suggestions that the second exon of HIV-1 Tat, in specific assays, is important for trans-activation (Jeang et al., 1993b) and for trans-repression (Howcroft et al., 1993).

There are two short motifs in the second exon of HIV-1 Tat that could have functional importance (see Fig. 2). The first is an RGD sequence that is used as a cell adhesion signal for binding to cellular integrins (Brake et al., 1990). This RGD motif is not found, however, in HIV-2 or SIV Tat proteins. In addition, a ESKKKVE motif is conserved across different HIV-1 Tat proteins and is partially preserved in HIV-2 and SIV Tats. However, because this motif falls beyond amino acid 86, commonly regarded as the C-terminal boundary of full-length Tat, its functional significance has not been examined.

CONCLUDING PERSPECTIVES

While research on Tat has been narrowly focused upon its trans-activation properties for the HIV LTR-promoter, there is increasing evidence that Tat has other pleiotropic effects on cellular genes, host cell metabolism, and viral infectivity/pathogenesis (see e.g. Drysdale and Pavlakis, 1991). For instance, Tat is reported to function as a secreted growth factor in stimulating the growth of Kaposi-like cells (Ensoli et al., 1990; Ensoli et al., 1993; Barillari et al., 1993). It can affect the organization of neurons and astrocytes (Kolson et al., 1993) and is neurotoxic at low concentrations (Sabatier et al., 1991). Some of these findings are partly explained by the ability of Tat to modulate expression of cellular genes (Roy et al., 1990b; Buonaguro et al., 1994; Scala et al., 1994), to perturb ambient levels of cytokines (Rautonen and Rautonen, 1992), and to protect cells from programmed cell death (Zauli et al., 1993).

Because the viral functions of Tat have been largely extrapolated from subgenomic experiments, one is unsure whether these findings reflect well the complete role of Tat in the setting of the whole virus. The limited number of experiments performed on Tat function in the context of infectious virions suggests that Tat has additional roles, beyond transcription, in affecting viral pathogenicity (Cheng-Mayer et al., 1991; Sakai et al., 1992; Huang et al., 1994).

Fig. 1. Interaction of Tat with DNA and RNA targets in the HIV-1 LTR. A) A schematic representation of the functional interactions between Tat, TAR-RNA-binding proteins and promoter elements. Biochemical evidence exists that Tat contacts directly SP1 (Jeang et al., 1993b) and TATAA-binding protein (TBP; Kashanchi et al., 1994). B) Secondary structure of TAR RNA. The crucial trinucleotide bulge and hexanucleotide loop elements are boxed.

Fig. 2. Domain classifications of Tat protein. The demarcation of domains is somewhat arbitrary. The first exon includes amino acids 1-72, while the second exon includes 73-101. Motifs and characteristics of each "domain" are indicated above or below each region.

Table I. Point mutations in Tat

Original Amino Acids(s) and Their Position(s)	Mutant Amino Acid(s)	Resulting Activity	Consensus Amino Acid(s)
Q2	A	++	E
P3	A	++	P
P3	Q	++	P
V4	A	++	V
D5	A	+	D
P6	A	++	P
P6	S	++	P
Δ3-6		+	
P6P10	LL	++	
R7	A	++	R
L8	A	++	L
E9	A	++	E
P10	A	++	P
P10P13	LL	++	
W11	A	++	W
K12	A	++	K
K12	N	++	K
P18	A	++	P
K19	R	++	K
A21	D	++	A
A21T23	VA	+	
T23	A	++	T
C22	S	-	C
C22	G	-	C
N23	T	++	T
N24	A	++	N
N24	K	++	N
C25	R	-	C
C25	G	-	C
Y26	A	+	Y
Y26	F	++	Y
C27	S	-	C
C27	G	-	C
K28K29	AA	+	KK
K28K29	EA	-	KK
C30	G	-	C
C31	S	++	C
C31	E	-	C
C31	G	++	C
F32	A	+	F
H33	A	-	H
C34	G	-	
C34	S	-	C
G35	A	+	Q

Key: Column 3, ++ > 50% wild type activity; + > 10% wild type activity; +/- or - indicate <10% wild type activity.

Table 1. (cont.) Point mutations in Tat

Original Amino Acids(s) and Their Position(s)	Mutant Amino Acid(s)	Resulting Activity	Consensus Amino Acid(s)
C37	G	-	C
C37	S	-	C
F38	A	-	F
F38	L	++	F
K40	D	-	T
K40	T	++	T
K41	A	-	K
K41*	T	-	K
K41*	T	++	K
L43	F	+	L
G44	S	++	G
S46	A	++	S
S46	P	-	S
Y47	H	++	Y
Y47	A	++	Y
G48	S	++	G
G48R49†	SG	++	
R49	T	++	R
K50	stop	-	K
K50K51	Y50Y51	+	KK
K50K51	S50G51	+/+	
K50	E	++	K
K50	T	++	K
R52	E	++	R
R53	I	++	R
Q54	N	++	Q
R55	G	+	R
R55R56	L55T56	+	
R56	E	++	R
R57Q63	SE	++	RQ
L69	I	++	L

Key: Column 3, ++ > 50% wild type activity; + > 10% wild type activity; +/- or - indicate <10% wild type activity.

* Different results reported for the same mutation from Kuppuswamy et al., 1989, and Myerhans et al., 1989.

† Amino acids beyond position 59 completely changed.

References

- [1] Barillari, G., R. Gendelman, R. C. Gallo, and B. Ensoli. 1993. The tat protein of human immunodeficiency virus type 1, a growth factor for AIDS Kaposi sarcoma and cytokine-activated vascular cells, induces adhesion of the same cell types by using integrin receptors recognizing the RGD amino acid structure. *Proc. Natl. Acad. Sci. USA* 90:7941-7945.
- [2] Berkhout, B., A. Gaignol, A. B. Rabson, and K.-T. Jeang. 1990. TAR-independent activation of the HIV-1 LTR: evidence that Tat requires specific regions of the promoter. *Cell* 62:7257-7267.
- [3] Berkhout, B., and K.-T. Jeang. 1989. Trans-activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. *J. Virol.* 63:5501-5504.
- [4] Berkhout, B., and K.-T. Jeang. 1991. A detailed mutational analysis of TAR RNA; critical spacing between the bulge and loop recognition domains. *Nucleic Acids Res.* 19:6169-6176.
- [5] Berkhout, B., and K.-T. Jeang. 1992. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the Human Immunodeficiency virus type 1 long terminal repeat. *J. Virol.* 66:139-149.
- [6] Berkhout, B., R. H. Silverman, and K.-T. Jeang. 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* 59:273-282.
- [7] Braddock, M., A. Thorburn, A. Chambers, A. Kingsman, and S. Kingman. 1990. A nuclear translational block imposed by the HIV-1 U3 region is relieved by the Tat-TAR interaction. *Cell* 62:1123-1133.
- [8] Brake, D. A., C. Debouk, and G. Biesecker. 1990. Identification of an Arg-Gly-Asp (RGD) cell adhesion site in human immunodeficiency virus type 1 transactivation protein tat. *J. Cell Biol.* 111:1275-1281.
- [9] Buratowski, S. 1994. The basics of basal transcription by RNA polymerase II. *Cell* 77:1-3.
- [10] Buonajuro, L., F. Buonajuro, G. Giraldo, and B. Ensoli. 1994. The human immunodeficiency virus type 1 tat protein transactivates tumor necrosis factor beta gene expression through a TAR-like structure. *J. Virol.* 68:2677-2682.
- [11] Calnan, B. J., B. Tidor, S. Biancalana, D. Hudson, and A. D. Frankel. 1991. Arginine-mediated RNA recognition: the arginine fork. *Science* 252:1167-1171.
- [12] Carroll, R., L. Martarano, and D. Derse. 1991. Identification of lentivirus Tat functional domains through generation of equine infectious anemia virus/human immunodeficiency virus type 1 tat gene chimera. *J. Virol.* 65:3460-3467.
- [13] Chang, Y.-N., and K.-T. Jeang. 1992. The basic RNA-binding domain of HIV-2 Tat determines preferential trans-activation of a TAR2 containing LTR. *Nuc. Acids Res.* 20:5465-5472.
- [14] Cheng-Mayer, C., T. Shioda, and J. A. Levy. 1991. Host range, replicative, and cytopathic properties of human immunodeficiency virus type 1 are determined by very few amino acid changes in tat and gp120. *J. Virol.* 65:6931-6941.
- [15] Churcher, M., C. Lamont, F. Hamy, C. Dingwall, S. Green, A. Lowe, J. Butler, M. Gait, and J. Karn. 1993. High affinity binding of TAR RNA by the human immunodeficiency virus type 1 tat protein requires base-pairs in the RNA stem and amino acids residues flanking the basic region. *J. Mol. Biol.* 230:90-110.
- [16] Cordingley, M. G., R. L. LaFemina, P. L. Callahan, J. H. Condra, V. V. Sardana, D. J. Graham, T. M. Nguyen, K. LeGrow, L. Gotlib, A. J. Schlachet, and R. J. Colonno. 1990. Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive sequence element of human immunodeficiency virus type 1 in vitro. *Proc. Natl. Acad. Sci. USA* 87:8985-8989.
- [17] Cullen, B. R. 1986. Trans-activation of human immunodeficiency virus occurs via a bimodal mechanism. *Cell* 46:973-982.
- [18] Cullen, B. R. 1993. Does HIV-1 Tat induce a change in viral initiation rights? *Cell* 73:417.

- [19] Dayton, A., J. Sodroski, C. Rosen, W. Goh, and W. Haseltine. 1986. The transactivator gene of the human T cell lymphotropic virus type III is required for replication. *Cell* 44:941-947.
- [20] Derse, D., M. Carvalho, R. Carroll, and B. M. Peterlin. 1991. A minimal lentivirus tat. *J. Virol.* 65:7012-7015.
- [21] Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Kam, A. D. Lowe, M. Singh, M. A. Skinner, and R. Vallerio. 1989. Human immunodeficiency virus 1 tat protein binds transactivation responsive region (TAR) RNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 86:6925-6929.
- [22] Drysdale, C. M., and G. N. Pavlakis. 1991. Rapid activation and subsequent down-regulation of the human immunodeficiency virus type 1 promoter in the presence of Tat: Possible mechanisms contributing to latency. *J. Virol.* 65:3044-3051.
- [23] Dynlacht, B. D., T. Hoey, and R. Tjian. 1991. Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* 66:563-576.
- [24] Edery, I., R. Petryshyn, and N. Sonenberg. 1989. Activation of double-stranded RNA-dependent kinase (dsI) by the TAR region of HIV-1 mRNA: a novel translational control mechanism. *Cell* 56:303-312.
- [25] Ensoli, B., G. Barillari, S. Z. Salahuddin, R. C. Gallo, and F. Wong-Staal. 1990. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* 345:84-86.
- [26] Ensoli, B., L. Buonaguro, G. Barillari, V. Fiorelli, R. Gendelman, R. A. Morgan, P. Wingfield, and R. C. Gallo. 1993. Release, uptake, and effects of extracellular human immunodeficiency virus type 1 tat protein on cell growth and viral transactivation. *J. Virol.* 67:277-287.
- [27] Felber, B., and G. Pavlakis. 1993. Molecular biology of HIV-1: positive and negative regulatory elements important for virus expression. *AIDS* 7:S51-62.
- [28] Fisher, A., M. Feinberg, S. Josephs, M. Harper, L. Marselle, G. Reyes, M. Gonda, A. Aldovini, C. Debouk, R. C. Gallo, and F. Wong-Staal. 1986. The transactivator gene of HTLV-III is essential for virus replication. *Nature* 320:367-371.
- [29] Frankel, A. D., D. S. Brecht, and C. O. Pabo. 1988. Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science* 240:70-73.
- [30] Garcia, J. A., D. Harrich, L. Pearson, R. Mitsuyasu, and R. Gaynor. 1988. Functional domains required for tat-induced transcriptional activation of the HIV-1 long terminal repeat. *Embo J.* 7:3143-3147.
- [31] Gatignol, A., C. Buckler, and K.-T. Jeang. 1993. Relatedness of an RNA binding motif in HIV-1 TAR RNA-binding protein TRBP to human P1/dsI kinase and *Drosophila* Staufen. *Mol. Cell. Biol.* 13:2193-2202.
- [32] Gatignol, A., A. Buckler-White, B. Berkhout, and K.-T. Jeang. 1991. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* 251:1597-1600.
- [33] Gaynor, R., E. Soultanakis, M. Kuwabara, J. Garcia, and D. S. Sigman. 1989. Specific binding of a HeLa cell nuclear protein to RNA sequences in the human immunodeficiency virus transactivating region. *Proc. Natl. Acad. Sci. U.S.A.* 86:4858-4862.
- [34] Harrich, D., J. Garcia, F. Wu, R. Mitsuyasu, J. Gonzalez, and R. B. Gaynor. 1989. Role of Sp1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* 63:2585-2591.
- [35] Hauber, J., M. H. Malim, and B. R. Cullen. 1989. Mutational analysis of the conserved basic domain of human immunodeficiency virus tat protein. *J. Virol.* 63:1181-1187.
- [36] Howcroft, T. K., K. Strebel, M. A. Martin, and D. S. Singer. 1993. Repression of MHC class 2 gene promoter activity by 2 exon Tat of HIV. *Science* 260:1320-1323.
- [37] Huang, L. M., and K.-T. Jeang. 1993. Increased spacing between Sp1 and TATAA renders HIV-1 replication defective: Implication for Tat function. *J. Virol.* 67:6937-6944.
- [38] Huang, L. M., A. Joshi, R. Willey, J. Orenstein, and K. T. Jeang. 1994. Human immunodeficiency

- viruses regulated by alternative trans-activators: genetic evidence for a novel non-transcriptional function of Tat in virion infectivity. *Embo J.* 13:in press.
- [39] Jeang, K.-T., and B. Berkhout. 1992. Kinetics of HIV-1 LTR trans-activation: use of intragenic ribozyme to assess rate limiting steps. *J. Biol. Chem.* 267:17891-17899.
- [40] Jeang, K.-T., B. Berkhout, and B. Dropic. 1993. Effects of integration and replication on the transcription of the HIV-1 LTR. *J. Biol. Chem.* 268:24940-24949.
- [41] Jeang, K.-T., Y. N. Chang, B. Berkhout, M.-L. Hammarikjold, and D. Rekosh. 1991. Regulation of HIV expression : mechanisms of action of Tat and Rev. *AIDS 1991. A year in Review. AIDS* 5:S3-S14.
- [42] Jeang, K.-T., R. Chun, N. H. Lin, A. Gatignol, C. G. Glabe, and H. Fan. 1993. In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor. *J. Virol.* 67:6224-6233.
- [43] Jeang, K.-T., and A. Gatignol. 1994. Comparisons of regulatory features among primate lentiviruses. *Current Topics in Microbiology and Immunology*, in press.
- [44] Jones, K. A., J. T. Kadonaga, P. A. Luciw, and R. Tjian. 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science* 232:755-759.
- [45] Jones, K. A., and R. Tjian. 1985. Sp1 binds to promoter sequences and activates herpes simplex virus "immediate-early" gene transcription in vitro. *Nature* 317:179-182.
- [46] Kamine, J., and G. Chinnadurai. 1992. Synergistic activation of human immunodeficiency virus type 1 promoter by the viral tat protein and cellular transcription factor Sp1. *J. Virol.* 66:3926-3932.
- [47] Kamine, J., T. Subramanian, and G. Chinnadurai. 1991. Sp1-dependent activation of a synthetic promoter by human immunodeficiency virus type I Tat protein. *Proc. Natl. Acad. Sci. USA* 88:8510-8514.
- [48] Kao, S. Y., A. E. Calman, P. A. Luciw, and B. M. Peterlin. 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature* 330:489-493.
- [49] Kashanchi, F., G. Piras, M. F. Radonovich, J. F. Duvall, R. Roeder, and J. N. Brady. 1994. Direct interaction of human TFIIID with the HIV-1 transactivator tat. *Nature* 367:295-299.
- [50] Koken, S. E., A. E. Greijer, K. Verhoef, J. vanWamel, and B. Berkhout. 1994. Intracellular analysis of in vitro modified HIV tat protein. *J. Biol. Chem.* 269:8366-8375.
- [51] Kolson, D., J. Buchhalter, R. Collman, B. Hellmig, C. F. Farrell, C. Debouk, and F. Gonzalez-Scarano. 1993. HIV-1 tat alters normal organization of neurons and astrocytes in primary rodent brain cell cultures: RGD sequence dependence. *AIDS Res. and Human Retroviruses* 9:677-685.
- [52] Kuppuswamy, M., T. Subramanian, A. Srinivasan, and G. Chinnadurai. 1989. Multiple functional domains of Tat, the trans-activator of HIV-1, defined by mutational analysis. *Nuc. Acids Res.* 17:3551-3561.
- [53] Laspija, M. F., A. P. Rice, and M. B. Mathews. 1989. HIV-1 tat protein increases transcriptional initiation and stabilizes elongation. *Cell* 59:283-292.
- [54] Laspija, M. F., A. P. Rice, and M. B. Mathews. 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* 59:283-292.
- [55] Luo, Y., S. J. Madore, T. G. Parslow, B. R. Cullen, and B. M. Peterlin. 1993. Functional analysis of interactions between Tat and the trans-activation response element of human immunodeficiency virus type 1 in cells. *J. Virol.* 67:5617-5622.
- [56] Marciniak, R. A., B. J. Calnan, A. D. Frankel, and P. A. Sharp. 1990. HIV-1 Tat protein trans-activates transcription in vitro. *Cell* 63:791-802.
- [57] Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfeldt-Manson, B. Asjo, and S. Wain-Hobson. 1989. Temporal fluctuation in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* 58:901-910.
- [58] Muesing, M., D. Smith, and D. Capon. 1987. Regulation of mRNA accumulation by human

- immunodeficiency virus trans-activator protein. *Cell* 48:691-701.
- [59] Parkin, N. T., E. Cohen, A. Darveau, C. Rosen, W. Haseltine, and N. Sonenberg. 1988. Mutational analysis of the 5' noncoding region of human immunodeficiency virus type 1: effects of secondary structure on translation. *Embo J.* 7:2831-2837.
- [60] Peterlin, B. M., P. A. Luciw, P. J. Barr, and M. D. Walker. 1986. Elevated levels of mRNA can account for the trans-activation of human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* 83:9734-9738.
- [61] Rappaport, J., S. J. Lee, K. Khalili, and F. Wong-Staal. 1989. The acidic amino-terminal region of HIV-1 tat protein constitutes an essential activating domain. *New Biologist* 1:101-110.
- [62] Rautonen, J., and N. Rautonen. 1992. Tat and Kawasaki disease. *Immunology Today* 13:190-191.
- [63] Rice, A., and M. B. Mathews. 1988. Transcriptional but not translational regulation of HIV-1 by the tat gene product. *Nature* 332:551-553.
- [64] Rice, A. P., and F. Carlotti. 1990. Mutational analysis of the conserved cysteine-rich region of the human immunodeficiency virus type 1 tat protein. *J. Virol.* 64:1864-1868.
- [65] Rice, A. P., and F. Carlotti. 1990. Structural analysis of wild-type and mutant human immunodeficiency virus type 1 Tat proteins. *J. Virol.* 64:6018-6026.
- [66] Rice, A. P., and F. Chan. 1991. Tat protein of human immunodeficiency virus type 1 is a monomer when expressed in mammalian cells. *Virology* 185:451-454.
- [67] Rosen, C. A., J. G. Sodroski, W. C. Goh, A. I. Dayton, J. Lippke, and W. A. Haseltine. 1986. Post-transcriptional regulation accounts for the trans-activation of the human T-lymphotropic virus type III. *Nature* 319:555-559.
- [68] Roy, S., U. Delling, C.-H. Chen, C. A. Rosen, and N. Sonenberg. 1990. A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev.* 4:1365-1373.
- [69] Roy, S., M. Katze, N. Parkin, I. Edery, and N. Sonenberg. 1990. Control of the interferon-induced 68-kilodalton protein kinase by the HIV-1 tat gene product. *Science* 247:1216-1219.
- [70] Roy, S., N. T. Parkin, C. Rosen, J. Itovitch, and N. Sonenberg. 1990. Structural requirements for trans-activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by Tat: importance of base pairing, loop sequence, and bulges in the Tat-responsive sequence. *J. Virol.* 64:1402-1406.
- [71] Ruben, S., A. Perkins, R. Purcell, K. Joung, R. Sia, R. Burghoff, W. A. Haseltine, and C. A. Rosen. 1989. Structural and functional characterization of human immunodeficiency virus tat protein. *J. Virol.* 63:1-8.
- [72] Sabatier, J., E. Vives, K. Mabrouk, A. Benjouad, H. Rochat, A. Duval, B. Hue, and E. Bahraoui. 1991. Evidence for neurotoxic activity of tat from human immunodeficiency virus type 1. *J. Virol.* 65:961-967.
- [73] Sadaie, M. R., J. Rappaport, T. Benter, S. F. Josephs, R. Willis, and F. Wong-Staal. 1988. Missense mutations in an infectious human immunodeficiency viral genome: functional mapping of tat and identification of the rev splice acceptor. *Proc. Natl. Acad. Sci. USA* 85:9224-9228.
- [74] Sakai, H., J. Sakuragi, S. Sayuri, R. Shibata, and A. Adachi. 1992. Functional analysis of biologically distinct genetic variants of simian immunodeficiency virus isolated from a mandrill. *Virology* 189:161-166.
- [75] Scala, G., M. R. Ruocco, C. Ambrosino, M. Mallardo, V. Giordano Baldassarre, F., E. Drag-onetti, I. Quinto, and S. Venuta. 1994. The expression of the interleukin 6 gene is induced by the human immunodeficiency virus 1 tat protein. *J. Exp. Med.* 179:961-971.
- [76] Seigel, L. J., L. Ratner, S. F. Josephs, D. Derse, M. Feinberg, G. A. Reyes, S. J. O'Brien, and F. Wong-Staal. 1986. Trans-activation induced by human T-lymphotropic virus type III (HTLV-III) maps to a viral sequence encoding 58 amino acids and lacks tissue specificity. *Virology*

148:226-231.

- [77] Selby, M. J., and B. M. Peterlin. 1990. Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. *Cell* 62:769-776.
- [78] SenGupta, D. N., B. Berkhout, A. Gaignol, A. Zhou, and R. H. Silverman. 1990. Direct evidence for translation regulation by leader RNA and Tat protein in human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* 87:7492-7496.
- [79] SenGupta, D. N., and R. H. Silverman. 1989. Activation of interferon-regulated, dsRNA-dependent enzymes by HIV-1 leader RNA. *Nucl. Acids. Res.* 17:969-978.
- [80] Sheline, C. T., L. H. Milocco, and K. A. Jones. 1991. Two distinct nuclear transcription factors recognize loop and bulge residues of the HIV-1 TAR RNA hairpin. *Genes Dev.* 5:2508-2520.
- [81] Siderovski, D. P., T. Matsuyama, E. Frigerio, S. Chui, X. Min, H. Erfle, M. Sumner-Smith, R. W. Barnett, and T. W. Mak. 1992. Random mutagenesis of the human immunodeficiency virus type-1 trans-activator of transcription (HIV-1 Tat). *Nuc. Acids Res.* 20:5311-5320.
- [82] Southgate, C., M. L. Zapp, and M. R. Green. 1990. Activation of transcription by HIV-1 Tat protein tethered to nascent RNA through another protein. *Nature* 345:640-642.
- [83] Southgate, C. D., and M. R. Green. 1991. The HIV-1 tat protein activates transcription from an upstream DNA-binding site: implications for tat function. *Genes Dev.* 5:2496-2507.
- [84] Tiley, L. S., P. H. Brown, and B. R. Cullen. 1990. Does the human immunodeficiency virus tat transactivator contain a discrete activation domain? *Virology* 178:560-567.
- [85] Tjian, R., and T. Maniatis. 1994. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 77:5-8.
- [86] Tong-Starksen, S., A. Baur, X. B. Lu, E. Peck, and B. M. Peterlin. 1993. Second exon of Tat of HIV-2 is required for optimal transactivation of HIV-1 and HIV-2 LTRs. *Virology* 195:826-830.
- [87] Viglianti, G., and J. L. Mullins. 1988. Functional comparison of transactivation by simian immunodeficiency virus from rhesus macaques and human immunodeficiency virus type 1. *J. Virol.* 62:4523-4532.
- [88] Weeks, K. M., and D. M. Crothers. 1991. RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* 66:577-588.
- [89] Wu, F., J. Garcia, D. Sigman, and R. Gaynor. 1991. tat regulates binding of the human immunodeficiency virus trans-activating region RNA loop-binding protein TRP-185. *Genes Dev.* 5:2128-2140.
- [90] Zauli, G., D. Gibellini, D. Milani, M. Massoni, P. Borgatti, M. Laplaca, and S. Capitani. 1993. Human immunodeficiency virus type 1 tat protein protects lymphoid, epithelial and neuronal cell lines from death by apoptosis. *Cancer Res.* 53:4481-4485.

Tat Structure and Function

HIV-1 Tat Consensus Sequences

```

                                \3'sj \3'sj
                                rev cds ->
CONSENSUS.A M?PVDpNLEFWnHPGSqPTTaCskCYCK?CCwHCqICFLnKGLGISYGrKKR...r?RRqtPQ.s?kDhQnp 64
CONSENSUS.B -e---r---k-----kt-ctn---k--f---v---tt-g-g-----...-Q-r rapg.dSqt--vs 68
CONSENSUS.C m--v--?-----s--K---t--yc-k-sY--lv--gt-----...-q--sa-?-SE----- 65
CONSENSUS.D md-v---i-p-----r-p-N--?-K--Y--v--it-----...-Q--rppq.g-Qa--v- 66
CONSENSUS.O -D---E?P--H---?-Q?P--NN---R--Y--YV--??-----?-.....?-PAAA?-P-?KD- 57
CONSENSUS.CPZ -D-?-????-?-??-?-?-NN-----Y--?-?-TK-----?-???--T???-?S?NN-D? 45

                                exon \ / exon
CONSENSUS.A ipKQp1Pgtqg?ptgpkESkkkVeSKteTDrif? 95
CONSENSUS.B Ls-q-?s-pr-D-----es-k---rE--tdP?dQ 100
CONSENSUS.C -s---p--r-d.---Ee-----t-p-D 98
CONSENSUS.D --k--SS-pr-d-----eQ-----kA-t-p-Dw 100
CONSENSUS.O V-?-S???-?RK.Q?RQE-QE??-K??GP?G?P???SC??CTR?S?Q 85
CONSENSUS.CPZ ??-????SR?-..????K-?-?-??--????G-C 56

```

HIV-2/SIV Tat Consensus Sequences

```

                                <- vpr
CONSENSUS.A METpLKaPEsSL?syNEPsS?TSeqdv??QelakqGeEiLsQLyrPLEaCtNsCYCK?CcyhCQICFlkKGLG 68
CONSENSUS.B --I--qEQ---k-Saep--S---pV-NT-G-DN-----k--d-t---K----- 73
CONSENSUS.D -----rEQ-n--e-sner--cis-a-a-tp-s-nlg-----pl---y-tc---k-----f----- 72

                                rev cds ->                                exon \ / exon
CONSENSUS.A IwY?RkgRRRRtPkKtK?hssasD.KSISTRTg?SQptKkQKKTle?tvctd?glGr 120
CONSENSUS.B -c-drS...k-Ss-RA-tTa---pd??-L-Ar--D-----keV-T-g?--I-P--SNTSTSRFA 125
CONSENSUS.D -c-Eqr?---p--a-ant----n?k?ish-tRhc--k-ak-etV--a-a-ap--g- 126

```

HIV-1 and HIV-2 Sequence Subtypes

1. Introduction

In addition to the six HIV-1 sequence subtypes A through F recognized in the 1993 database compendium (III-2 of APR 94), there are now two new subtypes to be considered, G and H, for what is now being called the M group of HIV-1 sequences (see Janssens et al., *AIDS Res. Hum. Retroviruses* 10:877-879, 1994). Furthermore, we are aware of several sequences that don't fit at this time into any of the defined sequence subtypes of group M. And, as the cover of this compendium reveals, the O group of HIV-1s includes sequences that are as different from one another as M group subtypes. (The O group CA9 and VI686 sequences have not been made public, but were kindly provided to us for the use in the cover figure by the Institute of Tropical Medicine, Antwerp.) Together, the M and O group sequences of HIV-1s appear to constitute a "double star" phylogeny (Myers, *AIDS Res. Hum. Retroviruses* 10:1317-1322, 1994).

Heretofore, the database has recognized two subtypes of HIV-2s, A and B. A recent study by Hahn's group at the University of Alabama argues for the existence of at least five HIV-2 sequence subtypes, A through E (Gao et al., *J. Virol.* 68:7433-7447, 1994; the criteria used were virtually identical to those used for defining HIV-1 group M sequence subtypes).

Both HIV-1 and HIV-2 data sets present some evidence for hybrid viruses (see Sharp et al., *AIDS* 8 (suppl 1):S27-S42, 1994).

2. HIV-1 Phylogenetic Trees

The availability of many new HIV-1 *env* gp120 and gp160 coding sequences (Part I) permits us to generate new phylogenetic trees that can be compared to i) HIV-1 *gag* cds trees and ii) partial *env* cds trees. Since the publication of the 1993 compendium, we have been experimenting with asymmetrically-weighted parsimony analysis (as implemented by Hillis et al., *Science* 264:671-677, 1994) as a way to reduce effects of homoplasy. The first step in this process is to run ordinary parsimony analysis in order to determine the substitution biases; all things being equal, the biases will tend to follow the base composition of the sequences. We use Macintosh versions of PAUP and MacClade to accomplish this step. A resulting substitution frequency matrix (for HIV-1 *env* gp120 sequences) is shown below:

to:	A	C	G	T
from: A		0.104	0.222	0.074
C	0.060		0.023	0.095
G	0.163	0.027		0.022
T	0.052	0.112	0.046	

We see from this table that some transversions (e.g., A to C) are as frequent as some transitions (e.g., C to T). Purine-purine transitions occur at almost twice the frequency of pyrimidine-pyrimidine transitions. These appear to be consequences of the unusual base composition of HIVs and SIVs.

The next step in the analysis is to re-execute the parsimony algorithm (PAUP) using an inverse weighting table generated from the substitution matrix. Hence, the least common changes (G to T) are given the greatest weight and the most common substitutions (A to G) are given the least weight. In applying the inverse weighting rule, it may be necessary to truncate values in order that the triangle inequality relationship be preserved (see Maddison and Maddison: *MacClade: Analysis of Phylogeny and Character Evolution*. Sinauer Associates, Sunderland MA, 1992). Because the branch lengths in the newly generated tree are derived from weighted terms, they are no longer indicative of the minimal number of base changes, as they would be in ordinary parsimony. However, the lengths do provide accurate relative distances. We continue to employ maximum likelihood analysis, neighbor-joining, and bootstrapping in order to strengthen inferences based upon parsimony tree analyses.

When the signal to noise ratio is sufficiently high, the weighted approach won't drastically change the tree analysis; however, it will invariably change one or two individual relationships, attachments to a major trunk. Weighted parsimony is especially warranted when analyzing "deeper" phylogenetic relationships that have attained "mutational saturation" (III.D).

In the following analyses, both unweighted and weighted trees are shown for HIV-1 *gag*, *env* and *tat* coding sequences in that order.

HIV Subtypes

HIV-1 GAG
unweighted

HIV Subtypes

HIV-1 ENV
unweighted

HIV Subtypes

HIV-1 TAT
unweighted

This page intentionally left blank.

3. HIV-1 Subtype Summary

From various phylogenetic analyses, including those shown on the previous pages of this section, we can update the table of HIV-1 sequence subtype representatives originally presented on page III-2 of APR 94. Not all sequences are listed in the table; this table emphasizes complete *env* and *gag* coding sequences and concordant and discordant *env-gag* pairs. Subtype assignments derived from V3 coding sequences, which in a few instances may not agree with the following assignments, are summarized in III.E. The subtype A sequences are particularly diverse, leading to uncertainties (The designation 'U' indicates uncertainty). Group O sequences have not been subtyped. HIV-2 *gag* subtypes are shown in the following phylogenetic trees, first an unweighted then a weighted tree.

TABLE I. HIV-1 Subtypes

<i>env</i> Subtype	<i>gag</i> Subtype	Sample Locale(s)
A U455, DJ258 DJ264, Z321, SF170, UG031, RW009, RW020, CAR423, CAR4054	A U455, DJ258 K112, K29, K88, V132, CI20, K7, K98, CI59, LBV2310, CI51, CI44, CI32	Central Africa
A VI191	G VI191	Central Africa
D K124(?), UG266	A K124, UG266	Central Africa
E TN243 (CM243), CAR4017, CAR4071, TH011, TH006	A TN243 (CM243)	Thailand, Central African Republic
B LA1, JRCSE, D31, SF2 TB132, BR014, TH014	B LA1, JRCSE, D31, SF2 TB132, PHI136, PHI153	North America, Europe, South America, Asia
C ZAM18, SM145, UG268 NOF, D747, BR025, DJ373	C ZAM18, SM145, UG268 DJ259, VI131, SE364	Central and Southern Africa, India, Brazil
D ELI, NDK, SE365, UG274 UG038, UG021, CAR4020	D ELI, NDK, SE365, UG274 VI120, G109, K31, G270	Central Africa
F BZ163, BZ126, BR7944	F VI174, BZ162, VI325	South America, Central Africa, Europe
G LBV217 CAR4067, CAR4081	G LBV217	Central Africa, Taiwan Russia
G VI525	H VI525	Central Africa
H VI557 (partial)	H VI557	Central Africa
U ZAM184, Z3	U	Central and Southern Africa
O Group ANT70, MVP5180, VAU	O Group ANT70, MVP5180	West Africa, France

HIV Subtypes

HIV-2 GAG
unweighted

HIV Synonymous and Nonsynonymous Substitution Frequencies

1. HIV-1 and HIV-2 *gag* and *env* substitution frequencies.

Of the possible kinds of nucleotide substitutions in a typical viral coding sequence, synonymous (or so-called "silent") substitutions are the most commonly observed. If "mutational saturation" has not been reached for any of the pairwise relationships in a homologous set of sequences, synonymous substitution frequencies can provide a linear measure of genetic variation under conditions of minimal selection pressure. Nonsynonymous substitutions (generating amino acid replacements) tell us something about negative and positive Darwinian selection; under certain circumstances, they also serve as a metric. We should keep in mind the possibility that subtle selective pressures may be exerted upon synonymous substitutions (we have seen in the previous section that base composition is a factor in this regard), as well as the possibility that amino acid replacements may be unselected (Kimura's neutral theory). The majority of sites in a coding sequence will be nonsynonymous targets, however the majority of changes observed will be synonymous, because these produce the least negative effects.

Nonsynonymous (p_n) and synonymous (p_s) substitution frequencies are shown in the accompanying plots for a large number of pairwise relationships over a) HIV-1 *env* SU, b) HIV-1 *gag* MA, c) HIV-2/SIV *env* SU, and d) HIV-2/SIV *gag* MA coding sequences. The Nei-Gojobori algorithm was used to determine these frequencies [1]. "Initial" slopes have been calculated, using a least-squares program, over the range of p_s and p_n values of 0.0 to 0.3; multiple hits are likely to be encountered for values greater than 0.3. Synonymous substitution frequencies approach saturation in the range of 0.6 to 0.8; the saturation levels for nonsynonymous substitutions are less certain due to the very nature of positive selection effects.

Of note in these plots is first, the uncanny similarity of *env* SU slopes to *gag* MA slopes—0.44 and 0.40 for HIV-1s, 0.28 and 0.27 for HIV-2s. Equivalently measured slopes for *gag* CA coding sequences (not shown)—0.15 for HIV-1s and 0.09 for HIV-2/SIVs—clearly indicate that the near-identities needn't have been the case. We don't have an explanation for this phenomenon at this time: it seems to suggest that either *env* SU and *gag* MA are "co-selected" in some sense, or that they are both demonstrating some fixed level of neutral mutation not seen in *gag* CA coding sequences. Synonymous substitution frequencies at the scale of complete coding sequences are generally uniform across the viral genomic molecule; that is to say no gradient of intrinsic base substitution has been observed for HIVs.

A second finding is that HIV-1s possess higher slopes of p_n to p_s than do HIV2/SIVs for *env* and *gag* coding sequences. This result is consistent with a study conducted by Shpaer and Mullins [2], which was based upon a smaller data set and a different method for determining substitution frequencies.

Finally, the extent of synonymous changes are approximately equivalent for the majority of HIV-1, HIV-2, and SIV sequence relationships— p_s of 0.4—although this cannot be taken as evidence of identical ages of the sequence sets. The clusters of more mutated sequence relationships, p_s greater than 0.4, come about through comparison of HIV-1 M group sequences to O group sequences, and through comparison of HIV-2 sequences of one subtype to those of another subtype and to SIVs. The former effect—the nearly identical extents of synonymous substitution in HIV-1s and HIV-2s—is probably merely a consequence of sampling: sequences tend to be reported for individuals who have been recently infected and therefore haven't diverged that much from other recently infected (and sampled) sources.

2. HIV-1 Intersubtype Comparisons

As new HIV-1 *env* coding sequences have become available, phylogenetic tree analyses (III-C) have pointed toward a "star" relationship among the M group of HIV-1s and separately among the O group of HIV-1s [3]. An analysis of synonymous and nonsynonymous substitution frequencies provides some corroboration of the hypothesis that HIV-1s are part of two star phylogenies, which are suggestive of two separate introductions of virus at about the same time into the human population.

In the following analysis, 340 nt from the C2-V3 coding regions of HIV-1 subtypes A through H (of the M group) were analyzed using the Nei-Gojobori algorithm [1]. The average frequencies of p_s (synonymous substitutions) and p_n (nonsynonymous substitutions) were determined for members of a given sequence subtype (e.g., A) in relation to all other subtype sequences (e.g., B through H). Ideally, such a study would involve equal numbers of each subtype; however, that is not possible with the current data sets. Notwithstanding this limitation, the averages are remarkably identical as shown in the schematic summary below. Preliminary results based upon unpublished O group sequences (not shown) point to highly similar averages among these viral sequences—moreover, values that are similar to the average value shared by group M subtype sequences [3].

Average intersubtype synonymous and nonsynonymous substitution frequencies are given for each subtype set relative to all other subtype sequences; thus '.38/.18' gives the average p_s and p_n values respectively for A group sequences. Although these averages were calculated over a shorter coding stretch and involved more sequences, they are comparable to the ones reported in refs 1 and 3 (slightly higher in every case as is to be expected for C2-V3 relative to *env* gp120, hence average p_n to p_s ratios closer to 0.5 than to 0.4, as shown in the previous section).

References

- [1] Korber et al., *J. Virol.* 68:6730-6744, 1994.
- [2] Shpaer, G and Mullins, J., *J. Mol. Evol.* 37:57-65, 1993).
- [3] Myers, *AIDS Res. Hum. Retroviruses* 10:1317-1323, 1994.

Global Variation in the HIV-1 V3 Region

Introduction

Due to the immunogenicity and functional importance of the V3 loop, there has been a great deal of interest in the V3 region of the envelope protein, resulting in a large international effort to obtain V3 region sequences. This section, which includes sequences taken from 707 individuals and references, provides an overview of the variation of sequences that span this region.

Sequences

To best summarize the spectrum of international HIV-1 variants, only one representative viral sequence was included per infected individual. A complete set of references accompanies the sequence alignments, and nomenclature was preserved from the original papers so individuals and isolates can be clearly identified. HIV-1 was deleted from the sequence names in this section, as all sequences included here are HIV-1. Included with the references when available are brief descriptions of critical features of the sequences. This includes the health status of the individual from whom the virus was derived, whether or not the virus was cultured, and the year the blood sample was taken. All sequences are prefaced by a two letter country code to identify the country that the individual resided in at the time that the blood sample was taken. If the person was a recent immigrant and this information was available, we included the country of origin in the references. The two letter code was developed for Internet (Copyright 1992, Lawrence H. Landwater and the Internet Society), and incorporated here based on a suggestion made by Dr. Francine McCutchan. The key to the country codes follows this introduction.

Sometimes only one viral sequence was available from a person: a clone from an isolate, or a direct sequence of PCR amplified peripheral blood DNA. For other individuals, up to 80 viral sequences from PCR amplified DNA or RNA from blood samples were available. Consequently, nearly 2000 sequences are represented by the 707 included in this section. When two sequences were available from a person, one of the two was randomly selected. When three or more sequences were generated from a person, all available sequences were aligned (without regard to different time points of sampling) and a consensus of the most common base found in each position in the alignment was generated. If there was a tie (e.g., 10 A's, 10 T's), the top base or amino acid in the alignment was used. If a set of sequences from two or more individuals was epidemiologically linked, and genetically very similar, only one sequence from the set was included. In the sequence description and references section, the short hand "PCR-direct, peripheral blood DNA" is used to signify that viral DNA was amplified from PBMCs, without culturing, and a single "direct" sequence was obtained from the amplification reaction products. The short hand "Consensus, PCR-clones, peripheral blood DNA" signifies that viral DNA was amplified from PBMCs and a set of clones was generated and sequenced from the PCR amplification products. The cloned sequences were aligned and a consensus was generated. In a handful of cases, a particular gp160 clone from an isolate was shown to be expressed and functional using a vaccinia virus T7 expression system. In these cases, the clone rather than the consensus of all sequences from a particular individual was included.

Phylogenetic clustering

Sequences have been organized according to their phylogenetic subtype association (A-H). These sequence subtypes (A-H) have been defined based on the phylogenetic relationships determined by using both *gag* and *env* genes (when possible), are approximately genetically equidistant in envelope, and have multiple members. The phylogenetic subtype designations and associations have generally been adopted by the HIV research community, and are now often presented with the publication of new sequences. We have either determined the subtype designations here at Los Alamos, if they were not specified in the original manuscript, or else confirmed the subtype designations of the

original manuscripts, and then used the subtypes to organize this section. Generally, confirmations were done by aligning a set HIV-1 V3 region sequences with longer *env* gene sequences (Part IIIC) that have clear subtype associations, and then using parsimony or neighbor joining trees to determine associations. Some of the shorter gene fragments from this region were given a subtype designation based on Hamming distances, using the similarity function of the MASE program (Faulkner DV, and Jurka J. *TIBS* 13:321-322 (1988)); these sequences have ".sh" appended to their name to indicate that they were too short for phylogenetic analysis. Parsimony trees were generated using PAUP (David Swofford, Illinois Natural History Survey), and neighbor-joining trees were generated with Kimura distances using PHYLIP (Joseph Felsenstein, University of Washington). All available nucleotide sequence information was used for phylogenetic analysis; longer protein sequences were trimmed to be approximately the same length as the majority of the PCR fragments in this region, for the purposes of presentation. Some sequences were difficult to classify, and are included in the "U", or unclassified, section.

The set of sequences used to help resolve subtype associations were selected based on being "typical" of the subtype they represent. Included were: A_2RW020W, A_UG273, B_W2BR020, B_D2US711, C_D3MA959, C_D760, D_D2UG021, D_HIVELI, E_W2TH006, E_ELOENV1, F_BZ126, F_BZ163, G_VI525, G_LBV21-7, H_CA13, H_VI557, O_MVP5180, and O_VAU (see references for more information about these sequences.)

Limitations of phylogenetic analyses

Most of the PCR derived sequences contain a sub-optimal length for phylogenetic analyses, given the level of variability in this region – typically on the order of 250 to 300 nucleotides. Due to this limitation, some of the classifications in this section are uncertain and are our best estimate given the available information. Control studies were performed to compare the phylogenetic clustering of the V3 region using available longer sequences, however, and these studies indicate that our subtype designations based on the V3 region are generally reliable. For 146 sequences, we had an approximately 700 base region of *env* available representing all of the subtypes A–H. (The limitation in length was due to including the H subtype sequences, which did not cover all of gp120.) After removing positions in the alignment which included gaps, 519 bases were left. When a 298 base V3 region fragment was excised from this set, and neighbor joining trees were constructed using both the 519 base and 298 base long sequences, the phylogenetic subtype designations were consistent in each case. Further, when a subset of longer gp120 sequences was analyzed (92 of the 146), including 935 bases after removing positions in the alignment which included gaps, the subtype designations were again clear in neighbor-joining trees. This indicates that the limited V3 region PCR fragments are adequate for subtype determination.

Without detailed analysis, genetic recombination between subtypes may obscure phylogenetic relationships between sequences. A characteristic of recombination is an indeterminate place in phylogenetic analyses, and some of the "Uncertain" category sequences may prove to be recombinant genomes on further inspection. Also, while a subtype designation based on a gene or gene fragment may be correct, recombination events outside the region examined may have occurred. Therefore, care should be taken to not overinterpret the subtype designations.

Limitations of V3 amino acid consensus sequences

The V3 amino acid consensus sequences generated for each subtype have interesting features; however, one should be wary about assuming that any of the consensus sequences may broadly represent their subtype. Certainly many V3 loop variants in each of the subtypes are extremely divergent from the consensus sequences. These divergent forms may have very different biological and immunological characteristics from viruses which are similar to the consensus. Additionally, because of the relatively small sample size of most of the subtypes, consensus sequences can be dominated by a small group of highly similar sequences, which may in turn be a sampling artifact. Hence, these consensus sequences are "evolving" as new sequences from each subtype become available.

V3 LOOP

```

ALL_CONSENSUS_95  LLNQSGLAEELIIRSENITDANKTIIVQNESVGIN  CTAPDNDIRSRIRIGGQAFYATSDITG...DIRQAWC  HLESTQONFTIQQVAKLRE.HF.MKT...IIFNPSSGGLLEITTSFNCNGEFFYCHT
A_CONSENSUS_95    VV-----N-----VKG-K-----V-----E-K-----TQ-K.Y-----M-----G-----
B_CONSENSUS_95    VV-----F-----V-----B-----R-----T-----D-----L-----A-----K-----I-----Q-----G-----V-----Q-----P-----V-----G-----
C_CONSENSUS_95    -----G-----I-N-V-----E-----V-----Y-----QRTN-----L-T-R-----KD-----R-GK-A-----P-----K-A-----
D_CONSENSUS_95    -----L-N-I-----T-----T-----GAE-K-----K-CD.L-----T-----K-----P-----G-----
E_CONSENSUS_95    -----L-N-H-----K-----S-----T-----T-----V-R-----K-Y-----E-MQ-----E-K-TE-K-----H-----Q-P-----N-H-----
F_CONSENSUS_95    -----Q-S-T-----H-T-Q-----S-----BL-----K-----V-G-Q-----M-V-R-RA-QS-----P-----T-----S-----N-----
G_CONSENSUS_95    -----R-----F-----V-----K-I-----T-----V-K-----KQM-N-QAQ-KIV-----S-A-----
H_CONSENSUS_95    MHV-----K-----T-N-----KSP-P-----S-----T-ED-KR-----RE-VQD-----RFN-Q-----E-----H-----M-T-----
I_CONSENSUS_95    -----C-----P-----Q-----N-----H-----TQ-K-----S-----S-----
O_CONSENSUS_95    -I---T-SKKK-R- -LESG-N---T-TIDNT -E--QIDTGE-----HA-N-SH-LGGNGLSS-A-Y-  NYTA-D-QTA-K-T-EDYL-LV.-Q-SEHNS-H-----A-V-RH-----

```

Subtype consensus sequences. This V3 region alignment shows a consensus sequence generated for each of the nine subtypes. The nine subtype consensus sequences indicate the most common amino acid found in each position among the sequences associated with each subtype. The sequences are aligned to a consensus based on the most common amino acid in the subtype consensus sequences, which approximates a "global" consensus. It was generated in this way (rather than by using all 707 sequences) to avoid over-representation of the B subtype, which has by far the largest number of available sequences. As is the convention in this compendium, a dash (-) indicates concurrence with the top sequence in the alignment; a period (.) indicates a deletion. The carets show where the N-linked glycosylation sites are found in the consensus. The V3 loop is set off from the surrounding sequence by a space on either side to facilitate viewing. Interesting features of the consensus alignment are: 1) Only in the B subtype is GPGR the most common tip of the V3 loop; globally, GPGQ is more prevalent. 2) A highly conserved N-linked glycosylation site is constitutively absent in the C subtype, proximal to the first cysteine (C) in loop. 3) The D subtype consensus has 34 amino acids from cysteine (C) to cysteine (C) rather than the more common 35; at the point where the deletion occurs, it is not uncommon to find insertions of 2 to 4 amino acids, as can be observed in the sequence alignments. 4) The D subtype has two arginine (R) residues in the V3 loop that are uncharacteristic relative to the other consensus sequences; positively charged amino acids in these positions may result in a syncytia inducing, non-monocytropic phenotype (Fouchier RAM, et al., J. Virol. 66:3183-3187 (1992)). 5) A higher degree of variation is seen in the region just downstream of the V3 loop than within it. This difference is also observed internally among the sequences of the different subtypes. 6) The A, C, G and H consensus sequences have very similar V3 loop sequences.

V3 loop amino acids - following 4 pages. The "global" consensus sequence for all of the subtypes is shown at the top. Directly beneath the global consensus are shown aligned the subtype alternative amino acids that can be found in the six subtype consensus sequences at each position. For each subtype, the number of sequences which are currently available is indicated. The V3 loop consensus sequences (the most common amino acid found in each position among the sequences associated with each subtype), are shown. Written just above each consensus is the percentage of sequences that had the most common amino acid at each position (if a particular amino acid is perfectly conserved in a given position in a subtype, the 100% is not shown). Directly beneath the most common amino acid in each position are the other amino acids found in that position, listed from most common to least common. A dash (-) indicates a deletion; unusual insertions are not shown.

V3 consensus for subtypes A - O

N C T R P N N N T R K S I R I G P G Q A F Y A T G D I I G D I R Q A H C N I S

A subtype (124 sequences)

89	76	98	71	98	96	97	96	60	82	56	48	93	98	96	89	60	94	98	94	86	88	65	94	90	78	80	99	80	87	87	86						
N	C	T	R	P	N	N	T	R	K	S	V	R	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	N	V	S
T	I	H	G	K	Y	I	L	R	G	I	H	L	A	S	R	R	T	I	H	T	R	D	E	V	T	N	T	K	K	V	Y			T	I	N	
D	S	L	S	R	K	K	N	T	R	M	P	M				S	S	L	I	-	-	-	A	M	V	S		R	E	Q			E	L	K		
S	A		T		-		V	S	Q	N	A	G	T			S	V	S	F	G	A	S	G	N	K									D		T	
I	M	V	Y					E	M	W	N	F				P			K	I	N	N	R	K	L												
								X		Y	C											S	-	Q													

B subtype (338 sequences)

94	94	99	86	96	95	97	94	83	68	96	61	75	97	88	99	83	89	83	89	55	94	86	43	97	90	84	97	99	88	92	94	56	94				
N	C	T	R	P	N	N	T	R	K	S	I	H	I	G	P	G	R	A	F	Y	T	T	G	E	I	I	G	D	I	R	Q	A	H	C	N	L	S
T	S	I	S	L	S	K	-	K	S	R	G	I	K	V	L	P	M	A	W	K	R	A	-	D	V	V	N	T	S	K	P	Y		T	I	N	
D	E	S			G	D	E	I	K	S	H	T			L	G	Q	S	A	F	T	V	-	K	L	A	G	L	M	-	A	H	I	T	S	X	-
H					F	I	K	S	H	T																											
I																																					
K																																					
Y																																					

C subtype (40 sequences)

64	91	98	81				83	95	69	93	91				67	98	93	98	98	71	78				78		88	85		93	98						
V	C	T	R	P	N	N	T	R	K	S	I	R	I	G	P	G	Q	T	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	N	I	S
E		A		Y	G				Q	G	V	G	F					A	V	F	-	N	N	G			N		R		Q				V	N	
N		V			S				R	R	M	S	L					V					K			K		E		Y							
M				A	T						L		V					P					A														
T											T																										
K																																					
R																																					

D subtype (73 sequences)

78	78		72	61	72	57	82	69	41	82	46	91	99	47	97	73	89	65	73	95	78	80	32	56	41	86	56	92	74	85	97	71		83	90	92			
N	C	T	R	P	Y	N	K	K	T	R	Q	R	T	H	I	G	P	G	Q	A	L	Y	T	T	-	R	I	I	G	D	I	R	Q	A	H	C	N	I	S
T		A			N	K	K	K	I	R	G	S	I	P	F	R	L	Q	R	S	Y	F	A	S	T	N	V	T	K	N	K	P	Y		E	V	N		
S		I			S	E	R	S	I	S	K	I	V	S	L	R	S	R	T	V	F	H	R	I	K	G	D	R	K	N	F	A	R		K	S	T		
D		S			A	D	R	S	E	W	H	Q	I	A	R	-				W	I	W	Y	N	K	G	D	R	K	N	F	A	R						
K		M			E	D	T	Q	I	-									Q	E	G	L	P	S	-														
R		V			G	K	R	H	I	Q	Y																												

E subtype (65 sequences)

99	99		77	79	80	73	96	89	88	59	55	88		97		77	89	92	94	71	97	92	65		61		83	97	97	83	99	88		97	97	97		
N	C	T	R	P	S	N	N	T	R	T	S	I	T	I	G	P	G	Q	V	F	Y	R	T	G	D	I	I	G	D	I	R	K	A	Y	C	E	I	N
K		Y			F	K	K	I	I	K	P	K	V	R	M	G		R	A	L	H	K	I	E	S	A	G	N	K	X		X		F		K	V	D
					A	D	I	M	X	Q		A	M	N	V	Q		H	I		G	S	X															
					T	F	R	S	K			T	L	S																								
					Y	R	S	T	T	Y																												
					-																																	

F Subtype (41 sequences)

98		98	98	95	95		98	90		85	70		95		83	90	95	98	80	98	90	75	95	93		93		95		78			80	78	73		
N	C	T	R	P	N	N	T	R	K	S	I	H	L	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	K	A	H	C	N	V	S
D				S	K	T	I	R	G		R	I	L		R	V	I	H	T	K	D	E	V	N		N		Q		Y			K	I	N		
					Y			R			P	F	V			P	X					S	G		T			R				V	H	X			
											X											V	N		V							X					
											Y												A										I				
																							K										S				
																							X														

G Subtype (8 sequences)

71		71									43	43	86	86		86	86	71				86	43		57		86		86				86				
N	C	T	R	P	N	N	T	R	K	S	I	T	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	N	V	S
T		I									H	F	A	T		R	V	L				S	A			N	K		Y					N			
V											K	L											E														
											P											N															
											R																										

H Subtype (2 sequences)

											50	50	50		50		50	50	50	50	50						50	50				50					
N	C	T	R	P	N	N	T	R	K	S	I	R	I	G	P	G	Q	A	F	H	A	I	G	A	I	I	G	D	I	R	K	A	H	C	N	I	S
											R	M	S		R				T	Y	T	D						Q					T				

O Sequences (3 sequences)

67		67		67	67	67	33	33	33		33	67	33	33		67		67		33	67	33	33	33	67	33	33	67		33		33	67						
T	C	E	R	P	G	I	Q	T	V	Q	E	I	R	I	G	P	M	A	W	Y	S	M	G	L	G	R	S	S	G	D	S	R	A	A	Y	C	N	Y	N
A		I		E	Q	N	D	E	I		D	M	Y	A			R	R			A	I	S	G	T	K	P	R	T		V			K	S				
						A	I	-		K	M	T									T	K	-	-	-	-	-	-	-	-	-	-	T						

Unclassified Sequences (13 sequences)

94	94	99	90	97	89	98	96	87	72	96	64	77	99	88	99	84	92	87	91	57	95	91	48	96	92	83	97	98	92	92	95	58	94				
N	C	T	R	P	N	N	T	R	K	S	I	R	I	G	P	G	Q	A	F	Y	A	T	G	D	I	I	G	D	I	R	Q	A	H	C	N	I	N
S				G	S	D	I	K	R	R	M	H	F	S		R	T	I	F	T	I	-	G	V	T	N	-	-	K	Y		D	V	S			
T				S		I	K	N	I		T	N	L			K	V			-	K	N	K	-	-					Q		I		T			
V						T	R		T		V	V								-	S	-											S		T		

HIV-1 V3 Region

Summary of variations in the tetrameric tip of the V3 loop. This table is a tally of the 38 different tetramers observed at the tip of the V3 loop among the viruses found in the 707 individuals analyzed. This tip is thought to form a turn, and is the focal point of the potent neutralizing antibody epitopes that have been mapped to the V3 loop, as well as of T cell epitopes. Each column shows the number of occurrences of a given tetramer in either the entire 707 sequences (combined), or in subsets consisting of subtypes A-O, the unclassified sequences (U). Underneath the column heading is the number of sequences in each category. The most common form found in each subtype is highlighted by bold lettering. In the B subtype, GPGR is the predominant form, however globally GPGQ tends to be more common.

	Combined 707	A 124	B 338	C 40	D 73	E 65	F 41	G 8	H 2	O 3	U 13
GPGR	341	9	235		15	10	6				4
GPGQ	289	105	11	40	19	50	33	6	1		7
GPGK	33	2	29								1
GWGR	22		22								
GLGQ	17				17						
GPGS	13	1	10								
GSGQ	12	3			6						1
APGR	11	1	10								
GPGG	8		7								
GQGQ	6		1		5						
GPMA	6									2	
GLGR	6		5				1				
GRGQ	5	1			3				1		
GVGR	4		1		1		1				
GQGR	4		2		1	1					
GTGQ	3				3						
GSGR	3	1	2								
GPGH	3					3					
GPLS	2										
GGGR	2		2								
APGQ	2	1						1			
GLGS	1		1								
AQGR	1				1						
GTGR	1							1			
GTGG	1		1								
GPMS	1										
GPMR	1									1	
GPLR	1										
GPLA	1										
GPKR	1		1								
GPGA	1				1						
GLRQ	1				1						
GGGQ	1					1					
GQRK	1		1								
GFGR	1		1								
GARR	1		1								
GAGR	1		1								
APGS	1		1								

Summary of octameric variation (cont.)

	Combined	A	B	C	D	E	F	G	H	O	U
	707	124	338	40	73	55	41	8	2	3	13
QRGPGRAF	2		2								
HMGP GKTF	2	1	1								
HMGP GKAF	2		2								
YIGP GRAV	2		2								
TRGP GHVF	2					2					
TMGP GRVW	2		2								
TMGP GQVF	2					2					
TMGP GKVF	2		2								
TIGP GQIF	2					2					
SIGQ GQAL	2				2						
SIGP GRVW	2		2								
SIGL GQAL	2				2						
RIGP GQSF	2	2									
RIGP GQAL	2				2						
RFGP GQAF	2										2
PIGR GQAL	2				2						
PIGP GKAF	2		2								
PIGL GQAY	2				2						
NTGP GRAF	2		2								
NIGP GQAF	2	2									
HMGP GRAL	2				2						
HMGL GRAF	2		2								
HLGP GRAW	2		2								
HLGP GQAL	2				2						
HIGS GRAF	2	1	1								
HIGS QAY	2				2						
HIGP GRTF	2		2								
HIGP GRSF	2	1	1								
HIGP GQAI	2				2						
HIGG GR TL	2		2								
AIGP GR TV	2		2								
unique	140	16	62	3	26	10	8	5	2	3	5

HIV-1 V3 Region

Summary of variations in the octameric tip of the V3 loop. This table is a tally of the 207 different octamers found at the tip of the V3 loop among the 707 individuals analyzed here. The table is structured the same as the tetramer table on the previous page. Amino acid changes proximal to the tip can influence the specificity of anti-V3 neutralizing antibodies. The 140 forms that were only found once in the entire data set are not shown here, to save space, and are summarized in the row called "unique."

	Combined 707	A 124	B 338	C 40	D 73	E 65	F 41	G 8	H 2	O 3	U 13
HIGPGRAF	116	3	110		1		2				
RIGPGQTF	55	26		27							2
RIGPGQAF	37	25		8		1		1			2
HIGPGQAF	32	27	1				4				
HLGPGQAF	28	2					26				
PIGPGRAF	27	1	26								
TIGPGQVF	23					23					
NIGPGRAF	17		16								1
RIGPGQVF	15	2		2		10	1				
SIGPGRAF	14	1	13								
PIGLQAL	11				11						
HMGWGRAAF	10		10								
HIAPGRAAF	10	1	9								
HIGPGKAF	9	1	8								
TIGPGRAF	9		9								
HIGPGRAY	9				9						
HIGPGQAL	8				7			1			
RIGPGRVF	7					7					
PIGPGQAF	6	2			1	2		1			
HIGPGSAF	5		5								
HLGWGRAAF	5		5								
HLGPGQAW	5		5								
HIGPGQTF	5	5									
PLGPGQAW	4		4								
PMGPGRAF	4		4								
NMGPGRAF	4		4								
QIGPGRAF	3		3								
HIGPGRAAF	3		3								
HIGPGGAF	3		3								
PMGPGKAF	3		3								
HMGPGRAF	3	1	2								
PIGPGQVF	3					3					
NIGPGQVF	3					3					
HMGWGRTF	3		3								
HLGPGRAF	3	1	1								1
GIGPGQTF	3	3									

COUNTRY CODES

It is becoming increasingly useful to name viral isolates and samples with a country code. The following code was captured from an Internet file.

This is a list based on the International Organization for Standardization (ISO) 3166:1988 standard, updated from a list prepared by Mark Horton. Note that the original standard has this same information sorted into about 6 different orders, both in English and French, therefore this is an abbreviated version not to be taken as the entire standard. While it has been checked against the standard, it may possibly contain errors; the standard and registration newsletters should be verified for any critical application.

This copy has been updated through registration Newsletter III-13 and is believed to be current through at least January 1991.

Table of Country Codes from ISO 3166

Country	A 2	A 3	Number
AFGHANISTAN	AF	AFG	004
ALBANIA	AL	ALB	008
ALGERIA	DZ	DZA	012
AMERICAN SAMOA	AS	ASM	016
ANDORRA	AD	AND	020
ANGOLA	AO	AGO	024
ANGUILLA	AI	AIA	660
ANTARCTICA	AQ	ATA	010
ANTIGUA AND BARBUDA	AG	ATG	028
ARGENTINA	AR	ARG	032
ARUBA	AW	ABW	533
AUSTRALIA	AU	AUS	036
AUSTRIA	AT	AUT	040
BAHAMAS	BS	BHS	044
BAHRAIN	BH	BHR	048
BANGLADESH	BD	BGD	050
BARBADOS	BB	BRB	052
BELGIUM	BE	BEL	056
BELIZE	BZ	BLZ	084
BENIN	BJ	BEN	204
BERMUDA	BM	BMU	060
BHUTAN	BT	BTN	064
BOLIVIA	BO	BOL	068
BOTSWANA	BW	BWA	072
BOUVET ISLAND	BV	BVT	074
BRAZIL	BR	BRA	076
BRITISH INDIAN OCEAN TERRITORY	IO	IOT	086
BRUNEI DARUSSALAM	BN	BRN	096
BULGARIA	BG	BGR	100
BURKINA FASO	BF	BFA	854
BURUNDI	BI	BDI	108
BYELORUSSIAN SSR	BY	BYS	112
CAMBODIA	KH	KHM	116
CAMEROON	CM	CMR	120
CANADA	CA	CAN	124
CAPE VERDE	CV	CPV	132

HIV-1 V3 Region

CAYMAN ISLANDS	KY	CYM	136
CENTRAL AFRICAN REPUBLIC	CF	CAF	140
CHAD	TD	TCD	148
CHILE	CL	CHL	152
CHINA	CN	CHN	156
CHRISTMAS ISLAND	CX	CXR	162
COCOS (KEELING) ISLANDS	CC	CCK	166
COLOMBIA	CO	COL	170
COMOROS	KM	COM	174
CONGO	CG	COG	178
COOK ISLANDS	CK	COK	184
COSTA RICA	CR	CRI	188
COTE D'IVOIRE	CI	CIV	384
CUBA	CU	CUB	192
CYPRUS	CY	CYP	196
CZECHOSLOVAKIA	CS	CSK	200
DENMARK	DK	DNK	208
DJIBOUTI	DJ	DJI	262
DOMINICA	DM	DMA	212
DOMINICAN REPUBLIC	DO	DOM	214
EAST TIMOR	TP	TMP	626
ECUADOR	EC	ECU	218
EGYPT	EG	EGY	818
EL SALVADOR	SV	SLV	222
EQUATORIAL GUINEA	GQ	GNQ	226
ETHIOPIA	ET	ETH	230
FALKLAND ISLANDS (MALVINAS)	FK	FLK	238
FAROE ISLANDS	FO	FRO	234
FUJI	FJ	FJI	242
FINLAND	FI	FIN	246
FRANCE	FR	FRA	250
FRENCH GUIANA	GF	GUF	254
FRENCH POLYNESIA	PF	PYF	258
FRENCH SOUTHERN TERRITORIES	TF	ATF	260
GABON	GA	GAB	266
GAMBIA	GM	GMB	270
GERMANY	DE	DEU	276
GHANA	GH	GHA	288
GIBRALTAR	GI	GIB	292
GREECE	GR	GRC	300
GREENLAND	GL	GRL	304
GRENADA	GD	GRD	308
GUADELOUPE	GP	GLP	312
GUAM	GU	GUM	316
GUATEMALA	GT	GTM	320
GUINEA	GN	GIN	324
GUINEA-BISSAU	GW	GNB	624
GUYANA	GY	GUY	328
HAITI	HT	HTI	332
HEARD AND MC DONALD ISLANDS	HM	HMD	334
HONDURAS	HN	HND	340
HONG KONG	HK	HKG	344
HUNGARY	HU	HUN	348
ICELAND	IS	ISL	352

INDIA	IN	IND	356
INDONESIA	ID	IDN	360
IRAN (ISLAMIC REPUBLIC OF)	IR	IRN	364
IRAQ	IQ	IRQ	368
IRELAND	IE	IRL	372
ISRAEL	IL	ISR	376
ITALY	IT	ITA	380
JAMAICA	JM	JAM	388
JAPAN	JP	JPN	392
JORDAN	JO	JOR	400
KENYA	KE	KEN	404
KIRIBATI	KI	KIR	296
KOREA, DEMOCRATIC PEOPLE'S REPUBLIC OF	KP	PRK	408
KOREA, REPUBLIC OF	KR	KOR	410
KUWAIT	KW	KWT	414
LAO PEOPLE'S DEMOCRATIC REPUBLIC	LA	LAO	418
LEBANON	LB	LBN	422
LESOTHO	LS	LSO	426
LIBERIA	LR	LBR	430
LIBYAN ARAB JAMAHIRIYA	LY	LBY	434
LIECHTENSTEIN	LI	LIE	438
LUXEMBOURG	LU	LUX	442
MACAU	MO	MAC	446
MADAGASCAR	MG	MDG	450
MALAWI	MW	MWI	454
MALAYSIA	MY	MYS	458
MALDIVES	MV	MDV	462
MALI	ML	MLI	466
MALTA	MT	MLT	470
MARSHALL ISLANDS	MH	MHL	584
MARTINIQUE	MQ	MTQ	474
MAURITANIA	MR	MRT	478
MAURITIUS	MU	MUS	480
MEXICO	MX	MEX	484
MICRONESIA	FM	FSM	583
MONACO	MC	MCO	492
MONGOLIA	MN	MNG	496
MONTSERRAT	MS	MSR	500
MOROCCO	MA	MAR	504
MOZAMBIQUE	MZ	MOZ	508
MYANMAR	MM	MMR	104
NAMIBIA	NA	NAM	516
NAURU	NR	NRU	520
NEPAL	NP	NPL	524
NETHERLANDS	NL	NLD	528
NETHERLANDS ANTILLES	AN	ANT	532
NEUTRAL ZONE	NT	NTZ	536
NEW CALEDONIA	NC	NCL	540
NEW ZEALAND	NZ	NZL	554
NICARAGUA	NI	NIC	558
NIGER	NE	NER	562
NIGERIA	NG	NGA	566
NIUE	NU	NIU	570
NORFOLK ISLAND	NF	NFK	574

HIV-1 V3 Region

NORTHERN MARIANA ISLANDS	MP	MNP	580
NORWAY	NO	NOR	578
OMAN	OM	OMN	512
PAKISTAN	PK	PAK	586
PALAU	PW	PLW	585
PANAMA	PA	PAN	590
PAPUA NEW GUINEA	PG	PNG	598
PARAGUAY	PY	PRY	600
PERU	PE	PER	604
PHILIPPINES	PH	PHL	608
PITCAIRN	PN	PCN	612
POLAND	PL	POL	616
PORTUGAL	PT	PRT	620
PUERTO RICO	PR	PRI	630
QATAR	QA	QAT	634
REUNION	RE	REU	638
ROMANIA	RO	ROM	642
RWANDA	RW	RWA	646
ST. HELENA	SH	SHN	654
SAINT KITTS AND NEVIS	KN	KNA	659
SAINT LUCIA	LC	LCA	662
ST. PIERRE AND MIQUELON	PM	SPM	666
SAINT VINCENT AND THE GRENADINES	VC	VCT	670
SAMOA	WS	WSM	882
SAN MARINO	SM	SMR	674
SAO TOME AND PRINCIPE	ST	STP	678
SAUDI ARABIA	SA	SAU	682
SENEGAL	SN	SEN	686
SEYCHELLES	SC	SYC	690
SIERRA LEONE	SL	SLE	694
SINGAPORE	SG	SGP	702
SOLOMON ISLANDS	SB	SLB	090
SOMALIA	SO	SOM	706
SOUTH AFRICA	ZA	ZAF	710
SPAIN	ES	ESP	724
SRI LANKA	LK	LKA	144
SUDAN	SD	SDN	736
SURINAME	SR	SUR	740
SVALBARD AND JAN MAYEN ISLANDS	SJ	SJM	744
SWAZILAND	SZ	SWZ	748
SWEDEN	SE	SWE	752
SWITZERLAND	CH	CHE	756
SYRIAN ARAB REPUBLIC	SY	SYR	760
TAIWAN, PROVINCE OF CHINA	TW	TWN	158
TANZANIA, UNITED REPUBLIC OF	TZ	TZA	834
THAILAND	TH	THA	764
TOGO	TG	TGO	768
TOKELAU	TK	TKL	772
TONGA	TO	TON	776
TRINIDAD AND TOBAGO	TT	TTO	780
TUNISIA	TN	TUN	788
TURKEY	TR	TUR	792
TURKS AND CAICOS ISLANDS	TC	TCA	796
TUVALU	TV	TUV	798

UGANDA	UG	UGA	800
UKRAINIAN SSR	UA	UKR	804
UNITED ARAB EMIRATES	AE	ARE	784
UNITED KINGDOM	GB	GBR	826
UNITED STATES	US	USA	840
UNITED STATES MINOR OUTLYING ISLANDS	UM	UMI	581
URUGUAY	UY	URY	858
USSR	SU	SUN	810
VANUATU	VU	VUT	548
VATICAN CITY STATE (HOLY SEE)	VA	VAT	336
VENEZUELA	VE	VEN	862
VIET NAM	VN	VNM	704
VIRGIN ISLANDS (BRITISH)	VG	VGB	092
VIRGIN ISLANDS (U.S.)	VI	VIR	850
WALLIS AND FUTUNA ISLANDS	WF	WLF	876
WESTERN SAHARA	EH	ESH	732
YEMEN, REPUBLIC O	YE	YEM	887
YUGOSLAVIA	YU	YUG	890
ZAIRE	ZR	ZAR	180
ZAMBIA	ZM	ZMB	894
ZIMBABWE	ZW	ZWE	716

List of changes applied, as specified in registration newsletters:

Newsletter III-1, 1989-12-5:

Burma deleted, Myanmar added (same numeric value, change of country name)

Newsletter III-2, 1990-07-16

Namibia, changing information not included in this file

Newsletter III-3, 1990-08-14

Afghanistan, changing information not included in this file

Newsletter III-4, 1990-08-14

Ethiopia, changing information not included in this file

Newsletter III-5, 1990-08-14

Fiji, changing information not included in this file

Newsletter III-6, 1990-08-14

Hungary, changing information not included in this file

Newsletter III-7, 1990-08-14

Unification of Yemen, under new numeric code

Newsletter III-8, 1990-08-14

Romania, changing information not included in this file

Newsletter III-9, 1990-08-14

Poland, changing information not included in this file

Newsletter III-10, 1990-08-14

Kampuchea deleted, Cambodia added (same numeric value, change of name)

Newsletter III-11, 1990-08-14

Benin, changing information not included in this file

Newsletter III-12, 1990-12-04

Czechoslovakia, changing information not included in this file

Newsletter III-13, 1990-10-30

Germany unified (DDR deleted, new name and numeric code for unified Germany)

V3 LOOP

A_C01	11423	K-R-A-A-F-A-A-A-A-A-D-S-	-R-A- - - - -V-A-M	IPT	- - - -G-O-V-I-	X-Y- - - - -	-X-Y- - - - -I-K-R- - -V-N-K-G-S-Y-T-S	K-H	- - - -D-	S-A-A-A	R
A_C01	11699	X-E-B- - -F-A-A-A-A-L-K-Q-	-R-A- - -S-A- - - - -	-H-M	-A-R- - - - -	-Y- - - - -I-K-R- - -N-	-Y- - - - -I-K-R- - -N-	-R- - - - -	- - - -D-	S-A-A-A	R
A_C01	2189	-I- - - - -F-A-A-A-A-D-S-	-G-A- - -T-A- - -G-A-A-A-X-R-M-S-	-A-A-R	P-A-A-R	-V-A-	- - - - -T-I-N-A-K-A-Q-A-K-K-K-G-I-	-H	- - - -	-A-T	K
A_C01	4010	-S- - - - -L- - - - -E-	-H-T- - -I- - - - -	I-B-	- - - - -	M- - - - -K-D- - - -V-Q-R-K-E-W	B-Q-	- - - -	- - - -	-T	S
A_C01	4016	G-R- - -P-S-S- - - -Q-	L-N-T- - -I- - - - -T-	-M-I	V-V-X	- - - - -I-K-M-T-N- - - -K- - - -E-N-L-R	-S	-S	A-T- - -	T	V
A_C01	4023	K- - - - -F-S- - - - -Q-	-I- - - - - - - - - -R-	M-A	- - - - -I- - - - -	- - - - -Y- - - - -K-E-D- - -T- - -K-V-E-K-G-T	I-S	- - - -	- - - -	- - - -	-
A_C01	4033	K-R-S-A-F-S- - - - -T-	- - - - - - - - - - -	- - - - -	- - - - -	- - - - -L-V-A- - -H-M-A- - - - -G-	-R	-K	-N	E-T-K- - -	K
A_C01	4044	-Q- - - -F-S-D-T-P- - - -T-	- - - - - - - - - -S- - - -K-T-	I-H	H-A-R	-A- - - - -V-A- - - - -I- - -K-A- - -D- - -D- - -K-K-G-T	-R	-N	-N	T-T	A
A_C01	4050	-S- - - - -D- - - - -L- - - -E-	-H-T- - -T- - - - -	-A-A	- - - - -	- - - - -H- - - - -V-Q-K- - -E-W	-T-U	- - - -	-S-T	-	A
A_C01	4054	R-R- - -F- - - - -N-E-T	-R- - -H- - - - -	I-L	- - - - -	- - - - -K- - - - -R- - - - -T-E- - -E-Q	I-U	- - - -	-Q- - -	-	v
A_C01	4055	K-R- - -F- - - - -N-E-T	-R- - -S- - - - -	I-L	- - - - -	- - - - -K- - - - -R- - - - -T-E- - -E-Q	I-H	- - - -	-Q- - -	-	v
A_C01	4058	- - - -F-D- - - - -N-E-T	-R- - - - - - - - - -R-	I-H	-S-R	- - - - -K- - - - -	- - - - -	- - - - -	- - - - -	- - - - -	-
A_C02	GAM	- - - - -I- - - - - - - - - -	-M-T- - -I- - - - -S-R-T-	M-I	-S- - - - -I	D-G-T- - - - -K-V- - -E-N-T- - -O- - - -V-K-K- - -T	-W	- - - -	-H-T	K-P- - - - -	-
A_C02	SAS	- - - - -G-K-I- - - -F-S- - - -N-E-S	-E- - - - - - - - - -R-R	N-H	-A-T	- - - - -D-A-V- - - - -Y- - - - -A-T- - - - -N- - - -E-E-	-Q-	- - - -	- - - -D-	-J-A-U- - -V-	-
A_C12	C1 14	- - - - -I-V- - - - - - - - - -	-R- - - - -S- - - - -	-P-	- - - - -	- - - - -Z- - - - -E- - - - -A- - - - -K-	H-K	-A-	- - - -	-P- - - - -V-V- - - - -	-
A_C12	C1 20	-K-I-A- - - -D- - - - -N- - - -	-R- - - - -S- - - - -	I-L	- - - - -T- - - - -	- - - - -R- - - - -	- - - - -	- - - - -	- - - -	-S- - - -A-	-
A_C12	C1 21	G-R- - - -D- - - - -N- - - -A-M	-T-D- - -I- - - - -	-N-	- - - - -	- - - - -A-A- - - -N- - - -E- - - -N- - - -K-I	K-R	- - - -	- - - -	-P- - - - -V-V- - - - -	-
A_C12	C1 3	G-V- - - - - - - - - -	-N-	- - - - -	- - - - -	- - - - -K-R- - - -S- - - - -R- - - - -Q	-S	-D-	T- - - -	-A- - - - -	-
A_C12	C1 326	G-V- - - - - - - - - -	-R- - - - -G- - - - -M-S-M	- - - - -	- - - - -	- - - - -K-V- - - -R- - - -V-K-D- - -X- - - -R- - - -W	-H	-M-	- - - -	-H- - - -V- - - - -	-
A_C12	C1 327	G-I-V- - - - -N-E-T	- - - - -T- - - - -G- - - - -	-H-	- - - - -T	- - - - -K-A-K- - -S- - - - -	-B	-D-	T- - - -	-A- - - - -T- - - - -	-
A_C12	C1 329	G-M-T- - - -K- - - - -N-	- - - - -G- - - - -	L-A	- - - - -T- - - - -	- - - - -K- - - - -E- - - - -	K-R	-D-	X- - - -	K-P- - - - -	-
A_C12	C1 330	G-D- - - -L- - - - -D-S-	-Y- - - - -	- - - - -	- - - - -	- - - - -M-M-D- - - -R- - - -K- - - -E- - - -Q-Q	K-	- - - -	-T-T	- - - - -I- - - - -M- - - - -	-
A_C12	C1 42	G-D- - - -S-L-D- - - -E-	- - - - -I- - - - -	-H-	- - - - -	- - - - -Y- - - - -M-V- - - -R- - - -K- - - -G- - - -Q-Q	H-	- - - -	I-N-D	K- - - - -	-
A_C12	C1 45	-K-A- - - -K-L- - - - -A-	- - - - -	- - - - -	- - - - -T- - - - -	- - - - -G- - - - -E- - - - -K- - - - -A-	J-	- - - -	-S-P- - - -V- - - - -	- - - - -	-
A_C12	C1 47	-G-I-V- - - - - - - - - -	- - - - -Q- - - - -Y- - - - -	-T	O-X- - - - -T- - - - -	-K- - - - -	- - - - -	- - - - -	- - - -D-	R- - - - -V- - - - -	-Y
A_C01	CA1	-K-I- - - -F-D- - - -T-J- - - -F-N-S	-R- - - - -	I-H	- - - - -S	- - - - -T- - - - -A-D- - - -N- - - -K- - - -R- - - -E	I-Y	- - - -	- - - -T	K-M- - - -V- - - - -	-
A_C01	CA11	G-K-V- - - - -V- - - - -A-B-	-Q- - - - -	I-B-	- - - - -T- - - - -	- - - - -N- - - - -G- - - - -A- - - - -V- - - - -E-	-G-M	- - - -	- - - -	- - - - -I- - - - -	-
A_C01	CA15	-A-V- - - - -S-D- - - - -E-	- - - - -R- - - - -	- - - - -	- - - - -T- - - - -	- - - - -A- - - - -G- - - - -E- - - - -	-L-K	-D-	T-T- - -	- - - - -	-
A_C01	CA17	G-K-I- - - - -D-T-	-H-T- - -T- - - - -	-A-A	- - - - -T- - - - -A-A	- - - - -I- - - - -Q-K-T- - -K- - - -E	K-	-C-	- - - -	S- - - - -	-
A_C01	CA18	K-B-V- -									

```

A_CONSENSUS_34      LLLGSLAEVHVIASIDTDAKTLIVQWCP..VAIN  CTRP..KMTYKKS..VEL..GPQQ..AETAY..GULIQ..DZAGQAC  NVSTNTHQNTLQZVATQLR..YF..MKT..IIFA..KSEGGOLELTHNFCGGEPFFCYH
A_KK1.KEN983        ---K-----M---FMTT---I-----L-G..IH-----S-----S-V..M-----
A_KK1.KEN984        -K-I-----H-----T---T-D---Y-----H-----SR-----M-----
A_KK1.KEN985        ---V-----Y-----E---I-----T-----E-----M-----
A_KK1.M1             -G-V-----H-----G..H-----H-----K-V---NT---K-T-----PL-V-----
A_KK1.SF1703        ---GE-K-----M-T---E-----Y-----AD-----G-M---KS---S-----S-----
A_KK1.W2BN008        ---V-----T---I-----S-R-----S---EA-R---VE---G-R-G-----M-T---
A_KK1.W2BN009        ---MET---Q---S-----E-----H-----V-----T---DQ-K---R---K---ER-SH---E-I---T---E---
A_KK1.W2BN016        ---D-A---M-T---I-----P-----R-----H-----EK---E---GE-R-R---V---
A_KK1.W2BN020        ---D-A-----G-----D-A-----G-----AK---D---RG---KH---E-M-K-----E-K-
A_KK1.W2BN021        ---D-A-----D-A-----AK---D---RG---KH---E-M-K-----E-K-
A_KK1.W2BN024        ---I---L---R-----T-----T-----R-V---TEN---M-----N---G---R-G..KMK---T---T-K
A_KK1.W2BN025        ---M-----V-----T-----T-----E-----R-F---T-NES---E---B-----O---K-E-----K-T
A_KK1.W2BN026        ---S-----S-----R---I-----S-M-----K-----S---M-----V---S-----H-S-
A_T2.TAM14.2        -R-----M---PAB---Y-----H-----AP-REA
A_T2.TAM1800H        -K-----M---YE---E---I-----H-----TS---K---T---K-VK---KQ---G-----
A_T2.TAM18.1        ---M---FI---E-----H-----VU---T---H-Q-----A-----R
A_T2.TAM18CON        -R-----M---FSR---R-----R-----A---T---XS---D---I---E-----
A_Q0.1033           -M---FTT---T---T---IHL---R---KB-VK---AD---KB-VK---
A_Q0.10455          -R-IR---F-----M---N---S-Y---MIXYS---S---V---K---ND---R-I---E---K-M---S---I-----
A_Q0.1066           -K-K---SD-----T---T---TFRV-RI---IH---R---S---SGL---Y---GSQ---V---R---E---N---T---H-S-KD-----V-----
A_Q01.W2BN029        -SD-----T---T---T---ISARR-IH---R---GHT---M---T---GSQ---T---VKB---R-FMQ---T---V-K-S-
A_Q01.W2BN031        -S-----M---M-S---I---S-----V-----M-----T---T---K-S-----T-VKB---M-N-M-----H-S-
A_Q01.W2BN037        ---V-N---NES---T---R---H-----T---R---KA-----VS-----
A_Q02.111           SP---GX-----L---M---T---T---H-----R-----KV-----VS-----
A_Q02.115           BSP---GX-----M---TS---I---IH---R-----H-----KA-----VS---T---T---T---P-
A_Q02.116           ---GX---M-L---PVTB---Q---G-----E---R-----G-R-----E---R---EGLH---T---D-X-----P-
A_Q02.117           ---I---M-L---ENED---Q---T-----T-----I---GAX---K-K-DL---T---H-T---
A_Q02.118           VNP---C-R---K---M---PKY---G---H-----H---E-S---R-V---E---K---T---S-G---T---T---P-
A_Q02.119           -DPK---II-----L---M-KI---I---G-----Q---H---K---L---G-V---Y---A-L---G---L---H-E---T---P-
A_Q02.124           ---SD-----FWS---T---H---H---R---A---R---S---R-REA-K-D---E-H-K-
A_Q02.71            ---K-----M---T---H---R---H---S---K---H---S-VK---T---Q-
A_Q02.78            X---A---K-K---SD---L---T---G---R---T---T---E---KE---K-VK---T-HM---P-
A_Q02.82            -CQ-V---F---M---FTT---M-T---T---T---E-MV---E---V---T---E-R---T---T---P-
A_Q02.84            -DKI---M---FTT---E---G---T---A---CM---G-K---E---X---EMFG---T---P-
A_Q04.U03273        ---T---K---SD---L---TT---G---T---K---K---K-VK---T-HQ---T---T---K-K-VK---T-HQ---
A_Q04.U0275        ---GX-K---M-L---TT---T---G---T---S---L---L---E---G---V---E---R-VS---T-HQ---G---R---D-
A_R1.6321           -O-R---F-D---I---M-T---K---I-S---R---F---D-SK-A---MFMV---D---V---
A_R1.6557.ph        ---S-----M---E---S---R-Q---S---A---N---A---T-N-K---N---D-KA---E---H-T-
A_R1.6559           ---D-----Q---H---S---BQ---H---V---R---DR---K-K-Y---E-M-V---M---R---A---G---S---Q---M-T-R---A-
A_R1.6563           ---F-D-Q-----E---M-S---H---T---T---H---R---R-E---T-TN---TT---M-S-
A_R1.6569           --A-----Q---R-V---I---S---IHL---IL---S---HKD---PA-HK---M-H---M-S-
A_R1.6571.ph        -A---D---M---PMS---Q---I---V---V-TN---I---XA---K-VK---H-P-
A_R1.6649           ---P-----L---D---S---I---V---N---A-S---D---T-H-N---M-NP---
A_R1.6653           ---P-----L---D---S---I---V---N---A-A---M---K---K---C-P---N-S-
A_R1.6657.ph        ---P-----L---D---S---I---V---N---E---E---E---H-K-VS---QT---R-T---N-T---S-
A_R1.6663           ---F-----H-T---Q---T---A-A---N---H-K---VB---T-L-A---A---V-T-S-V-

```

[illegible]

III-55
DEC 94

[illegible]

Chen 15

III-56
DEC 94

III-57
DEC 94

B_CONSUME_94	WIRSDONTMANTIVOLNES..VEIM	CTRP..NDTAKS..ZNI..GPCR..APYTT..GRIQ..DIRQMC	N1GRANDSHTKQIVKRE..QEG..NKT..	IVFM.QE2GOGPEIWMSPACGGEPFVCH
B_NU4.13064_90	-A-S-I-----T-----	-----A-----D-----	-X-----DD-RR-----	-----
B_NU4.13073_90	-K-S-T-S-----V-----	-----H-----F-----K-----	-----D-----K-----	-A-S-P-----
B_NU4.13077_91	-A-N-I-----F-----	-----E-----A-----	-----D-----K-----	-----
B_NU4.13084_90	-N-N-I-----S-O-----	-----D-----	-----T-----RV-V-----	-----
B_NU4.13091_91	-N-N-I-----O-----	-----R-----	-----	-----
B_NU4.131_90	-A-N-I-----V-----	G-SRO-----	-----	-----
B_NU4.139_91	-A-S-T-----T-----	-----H-----D-----	-----Q-----K-G-----	-----
B_NU4.14004_86	-D-S-----Y-----	-----PM-----D-----	-----D-R-----	-----
B_NU4.14004_86	-D-M-I-----Q-----	-----H-----P-----	-----E-----K-----	-K-T-K-----
B_NU4.15020_89	-TA-----A-----	-----G-----P-----	-----D-----	-----
B_NU4.15032_87	-----N-----A-----	-----A-----A-----	-----E-----	-----
B_NU4.15037_89	-----I-----R-----	-----K-----A-----	-----T-----R-----	-----
B_NU4.15038_89	-H-S-T-S-----V-----	-----H-----D-----	-----D-----K-----	-A-S-P-----
B_NU4.17025_91	-----N-----I-----	-----P-----Q-----	-----I-----TD-----	-----
B_NU4.170_88	-A-T-----K-----I-----	-----G-----H-----	-----SD-EA-----V-N-----	-----
B_NU4.190_89	-A-SH-V-----	-----G-----S-----	-----Q-----R-----D-----	-----
B_NU4.1999_86	-D-H-I-H-----	-----S-----	-----B-----R-----	-----
B_PRL.D287732	-----N-----X-----	-----D-----V-----Y-----	-----I-----VA-----	-K-----
B_PRL.D287726	-----NDP-----IX-----	-----A-----	-----S-----H-----	-----
B_PRL.D287729	-----NDP-----IX-----	-----A-----	-----S-----H-----	-----
B_PRL.D287733	-----NDP-----IX-----	-----A-----	-----S-----H-----	-----
B_SE1.1032	-----V-----H-----K-----	-----G-----A-----D-----	-----T-----K-----	-----
B_SE1.1433	-A-SH-----	-----I-----Q-----	-----I-----A-----E-----	-----
B_SE1.1846	-D-----I-----T-----	-----G-----P-----	-----?-----D-----ADG-----	-----
B_SE1.2815	-I-M-----P-----	-----T-----D-----	-----E-----T-----	-----
B_SE1.2951	-D-----KDP-----IX-----	-----R-----M-----	-----E-----K-----ADG-----	-----
B_SE1.930	-I-T-----KT-----T-----	-----A-----A-----Y-----	-----I-----T-----G-----K-----	-----
B_SE1.931	-M-----K-P-----	-----A-----A-----N-----	-----I-----Y-----K-----	-----
B_TH1.7747	-----TRL-----	-----L-----M-----Q-----	-----ST-----	-----
B_TH1.8641	-V-S-----V-----	-----L-----Q-----M-A-----	-----ST-----R-TK-----	-----
B_TH1.8643	-S-----RV-----	-----L-----Q-----M-A-----	-----ST-----T-----D-----	-----
B_TH1.8645	-S-----V-----	-----L-----Q-----M-A-----	-----ST-----T-----T-----	-P-Q-----
B_TH1.8647	-S-----RV-----	-----PL-----M-A-----	-----ST-----R-T-----	-----
B_TH1.8649	-S-----RV-----	-----PL-----Q-----M-A-----	-----ST-----S-----A-----	-----
B_TH1.8651	-S-----RV-----	-----PL-----Q-----M-A-----	-----ST-----K-----E-----	-----
B_TH1.8653	-S-----RV-----	-----PL-----Q-----M-A-----	-----ST-----T-----T-----	-----
B_TH1.8655	-S-H-RV-----	-----PL-----Q-----M-A-----	-----ST-----T-----T-----	-F-----
B_TH1.8659	-----RI-----	-----L-----Q-----M-A-----	-----ST-----T-----T-----	-----
B_TH2.79132	-X-----K-F-----	-----Y-----X-----M-----	-----R-----I-ST-----K-----R-----	-B-M-----
B_TH3.M2TH014	-----RV-----	-----L-----M-----Q-----	-----ST-----H-T-----	-----
B_TH3.M2TH026	-N-N-I-----	-----L-----Q-----M-----	-----ST-----R-T-----	-----
B_TH4.M752	-T-----I-----	-----G-----P-----	-----TQ-D-A-----AK-G-----	-L-M-----
B_TH4.M753	-T-----I-----D-----	-----G-----P-----	-----I-TD-KKA-----S-Q-----YR-----	-A-----
B_TH4.M754	-T-----I-----	-----G-----P-----	-----I-TD-A-----AQ-----	-L-M-----
B_TH4.M755	-T-----I-----	-----G-----P-----	-----I-TD-A-----EQ-----	-L-M-----
B_TH4.M756	-T-----A-----J-----	-----G-----P-----	-----I-TD-A-----AK-----	-L-M-----
B_TH4.M757	-T-----A-----I-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M758	-T-----Q-----I-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M759	-T-----I-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M760	-T-----KFP-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M761	-N-V-M-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M762	-N-RV-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_TH4.M763	-N-RV-----	-----G-----P-----	-----I-TD-A-----K-----	-L-M-----
B_UK.CM1	-----N-----K-F-----	-----L-----M-----Q-----	-----ST-----T-----T-----	-----
B_UK.V12	-----N-----	-----S-R-I-R-----	-----S-----I-----	-----
B_UK.V74	-N-V-----V-----	-----S-R-I-R-----	-----S-----I-----	-----
B_UK.V77	-----KDP-----M-T-----	-----S-----R-----	-----S-----I-----	-----
B_UK.V87	-----KDP-----M-T-----	-----S-----R-----	-----S-----I-----	-----
B_UK.V88	-----KDP-----M-T-----	-----S-----R-----	-----S-----I-----	-----
B_UK.V89	-----KDP-----M-T-----	-----S-----R-----	-----S-----I-----	-----

3_CONSENSUS_94
B_US1.HC40
B_US1.HC41
B_US1.HC42
B_US1.HC43
B_US1.HC44
B_US1.HC45
B_US1.HC46
B_US2.D208536
B_US2.D208557
B_US2.D208560
B_US2.D208562
B_US2.D208564
B_US2.D208565
B_US2.D208704
B_US2.D208705
B_US2.D208706
B_US2.D208714
B_US2.D208717
B_US2.D208724
B_US2.D208725
B_US2.D208726
B_US2.D208727
B_US2.D208711
B_US2.D208712
B_US2.D208715
B_US2.D208716
B_US4.ZhuPtA
B_US4.ZhuPtF
B_US4.ZhuPtL
B_US4.ZhuPtR
B_US4.ZhuPtV
B_US5.pt5.b1
B_US5.pt6.b1

V3 LOOP
YVIRSEPTDQAKTIVQLMES..VEM CYRP.NNYTRKS..INX..GGRH..AFYTP..GQEEG..DEIQARD NLSRAQAGVYLMQIVKLAE.QFG.MKY...IYFN.QSSGGDMETVMSFNCGGEFFYCHT
-----M-----A-----D-----D-----L-----T-----I-----
-----T-----P-----S-----A-----D-----L-----M-----G-----HL-----HQ-----
-----L-----H-----K-----P-----G-----R-----A-----M-----L-----I-----AD-----E-----K-----P-----T-----
-----H-----V-----P-----D-----A-----Q-----V-----A-----K-----I-----H-----T-----
-----H-----T-----Q-----D-----A-----D-----I-----A-----M-----T-----A-----K-----M-----I-----R-----P-----
-----L-----D-----V-----G-----A-----N-----T-----Q-----N-----
-----M-----T-----IK-----G-----V-----H-----I-----STD-----GK-----EK-----E-----K-----S-----M-----H-----
-----L-----T-----KMP-----V-----PM-----K-----M-----A-----K-----Y-----I-----TE-----K-----EG-----K-----D-----R-----AD-----L-----T-----
-----N-----K-----A-----V-----V-----I-----ST-----G-----K-----D-----R-----D-----K-----AD-----L-----T-----
-----I-----V-----H-----Y-----I-----A-----D-----K-----D-----SEA-----N-----TAR-----K-----L-----T-----T-----L-----M-----I-----S-----
-----I-----H-----I-----FM-----V-----M-----Y-----I-----K-----D-----T-----H-----K-----I-----GE-----EK-----K-----K-----R-----D-----K-----H-----Y-----
-----M-----T-----V-----H-----V-----I-----M-----D-----GH-----E-----A-----C-----K-----K-----
-----I-----I-----K-----N-----M-----K-----E-----E-----T-----AD-----K-----Y-----K-----
-----S-----H-----L-----M-----T-----I-----R-----K-----A-----D-----I-----T-----E-----QR-----K-----G-----P-----
-----P-----KMP-----I-----H-----N-----D-----M-----M-----T-----E-----R-----V-----K-----EK-----A-----K-----P-----T-----SS-----
-----SM-----A-----A-----Q-----K-----IR-----SM-----R-----Q-----N-----M-----T-----E-----R-----V-----K-----EK-----A-----K-----P-----T-----
-----K-----V-----P-----R-----A-----H-----R-----D-----I-----TC-----D-----H-----T-----K-----AD-----
-----N-----A-----S-----V-----V-----D-----K-----I-----AT-----D-----K-----K-----
-----M-----K-----A-----L-----A-----K-----T-----NKT-----EA-----N-----G-----I-----K-----EQ-----
-----A-----L-----I-----M-----I-----D-----K-----I-----M-----T-----E-----R-----E-----E-----K-----L-----L-----S-----
-----A-----L-----QM-----Q-----Q-----Q-----S-----G-----EN-----T-----K-----P-----L-----L-----S-----
II-----A-----S-----A-----P-----D-----K-----I-----E-----G-----G-----E-----A-----T-----R-----T-----
-----K-----T-----I-----T-----S-----R-----H-----L-----D-----K-----I-----E-----K-----T-----YE-----I-----I-----K-----P-----
-----A-----Q-----I-----I-----Q-----K-----E-----I-----C-----G-----G-----EK-----R-----R-----H-----M-----V-----
-----N-----Q-----V-----R-----P-----D-----Q-----GIC-----E-----R-----K-----H-----I-----T-----T-----
L-----D-----I-----TH-----K-----V-----V-----D-----Q-----T-----N-----T-----E-----AK-----
-----L-----N-----T-----T-----S-----R-----A-----V-----EQ-----R-----I-----K-----R-----R-----VA-----HL-----K-----H-----T-----P-----
-----M-----K-----A-----M-----T-----E-----R-----A-----D-----I-----GT-----N-----Q-----P-----T-----F-----

C_CONSENSUS_94 LLNGSLADGSIIRSENLTAAKTIIVHLNES..VELV CTRP..NRRTKRS..IRI...GPGQ..TYYAT...QDIIO...DIRQAHG NISKDIGNMTLQKVGKRLAE..HFF..MKT...IKFA..PSGGGLEIITTHSTNCRGEFFYGM
 C_M1.N2BR025 -----E-----K-----D-----G-N -----A-----E-----L-----A-----X-----
 C_CF1.15166 -----E-D-T-I-----Q-K-----SL-----A-----M-----V-TD-----KH-----TK.1-----H-----T-----
 C_DJ1.DJ259 -----DD-----A-I-----Q-N-----E-----A-----A-----RKE-----RK-RE-----KT.L-RTAN-----T-Q-----Q-----
 C_DJ1.DJ373 -----T-ED-----A-I-----Q-Q-----D-----Q-----A-----RQ-----K-----Q-KG-----K-----I-----G-----
 C_IN1.D1044 -----RS-R-----L-----P-----Q-----H-----A-----M-----RS-----G-----
 C_IN1.D747 -----A-----Q-----Y-----O-----V-----A-----H-----R-----S-----
 C_IN1.D757 -----A-----Q-----R-----Y-----A-----E-----BQ-----V-----S-----G-----
 C_IN1.D760 -----A-----Q-----V-----A-----E-----E-----E-----M-----S-----
 C_IN2.D1024 -----A-----Q-----E-----S-----A-----E-----A-----G-T-R-----S-----S-----S-----
 C_IN2.D744 -----A-----Q-----E-----S-----A-----E-----A-----G-T-R-----S-----S-----S-----
 C_IN2.D766 -----R-----L-----Q-----L-----A-----R-----E-----S-----S-----S-----
 C_IN2.D808 -----R-----Y-----Q-----T-----A-----E-----E-----S-----S-----S-----
 C_IN2.D868 -----D-----EQ-----A-I-----DGP-----R-----A-----R-----R-Y-----E-----S-----S-----
 C_M1.12199 -----A-----M-----D-----A-----A-----Y-----A-----E-----K-----EQ-----R-----E-----T-----E-----H-----
 C_M1.12203 -----D-A-----P-----A-----A-----A-----N-----M-----M-----K-----Q-----C-----L-----R-----C-----V-----
 C_M1.12205 -----A-----D-----A-----Q-----A-----V-----K-----A-----RNR-----T-----N-----G-----I-----P-----
 C_M1.12209 -----K-----D-----T-----T-----Q-----A-----F-----RG-----A-----E-----R-----T-----N-----G-----I-----P-----
 C_M1.12213 -----A-----D-----ID-----A-----A-----G-----A-----H-----H-----E-----E-----G-----O-----O-----
 C_M1.12215 -----A-----B-----R-----B-----A-----G-----A-----H-----H-----E-----E-----G-----O-----O-----
 C_M1.12225 -----KP-----V-----A-----Y-----A-----C-----N-----H-----E-----K-----E-----X-----S-----G-----H-----
 C_M1.12227 -----D-----A-----A-----N-----A-----N-----A-----AG-----N-----Q-----A-----QNN-----S-----S-----
 C_M1.12229 -----B-----V-----M-----M-----M-----ER-----SE-----X-----H-----R-----E-----P-----
 C_M1.12233 -----D-----P-----G-----M-----P-----M-----RE-----H-----R-----E-----E-----E-----P-----
 C_M1.6506 -----A-----E-----M-----M-----M-----H-----BK-----SE-----K-----N-----
 C_M1.6508 -----D-A-I-----M-----V-----A-----M-----EQT-----KI-----MK-----D-----OK-----LP-----S-----VNS-----M-----PESMS
 C_M1.6510.sh -----A-----D-----V-----A-----N-----V-----BH-----S-----B-----S-----V-----HA-----
 C_M1.6512 -----E-T-----I-----K-----Q-----V-----A-----R-----E-----K-----SE-----Q-----A-----D-----K-----
 C_M1.6518 -----A-----E-----A-----G-----R-----V-----N-----N-----A-----V-----S-----K-----K-----T-----E-----H-----
 C_M2.D3MA959 -----A-----F-----A-----A-----V-----N-----NE-----M-----E-----G-----M-----A-----EK-----K-----G-----G-----S-----E-----
 C_M2.D3MA960 -----A-----F-----A-----A-----V-----N-----NE-----M-----E-----G-----M-----A-----EK-----K-----G-----G-----S-----E-----
 C_SN.SB304 -----A-----Q-----Q-----A-----A-----M-----V-----A-----E-----S-----D-----G-----V-----
 C_SC.1574 -----A-----Q-----Q-----A-----A-----M-----V-----A-----E-----S-----D-----G-----V-----
 C_SC.SM145 -----A-----A-----Q-----Q-----B-----Y-----A-----V-----H-----A-----G-----R-----A-----E-----E-----
 C_DG.DG268 -----I-----E-----A-----I-----Q-----K-----N-----A-----B-----V-----Y-----RHE-----T-----M-----RE-----K-----M-----N-----T-----Q-----
 C_ZA.HOF -----D-----I-----V-----K-----D-----A-----R-----Y-----V-----N-----KL-----EQ-----K-----Y-----A-----N-----S-----
 C_ZH1.ZAN18 -----E-----A-----A-----S-----A-----A-----A-----G-----N-----BH-----K-----K-----K-----D-----OH-----
 C_ZH1.ZAN20 -----S-----E3-----A-----Q-----Q-----K-----I-----A-----G-----A-----P-----A-----RSE-----PH-----SR-----G-----R-----I-----E-----
 C_ZN.2647 -----D-----Q-----KP-----T-----A-----G-----A-----P-----A-----N-----TE-----R-----E-----KE-----


```

      100      200      300      400      500      600      700      800      900
D_CONSENSUS_94  LGLAGSLADEEIIIRSEMLTRAKKIIIVQLNES..VTIM CTLP YNM7BQR...THI...GPGQ..ALYTY...RIIG..DIBQANC MISGAZMNTLQGVAKLGD.LL..NCT...TIIFK.PSGGGPEITTHISPMQGGPFFYCN
D_2R.JY1        -----V-T-----H-----E-----DNKI---S...P-----L-----K-----Y-----A-R-----V-----Q-----
D_2R.MAL        -----M-----D-T-H-----T-----G-----NG...I-P-----G-V-----R-Y- T-NET-D-----V---S---K---N.S
D_2R.NOK        -----V-T-----A---IV-----KV-----S---LA---S---ITGNKKKT---Y-G---X-R-----A-----T---W---T---
D_2R.E226       -----A-----R-I-----S---L-----KV-----RTS...I-G---Y- ---KN-----I---N---
D_2R1.6553      -----K-I---VMM-----A---I---N-----S---L-----KV-----Y- ---T---A---I---N---P---M-D-I-----
D_2R1.6565.sh   -----D-L-N-----I---N-----KS...I-----I---DV-----V- E-N-PK-----
D_2R1.6551.sh   -----P-----AT---A---N---KG...L-----S---GD-V-----
D_2R1.6559.sh   -----D---T-----N-V---P---F---B

```

HIV-1 V3 Region

III-64
DEC 94

V3 LOOP

Q_CONSENSUS_94	LLANGSLASERIAISSTWYDAAKVIIVQAKNS..IEM	CTRP.MDYTRKS...ITL...GPGQ..AFYAT...GDIIG...DIRQAKC	NVSTKWKQMLQVQAQAKIV...NKT...IFN.SSAGGLEITTHSPNCRGFFYQNT
Q_CF.4067	---G--H---L--N--I---V--I---	---SF---...--L---...A---	---R--A--N---TIR-KETENG-I...LEN-TYN-P---V--K---C---
Q_GA.LBV21.7	---N---N---N---F---D-V---	---H---...--L---...A---	---E--D--RD---K--K---Q3---...S---...P---V--K---
Q_GA.V2525	---K--I---N---N---BT...V---	---KF---T-R...VL---...A---N---	---TE--QK---A--K--Q-EFN.K-I---...S---...K---
Q_M31.G3	---Q--V---L---T---D---Q---	---R---...--L---...N---	---QDS-QE---K---EQVF...S---...IT---
Q_M31.G5	---SL--T---D---...---	---L---...--L---...N---	---I---...MVB--E--Q...I---...IN---
Q_M31.JP862	---D---...--D---...---	---F---...--L---...N---	---EN-EK---B--VB---D...D---...IS-S---
Q_M31.JV832	---D---...--T---...N---...-	---I---...--P---...--L---...N---	---BI--E--K--T---...DMS-I---

V3 LOOP

N_CONSENSUS_93	NRVIRSKHIZTNTDRIIVKASP..VPIN	CTRP.MDYTRKS...ISI...GPGQ..AFYAT...GDIIG...DIRQAKC	NITREDNRHITLAEVQAQAKIV...NKT...IIFN.SSAGGLEITTHSPNCRGFFYQNT
N_ER.V1557	---	---	---
N_CH.CA13	ER-H--E---N---N---V---	---R---...--L---...A---	---SEAK-NEN-TE-AT--GK-L...KS...K-Q-K---I--IT-N--A---

V3 LOOP

Q_CONSENSUS_95	LILNGTLAKGKIRKSKNLSGKNIIVTMIT..INMT	CERPGHNTV.QS...IRI...GEMA..MTEN...GLAR.MINSKAAVC	INNTITWEKALNOYAEVYLELV...NOTMSVTWTFN.HSAGGLEITTHSPNCRGFFYQNT
Q_XMT70	---N-A-D-LDG---S...L---	---QIDI...N---...--L---...A---	---K--A--G-I---...N-G-IN---...L---
Q_MVP5180	---RE---TE-A---P---...-	---I-E-IAM--D---YT---R--T...LKRSNTSPR-V---	---T--K-V--N--Q---I---N---...EN--II-S-RT---S---
Q_VAU	---N-T---E--LI---N---TIA	---I...--MA---...--L---...A--S-MTGDT---	---SA--N---MIT---...BY---...DVTIKFOMH--S---HPT---

V3 LOOP

U_CONSENSUS_95	LLANGSLABGEIIRSDNTDAAKVIIVQAKNS..IEM	CTRP.MDYTRKS...ITL...GPGQ..AFYAT...GDIIG...DIRQAKC	NVSTKWKQMLQVQAQAKIV...NKT...IFN.SSAGGLEITTHSPNCRGFFYQNT
U_B83.RJ101	---Q--S---I---H---S---E---	---H---...--L---...A---	---AQ---E-I-K-GEGEK...K---...A-T-Q---P---
U_CF1.4040	---L-S-H---SYSF--S...I---	---IRNIR-QTH---...S---...IFT---KV---...K-Y-	---T-NA---K--N---NK-KS.LL...T---...Q---
U_CF1.4056	---TK--S--T-N---KV---N-T---	---T---...--L---...A---	---D--K--N--VR---
U_CF1.4067	---N---N---DS---C---...-	---G-T--R...N---...R...T--TI---...G--K--	---VMS-E--E---N---E-E
U_CF1.4081	---I---E---S---S---B---	---V---...--L---...A---	---V---Q---I---Y-G--T---...T-NF--K---
U_XE.K124	---VK--Q-L-N---VK---R---	---I---...--L---...A---	---VMS-E--E---IS--K---N---...--K-P---P---
U_XE.KEM976	---K---N---PT---...I---	---F---...--L---...A---	---V---K---K---K---E---
U_ML.A21	---Q--S--T---H---T---...-	---R---...--L---...A---	---D-MU-Q--K--E--YMK-RT--P---...-Q--Q---
U_EN.SAM184	---VK---N---PT---...S---	---S---R---...--L---...A---	---S-ME---A---K---GM,Y-P---...S-A-N---C---
U_ZA.23	---V---...--T---...K---	---GEMKIR-QS---...K--V---K---Q-T---	---EDGE-R---IA--R-Q-N--S---...--I---T--Q---

A Subtype

At this time there are viral sequences from 124 HIV-1 infected individuals associated with HIV-1 subtype A. The A subtype consensus sequence (A_CONSENSUS_94) generated from these sequences was based on the most common amino acid found in each position in an alignment of these sequences. 105 of these sequences have been published and/or have been made available for printing in the database by their authors.

1. **CF1.ID#:** These twelve sequences are from a set of sequences obtained from 27 symptomatic patients from the Central African Republic, from whom blood was drawn in 1990-1991. Consensus, PCR-clones, cell culture, DNA. Murphy E, et al., AIDS Res. and Human Retroviruses 9:997-1006 (1993). GenBank accession numbers L11457-L11458, L11461-L11463, L11469-L11471, L11474-L11475, L11477-L11479, L11484-L11496, L11498, L11518, and L11533-L11524.
2. **CF2.GAN and SAS:** These sequences were kindly provided prior to publication by Dr. MP Kieny of Transgene, Strasbourg, France. They are a part of a set of 14 HIV-1 gp120 isolates from Bangui, Central African Republic. The gp120 proteins have been expressed in a hybrid gp160 vaccinia virus expression system. Year of isolation and health status of individuals from which the virus was isolated were not provided. Viruses were isolated in the CAR by C Mathiot and B You (Pasteur Inst., Bangui), grown by F Barre-Sinoussi and A Deslandres (Pasteur Inst., Paris), and cloned and sequenced by D Schmitt and MP Kieny.
3. **CI2.CI-ID#:** These sequences are from 11 of 13 isolates from individuals from Abidjan, Cote d'Ivoire. CI-14 and CI-20 were symptomatic, and the others were asymptomatic. CI-14, CI-45 and CI-47 were serologically dually reactive for HIV-1 and HIV-2. The C2V3 region is part of a 900 bp sequence. Samples were collected between May 1990 and Sept. 1991. Virus was cultured with donor PBMCs, nested PCR amplified, 3-4 clones were sequenced, and the consensus of those clones is presented here. Janssens W, et al. AIDS 8:21-26 (1994). GenBank accession numbers X72024-X72027, X72030-X72039, X72043-X72056, and X72059-X72065.
4. **CM1.CA-ID#:** These sequences are 11 of 17 sequences from a very diverse set of isolates from Yaounde and Douala, Cameroon. The sequences were derived from asymptomatic (CA7, CA11, CA15, CA17, CA18, and CA21) and symptomatic (CA1, CA2, CA6, CA19 and CA22) individuals. Virus was isolated by culture with donor PBMCs, and nested PCR amplified. A single clone was sequenced representing each HIV-1 isolate. Nkengasong JN, et al., AIDS 8:1405-12 (1994). GenBank accession numbers for the entire set of 17 envelope sequences: X80438-X80454.
5. **DJ.DJ-ID#:** These three sequences from Djibouti were from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession numbers L22939, L22941, and L23064.
6. **GA.VI191:** This sequence from Gabon was from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22952.
7. **GH.D687:** A single sequence from an individual from Ghana, provided by Georg-Speyer Haus, Frankfurt, Germany (Dr. Ursula Dietrich). GenBank accession number L07652.
8. **KE.K89:** This sequence is named "KENYA" in the GenBank entry, but is identified as K89 in the original manuscript. It is a Kenyan sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22943.
9. **KE.KEN-ID#:** These 19 sequences were derived from patients who were part of a 1990-1992 cohort study of maternal risk factors in mother to child transmission, including 22 pregnant women and an infant from Kenya. The C2V3 region was sequenced. Janssens W et al., in press AIDS Res. Hum. Retroviruses (1994). GenBank accession numbers for the entire set of 23 patients surveyed in this study: U12984-U13006.
10. **NG.NI:** A full gp120 envelope sequence from Nigeria was kindly provided by Dr. Tom Howard from the University of Southern California (USC) Howard et al. AIDS Res. and Hum. in press (1994).
11. **RW.SF1703:** This sequence is from Rwandan isolate sf170, a biologically active clone reported to be macrophage-tropic. Cheng-Mayer C, Homsy J, Evans LA, and Levy J. Proc. Natl. Acad. Sci. USA, 85:2815-19 (1988); and, Evans, L, Higgins, D, Cheng-Mayer C and Levy J. Unpublished (1990). GenBank accession number M66533.

12. **RW.W2RW-ID#:** Eight sequences from asymptomatic individuals from Rwanda sampled in 1992. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the Human Retroviruses and AIDS 1993 database. Relevant papers are in press: de Wolf F, et al. AIDS Res. and Human Retroviruses (1994); Osmanov S, et al. AIDS Res. and Human Retroviruses (1994); Gao F, et al. AIDS Res. and Human Retroviruses (1994), all in the November issue. GenBank accession numbers U08630-U08641, U08647-U08665, U08763-U08766, and U08793-U08794.
13. **TZ.TAN-ID#:** These four sequences are from a set of 14 Tanzanian samples from symptomatic individuals, using serum samples taken in 1988 to generate PCR clones from viral RNA for sequencing. Zwart G, et al., AIDS 7:467-474 (1993). GenBank accession numbers L01313, L01315-L01316, L01335, L01337-L01339.
14. **UG.1033:** This sequence is a consensus sequence of blood and CSF samples taken from a Ugandan patient 1033, CDC class IV-A. sequences which compose this consensus are: Z23182-Z23184, Z23220 Keys et al., Virology 196:475-483 (1993). GenBank accession numbers Z23177, Z23182-Z23184, and Z23220-Z23223.
15. **UG.U465:** This sequence is from the 1985 Ugandan isolate U455; the complete genomic sequence is available. Oram JD, et al., AIDS Res. and Human Retroviruses 6:1073-1078 (1990). GenBank accession number M62320.
16. **UG.UG08:** This sequence is from a Ugandan isolate. Atkin A, et al., AIDS Res. Human Retroviruses 9:351-356 (1993). GenBank accession number M98503.
17. **UG1.W2UG-ID#:** Three sequences from asymptomatic individuals from Uganda in 1992. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the Human Retroviruses and AIDS 1993 database. Relevant papers are in press: de Wolf F, et al. AIDS Res. and Human Retroviruses (1994); Osmanov S, et al. AIDS Res. and Human Retroviruses (1994); Gao F, et al. AIDS Res. and Human Retroviruses, November issue (1994). GenBank accession numbers U08666-U08669 and U08767-U08770, U08788-U08792 and U08795.
18. **UG2-ID#:** These 11 sequences are part of a set of sequences derived from 22 Ugandans who were attending an AIDS clinic, sampled in 1990. Consensus, PCR-clones, peripheral blood DNA. Albert J, et al., Virology 190:674-681 (1992). GenBank accession numbers M98902-M98905, M98908-M98910, M98914-M98917, M98919, M98924-M98928, M98933-M98941, M98946-M98966, and M98976-M98978.
19. **UG4.UG-ID#:** Two Ugandan sequences from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession numbers L22957 and L22951.
20. **ZR.Z321:** This sequence is from the 1976 Zairean isolate Z321. Srinivasan A, et al., AIDS Res. and Human Retroviruses 5:121-9 (1989). GenBank accession number M15896.
21. **ZR1.ID#:** These ten sequences are part of a set of 14 A and D sequences from women from Zaire; 8 were healthy, 4 showed minor signs of illness, and 2 had AIDS. PCR-direct, peripheral blood DNA. Potts KE, et al., AIDS Res. and Human Retroviruses 9:613-618 (1993). GenBank accession numbers L19624-L19626, L19628-L19630, L19632-L19634, and L19636.

The A subtype sequences which are not yet published, and for which the authors have not yet given permission for release, are from the following set:

1. **CH1.ID#:** Ten sequences from individuals from the Cote d'Ivoire. PCR-direct, peripheral blood DNA. These sequences were provided by the Centers for Disease Control, Atlanta GA, USA (Dr. Chin-Yih Ou and Dr. Marcia Kalish).
2. **UG3.ID#:** Six sequences from individuals from Uganda. PCR-clone, peripheral blood DNA. These sequences were provided by the Centers for Disease Control, Atlanta GA, USA (Dr. Chin-Yih Ou and Dr. Marcia Kalish).
3. **ZR2.ID#:** Three sequences from individuals from Zaire, provided by the LTCB, NCI, NIH, Bethesda, MD, USA (Dr. Marvin Reits).

B Subtype

At this time we have included viral sequences from 338 HIV-1 infected individuals associated with HIV-1 subtype B. The B subtype consensus sequence (B_CONSENSUS_94) generated from these sequences was based on the most common amino acid found in each position in an alignment of these sequences. Please note that none of the studies which have published sequences of only the V3 loop sequences are included here, as the DNA sequences were deemed too short for phylogenetic analyses. (For example, LaRosa G, et al., *Science* 249:932-935 (1990) and Fouchier RAM, et al., *J. Virol.* 66:3183-3187 (1992).) 328 of these sequences have been published and/or have been made available for printing in the database by their authors.

1. **BE.SIMI84** One of two cloned env sequences from a patient with AIDS from Belgium. A vaccinia construct that expresses this gene was created to vaccinate the patient's non-infected brother with the goal of immune therapy by adoptive transfer of lymphocytes. Bex F, C Vanhulle, S Sprecher, M Geissendorfer, H Rubsamen-Waigmann and A Burny. Unpublished (1992). GenBank accession number L07421.
2. **BR1.ID#:** These 21 sequences represent the B env subtype sequences found among 22 Brazilian outpatients with varying degrees of disease progression. Consensus, PCR clones, peripheral blood PBMC DNA. Potts KE, et al., *AIDS* 7:1191-1197 (1993). GenBank accession numbers for 20 of the 56 clones from which consensus sequences were calculated: L19225-L19236, L19240-L19246 and L20963.
3. **BR2.W2BR-ID#:** 14 sequences from individuals from Brazil. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the Human Retroviruses and AIDS 1993 database. Relevant papers are in press: de Wolf F, et al. *AIDS Res. and Human Retroviruses* (1994); Osmanov S, et al. *AIDS Res. and Human Retroviruses* (1994); Gao F, et al. *AIDS Res. and Human Retroviruses*, November issue (1994). GenBank accession numbers U08670-U08714, U08771-U08778, U08780-U08782, U08792, U08796, U08797-U08798 and U08800.
4. **BR3.RJ- or SP-ID#:** These sequences came from the Brazilian cities Rio de Janeiro and Sao Paulo. The sequences that are very short, containing V3 loop fragments insufficient for phylogenetic analysis, are not included here (5 of the 26). The included 21 viral sequences are of the B subtype. An F subtype and a B-F recombinant were also observed in this set. Year of isolation for the sequences range from 1990-1992 for Rio de Janeiro, and 1992 for Sao Paulo. The only two with CD4+ cells < 200 were RJ636 and RJ27. The CDC clinical class ranged from II-IV. DNA extracted from PBMCs of HIV infected individuals was amplified, and the PCR product was directly sequenced. Morgado et al., *AIDS Res. and Human Retroviruses* 10:569-576 (1994) and Sabino EC et al. *J. Virol.* 68:6340-6346 (1994). GenBank accession numbers U00400-U00401, U00403, U00405, U00407-U00414, U00416-U00418, U00421, U00424-U00425, and U00427.
5. **CI.CI-22:** A single B subtype sequence from a set of 13 isolates from individuals from Abidjan, Cote d'Ivoire. CI-22 was symptomatic. The C2V3 region is part of a 900 bp fragment that was sequenced for each individual. Samples were collected between May 1990 and Sept. 1991. Virus was cultured with donor PBMCs, nested PCR amplified, 3-4 clones were sequenced, and the consensus of those clones is presented here. Janssens W, et al. *AIDS* 8:21-26 (1994). GenBank accession numbers X72040-X72042.
6. **CM.CA5:** A single B subtype sequence from a set of 17 sequences from a very diverse set of isolates from Yaounde and Douala, Cameroon. The sequences were derived from asymptomatic and symptomatic HIV infected individuals, specifically, patient CA-5 was asymptomatic. Virus was isolated by culture with donor PBMCs, and nested PCR amplified. A single clone was sequenced representing each HIV-1 isolate. Nkengasong JN, et al., *AIDS* 8:1405-12 (1994). GenBank accession numbers for the full set of 17 sequences which includes CA5 are X80438-X80454.
7. **DE.D31:** This sequence is from isolate D31. Kreutz A, et al., *AIDS Res. and Hum. Retroviruses* 8:1619-1629 (1992). GenBank accession number X61240.
8. **DE.HAN:** This sequence is from an infectious clone from the German isolate DE.HAN-2. Saenermann U, et al., *AIDS Res. and Hum. Retroviruses* 6:813-823 (1990).
9. **FR.J61:** This sequence is from one of the JBB clones from the French patient Bru. Wain-Hobson S, et al. *Science* 252:961-64 (1991) and Guo et al. *Nature* 349:745-6 (1991).
10. **FR.LAI:** This sequence is from the French isolate LAI (formerly BRU). Wain-Hobson S, Sonigo P, Danos O, Cole S, and Alison M. *Cell* 40:9-17 (1985). Also see: Alison M, Wain-Hobson S, Montagnier L, and Sonigo P, *Cell* 46:63-74 (1986) and Wain-Hobson S, et al. *Science* 252:961-964 (1991). GenBank accession number K02013.

11. **GA.OYI:** This sequence is from the Gabonese isolate OYI (designated elsewhere as isolate 397), isolated from a healthy HIV-1 infected individual. GA.OYI appears to have been the first viral sequence from Africa that phylogenetically clustered with North American viruses. Huet T, et al., *AIDS* 3:707-715 (1992). GenBank accession number M26727.
12. **HT.RF:** This sequence is from the clone HAT-3, from Haitian isolate RF. Starcich BR, et al., *Cell* 45:637-648 (1986). GenBank accession number M17451.
13. **HT.WMJ22:** This sequence is from the isolate WMJ22, isolated from a person of Haitian descent living in the U.S. Hahn BH, et al., *Science* 232:1548-1553 (1986). GenBank accession number M12607.
14. **HT1.D-ID#:** These seven sequences are from Haitians, and are part of a set of sequences generated as part of the DAIDS variation program in the laboratories of Dr. Beatrice Hahn at the University of Alabama, and Dr. Marcia Kalish at the the Centers for Disease Control, Atlanta, GA. Except for D2HA590, the full gp160 was sequenced from clones derived from expanded culture stocks. D2HA590 is a direct sequence from PCR amplified DNA from expanded culture. The sequence ID numbers are abbreviated, for example D2HA590 can be read as DAIDS sequence (D), isolated in 1992 (2), Haitian (HA), patient 301590 (590). GenBank accession numbers: U08441-U08447, U04900. Both U08441 and U08442 are sequences from patient HT1.D1HA651 and are identical over the region of interest. GenBank accession numbers for additional clones derived from these patients: U04905-U04906, U04901 U04902 U04903 U04904
15. **HT2.H-ID#:** These 25 sequences are from Haitians. All sequences were PCR amplified from the infected individuals PBMCs, and this set includes direct sequences of PCR amplification products, consensus sequences of multiple clones of PCR products plus one direct sequence, and single clones of PCR products. These sequences were provided by the Centers for Disease Control, Atlanta, GA USA (Dr. Chin-Yih Ou), and Johns Hopkins University School of Hygiene and Public Health, Baltimore, MD USA (Dr. Neal Halsey), and the Centers for Development and Health, Complexe Medico Sociale de la Cite Soleil, Port-au-Prince, Haiti (Dr. Reginald Boulos). GenBank accession numbers L07145-L07161, L07163-L07165, L07167-L07207, L07209-L07239, L07241-L07246.
16. **IN1.ID#:** These four sequences were isolated in Hyderabad, Andhra Pradesh, in southern India. The C2V3 region of env was amplified by nested priming from DNA from PBs from fresh blood samples. Date of sampling and health status of HIV-1 infected individuals is unknown. Baskar PV, et al., *AIDS Res. and Human Retroviruses* 10:1039-1041 (1994). GenBank accession numbers L29091-L29094.
17. **JP.GUNA:** A Japanese 1989 isolate HIVGUN, infectious to T cells, was adapted to growth on fibroblast-like BT cells. A single amino acid change at the tip of the V3 loop was shown to be responsible for the change in tropism, GPGR to GSGR. Takenchi Y, Akutsu M, Murayama K, Shimizu N, and Hoshino H. *J. Virol.* 65:1710-1718 (1991). GenBank accession number M59192.
18. **JP.JH32:** This is a sequence from a lambda clone of Japanese isolate JH3. Komiyama N, et al., *AIDS Res. Hum. Retroviruses* 5:411-419 (1989). GenBank accession number M21138.
19. **MFA:** This sequence is from the isolate MFA derived from an asymptomatic patient with hemophilia. Stevenson M, et al., *J. Virol.* 64:3792-3803 (1990). GenBank accession number M33943.
20. **NL1.ID#:** These nine sequences are part of a study of presumed donor-recipient pairs from an HIV-1 transmission study conducted in the Netherlands. If pairs were extremely close or identical, only the "recipient" is included here. Recipient samples were from the first sample to be antibody positive, and are numbers 1,3,5,7,9, and 13. These sequences are consensus sequences of multiple clones from PCR amplified serum RNA. Wolfs TFW, Zwart G, Bakker M, and Goudsmit J. *Virology* 189:103-110 (1992). GenBank accession numbers M91819-M91827, M91829, M91831-M91832, M91839, M91857-M91870, M91872, M91874, M91881-M91884, M91891, M91893, 91895-M91908, M91910, M91911-M91926.
21. **NL2.ID#:** These two sequences are part of a Dutch study of mutations occurring over a five year period (starting in 1985) in two patients. Serum RNA was PCR amplified and multiple clones were sequenced. The consensus for each patient is shown. Wolfs TFW, Zwart G, Bakker M, Valk M, Kuiken CL, and Goudsmit J. *Virology* 185:195-205 (1991). GenBank accession numbers M74591-M74684.
22. **NL3.NET-ID#:** These six consensus sequences from the Netherlands are samples from AIDS patients, using serum samples to generate PCR clones from viral RNA for sequencing. Zwart G, et al., *AIDS* 7:467-474 (1993). GenBank accession numbers L01282-L01297.
23. **NL4.ID#:** These 74 sequences represent a study of early seroconverters from different times with different risk factors for transmission during the AIDS epidemic in the Netherlands. The year the sample was taken is indicated in the last part of the sequence name. The risk group of the individual from whom the virus is derived is indicated in the first letter of the sequence name (D, B and H for IVDUs, hemophiliacs, and

homosexuals, respectively). Viral genomic RNA from sera was PCR amplified and amplification product was direct sequenced. Kuiken CL, et al., *Proc. Natl. Acad. Sci. USA* 90:9061-9065 (1993). GenBank accession numbers Z29219-Z29225, Z29256-Z29258, and Z29262-Z29325.

24. **PR1.D-ID#:** These four sequences are from Puerto Rico, and were generated as part of the DAIDS variation program in the laboratory of Dr. Marcia Kalish at the the Centers for Disease Control, Atlanta, GA. The C2V3 region was directly sequenced from PCR amplification of DNA from viral culture. The sequence ID numbers are abbreviated; for example D2PR732 can be read as DAIDS sequence (D), isolated in 1992 (2), Puerto Rico (PR), patient 301732 (732). A full description of these sequences can be found in the April 1994 supplement to the HIV database, part III. GenBank accession numbers U04926-U04929.
25. **SE1.ID#:** Seven sequences that are consensus sequences of blood and CSF samples taken from each patient. The CDC disease stage class for the patients are as follows: II pts 930, 2815; III pt 931; IV-E pt 2951; IV-A pt 1032; and IV-C2 pts 1433, 1866. Keys et al., *Virology* 196:475-483 (1993). GenBank accession numbers Z23178-Z23181, Z23185-Z23187, Z23192-Z23194, Z23200-Z23219, Z23224-Z23227, Z23232-Z23235, and Z23240-Z23255.
26. **TH1.ID#:** These ten sequences are from individuals from Thailand. PCR-direct, peripheral blood PBMC DNA. Referred to as Thai subtype B in Ou et al. Ou C-Y et al., *AIDS Res. and Human Retroviruses* 8:1471-1472 (1992) and Ou C-Y et al. *Lancet* 341:1171-1174 (1993). (Published erratum appears in *Lancet* 342:250 (1993).) GenBank accession numbers L07442, L07449-L07456 and L07460.
27. **TH2.TB132:** This sequence is from a set of isolates from HIV seropositive individuals from Thailand. PCR, PBMC co-culture, DNA. Full env sequence is available. McCutchan FE, et al., *AIDS Res. and Human Retroviruses* 8:1887-1895 (1992). Please note: the TB132 locus name in the database corresponds to the McCutchan et al.'s "BK132" isolate. GenBank accession number L03697.
28. **TH3.W2TH-ID#:** 2 sequences from Thailand from asymptomatic individuals. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the Human Retroviruses and AIDS 1993 database. Relevant papers are in press: de Wolf F, et al. *AIDS Res. and Human Retroviruses* (1994); Osmanov S, et al. *AIDS Res. and Human Retroviruses* (1994); Gao F, et al. *AIDS Res. and Human Retroviruses*, November issue (1994). GenBank accession numbers U08715-U08718, U08801-U08802, and U08783-U08784.
29. **TH4.ID#:** These twelve sequences are B subtype sequences from Thailand. Ten were genetically most similar to HIV1 found in the Americas and Europe; these sequences were derived from people infected prior to 1988 (diagnosed in 1986 or 1987). The other two (N762 and N763) were designated B' and were isolated from people with more recent infections, 1988 and 1992. The sequences were obtained from PCR amplified PBMC DNA. The naming of the sequences submitted to GenBank does not correspond with the naming of the sequences in the paper. Kalish ML, et al, *AIDS Res. and Human Retroviruses* Nov., in press (1994). GenBank accession numbers for the entire set of thirteen sequences studied in this publication: U15576-U15588.
30. **UK.CAM1:** This sequence is from the British isolate CAM1. McIntosh A, and Karpas A, Thesis (1991), Cambridge University, England. GenBank accession number D10112.
31. **UK.V-ID#:** A set of six sequences from a study of hemophiliacs from Scotland who were originally thought to have been infected by the same batch of factor VIII. (ScV12 is a sequence from a hemophiliac from the U.S., included as a control). All are consensus sequences of multiple direct PCR sequences obtained from limiting dilution of PBMCs. The Scottish hemophiliacs were infected in 1984 and the PBMCs were obtained for analysis in 1989. Although the samples were potentially related, they were deemed sufficiently divergent in this region for inclusion in this set. Simmonds P, Balfe P, Ludlam CA, Bishop JO, and Leigh Brown, AJ. *J. Virol.* 64:5840-5850 (1990), and Balfe P, Simmonds P, Ludlam CA, Bishop JO, and Leigh Brown, AJ. *J. Virol.* 64:6221-6233 (1990). GenBank accession numbers M61327-M61346 and M61391-M61407.
32. **US.ACP1:** This virus was cultured from a seronegative man with Kaposi's sarcoma. (See: Ho D, *The American J. of Med.*, 86:349-351 (1989). ACP1 was the sequenced after one passage in PBMCs. The sequence ACH9 was also derived from this patient. Asikhanazi, A, et al., *Proc. Natl. Acad. Sci., USA* 88:7056-7060 (1990). GenBank accession number M60660.
33. **US.ADA:** This sequence is from the monocytotropic U.S. isolate ADA. Westervelt P, Gendelman HE, and Ratner L, *Proc. Natl. Acad. Med. Sci. U.S.A.* 88:3097-3101 (1991). GenBank accession number M60472.
34. **US.ALA1:** This sequence is from an infectious clone of the 1985 U.S. isolate AL-1, taken from a patient with AIDS. Buckler-White A, Theodore T, et al., Unpublished (1988). GenBank accession number M38430.
35. **US.BAL1:** This sequence is from the macrophage tropic U.S. isolate BAL, harvested from lung alveolar tissue. Reitz M, et al., Unpublished (1990). GenBank accession number M68803.

36. **US.BCSG3:** This is a fragment of a full genomic sequence from the provirus SG3, cloned as a single proviral unit. This clone replicates more efficiently in chimpanzee than in human lymphocytes, and is extremely cytopathic in immortalized human T-cell lines. Ghosh SK, et al., *Virology* 194: 858-864 (1993). GenBank accession number L02317.
37. **US.BRVA:** This sequence is from U.S. isolate BRVA, which was taken from the brain tissue of a AIDS patient with neurological disorders. Anand R, et al., *Virology* 168:79-89 (1989). GenBank accession number M21098.
38. **US.BWB11A:** This sequence is from a study of variation in brain derived HIV-1 isolates by Monken CE, B Wu and A Srinivasan, and is part of a 32 sequence set, GenBank Accession L17097. A clone of proviral DNA was sequenced. (Manuscript in preparation, 1994).
39. **US.CDC42:** This sequence is from an infectious clone of the U.S. isolate CDC-451. Desai SM, et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:8380-8384 (1986). GenBank accession number M13137.
40. **US.JFL:** This sequence is from a non-infectious clone from the monocytotropic U.S. isolate JFL. McNearney T, et al., *Proc. Natl. Acad. Sci. U.S.A.* 87:1917-21 (1990). GenBank accession number M31451.
41. **US.JM:** This sequence, along with B.US.WM, came from viral isolates after short term culture in PBMCs, PCR amplification, and cloning of PCR products. Both are from asymptomatic, seropositive individuals. Ashkanazi, A, et al., *Proc. Natl. Acad. Sci., USA* 88:7056-7060 (1990). GenBank accession number M80662.
42. **US.JRCSE:** This sequence is from an infectious clone of the 1986 U.S. isolate JRCSE, derived from from the CSF of a patient who died with Kaposi's sarcoma and severe AIDS encephalopathy. The infectious clone JRFL was isolated from the brain of the same patient. O'Brien WA, et al., *Nature* 348, 69-73 (1990). Also see: Pang S, Vinters HV, Akashi T, O'Brien WA, and Chen ISY, *J AIDS* 4:1082-1092 (1991). GenBank accession number M38429.
43. **US.MA:** This is a consensus of 35 sequences obtained either after viral culture or from PCR amplification products from patient PBMCs. Patient MA, from the US, was infected in 1984 or 1985, and had been experiencing neurological disorders prior to 1989 when the sample was taken. Kusumi K, et al., *J. Virol.* 66:875-885 (1992). GenBank accession numbers M79342-M79354 and M90025-M90046.
44. **US.MN:** This sequence is from 1984 U.S. isolate MN, taken from a pediatric AIDS patient. Gargo C, et al., *Virology* 164:531-536 (1988). GenBank accession number M17449.
45. **US.NY5CG:** This sequence is from the 1984 U.S. isolate NY5. Willey R.W., et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:5038-5042 (1986). GenBank accession number M38431.
46. **US.P896:** This sequence represents a molecular clone from an primary isolate derived from a Jamaican man who immigrated to Philadelphia 15 years earlier. At the time of viral isolation, he had no antiviral therapy, but was an AIDS patient with < 10 CD4 cells per mm³. The infectious molecular clone from which this sequence was derived is both macrophage-tropic and extremely cytopathic in lymphocytes. Collman R, et al. *J. Virol.* 66:7517-21 (1992). GenBank accession number M96155.
47. **US.RJS:** This is a consensus of six biologically characterized clones from isolate RJS, isolate 4. The HIV-1 infected individual had been infected for five years at the time of isolation. Complete env sequence is available. Daniels RS, Smith MH, and Fisher AG. *J. Virol.* 65:5574-5578 (1991) and Fisher AG et al., *Nature* 334:444-447 (1988). GenBank accession numbers M37491 and M37573-M37577.
48. **US.SB(A-C):** These three sequences are from 1988 U.S. isolates taken from a woman, her daughter and her sexual partner. The three viruses are epidemiologically linked, however the amino acids sequences appeared sufficiently divergent in this region to merit the inclusion of all three samples. Sequences were directly sequenced from PCR amplification products after the virus was briefly cultured. Burger H, et al., *Proc. Natl. Acad. Sci. U.S.A.* 88:11236-11240 (1991). GenBank accession numbers M77228-M77230.
49. **US.SC:** This sequence is from the 1984 U.S. isolate SC, from an AIDS patient. Gargo C, et al., *Virology* 164:531-536 (1988). GenBank accession number M17450.
50. **US.SF162:** This sequence is from an infectious clone from the U.S. isolate SF162, cultured from cerebrospinal fluid. Cheng-Mayer C, Quiroga M, Tung JW, Dina D, and Levy J, *J. Virol.* 64:4390-4398 (1990). GenBank accession number M38428.
51. **US.SF2:** This sequence is from an infectious clone from the U.S. isolate ARV-2. Sanchez-Pescador R, et al., *Science* 227:484-492 (1985). HIVSF13 is a more infectious virus taken from the same patient (C. Cheng-Mayer, UCSF, San Francisco, CA, USA). GenBank accession number K02007.
52. **US.SF33:** This sequence is from an infectious clone from the 1984 U.S. isolate SF33. York-Higgins D, Cheng-Mayer C, Bauer D, Levy JA, and Dina D, *J. Virol.* 64:4016-4020 (1990). GenBank accession number M38427.

HIV-1 V3 Region

53. **US.WM:** This sequence, along with B-US.JM, came from viral isolates after short term culture in PBMCs, PCR amplification, and cloning of PCR products. Both are from asymptomatic, seropositive individuals. Ashkanazi, A, et al., *Proc. Natl. Acad. Sci., USA* 88:7056-7060 (1990). GenBank accession number M80663.
54. **US.YU:** This is a consensus sequence of eight lambda phage clones and 12 PCR amplified clones derived from the uncultured brain tissue of a patient with AIDS dementia complex. A macrophage tropic clone (YU-2) is almost identical to the consensus sequence of YU in this region, with only a single amino acid change (K to N) 10 amino acids from the carboxy-terminal end of the sequence. Li Y, et al., *J. Virol.* 65:3973-3985 (1991). Complete genomic sequences are available for two of the HIV1YU clones, along with biological characterizations of four of the HIV1YU clones: Li Y, et al., *J. Virol.* 66:6587-6600 (1991). The GenBank accession numbers for the YU-10 and YU-2 complete genomes are M93259 and M93258, respectively.
55. **US1.HC-ID#:** These are forty 1990-1991 U.S. samples, from the study of the dentist who was thought to have been the source of HIV-1 infection of six of his patients. Only the dentist's viral sequence and the Florida control sequences are shown here; the six epidemiologically and genetically linked patients are excluded from this alignment as their viral sequences were very similar to the dentist's. All sequences were PCR amplified from patient PBMCs. Most are direct sequences from the amplification products, although some are consensus sequences of multiple clones of PCR products and one direct sequence, Ou, C-Y, et al., *Science* 256:1165-1171 (1992). GenBank accession numbers for 70 of 75 sequences in this set: M90847-M90853, M90881-M90886, M90894-M90900, M90914-M90956, and M90958-M90964.
56. **US2.D-ID#:** These 15 sequences are from the USA, and are part of a set of sequences generated as part of the DAIDS variation program in the laboratory of Dr. Marcia Kalish at the the Centers for Disease Control, Atlanta, GA. The C2V3 region was directly sequenced from PCR amplification products of DNA from viral culture. The sequence ID numbers are abbreviated, for example D2US711 can be read as DAIDS sequence (D), isolated in 1992 (2), United States (US), patient 301711 (711). A full description of these sequences can be found in the April 1994 supplement to the HIV database, part III. GenBank accession numbers U04907-U04915, U04918 and U04921-U04925.
57. **US3.D-ID#:** These four sequences are from the US, and are part of a set of sequences generated as part of the DAIDS variation program in the laboratory of Dr. Beatrice Hahn at the University of Alabama. gp160 sequences of clones from expanded culture stocks are available. The sequence ID numbers are abbreviated, for example D2US711 can be read as DAIDS sequence (D), isolated in 1992 (2), United States (US), patient 301711 (711). A full description of these sequences can be found in the April 1994 supplement to the HIV database, part III. GenBank accession numbers U08448-U08449 and U08451-U08452. GenBank accession numbers for additional clones from these patients: U04916-U04917 and U04919-U04920.
58. **US4.ZhuPt-ID#:** Sequences from five primary seroconverters. Consensus, PCR-clones, peripheral blood PBMC DNA. Although sequences were also available from two of the donors, only sequences from five recipients are shown here. Zhu et al., *Science* 261:1179-1181 (1993). GenBank accession numbers L21224-L21324, L21372-L21373, L21376-L21377, L21380-L21381, L21384-L21389, L21393-L21394, L21397-L21398, L21400-L21401, L21404, L21405, L21408, L21410-L21411, L21414, L21417-L21418, L21426-L21427, L21430-L21432, L21434-L21437, L21440-L21442, L21445, L21447, L21449-L21450, L21453-L21454, L21457, L21459, L21462-L21463, L21519-L21520, L21522-L21523, L21525-L21526, L21528-L21533, L21535-L21536, L21538-L21539, L21541, L21543, L21545-L21546, L21548, L21550-L21551, L21553-L21554, L21556, L21558-L21559, L21561-L21569, L21571-L21578, L21580, L21582, L21584, L21586, L21588, and L21590.
59. **US6.pt-ID#:** A consensus of PCR-clones, peripheral blood DNA. These sequences were part of a study of blood sequences compared to brain sequences from six individuals. Korber BTM, et al., *J. Virol.* 68:7467-7481 (1994). GenBank accession numbers U05360-U05568.

The 10 B subtype sequences which are not yet published, and for which the authors have not yet given permission for release, are from the following sets:

1. **PE:** A single envelope sequence from Peru, provided by the Centers for Disease Control, Atlanta, GA, USA, and the Walter Reed Research Institute.
2. **US:** A sequence from the United States, provided by Dr. Marvin Reitz of the NIH.
3. **US:** Six consensus consensus of PCR-clones, from peripheral blood DNA. These sequences were provided by Dr. Steven Wolinsky of Northwestern University Medical School, Chicago, IL, USA.
4. Two sequences of unspecified origin provided by Dr. Karen Coates Fryer, Macfarlane Burnet Centre for Medical Research, Fairfield, Victoria, Australia.

C Subtype

At this time there are viral sequences from 40 HIV-1 infected individuals associated with HIV-1 subtype C. The C subtype consensus sequence (C.CONSENSUS.94) generated from these sequences was based on the most common amino acid found in each position of an alignment. 38 of these sequences have been published and/or have been made available for printing in the database by their authors.

1. **BR.W2BR025:** This sequence is part of a gp160 sequence from an asymptomatic individual from Brazil, sampled in 1992. A clone was derived from an expanded viral culture, expressed and sequenced. This sequence was provided by the WEO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences from this patient can be found in the April supplement to the Human Retroviruses and AIDS 1993 database. Relevant papers are in press: de Wolf F, et al. AIDS Res. and Human Retroviruses (1994); Osmanov S, et al. AIDS Res. and Human Retroviruses (1994); Gao F, et al. AIDS Res. and Human Retroviruses (1994). GenBank accession number U15121.
2. **CF1.15166:** A single sequence from a set of sequences obtained from 27 symptomatic patients from the Central African Republic, from whom blood was drawn in 1990-1991. Consensus, PCR clones, cultured virus, DNA. Murphy E, et al., AIDS Res. and Human Retroviruses 9:997-1006 (1993). GenBank accession number L11525.
3. **DJ1.DJ-ID#:** These two sequences from Djibouti were from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession numbers L22940 and L23065.
4. **IN1.D-ID#:** These four sequences are from samples from high risk patients in India, PCR clones, DNA, PBMC culture. Dietrich U et al., AIDS 7:23-27 (1993). GenBank accession numbers L07651 and L07653-L07655.
5. **IN2.D-ID#:** These five sequences are from samples from high risk patients in India, primarily stage I. They were nested PCR amplified from DNA obtained from uncultured PBMC from patients serologically defined as HIV-1/HIV-2 mixed infections. Grez M et al., J Virol 68: 2161-8 (1994). GenBank accession numbers U07098 and U07100-U07103.
6. **MW1.ID#:** These 15 sequences are from pregnant women with risk factors from Malawi. PCR-direct, peripheral blood DNA. Orloff GM et al., AIDS Res. and Human Retroviruses 9:705-706 (1993). GenBank accession numbers L07427-L07441, L15721-L15733 and L15735.
7. **MW2.D-ID#:** These two sequences are from individuals from Malawi, generated as part of the DAIDS variation program in the laboratory of Dr. Beatrice Hahn at the University of Alabama. The C2V3 was excised from full gp160 sequences, derived from clones from expanded culture stocks. The sequence ID numbers are abbreviated, for example D3MA959 can be read as DAIDS sequence (D), isolated in 1993 (3), Malawi (MA), patient 301959 (959). GenBank accession numbers U08453-U08454.
8. **SN.SE364:** A Senegalese sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22944.
9. **SO.1574:** This sequence is a consensus sequence of blood and CSF samples taken from the Somali patient 1574, CDC classification II. Keys et al., Virology 196:475-483 (1993). GenBank accession numbers Z23188, Z23190-Z23191, and Z23228-Z23231.
10. **SO.SM145:** A Somali sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22946.
11. **UG.UG268:** A Ugandan sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22948.
12. **ZA.NOF:** This sequence is from a South African individual, Dietrich U et al., AIDS 7:23-27 (1993). GenBank accession number L07426.

HIV-1 V3 Region

13. **ZM1.ZAM-ID#:** Two sequences from Zambia, from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession numbers L22954 and L22956.
 14. **ZW.2647:** This sequence is a consensus sequence taken from blood and CSF samples taken from Zimbabwe patient 2647, CDC classification II. Keys et al., Virology 196:475-483 (1993). GenBank accession numbers Z23196-Z23199 and Z23236-Z23239.
-

The two C subtype sequences which are not yet published, and for which the authors have not yet given permission for release, are from S. Africa, and Zambia, provided by the Dept. of Med. Virology, University of Stellenbosch, Tygerberg, South Africa (Dr. Michael Becker).

D Subtype

At this time there are viral sequences from 73 HIV-1 infected individuals associated with HIV-1 subtype D. The D subtype consensus sequence (D.CONSENSUS.94) generated from these sequences was based on the most common amino acid found in each position of an alignment. All but 7 of these sequences have been published and/or have been made available for printing in the database by their authors.

1. **CF.4020:** This sample was the only D subtype from a set of sequences obtained from 27 symptomatic patients from the Central African Republic, from whom blood was drawn in 1990-1991. Consensus, PCR-clones, cell culture, DNA. Murphy E, et al., *AIDS Res. and Human Retroviruses* 9:997-1006 (1993). A full gp120 sequence from this isolate was kindly provided prior to publication by Dr. MP Kieny of Transgene, Strasbourg, France. It is a part of a set of 14 HIV-1 gp120 isolates from Bangui, Central African Republic. The gp120 proteins have been expressed in a hybrid gp160 vaccinia virus expression system. Year of isolation and health status of individuals from which the virus was isolated were not provided. Viruses were isolated in the CAR by C Mathiot and B You (Pasteur Inst., Bangui), grown by F Barre-Sinoussi and A Deslandres (Pasteur Inst., Paris), and cloned and sequenced by D Schmitt and MP Kieny. GenBank accession numbers L11472-L11473.
2. **CL.CI-13:** A single D subtype sequence from a set of 13 isolates from individuals from Abidjan, Cote d'Ivoire. CI-13 was symptomatic, and serologically dually reactive for HIV-1 and HIV-2. The C2V3 region is part of a 900 bp fragment that was sequenced for each individual. Samples were collected between May 1990 and Sept. 1991. Virus was cultured with donor PBMCs, nested PCR amplified, 3-4 clones were sequenced, and the consensus of those clones is presented here. Janssens W, et al. *AIDS* 8:21-26 (1994). GenBank accession numbers X72028-X72029.
3. **KE1.KEN-ID#:** These three patients were part of a 1990-1992 cohort study of maternal risk factors in mother to child transmission, including 22 pregnant women and an infant from Kenya. The C2V3 region was sequenced. W Janssens et al., in press *AIDS Res. Hum. Retroviruses* (1994). GenBank accession numbers for sequences from the entire set of 23 patients studied in this publication: U12984-U13006.
4. **SN.SE365:** A Senegalese sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22945.
5. **TZ.TAN-ID#:** These ten sequences are from from a set of 14 Tanzanian samples from symptomatic individuals, using serum samples taken in 1988 to generate PCR clones from viral RNA for sequencing. Zwart G, et al., *AIDS* 7:467-474 (1993). GenBank accession numbers L01298-L01339.
6. **UG.U44342:** This sequence is from a Ugandan. Consensus of PCR-clones, peripheral blood DNA. Intact env sequences are available from this sample. Bruce C, C Clegg, A Featherstone, J Smith and J Oram. *AIDS Res. and Human Retroviruses* 9:357-368 (1992). GenBank accession numbers M98408-M98416.
7. **UG.UG23:** This sequence is from a Ugandan isolate. Atkin A, et al. and Pestano G, et al., unpublished. Andrew Atkin is at CUNY, NY, NY, USA. GenBank accession number M98504.
8. **UG1.W2UG-ID#:** Twelve sequences from asymptomatic individuals from Uganda in 1992. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the *Human Retroviruses and AIDS 1993* database. Relevant papers are in press: de Wolf F, et al. *AIDS Res. and Human Retroviruses* (1994); Osmanov S, et al. *AIDS Res. and Human Retroviruses* (1994); Gao F, et al. *AIDS Res. and Human Retroviruses* (1994). GenBank accession numbers HU08721-U08741, HU08786-U08787, U08803-U08809 and U08821-U08824.
9. **UG2.ID#:** These twelve sequences are from 1986-1987 Ugandan samples. Consensus, PCR-clones, peripheral blood DNA. Oram JD, et al., *AIDS Res. and Hum. Retroviruses* 7:605-614 (1991). GenBank accession number for one of the 161 sequences used: M62320.
10. **UG3.ID#:** These 11 sequences are part of a set of sequences derived from 22 Ugandans who were attending an AIDS clinic, obtained in 1990. Consensus, PCR-clones, peripheral blood DNA. Albert J, et al., *Virology* 190:674-681 (1992). GenBank accession numbers L00614-L00618, L00733-L00737, M98894-M98899, M98901, M98906-M98907, M98911-M98913, M98918, M98920-M98923, M98929-M98937, M98942-M98945, and M98967-M98975.
11. **UG5.UG-ID#:** Three Ugandan sequences from a set of HIV-1 viral isolates from Africa. Health status of the individuals from which the virus was cultured was unspecified, and the year of viral isolation was probably

HIV-1 V3 Region

between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession numbers L22947 and L22949-L22950.

12. **US.AMK:** This sequence comes from a student living in Alabama, who moved from Zaire to the US in 1988. Virus was isolated from the patients PBMCs; this isolate was PCR amplified, and amplification products from both gag and env were subcloned and sequenced. His CD4 count was < 5 cells/mm³, and he was symptomatic at the time of viral isolation. Gao F, et al., *AIDS Res. and Human Retroviruses* (1994). GenBank accession number U08193.
13. **ZR.ELI:** This sequence is from the Zairean isolate ELI. Alison M, Wain-Hobson S, Montagnier L, and Sonigo P. *Cell* 46:63-74 (1986). The complete genomic sequence and an infectious clone are available. GenBank accession number M27949.
14. **ZR.JY1:** This sequence is from Zairean isolate Z-84, clone JY1. Yourho J, et al., *AIDS Res. and Human Retroviruses* 4:165-173 (1988). GenBank accession number J03653.
15. **ZR.MAL:** This sequence is from a non-infectious clone of the Zairean isolate MAL. Alison M, Wain-Hobson S, Montagnier L, and Sonigo P. *Cell* 46:63-74 (1986). The complete genomic sequence and an infectious clone from the isolate MAL are available. GenBank accession number K03456.
16. **ZR.NDK:** This sequence is from an infectious clone of the Zairean isolate NDK. The molecular clone is highly cytopathic in vitro. Spire B, et al., *Gene* 81:275-284 (1989). The complete genomic sequence is available. GenBank accession number M27323.
17. **ZR.Z2Z6:** This sequence is from an infectious clone of Zairean isolate Z2. Theodore T, and Buckler-White A, unpublished, and Srinivasan A, et al., *Gene* 52:71-82 (1987). The complete genomic sequence is available. GenBank accession number M22639.
18. **ZR1.ID#:** These four sequences are part of a set of 14 A and D sequences from women from Zaire. 8 were healthy, 4 showed minor signs of illness, and 2 had AIDS. PCR-direct, peripheral blood DNA. Potts KE, et al., *AIDS Res. and Human Retroviruses* 9:613-618 (1993). GenBank accession numbers L19623, L19627, L19631 and L19635.

The seven D subtype sequences which are not yet published, and for which the authors have not yet given permission for release, are from the following set:

UG4.ID#: Seven sequences from individuals from Uganda. PCR-direct, peripheral blood DNA. These sequences were provided by the Centers for Disease Control, Atlanta GA, USA (Dr. Chin-Yih Ou).

E Subtype

At this time there are viral sequences from 65 HIV-1 infected individuals associated with HIV-1 subtype E. The E subtype consensus sequence (E-CONSENSUS.94) generated from these sequences was based on the most common amino acid found in each position of an alignment. All but 17 these sequences have been published and/or have been made available for printing in the database by their authors.

1. **CF1.ID#:** These eight sequences are from a set of sequences obtained from 27 symptomatic patients from the Central African Republic, from whom blood was drawn in 1990-1991. Consensus, PCR-clones, cell culture, DNA. Murphy E, et al., *AIDS Res. and Human Retroviruses* 9:997-1006 (1993). GenBank accession numbers L11459-L11460, L11463-L11468, L11476, L11480-L11481, L11505-L11504, L11511-L11513 and L11519-L11521.
2. **CF2.ID#:** These three sequences were kindly provided prior to publication by Dr. MP Kieny of Transgene, Strasbourg Cedex, France. It is a part of a set of 14 HIV-1 gp120 isolates from Bangui, Central African Republic. The gp120 proteins have been expressed in a hybrid gp160 vaccinia virus expression system. Year of isolation and health status of individuals from which the virus was isolated were not provided. Viruses were isolated in the CAR by C Mathiot and B You (Pasteur Inst., Bangui), grown by F Barre-Sinoussi and A Deslandres (Pasteur Inst., Paris), and cloned and sequenced by D Schmitt and MP Kieny.
3. **CM.CA10:** A single E subtype sequence from a set of 17 sequences from a very diverse set of isolates from from Yaounde and Douala, Cameroon. The sequences were derived from asymptomatic and symptomatic HIV infected individuals, specifically, CA10 was symptomatic. Virus was isolated by culture with donor PBMCs, and nested PCR amplified. A single clone was sequenced representing each HIV-1 isolate. Nkengasong IN, et al., *AIDS* 8:1405-12 (1994). GenBank accession numbers for the entire set of 17 sequences in this publication: X80438-X80454.
4. **TH1.ID#:** These twelve sequences are from a set of 23 individuals from Thailand. PCR-direct, peripheral blood PBMC DNA. Referred to as Thai subtype A in Ou et al. Ou C-Y et al., *AIDS Res. and Human Retroviruses* 8:1471-1472 (1992) and Ou C-Y et al. *Lancet* 341:1171-1174 (1993). (Published erratum appears in *Lancet* 342:250 (1993).) GenBank accession numbers L07443-L07445, L07447-L07448, L07457-L07459 and L07461-L07464.
5. **TH2.ID#:** These six sequences are from 16 isolates from HIV seropositive individuals from Thailand. PCR, PBMC co-culture, DNA. Full env sequence is available. McCutchan FE, et al., *AIDS Res. and Human Retroviruses* 8:1887-1895 (1992). Please note: the "TN-ID#" locus names in the database correspond to the McCutchan et al.'s "CM-ID#" isolates. GenBank accession numbers L03698-L03701 and L03703-L03704.
6. **TH3.W2TH-ID#:** Fifteen sequences from asymptomatic individuals from Thailand in 1992. Consensus, PCR-clones, cell-culture DNA and RNA. These sequences were provided by the WHO Global Programme on AIDS Vaccine Development Study. The complete set of C2V3 region WHO sequences can be found in the April supplement to the *Human Retroviruses* and *AIDS* 1993 database. Relevant papers are in press: de Wolf F, et al. *AIDS Res. and Human Retroviruses* in press (1994); Osmanov S, et al. *AIDS Res. and Human Retroviruses* in press (1994); Gao F, et al. *AIDS Res. and Human Retroviruses*, in press (1994). GenBank accession numbers U08810-U08811, U08825-U08836 and U08742-U08761.
7. **TH4.D-ID#:** These two sequences are are part of a set of sequences generated for the DAIDS variation program in the laboratory of Dr. Beatrice Hahn at the University of Alabama. They are clones from expanded culture stocks, and are excised from full gp160 sequences. The sequence ID numbers are abbreviated, for example D3TH966 can be read as DAIDS sequence (D), isolated in 1993 (2), Thai (TH), patient 301966 (966). GenBank accession numbers U08456 and U08457.
8. **TH.N764** This sequence is a E subtype sequence from Thailand (THP13). 12 of 13 sequences from Thai prisoners were of subtype B; N764 represents the only subtype E sequence identified in this set, from a prisoner infected in 1989. The sequences were obtained from PCR amplified PBMC DNA. Kalish ML, et al., *AIDS Res. and Human Retroviruses* Nov., in press (1994). GenBank accession number U15588.

The seventeen E subtype sequences not yet published, and for which the authors have not yet given permission for release, were provided by Dr. Xiao-Fang Yu of John Hopkins University, Dept. of Med. Microbiology, Baltimore, Maryland. The sequences from these individuals were consensus of clones derived from PCR amplified primary isolate DNA from Thai AIDS patients.

NOTE:

1. While the sequences in this subtype were distinct over this region of env from the other four env subtypes, in the gag gene it is not possible to make a distinction between this subtype and subtype A. What this means is that the isolates for which both gag and env are sequenced which cluster together as the "A" subtype in gag, are very distinctive in env and are broken down into two subtypes, env "A" and env "E". This holds true for the E subtypes sequences that originated in Thailand, as well as the E subtype isolate from the Central African Republic for which gag sequence was obtained. McCutchan FE, et al., AIDS Res. and Human Retroviruses 8:1887-1895 (1992); McCutchan FE, et al., AIDS 7:769-780 (1993); and Murphy E, et al., AIDS Res. and Human Retroviruses 9:997-1006 (1993).
2. The relative lack of diversity in the Thai sequences in this subtype relative to the other subtypes is likely to be a consequence of the short time span of the HIV-1 Subtype E epidemic in Thailand. McCutchan FE, et al., AIDS Res. and Human Retroviruses 8:1887-1895 (1992), and Ou C-Y et al., AIDS Res. and Human Retroviruses 8:1471-1472 (1992).

F Subtype

At this time there are viral sequences from 41 HIV-1 infected individuals associated with HIV-1 subtype F. The F subtype consensus sequence (F_CONSENSUS_94) generated from these sequences was based on the most common amino acid found in each position of an alignment. All of these sequences have been published.

1. **BR1.7944:** This sequence represents a single env F subtype sequence found among 22 Brazilian outpatients with varying degrees of disease progression. It was identified by Potts et al. as the single sequence which did not cluster with North American sequences in phylogenetic analysis. Consensus, PCR clones, peripheral blood PBMC DNA. Potts KE, et al., AIDS 7:1191-1197 (1993). GenBank accession number L19237.
2. **BR2.BZ-ID#:** Two sequences from Brazil of the F subtype. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994); Louwagie J, E Delwart, JJ Mullins, PE McCutchan, G Eddy, and DS Burke Unpublished (1993). GenBank accession numbers L22085 and L22082.
3. **BR3.RJI03:** An F subtype sequence from Rio de Janeiro, Brazil. 26 additional B and a B-F recombinant were also observed in this set. Year of isolation for was 1993, from an individual of CDC clinical stage II. Morgado et al., AIDS Res. and Human Retroviruses 10:569-576 (1994). DNA was amplified directly PBMCs of HIV infected individual, and the PCR product was directly sequenced. GenBank accession number U00422.
4. **CM.CA-ID#:** These sequences are 3 of 17 sequences from a very diverse set of isolates from Yaounde and Douala, Cameroon. The sequences were derived from asymptomatic and symptomatic HIV infected individuals; specifically patients CA16 and CA20 were asymptomatic and patient CA4 was symptomatic. Virus was isolated by culture with donor PBMCs, and nested PCR amplified. A single clone was sequenced representing each HIV-1 isolate. The F subtype designation of these sequences is tentative. Although the F subtype sequences from Cameroon and Brazil consistently form a clade in phylogenetic analyses, the branch lengths between isolates from the two countries are typical of inter-subtype distances, and sequences from the two countries each form their own distinct clade within the F subtype (HIV database and Wouter Janssens, personal communication). Nkengasong JN, et al. AIDS 8:1405-12 (1994). GenBank accession numbers for the entire set of 17 sequences studied in this publication: X80438-X80454.
5. **RO1.ID#:** These nine sequences are from isolates from Romanian children, in different clinical stages. All isolates showed cytopathic properties in peripheral blood mononuclear cells. They are also known as RM(A-J). Dumitrescu O, ML Kalish, SC Kliks, CI Bandea, JA Levy. J Infect Dis 169: 281-8 (1994). GenBank accession numbers L19571-L19579.
6. **RO2.RM-ID#:** These 24 sequences are from isolates from Romanian children. Unpublished: Holm-Hansen C, G Grothues, B Rosok, S Rustad, R Pascu and B Asjo, University of Bergen, Norway. GenBank accession numbers X77984-X77987.

G Subtype

At this time there are viral sequences from 8 HIV-1 infected individuals associated with HIV-1 subtype G. The G subtype consensus sequence (G_CONSENSUS_94) generated from these 8 sequences was based on the most common amino acid found in each position of an alignment. 7 of these sequences have been published and/or have been made available for printing in the database by their authors.

1. **CF.4067:** This sequence was associated with the C subtype in first analysis of the C2V3 region (Murphy E, et al., *AIDS Res. and Human Retroviruses* 9:997-1006 (1993), but when a full gp120 sequence became available from this isolate, and phylogenetic analysis was performed including some of the new subtype G sequences, it was more closely associated with G. The full length sequence was kindly provided prior to publication by Dr. MP Kieny of Transgene, Strasbourg Cedex, France. It is a part of a set of 14 HIV-1 gp120 isolates from Bangui, Central African Republic. The gp120 proteins have been expressed in a hybrid gp160 vaccinia virus expression system. Year of isolation and health status of individuals from which the virus was isolated were not provided. Viruses were isolated in the CAR by C Mathiot and B You (Pasteur Inst., Bangui), grown by F Barre-Sinoussi and A Deslandres (Pasteur Inst., Paris), and cloned and sequenced by D Schmitt and MP Kieny. GenBank accession numbers L11499 and L11500.
2. **GA.LBV21-7:** A sequence from Gabon from a set of HIV-1 viral isolates from Africa. This sequence was derived from a clone of PCR amplified DNA from cultured PBMCs. It represents a fragment of a full length env sequence. Janssens W, et al. *AIDS Research and Human Retroviruses* 10:877-878 (1994). GenBank accession number U09664.
3. **GA.VI525:** A sequence from Gabon from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al (in press, 1994), and Janssens W, et al. *AIDS Research and Human Retroviruses* 10:877-878 (1994). GenBank accession number L22953.
4. **NG1.ID#:** These four sequences represent G subtype sequences from Nigeria. Abimiku AG, et al., in press, *AIDS Res. Hum. Retroviruses* (1994). JP882 and JV832 were derived from AIDS patients, and G3 and G9 from healthy women. G9 was cultured on the T cell line CEM-SS, and the other three isolates were cocultured with uninfected donor PBMCs. DNA from viral cultures was PCR amplified, cloned and sequenced. GenBank accession numbers U13208-U13209, U13211 and U13213.

A G subtype sequence from Ghana was provided by Jaap Goudsmit's laboratory at the Dept. of Virology, Academic Medical Center, Amsterdam, The Netherlands. This sequence is not yet published (manuscript in preparation, 1994).

H Subtype

At this time there are viral sequences from 2 HIV-1 infected individuals associated with HIV-1 subtype H. The H subtype consensus sequence (H_CONSENSUS_94) generated from these 2 sequences was based on the most common amino acid found in each position of an alignment. Both of these sequences have been published and/or have been made available for printing in the database by their authors. Eight sequences that are too short for classification are closer to H than to other subtypes. The locus names (ID's) and sources of the sequences are:

1. CM.CA13: A sequence from Cameroon from a set of HIV-1 viral isolates from Africa used to define the prototype G and H env sequences. This sequence was derived from a clone of PCR amplified DNA from cultured PBMCs. It represents a fragment of a 900 base pair sequence. Janssens W, et al. AIDS Research and Human Retroviruses 10:877-878 (1994) and Nkengasong JN, et al., AIDS 8:1405-12 (1994). The H subtype association is not always clearly apparent using some sets of background sequences for comparison, and neighbor joining trees (HIV database, Wouter Janssens, personal communication), although parsimony trees confirmed the original association documented in Janssens W, et al. AIDS Research and Human Retroviruses 10:877-878 (1994). GenBank accession number U09667.
2. ZR.VI557: A sequence from Zaire from a set of HIV-1 viral isolates from Africa used to define the prototype G and H env sequences. This sequence was derived from a clone of PCR amplified DNA from cultured PBMCs. It represents a fragment of a 900 base pair sequence. Janssens W, et al. AIDS Research and Human Retroviruses 10:877-878 (1994). GenBank accession number U09666.

O Sequences

At this time there are viral sequences 3 HIV-1 infected individuals associated with HIV-1 subtype O that have been published and/or have been made available for printing in the database by their authors. The O subtype consensus sequence (O_CONSENSUS_94) generated from these 3 sequences was based on the most common amino acid found in each position of the alignment; when there was no consensus in a position the amino acid from ANT70 was used. These sequences represent a set of sequences that are extremely divergent relative to other HIV-1's. The subtypes A-H have been grouped together under the heading "M" for main. "O" sequences are as different from one another as are sequences from different "M" subtypes.

1. **CM.ANT70:** The complete viral genome has been sequenced from this viral isolate derived from a symptomatic Cameroonian, CDC stage III. R De Leys, et al., *J. Virol.* 64:1207-1216 (1990) and V Haesevelde *J. Virol.*, 68:1586-1596 (1994). GenBank accession number L20587.
2. **CM.MVP5180:** The complete viral genome has been sequenced from an isolate derived from a Cameroonian woman, sampled in 1991; the donor died of AIDS in 1992. The viral isolate MVP-5180 was grown in several human T-cell lines and the monocytic U937 line. Gurtler L, et al. *J. Virol.*, 68:1581-5 (1994). GenBank accession number L20571.
3. **FR.VAU:** This sequence was derived from an isolate from a French woman who died of AIDS in 1992. DNA was extracted from VAU infected PBMCs, PCR amplified, cloned, and gp160 env was sequenced. The viral isolate was highly cytopathic. P Charneau et al., *Virology*, in press (1994). GenBank accession number X80020.

Uncertain Classification

At this time there are viral sequences from 13 HIV-1 infected individuals that are not clearly associated with any of the HIV-1 genetic subtypes A through H. They either appeared distinct from the subtypes A-H in phylogenetic analysis, or else the subtype association was unclear, with different associations in different analyses. For some of the shorter gene fragments, subtype associations might have been established if more sequence information was available or if a different set of sequences was included in the background set used to define subtype associations. Some of these sequences may be representatives of subtypes as divergent as A-H, but only a single limited sample is yet available. Still others may represent recombinant genomes.

1. **BR3.RJI01** This sequence is B-F recombinant in the V3 region. DNA was amplified directly from PBMCs of HIV infected individual, and the PCR product was directly sequenced. Sabino EC et al. *J. Virol.* 68:6340-6346 (1994). GenBank accession number U00420.
2. **CF.ID#:** These four sequences are from a set of sequences obtained from 27 symptomatic patients from the Central African Republic, from whom blood was drawn in 1990-1991. Consensus, PCR-clones, cell culture, DNA. Murphy E, et al., *AIDS Res. and Human Retroviruses* 9:997-1006 (1993). GenBank accession numbers L11482-L11483, L11497, L11508-L11510 and L11514-L11515.
3. **KE.K124:** A Kenyan sequence from a set of HIV-1 viral isolates from Africa. Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). This isolate is not clearly associated with D subtype, however Louwagie and colleagues found that it associated with the D subtype in env, and the A subtype in gag. Using parsimony analysis, we found that it was difficult to determine a clear association, and this observation was confirmed by Wouter Janssens (personal communication). GenBank accession number L22942.
4. **KE.KEN976** This is a single unclassified sequence from a set of patients who were part of a 1990-1992 cohort study of maternal risk factors in mother to child transmission, including 22 pregnant women and an infant from Kenya. The C2V3 region was sequenced. W Janssens et al., in press *AIDS Res. Hum. Retroviruses* (1994). GenBank accession number U12992.
5. **NL.A11:** This sequence is from a Dutch study of presumed HIV-1 donor-recipient pairs. This sequence is from a recipient at the time of seroconversion; the donor was a Zairean woman living in the Netherlands. The sequences from both donor and recipient were extremely similar, so only the recipient is shown here. This sequence is a consensus sequences of multiple clones from PCR amplified serum RNA. Wolfs TFW, G Zwart, M Bakker, and J Goudmsit. *Virology* 189:103-110 (1992). GenBank accession numbers M91849-M91856.
6. **ZM.ZAM184:** This Zambian sequence is an outlier, though in some phylogenetic analysis it appears most closely associated with the A subtype. In particular it is closely associated with A.CF.SAS (100/100 replicates in parsimony analysis of gp120). Health status of the individual from which the virus was cultured was unspecified, and the year of viral isolation was probably between 1989-1992. Viruses were cultured with donor PBMCs for 2 to 3 weeks. Full length env (gp160) was amplified, cloned and sequenced. Louwagie J, et al. in press (1994). GenBank accession number L22955.
7. **ZR.Z3:** This sequence is from the 1983 Zairean isolate Z-3 (non-infectious, possibly due to frame-shift). Willey RW, et al., *Proc. Natl. Acad. Sci. USA* 83:5038-5042 (1986). GenBank accession number K03347.

The three sequences which are not associated with any subtype and are not yet published, and for which the authors have not yet given permission for release, are from the following sets:

1. **ZR:** Two sequences from individuals from Zaire, provided by the LTCB, NCI, NIH, Bethesda, MD, USA (Dr. Marvin Reitz).
2. **GA:** A sequence provided by Dr. Wouter Janssens of the Prince Leopold Institute of Tropical Medicine, Antwerp, Belgium. It is a Gabonese isolate that is phylogenetically linked to F subtype sequences, but the relationship is distant.

Phenogram of the 14 amino acid tip of the V3 loop.

The immunogenic tip of the V3 loop has been the focus of a large number of serological studies. Here we have organized the 14 amino acid tip of the V3 sequences according to their amino acid similarities. The four amino acid tip of the loop (the GPGR motif) plus the 5 amino acids on either side were clustered using maximum linkage based on pairwise distance scores using the STR matrix (S. Henikoff and J. G. Henikoff, *Proteins: Structure, Function and Genetics* 17: 49-61 (1993)). The resulting phenogram (B. T. M. Korber et al., *J. Virol.* 68:6730-6744 (1994)) of all 660 non-embargoed sequences (660/707) is shown, with related sequences grouped into 20 clusters. The y-axis represents the number of sequences included in the phenogram, and the x-axis indicates the level of similarity between amino acids; the higher the score, the more similar the sequences. The twenty clusters were grouped so that the sequences within a cluster all are related by scores greater than 35. The alignment of these sequences corresponding to the phenogram is included, with a consensus from each cluster heading each group. The order of the sequences in the alignment corresponds to the order that the sequences appear in the phenogram, from top to bottom. There are two things worth noting in the alignment. First is the intermixing of amino acid sequences from different clades—often sequences from phylogenetically distinct subtypes have identical or nearly identical sequences in this region. Second is the diversity of the D subtype sequences. The two letter country codes can be used to follow geographic associations of similar sequences, or alternatively, the range of variation of a given subtype in a country of interest.

V3 Phenogram

HIV-1 V3 Region

ALL_CONSENSUS	KSIHIGPGRAFYT	ALL_CONSENSUS	KSIHIGPGRAFYT
CONSENSUS.1	---R---QT---	C_MW1.12215	---R---Q-----
C_SO.1574	--VR---QT---	C_MW1.12203	---R---Q-----
U_CF1.4081	--VR---QT---	C_MW1.12199	---R---Q-----
C_MW1.6508	--VR---QT---	C_BR.W2BR025	---R---Q-----
C_MW1.12229	--VR---QT---	A_ZR1.6653	---R---Q-----
C_MW1.12225	--VR---QT---	A_CF1.4055	---R---Q-----
A_ZR1.6657	--VR---QT---	A_CF1.4054	---R---Q-----
A_ZR1.6563	--VR---QT---	A_CF1.4033	---R---Q-----
A_UG2.117	--VR---QT---	A_RW.W2RW021	--VR---Q-----
A_DJ.DJ264	--VR---QT---	A_UG1.W2UG031	--VR---Q-----
A_DJ.DJ263	--VR---QT---	A_RW.SF1703	--VR---Q-----
A_DJ.DJ258	--VR---QT---	A_CONSENSUS_94	--VR---Q-----
A_CM1.CA2	--VR---QT---	A_CI2.CI330	--VR---Q-----
A_CM1.CA19	--VR---QT---	A_CI2.CI3	--VR---Q-----
A_CI2.CI45	--VR---QT---	A_KE.KEN967	R--R---Q-----
A_CF1.4081	--VR---QT---	A_RW.W2RW026	R--R---Q-----
A_UG1.W2UG037	R-VR---QT---	A_CF1.4050	R-VR---Q-----
A_UG2.78	R-VR---QT---	A_CM1.CA18	R-VR---Q-----
A_CM1.CA15	R-VR---QT---	A_ZR1.6557.sh	Q-VR---Q-----
A_CM1.CA17	--VR---QT---A	C_MW1.6512	Q-VR---Q-----
C_IN2.D808	--TR---QT---	A_KE.K89	M--R---Q-----
C_MW2.D3MA960	---R---QT---	A_ZR1.6655.sh	---R---QV----
C_ZW.2647	---R---QT---	F_CM.CA4	---R---QV----
C_MW1.6518	---R---QT---	C_MW1.12205	Q--R---QV----
C_MW1.6510.sh	---R---QT---	C_MW1.12209	Q--R---Q--F--
C_MW1.12213	---R---QT---	U_KE.K124	---R---Q--F--
C_IN2.D868	---R---QT---	B_BR3.RJ626	---R---S-----
C_IN2.D744	---R---QT---	A_CM1.CA22	R-VR---Q-I--
C_IN2.D1024	---R---QT---	A_CI2.CI21	--VN---Q-----
C_IN1.D760	---R---QT---	A_KE.KEN982	--AN---Q-----
C_IN1.D757	---R---QT---	A_KE.KEN965	T--R---Q---GR
C_IN1.D1044	---R---QT---	E_CM.CA10	T-SR---Q---KI
C_CONSENSUS_94	---R---QT---		
A_CI2.CI20	---R---QT---	CONSENSUS.2	T--T---QV--R-
C_ZM1.ZAM20	---R---QT-F--	E_TH3.W2TH019	T--T---QV--R-
C_DJ1.DJ259	E--R---QT---	E_TH3.W2TH022	T--T---QV--R-
C_UG.UG268	E--R---QT---	E_TH3.W2TH011	T--T---QV--R-
C_DJ1.DJ373	Q--R---QT---	E_TH3.W2TH009	T--T---QV--R-
A_KE.KEN985	---R---QT---G	E_TH3.W2TH006	T--T---QV--R-
C_SO.SM145	--VR---QT---TN	E_TH2.TN244	T--T---QV--R-
A_UG2.84	---G---QT---A	E_TH2.TN242	T--T---QV--R-
A_CI2.CI47	VRTG---QT---	E_TH2.TN241	T--T---QV--R-
A_UG4.UG273	T-VR---Q-----	E_TH2.TN239	T--T---QV--R-
A_ZR1.6663	T-VR---Q-----	E_TH1.TA8176	T--T---QV--R-
A_UG2.82	T-VR---Q-----	E_TH1.8673	T--T---QV--R-
A_CM1.CA21	T-VR---Q-----	E_TH1.8671	T--T---QV--R-
A_KE.KEN977	T--R---Q-----	E_TH1.8663	T--T---QV--R-
A_KE.KEN981	T--R---Q-----	E_TH1.8659	T--T---QV--R-
A_RW.W2RW024	T-VR---QT---	E_TH.N764	T--T---QV--R-
A_RW.W2RW025	T-VR---QT---	E_CONSENSUS_94	T--T---QV--R-
A_KE.KEN88	T--R---QT---	E_TH1.8657	T--T---QI--R-
A_UG4.UG275	T-VR---QS----	E_TH3.W2TH018	T--T---QI--R-
A_RW.W2RW008	T--R---QS-H--	E_TH3.W2TH020	T-LT---QV--R-
A_KE.KEN979	T--C---QT---G	E_TH3.W2TH021	T-LT---QV--R-
G_NG1.G3	---R---Q-----	E_TH1.8639	TR-T---QV--R-
U_CONSENSUS_94	---R---Q-----	E_TH3.W2TH015	TG-T---QV--R-
C_ZM1.ZAM18	---R---Q-----	E_TH3.W2TH007	T--N---QV--R-
C_MW1.1227	---R---Q-----	E_TH3.W2TH023	T--N---QV--R-

ALL_CONSENSUS	KSINIGPGRAFYAT	ALL_CONSENSUS	KSINIGPGRAFYAT
E_TH3.W2TH001	T--N---QV--R-	C_CF1.4067.CON	---GF---Q-----
E_TH1.TA10671	T--TM---QV--R-	C_CF1.4068.CON	---GF---Q-----
E_TH3.W2TH003	T--TV---QV--R-	U_CF.4067	---SF---Q-----
E_TH4.D3TH966	T--TT---QV--R-	G_NG1.JP882	---TFA---Q-----
E_TH1.TA7794	T--T---QV--K-	F_RO2.RM53024	---L---Q---T-
E_TH3.W2TH024	T--T---QVL-K-	F_RO2.RM5308	---L---Q---T-
E_TH1.TA8173	T--P---QV--R-	F_RO2.RM53023	---L---Q---T-
E_TH3.W2TH005	T--P---QV--R-	F_RO1.14034	---L---Q---T-
E_TH1.8683	T--P---QV--R-	F_RO2.RM53014	---XL---Q---T-
E_TH2.TN235	T--P---Q---R-	A_KE.KEN973	---R---Q---T-
E_TH1.TA7792	P--T---QV--R-	U_KE.KEN976	---RF---Q---TS
E_TH2.TN243	P--TX---QV--R-	U_ZM.ZAM184	R--RF---Q---TN
CONSENSUS.3	---R---QTS---	B_TH2.TB7787	---L---W-T-
A_CM1.CA7	R--R--S-QTS---	B_TH3.W2TH014	---L---W-T-
A_T2.TAM15COM	--V---QTS---	B_TH1.8645	---L---Q-W-T-
C_ZA.NOF	-R-RV---QTV---	B_TH1.8669	---L---Q-W-T-
CONSENSUS.4	---P-----	B_TH1.8651	-G-L---Q-W-T-
B_US1.HC44	---P-----	B_TH4.N762	-G-L---Q-W-T-
B_US2.D2US725	---P-----	B_TH1.8641	---L---Q-W---
B_US1.HC24	---P-----	B_TH1.8643	---L-Q-Q-W---
B_NL4.H186_85	---P-----	B_TH1.8653	---PL---Q-W-T-
B_NL2.Wolfsh1	---P-----	B_TH3.W2TH026	---PL---Q-W-T-
B_HT2.H6012	---P-----	B_TH1.8649	---PL---Q-W-T-
B_HT2.H6010	---P-----	B_TH4.N763	-G-PL---Q-W-T-
B_HT1.D1HA652	---P-----	B_TH1.8655	---PL---QWWT-
B_NL4.I1025_89	R--P-----	CONSENSUS.6	-G-----Q-----
A_RW.W2RW016	--VP-----	B_NL4.H23_82	-G-----K---T-
B_UK.V91	---P-----I-	B_US.WM	-G-----K---T-
B_US.MA	---PM-----	B_NL4.I70_88	-G-----K-----
F_BR3.RJ103	---PL-----	A_NG.NI	-GV-----Q-----
B_TH1.8647	---PL---W---	A_UG2.118	-GV-----Q-----
A_CI2.CI14	--VP---Q-----	A_KE.KEN983	-G-----QS---
A_KE.KEN978	--VP---Q-----	B_BR1.8625	-----S-----
G_NG1.JV832	---P---Q-----	B_BR2.W2BR018	-----S-----
B_US1.HC09	Q--P---K-----	B_NL4.I90_89	-G-----S-----
B_BR3.RJ636	---P-A--S-WF--	A_CM1.CA1	-G-----S-I---
B_TH4.N758	---P-----T-	B_HT2.H5996	-G-----G-----
B_US4.ZhuPEA	---P-----T-	B_HT2.H6022	-G-----G---T-
B_NL4.I7025_91	---P-----T-	B_NL4.H434_85	-----G-----
B_NL4.I4008_86	---P-----T-	B_BR1.10561	-G-M---ST---
B_NL4.I1057_87	---P-----T-	B_NL1.A2	-G-L---GT---
B_NL4.H495_85	---P-----T-	A_CM1.CA11	TG-----QT---
B_NL4.H188_90	---P-----T-	B_BR1.10581	-G-----T-F--
B_NL4.H164_90	---P-----T-	B_US1.HC19	-G-P-----T-
B_NL4.H1091_90	---P-----T-	B_US1.HC40	-G-P-----T-
B_NL2.Wolfsh495	---P-----T-	B_TH4.N752	-G-P-----T-
B_US4.ZhuPcR	R--P-----T-	B_TH4.N753	-G-P-----T-
B_HT2.H6018	---P-----FT-	B_HT2.H6008	-G-T-----T-
B_NL1.A7	R--P---K--FT-	B_BR2.W2BR004	-G-P---GS---
B_NL4.H1112_90	R--P--L-S--T-	A_RW.W2RW020	-GVR---Q-----
CONSENSUS.5	---L---Q---T-	C_IN1.D747	-GVR---QT---
G_CONSENSUS_94	---T---Q-----	F_GA.VI354	-G-R-----VI---
H_ZR.VI557	---S---Q-----	B_NL4.H187_90	-G-R-----I-A
C_CF1.15166	---SL---Q-----	A_UG2.111	T-V-----Q---R
G_NG1.G9	---TL---Q-----	A_UG2.115	-----Q---R
		A_KE.KEN29	TG-----Q---R
		A_CI2.CI327	-GV-----Q---R

HIV-1 V3 Region

ALL_CONSENSUS	KSINIGPGRAFYAT	ALL_CONSENSUS	KSINIGPGRAFYAT
A_GH.D687	-NV-----Q---R	B_US.SF33	RR-TS---KVL-T-
A_ZR1.6559	QGV-----QV---R	B_HT.RF	---TK---VI---
A_CF1.1189	RIM-----P---R	G_GA.VI525	---KF-T---VL---
CONSENSUS.7	---M---K-----	CONSENSUS.10	-----
B_US1.HC08	---PM-----T-	B_US2.D2US664	-----
B_US2.D2US665	---PM-----T-	F_CM.CA20	-----
B_NL4.I4004_88	---PM-----T-	B_US2.D2US656	-----
B_HT2.H5986	---PM---K---T-	B_US1.HC39	-----
B_US1.HC05	R---PM---K---T-	B_US1.HC34	-----
B_HT2.H6004	---PM---K-----	B_US1.HC26	-----
B_US2.D2US660	---PM---K-M---	B_SE1.931	-----
B_US1.HC07	---TM---KV---G	B_PR1.D2PR733	-----
B_US4.ZhuPteV	---TM---KV---V-	B_PR1.D2PR728	-----
B_HT2.H5994	---M---K---T-	B_NL4.I3064_90	-----
B_NL4.H135_90	---M---KT---T-	B_NL4.I3008_88	-----
B_BR1.10591	-G-M---G---T-	B_NL4.I1020_86	-----
B_US1.HC27	-G-M-----T-	B_NL4.H594_85	-----
B_CI.CI22	---N---S-I---	B_NL4.H552_90	-----
B_HT2.H6016	---M---S-----	B_NL4.H17_82	-----
A_CF1.11699	--V-M-----	B_NL4.H169_85	-----
B_TH4.N761	---N-Q-----	B_NL4.BN18_84	-----
B_BE.SIMI84	-TL-M---K-----	B_NL4.BN130_85	-----
A_CM1.CA6	R---F---QTL---	B_HT2.H6024	-----
A_ZR1.6569	RG---M---QIL---	B_BR2.W2BR020	-----
A_CF1.4018	T-V-M---KT---	B_BR2.W2BR017	-----
CONSENSUS.8	T-VR---QV---K-	A_U9.1033	---L-----
E_CF1.4039	T-VR---QV---K-	B_US3.D2US711	---L-----
E_CF1.4084	T-VR---QV---K-	B_US.SC	R-----
E_CF1.4031	T-VR---QV---K-	B_US1.HC29	R-----
E_CF1.4017	T-VR---QV---K-	B_HT2.H6020	R-----
E_TH4.D3TH975	T-VR---QV---R-	B_US2.D2US714	R---M-----
E_CF2.NBA	T-VR---QV---G-	U_CF1.4056	T---L-----
E_CF1.1697	T-AR---QV---K-	B_US1.HC20	---A-----
E_CF2.ELO	T-AR---QV---K-	B_US1.HC41	---A-----
E_CF1.4013	T-VR---V-HK-	B_US1.HC16	---A-----
E_CF2.4071TG2	T-VR---V---K-	B_SE1.930	---A-----
E_CF1.4002	I-AR---V-HT-	B_NL4.I5032_B7	---A-----
E_CF1.4069	I-AR---V-HK-	B_US1.HC30	R---A-----
C_MW1.6506	--NR---QT---	A_CF1.4044	T---A-----
C_SN.SE364	--NR---QT---	B_US1.HC46	-G-----
A_CF1.4023	R-NR---QT---	F_BR2.B2163	-G-----
A_CI2.CI329	--LR---QT---	B_US1.HC42	-G-----
C_IN2.D766	--LR---QT---	B_US1.HC18	-G-----
C_MW1.12233	--MR---QP---	B_US1.HC01	-G-----
U_CF1.4087	R-NR---T---TI	B_SE1.1032	-G-----
A_ZR1.6571.sh	---R---QV---TN	B_NL4.I3091_91	-G-----
C_MW2.D3MA959	R---R---QV---N	B_NL4.I142_87	-G-----
U_ZR.Z3	Q---R---KV---K	B_NL4.H450_85	-G-----
CONSENSUS.9	-R-TM---VL-T-	B_NL4.H228_90	-G-----
B_NL1.A13	-RVTM---VW-T-	B_NL3.NET4	-G-----
B_US.CDC42	-RVTL---VW-T-	B_BR2.W2BR030	-G-----
B_US.BRVA	-R-TM---VY-T-	B_BR1.8633	-G-----
B_US1.HC13	-R-TM---VL-T-	A_GA.VI191	-G-----
B_US.BCSG3	-R-TT---VY-T-	B_UK.V74	RG-----
B_NL4.BN132_84	-RMTL---KV-T-	B_NL4.I117_90	IG-----
		B_NL4.I3077_91	-X-----
		B_US.RJ5	RH-----

ALL_CONSENSUS	KSIHIGPGRAYAT	ALL_CONSENSUS	KSIHIGPGRAYAT
B_IN1.20023	RR-----F--	B_HT2.H5990	---N-----
B_IN1.30008	R-----F--	B_BR3.RJ101	---N-----
F_RO2.RM5306	---L---Q----	B_HT1.D1HA651	---N-----W--
F_RO2.RM5307	---L---Q----	B_BR2.W2BR026	---T-----
F_RO2.RM53035	---L---Q----	B_US.SF162	---T-----
F_RO2.RM53034	---L---Q----	B_BR1.10593	---S-----
F_RO2.RM53032	---L---Q----	B_NL3.NET3	---S-----
F_RO2.RM53031	---L---Q----	B_NL4.I5037_89	---X-----
F_RO2.RM5303	---L---Q----	A_2R.Z321	---S-----F--
F_RO2.RM53029	---L---Q----	B_SE1.2951	R--N-----
F_RO2.RM53027	---L---Q----	B_US1.HC28	R--T-----
F_RO2.RM53022	---L---Q----	B_BR2.W2BR014	R--I-----
F_RO2.RM53018	---L---Q----	B_NL3.NET2	R--G-----I--
F_RO2.RM53013	---L---Q----	B_US.SF2	---Y-----HT-
F_RO2.RM53011	---L---Q----	B_US2.D2US704	-----HT-
F_RO2.RM53002	---L---Q----	F_BR2.BZ126	---YF-----HTA
F_RO1.14046	---L---Q----	B_GA.OYI	NR-S-----HT-
F_RO1.14041	---L---Q----	B_US.ALA1	GR-----HT-
F_RO1.14036	---L---Q----	B_US1.HC37	---LA---G-H--
F_RO1.14028	---L---Q----	B_US2.D2US726	R---A-----H--
F_RO1.14027	---L---Q----	B_UK.V77	-----V-H--
F_RO1.14024	---L---Q----	B_US1.HC25	-----K--D--
F_RO1.14020	---L---Q----	B_JP.GUNA	---T--S---H-I
F_RO1.14018	---L---Q----		
F_CONSENSUS_94	---L---Q----	CONSENSUS.11	---M-W-----
F_RO2.RM53040	-----Q-----	B_BR2.W2BR021	---M-W-----
F_RO2.RM53043	-----Q-----	B_BR3.SPB1	---M-W-----
F_RO2.RM53037	-----Q-----	B_BR2.W2BR003	---M-W-----
F_RO2.RM53015	-----Q-----	B_BR1.8635	---M-W-----
A_KE.KEN31	-----Q-----	B_BR1.7936	---M-W-----
A_KE.KEN112	-----Q-----	B_BR1.7932	---M-W-----
A_CF1.4010	-----Q-----	B_BR2.W2BR019	R--M-W-----
U_NL.A11	R--V---Q----	B_BR3.RJ59	---L-W-----
A_KE.KEN980	X-----Q-----	B_BR3.RJ62	---L-W-----
B_BR1.8623	T-----Q-----	B_BR2.W2BR024	---L-W-----
A_KE.KEN968	R-----Q---S	B_BR1.8621	---L-W-----
A_KE.KEN984	-----Q---S	B_BR1.8629	---M-W---T---
A_UG2.124	--V---Q-----	B_BR2.W2BR028	---M-W---T---
A_2R1.6649	--V---Q-----	B_BR1.7930	-G-M-W---T---
A_UG2.116	--V---Q-----	B_BR3.RJ12	-----W---I---
A_TZ.TAN81	--V---Q-----	B_BR3.RJ484	-G---W-----
A_TZ.TAN142	--V---Q-----	B_BR3.RJ478	---M-W---HTM
A_RW.W2RW009	--V---Q-----	B_BR3.RJ49	---M-W---T---
A_KE.KEN98	--V---Q-----	B_BR3.SPB2	---M-W---T---
A_KE.KEN970	--V---Q-----	B_BR3.RJ64	---L-W---T---
A_C12.CI42	--V---Q-----	B_BR3.RJ14	---V-W---L-T-
A_TZ.TAN9CON	R-V---Q-----	B_BR1.7934	---V-W---SLEP-
A_UG2.72	R-V---Q-----		
Q_GA.LBV217	-----Q-L---	CONSENSUS.12	---N-----T-
B_NL1.A2	-----K-----	B_NL4.I5038_89	---N-----T-
B_NL4.H1_85	-----K-----	B_US2.D2US706	---N-----T-
B_BR2.W2BR023	-----K-----	B_NL4.I39_91	---N-----T-
B_US1.HC32	R-----K-----	B_NL4.I1035_87	---N-----T-
A_UG2.119	Q-V---K-----	B_NL4.H29_82	---N-----T-
B_US2.D2US705	---N-----	B_NL1.A3	---N-----T-
B_US5.pt6b1	---N-----	B_BR1.7942	---N-----T-
B_NL4.H748_90	---N-----	B_NL4.I3072_90	---N-----FT-
B_HT2.H5998	---N-----	B_US.YU	---N-----L-T-

HIV-1 V3 Region

ALL_CONSENSUS	KSIHIGPGRAFYAT	ALL_CONSENSUS	KSIHIGPGRAFYAT
B_US1.HC11	R--NM-----T-	CONSENSUS.15	-R-R-??-----
B_US1.HC21	R--NM-----T-	B_DE.D31	-R-R--AR---TK
B_US.SBA	R--N-----T-	F_CM.CA16	RR-R--L--V----
B_US2.D2US717	R--N-----T-	B_BR3.RJ19	---QY-T-GDAFYA
B_NL4.I138_89	---NT-----T-	CONSENSUS.16	-----T-
B_SE1.2815	---T-----T-	B_US3.D2US715	-----T-
B_US1.HC22	---T-----T-	B_US3.D2US716	-----T-
B_NL4.I221_90	---T-----T-	B_US3.D2US712	-----T-
B_BR1.8615	R--T-----T-	B_US2.D2US662	-----T-
B_US.SBC	---S-----T-	B_US1.HC45	-----T-
B_US.JPL	---TL-----T-	B_US1.HC35	-----T-
B_BR1.10553	---Q-----T-	B_US1.HC33	-----T-
B_NL4.B63_84	---Q-----T-	B_US.JRCSE	-----T-
P_BR.7944	---QL-----T-	B_US.BAL	-----T-
B_NL4.I141_91	-G-Q-----T-	B_US.ADA	-----T-
B_BR1.10575	--LSL-----T-	B_SE1.1433	-----T-
B_HT1.D2HA594	-R-S-----VM-T-	B_PRI.D2PR732	-----T-
B_HT1.D2HA596	-R-S-----VM-T-	B_PRI.D2PR729	-----T-
B_HT2.H13962	-R-S-----W-T-	B_NL4.K999_86	-----T-
B_HT2.H13960	-R-S-----S-T-	B_NL4.I3048_89	-----T-
B_BR3.SPB4	-RKS-----T-	B_NL4.I3007_86	-----T-
B_BR1.7940	-R-Y-----V-T-	B_NL4.I257_91	-----T-
B_UK.V82	-R-Y-----V-T-	B_NL4.I1164_90	-----T-
B_US5.pt5b1	-R-A-----V-T-	B_NL4.H724_81	-----T-
B_TH2.TB132	-X-XM---XX-T-	B_NL4.H557_82	-----T-
B_UK.CAM1	---A-----TV---	B_NL4.H466_82	-----T-
B_US1.HC38	-G-A-----TV---	B_NL4.H457_85	-----T-
B_US.NY5CG	-G-A-----TL--R	B_NL4.H36_B2	-----T-
B_US.ACP1	-G-G-----TV-TA	B_NL4.H35_80	-----T-
CONSENSUS.13	RR-S-----V--	B_NL4.H339_82	-----T-
B_HT1.D2HA593	-R-S-----R--	B_NL4.H293_85	-----T-
B_HT2.H13958	-R-S-----R--	B_NL4.H153_82	-----T-
B_US.SBB	-R-S-----V-A	B_NL4.BM93_85	-----T-
B_US1.HC06	-RLS-----V--	B_NL4.BM90_85	-----T-
B_US.P896	RRLS-----R	B_NL4.BM152_84	-----T-
B_US1.HC31	RR-S-----T-R	B_NL4.BM142_84	-----T-
B_UK.V87	ER-S-----I-R	B_NL4.BA206_84	-----T-
A_CF2.SAS	RRM-----I--	B_NL4.BA1_84	-----T-
B_IN1.20016	RR-K-----V--	B_NL3.NET1	-----T-
B_HT.WM322	R-LS-----XR-	B_NL1.A5	-----T-
B_HT2.H5992	R--T-----D-Y	B_HT2.H6014	-----T-
B_US2.D2US724	R--SN-----RTG	B_HT2.H6006	-----T-
B_FR.LAI	IR-QR-----VTI	B_HT2.H6002	-----T-
B_MFA	IR-QR-----VTI	B_CONSENSUS.94	-----T-
B_IN1.30005	SRLQV-----V-A	B_BR3.RJ54	-----T-
CONSENSUS.14	-----L-----T-	B_BR3.RJ379	-----T-
B_BR1.7946	----L-L-----T-	B_US2.D2US727	--V-----T-
B_NL1.A9	-----L-----T-	D_UG3.109	Q-V-----T-
B_US4.2hupcF	R--M-L-----T-	B_NL4.I1083_87	R-----T-
B_NL1.A4	AG---L--R--TR	B_US1.HC43	R-----T-
A_CF1.4058	R-----S-----II	B_NL3.NET6	R-----T-
B_UK.V12	R-----S-----TI	B_US1.HC12	R-----W-T-
B_BR1.10565	-G--L-F--L---	A_CT2.CI326	R-W-----T-
B_BR3.RJ483	-G--M-L-----	B_JP.JH32	RR-----T-
B_BR3.RJ102	----Y-V-----I--	B_US.MN	-R-----T-
P_RO2.RM53012	GIHI--V--P--R-		

ALL_CONSENSUS	KSIHIGPGRAFYAT	ALL_CONSENSUS	KSIHIGPGRAFYAT
A_UG.UG06	RR-----S--TS	D_KE1.KEN966	Q-TR----QTLFT-
B_HT1.D2HA590	-----TG	D_UG.044342	Q-TQ----Q-LFTK
B_HT2.H13954	-----TG		
B_US.JM	-----A-----T-	B_HT1.D2HA599	H-G---G--TLFT-
B_NL1.A10	-----S---T-	B_HT2.H13968	H-G---G--TLFT-
B_US1.HC36	-----S---T-		
B_HT2.H5988	-----T-FT-	CONSENSUS.17	QRT---S-Q-?-T-
B_HT2.H6000	-----S---T-	D_CONSENSUS_94	QRT-----Q-L-T-
B_BR3.RJ623	R---L---G---T-	D_UG2.G1	QRT---S-Q-L-T-
B_US1.HC10	R-V-----S-L-T-	U_CF1.4040	QRT---S-Q-IFT-
B_US2.D2US657	-Q-----T-	D_UG2.2999	QRT---T-Q-LHT-
B_US4.ZhuPcL	-G-----T-	D_UG1.W2UG059	QRT---T-Q-Y-TK
B_TH4.N750	-G-----T-	D_UG1.W2UG070	RRT---S-Q-YSTK
B_TH4.N759	-G-----T-	D_UG2.4133	IRT---S-Q-Y-R-
B_TH4.N757	-G-----T-	A_CF2.GAN	-RW---S-Q----I
B_TH4.N756	-G-----T-	A_UG1.W2UG029	RR-----R-Q----
B_TH4.N755	-G-----T-	A_UG.0455	RRYS--S-Q---V-
B_TH4.N754	-G-----T-		
B_SE1.1866	-G-----T-	CONSENSUS.18	QRTF--L-Q-L-T-
B_NL4.I5020_89	-G-----T-	D_UG2.1665	QRTF--L-Q-L-T-
B_NL4.I3084_90	-G-----T-	D_UG2.31	QRTF--L-Q-L-T-
B_NL4.BN29_85	-G-----T-	D_UG1.W2UG024	QRTF--L-Q-L-T-
B_NL4.BA65_85	-G-----T-	D_ZR.ELT	QRTF--L-QSL-T-
B_CM.CA5	-G-----T-	D_UG2.5055	Q-TF--L-Q-L-T-
B_NL4.I33_90	RG-----T-	D_ZR.JY1	Q-TF--L-Q-L-T-
B_UG1.HC23	RG-----T-	D_CF.402019	QQTF--L-Q-L-T-
B_NL4.H43_90	RG-----T-	D_UG1.W2UG046	RRTF--L-Q-Y-T-
D_UG1.W2UG005	-G-----Y-T-	D_UG3.83	RRTF--L-Q-L-T-
B_NL3.NET5	RG-----V-T-	D_UG3.70	QRTF--R-Q-LFT-
B_NL4.I3057_91	RG-----V-T-	D_UG5.UG266	QRTF--R-Q-LFT-
B_DE.HAM	-G-----V-T-	D_UG1.W2UG038	QRTF--Q-QVL-T-
B_US.BWB11A	-G-MV----L-T-	D_UG2.C2	QRTF--Q-Q-L-T-
D_ZR.MAL	RG--F---Q-L-T-	D_T21.TAN5	QRTF--S-Q-L-T-
D_ZR1.6651.sh	-G-----Q-L-TS	D_UG5.UG274	QRTF--T-Q-L-T-
D_ZR1.6665.sh	-----Q-I-T-	D_ZR1.6659.sh	QRTF--Q---T-
F_RO2.RM53021	-----Q-X-T-	D_SN.SE365	QRTF--L-QVLAT-
B_NL4.H490_85	----L---K-V-T-	D_UG1.W2UG021	YRTF--V---S-T-
D_UG3.120	-----Q-LFTI	D_T21.TAN13	QRTF--L-Q-LVTS
D_T21.TAN3	QGT-----Y-T-	D_UG3.UG23	HRTF--L-Q-LITN
D_T21.TAN6	QGT-----Y-T-	D_UG2.5059	RRTF--L-Q-LITS
D_T21.TAN2	QGT-----Y-T-	D_UG1.W2UG053	RQTF--L-Q-VLT-
D_KE1.KEN971	QGT-----Y-T-		
D_UG3.110	QGT-----YWT-	CONSENSUS.19	QRTS--Q-Q-L-T-
D_T21.TAN7	Q-T-F---Y-T-	D_UG1.W2UG001	QRTS--Q-Q-L-T-
D_T21.TAN1	QGA-----Y---	D_UG2.4132	QRTS--Q---L-T-
D_UG3.114	QG-----Y-TD	D_T21.TAN4	QKTS--R-Q-L-T-
D_CI.C113	QGT-----Q-L-T-	D_UG2.1685	QRTS--S-Q-L-T-
D_KE1.KEN986	QGT-----QTLFT-	D_UG2.462	RRTS--Q-Q-Y-T-
D_UG1.W2UG035	SGT-----LFT-	D_US.ANK	LKTS--Q-Q-L-T-
D_UG3.74	QGT-M-----L-TI	D_UG1.W2UG065	QRTS--L-Q-L-TN
D_UG3.121	Q-T-----Q-L-T-	D_ZR.Z226	QRTS--L-Q-L-T-
D_UG3.79	Q-T-----Q-L-T-	D_ZR.NDK	QRTS--LRQSL-TI
D_ZR1.6555	Q-T-L---Q-L-T-		
D_UG3.122	Q-V-L---Q-L-T-	H_CM.CA13	R-MR--I--GGTFH
D_UG3.71	Q-----Q-L-T-	H_CONSENSUS_94	R-MR--I--GGTFH
D_UG1.W2UG040	Q-T-----Q-I-TL	D_UG2.653	SIRR---WQPYI
D_T21.TAN11	Q-TR---Q-L-TN	A_CF1.11423	-M-PT---QGQVIY
D_T21.TAN12	Q-TR---Q-LFTS		

HIV-1 V3 Region

ALL_CONSENSUS	RSIHIGPGRAFYAT
CONSENSUS_20	QE-R---MAWYSMG
O_ANT70	QEMR---MAWYSMG
O_CONSENSUS_94	QE-R---MAWYSMG
O_VAU	QK-MA--MAWYSMA
O_MVP5100	QD-YT--M-WRSN-

Mutations in HIV-1 Reverse Transcriptase and Protease Associated with Drug Resistance

John W. Mellors,¹ Brendan A. Larder,² Raymond F. Schinazi³

¹ *University of Pittsburgh/VAMC, Graduate School of Public Health, 403 Parran Hall, 130 Desoto St, Pittsburgh, PA 15261,*

² *The Wellcome Foundation, Ltd., Langley Court, Beckenham, Kent BR3 3BS, UK.*

³ *Emory University/VAMC, 1670 Clairmont Road, Decatur, GA 30033.*

INTRODUCTION

The knowledge of HIV-1 drug resistance continues to increase at an exponential rate. This is particularly true for resistance to HIV-1 protease inhibitors, for which the first resistant variants were described less than two years ago (Otto et al., *PNAS* 90:7453-7457, 1993). Now resistance to virtually all classes of protease inhibitors has been documented. To keep our readership informed of newly described mutations, we have updated the tables below, which first appeared in *IAVN* in May 1994 (Vol 2; No 5). Following the tables is a list of abbreviations used. We hope that the tables' revised content will be useful to virologists and clinicians active in the field. We urge all investigators to provide additions or amendments of the table to any of the authors. Data formatted as in the table with an appropriate reference (abstract, manuscripts or paper) would be welcomed.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Margaret Tisdale, Michael J. Otto, Martin L. Bryant, Daniel W. Norbeck, Pin-Fang Lin and Emilio A. Emini for contributing essential information to the table.

Mutations in RT and Protease

NUCLEOSIDE RT INHIBITORS

Compound	Amino Acid Change	Codon Change	In Vitro	In Vivo	Comments	Confirmed by Site-directed Mutagenesis	Refs.
AZT	M41L	ATG to TTG	n.d.	Yes	M41L: 4-fold resistance; M41L/T215Y: Yes	Yes	[1, 2, 3]
	D67N	GAC to AAC	Yes	Yes	60- to 70-fold; K67N/K70R/T215Y/	Yes	
	K70R	AAA to AGA	Yes	Yes	K219Q: 120-fold; M41L/K67N/	Yes	
	T215Y	ACC to TAC	Yes	Yes	K70R/T215Y: 180-fold. Effect of	Yes	
	T215F	ACC to TTC	n.d.	Yes	T215Y is reversed by a ddl resistance	Yes	
	K219Q	AAA to CAA	n.d.	Yes	mutation (L74V), NNRTI	Yes	
	K219E	AAA to GAA	n.d.	No	mutations (L100I, Y181C) or a (-)-FTC/3TC mutation (M184V)	Yes	
ddl	L74V	TTA to GTA	No	Yes	5- to 10-fold resistance; cross-resistance Yes to ddC; can reverse the effect of the T215Y AZT resistance mutation	Yes	[4]
	V75T	GTA to ACA	Yes	No	observed with D4T selection; cross-resistance to ddl/ddC	Yes	[5]
	M184V	ATG to GTG	Yes	Yes	3- to 4-fold resistance; cross-resistance to ddC	Yes	[6]
ddC	K65R	AAA to AGA	Yes	Yes	4- to 10-fold resistance; observed in pts on receiving ddl or ddC therapy	Yes	[7, 8]
	T69D	ACT to GAT	No	Yes	5-fold resistance	Yes	[9]
	L74V	TTA to GTA			observed with ddl therapy	Yes	[2]
	V75T	GTA to ACA			observed with D4T selection in vitro	Yes	[4]
	M184V	ATG to GTG			observed with ddl, 3TC therapy	Yes	[6]
	Y215C	TTC to TGC	No	Yes	4-fold resistance; arises on background of T215YAZT resistance mutation	Yes	[10]
D4T	I50T	ATT to ACT	Yes	?	30-fold resistance	Yes	[11]
	V75T	GTA to ACA	Yes	Yes	7-fold resistance; cross-resistance to ddl, ddC, d4C, and (-)-FTC	Yes	[5]
3TC or (-)-FTC	M184V	ATG to GTG or GTA	Yes	Yes	>100-fold resistance; M184V and M184I can suppress effects of AZT resistance mutations	Yes	[12-14]
	M184I	ATG to ATA	Yes	Yes		Yes	[12-14]
Foscarnet	W88S	TGG to TCG	No	Yes	4-fold resistance	Yes	[15]
	E89G	GAA to GGA	Yes	No	isolated by screening RT clones for ddGTP resistance; 14-fold viral resistance	Yes	[16]
	E89K	GAA to GGA	Yes	No	E89K and L92I cause increased		[17]
	L92I	TTA to ATA	Yes	No	susceptibility to AZT		[17]
	S156A	TCA to GCA	Yes	No	and HIV-1 specific RTI		[17]
	Q161L	CAA to CTA	Yes	Yes	10-fold resistance	Yes	[15]
	H208Y	CAT to TAT	Yes	Yes	2-fold resistance	Yes	[15]
					Q161L + H208Y cause increased susceptibility to AZT (100-fold), nevirapine (20-fold) and TIBO R82150 (30-fold)		
1592U89	K65R	AAA to AGA	Yes	No	3-fold resistance	Yes	[18]
	L74V	TTA to GTA	Yes	No	4-fold resistance	Yes	
	Y115F	TAT to TTT	Yes	No	2-fold resistance	Yes	
	M184V	ATG to GTG	Yes	No	3-fold resistance	Yes	
					K65R/M184V: 8-fold resistance L74V/M184V: 9-fold resistance L74V/Y115F/M184V: 11-fold resistance		

HIV-1-SPECIFIC RT INHIBITORS

Compound	Amino Acid Change	Codon Change	In Vitro	In Vivo	Comments	Confirmed by Site-directed Mutagenesis	Ref.
Nevirapine	A98G	GCA to GGA	No	Yes		Yes	[19]
	L100I	TTA to ATA	No	Yes		Yes	[20]
	K103N	AAA to AAC	No	Yes		Yes	[20]
	V106A	GTA to GCA	Yes	Yes	~100-fold resistance; no effect on AZT resistance	Yes	[19-22]
	V108I	GTA to ATA	No	Yes		Yes	[20]
	Y181C	TAT to TGT	Yes	Yes	>100-fold resistance; cross-resistance to other NRTI; can suppress effects of AZT resistance mutations	Yes	[19, 23-25]
	Y181I	TGT to ATT	No	Yes	High-level resistance		[26]
	Y188C	TAT to TGT	No	Yes		Yes	[20]
	G190A	GGA to GCA	No	Yes		Yes	[19]
TIBO R82150	L100I	TTA to ATA	Yes	?	>100-fold resistance; can reverse effects of AZT resistance mutations	Yes	[27-29]
TIBO R82913	L100I	TTA to ATA	Yes	?		Yes	[31]
	K103N	AAA to AAC	Yes	?	>100-fold resistance	Yes	[22]
	V106A	GTA to GCA	Yes	?	~100-fold resistance	Yes	[21]
	E138K	GAG to AAG	Yes	?			[28]
	Y181C	TAT to TGT	Yes	?	>100-fold resistance	Yes	[21]
	Y188H	TAT to CAT	Yes	?		Yes	[28]
	Y188L	TAT to TTA	No	Yes		Yes	[30]
L'697,593	K103N	AAA to AAC	Yes	?	20-fold resistance	Yes	[23]
	Y181C	TAT to TGT	Yes	?	>100-fold resistance	Yes	[23]
L'697,661	A98G	GCA to GGA	No	Yes	8-fold resistance	Yes	[31]
	L100I	TTA to ATA	Yes	No	2-fold resistance	Yes	[31]
	K101E	AAA to GAA	No	Yes	8-fold resistance	Yes	[31]
	K103N	AAA to AAC	Yes	Yes	8-fold resistance	Yes	[31]
	K103Q	AAA to CAA	No	Yes	8-fold resistance	Yes	[32]
	V108I	GTA to GCA	Yes	Yes	4-fold resistance	Yes	[31]
	V179D	GTT to GAT	No	Yes	4-fold resistance	Yes	[31]
	V179E	GTT to GAG	No	Yes	8-fold resistance	Yes	[31]
	Y181C	TAT to TGT	Yes	Yes	>30-fold resistance	Yes	[31]
BHAP U-90152 (delavirdine)	P236L	CCT to CTT	Yes	?	P236L sensitizes RT ~10-fold to Nevirapine, TIBO R82193, and L'697,661	Yes	[33]
BHAP U-87201 (ateviridine)	K101E	AAA to GAA	No	Yes	K103N and Y181C observed with U-87201 monotherapy; K101E, Y188H, E233Y and K238T observed with U87201/AZT combination therapy		[34]
	K103N	AAA to AAC	No	Yes			[34]
	Y181C	TAT to TGT	No	Yes			[34]
	Y188H	TAT to CAT	No	Yes			[34]
	E233V	GAA to GTA	No	Yes			[34]
	P236L	CCT to CTT	Yes	No		Yes	[33]
	K238T	AAA to ACA	No	Yes			[34]
BHAP U 88204	L100I	TTA to ATA	Yes	?			[22, 35]
	V106A	GTA to GCA	Yes	?		Yes	[35]
	Y181C	TAT to TGT	Yes	?		Yes	[35]
	Y181I	TGT to ATT	Yes	Yes	High-level resistance to BHAP, Nevirapine, and TIBO. Observed in one Nevirapine-treated patient		[36]
HEPT	Y188C	TAT to TGT	Yes	?			[37]
E-EBU	Y181C	TAT to TGT	Yes	?			[37]
E-EBU-dM	Y106A	GTA to GCA	Yes	?			[37]

Mutations in RT and Protease

HIV-1-SPECIFIC RT INHIBITORS (cont.)

Compound	Amino Acid Change	Codon Change	In Vitro	In Vivo	Comments	Confirmed by Site-directed Mutagenesis	Ref.
E-EPU and E-EPSeU	Y181C Y188C	TAT to TGT TAT to TGT	Yes Yes	? ?	Y188C is the predominant mutation for E-EPSeU; Y188C confers greater resistance to E-EPSeU/E-EPU than Y181C	Yes Yes	[38] [38]
α -APA R18893	Y181C	TAT to TGT	Yes	?		Yes	[39]
S-2720	G190E	GGA to GAA	Yes	?	Mutation decreases RT activity and viral replication competency	Yes	[40]
TSAO	B138K	GAG to AAG	Yes	?	>100-fold resistance	Yes	[41, 42]
BM+51.0836	Y181C	TAT to TGT	Yes	?		Yes	[43]

PROTEASE INHIBITORS

Compound	Amino Acid Change	Codon Change	In Vitro	In Vivo	Comments	Confirmed by Site-directed Mutagenesis	Refs.
A-77003	R8Q	CGA to CAA	Yes	?	10-fold viral resistance	Yes	[44, 45]
	R8K	CGA to AAA	Yes	?	10-fold viral resistance	Yes	[44]
	V32I	GTA to ATA	Yes	?	7-fold enzyme resistance	Yes	[45]
	M46I	ATG to ATA	Yes	?	No affect on susceptibility but restores replication kinetics of R8Q mutant	Yes	[44]
	M46L	ATG to TTG	Yes	?	2-3-fold enzyme resistance	Yes	[45]
	M46F	ATG to TTC	Yes	?	4-fold enzyme resistance	Yes	[45]
	G48V	GGG to GTG	Yes	?	R8K/M46I/G48V: 20-fold viral resistance		[46]
	V82I	GTC to ATC	Yes	?	no resistance alone but V32I and V82I are synergistic mutations yielding 20-fold enzyme resistance	Yes	[45]
	V82A	GTC to GCC	Yes	?	rare; seen with M46F		[47]
	I82T	ATC to ACC	Yes	?	G48V/I82T combined produce 100-fold viral resistance		[47]
A75925	V32I	GTA to ATA	Yes	?	40-fold viral resistance		[48]
ABT-538	V82F	GTC to TTC	Yes	?	V82F/I84V: 8- to 10-fold viral resistance	Yes	[49]
	I84V	ATA to GTA	Yes	?	M46I/L63P/A71V/V82F/I84V: 27-fold resistance	Yes	
BILA 1906 BS	V32I	GTA to ATA	Yes	?			[50]
	M46L	ATG to TTG	Yes	?	V32I/A71V: 3-fold viral resistance		
	A71V	GCT to GTT	Yes	?	V32I/A71V/M46L/I84V: 5-fold		
	I84V	ATA to GTA	Yes		V32I/A71V/M46L/I84V: 1000-fold (mutation also detected in p6/p7 cleavage site)		
BMS 186,318	A71T	GCT to ACT	Yes	?	A71T/V82A: 15-fold viral resistance;	Yes	[51]
	V82A		Yes	?	4-fold cross resistance to A77003	Yes	
L-735, 524	L10R	CTC? to CGC?	No	Yes	M46I/L63P/V82T: 4-fold viral resistance	Yes	[52]
	M46I	ATG to ATA	No	Yes	L10R/M46I/L63P/V82T: 4-fold viral resistance		
	L63P	CTC? to CCC?	No	Yes	L10R/M46I/L63P/V82T/I84V: 8-fold viral resistance; cross resistance to XM-323 (15-fold), A-80987 (4-fold), Ro-31-8959 (8-fold), VX-478 (8-fold), SC-52151 (8-fold)		
	V82T	GTC to ACC	No	Yes			
	I84V	ATA to CTA	No	Yes			
	V32I	GTA to ATA	Yes	?	V32I/M46L/V82A: 3-fold viral resistance		[46]
	M46I	ATG to ATA	Yes	?	V32I/M46L/A71V/V82A: 14-fold viral resistance		
	A71V	GCT to GTT	Yes	?			
	V82A	GTC to GCC	Yes	?			
P9941	V82A	GTC to GCC	Yes	?	6-8-fold resistance	Yes	[53]
Ro 31-8959	G48V	GGG to GTG	Yes	Yes		Yes	[54]
	I84V	ATA to GTA	Yes	?			[46]
	L90M	TTG to ATG	Yes	Yes	G48V/L90M combined yield >100-fold enzyme resistance, but double mutant rare in vivo; L90M most common in vivo		[54]
					G48V/I84V/L90M: 90-fold viral resistance		[46]
RPI-312	I84V	ATA to GTA	Yes	?	5-fold viral resistance	Yes	[55]
SC-52151	L24V	TTA to GTA	Yes	?	N88D: 10-fold viral resistance		[56, 57]
	G48V	GGG to GTG	Yes	?	G48V/V82A: 15-fold viral resistance		
	A71V	GCT to GTT	Yes	?	G48V/A71V/V75I/P81T: 20- to 30-fold		
	V75I	GTA to ATA	Yes	?	A71V/V75I/P81T: 30-fold viral resistance		
	P81T	CCT to ACT	Yes	?	L24V/G48V/A71V/V75I/P81T: 1000-fold		

Mutations in RT and Protease

PROTEASE INHIBITORS (cont.)

Compound	Amino Acid Change	Codon Change	In Vitro	In Vivo	Comments	Confirmed by Site-directed Mutagenesis	Refs.
	V82A	GTC to GCC	Yes	?	viral resistance; some cross resistance to SC55389a and Ro31-8959, but not to L-735,524		
	N88D	AAT to GAT	Yes	?			
SC-55389	L10F	CTC to CGC	Yes	?	N88S seen alone with L10F; 10- to 20-fold viral resistance		[56, 57]
	N88S	AAT to AGT	Yes	?			
VB 11,328	L10F	CTC to GGC	Yes	?	L10F/I84V: 10-fold viral resistance		[58]
	M46I	ATG to ATA	Yes	?	150V/M46I/I47V: 20-fold viral resistance	Yes	[46, 58]
	I47V	ATA to CTA	Yes	?			
	I50V	ATT to GTT	Yes	?	3-fold viral resistance	Yes	[46]
	I84V	ATA to GTA	Yes	?			
XM323	L10F	CTC to CGC			L10F/V82A: 2.0-fold viral resistance; L10F/K45I/I84V: 50-fold	Yes	[59]
	K45I	AAA to ATA					[46]
	Q48V	GGG to GTG	Yes	?		Yes	[59]
	V82A	GTC to GCC	Yes	?	V82A/M46L: 7-fold resistance	Yes	[59]
					V82A/M46L/L97V: 11-fold resistance		
	V82I	GTC to ATC	Yes	?	<2-fold viral resistance	Yes	[59]
	V82F	GTC to TTC	Yes	?	see below	Yes	[59]
	I84V	ATA to GTA	Yes	?	12-fold resistance alone; V82F/I84V: 92-fold resistance; cross resistant to P9941, but not A77003 or Ro 31-8959	Yes	[46, 59]
	L97V	TTA to GTA	Yes	?	no resistance alone; V82A/L97V: 3.0 fold resistance	Yes	[59]

ABBREVIATIONS

n.d.: not determined ?; unknown

Amino acids: A, alanine; C, cysteine; D, aspartate; E, glutamate; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

1592U89:	(1S, 4R)-4-[2-amino-6-cyclopropylamino]-9H-purin-9-yl]-2-cyclopentene-1-methanol succinate
3TC:	(-)- β -L-2',3'-dideoxy-3'-thiacytidine
α -APA 18893:	alpha-nitroanilino-phenylacetamide
A-77003:	C2 symmetry-based protease inhibitor (Abbott)
A-75925:	C2 symmetry-based protease inhibitor (Abbott)
ABT-538:	C2 symmetry-based protease inhibitor (Abbott)
BM+51.0836:	thiazolo-isoxindolinone derivative
AzddU:	3'-azido-2',3'-dideoxyuridine
AZT-p-DDI:	3'-azido-3'-deoxythymidyl-(5',5')-2,3,-dideoxy-5'-inosinic acid
AZT:	3'-azido-2',3'-dideoxythymidine
BHAP:	bisheteroaryl piperazine
BILA 1906:	protease inhibitor (Boehringer-Ingelheim)
BMS 186,318	aminodiol derivative HIV-1 protease inhibitor (Bristol-Myer-Squibb)
d4C:	2',3'-didehydro, 2',3'-deoxycytidine
d4T:	2',3'-didehydro,3'-deoxythymidine
ddC:	2',3'-dideoxycytidine
ddI:	2',3'-dideoxyinosine
EBU-dM:	5-ethyl-1-thioxymethyl-6-(3,5-dimethylbenzyl)-uracil
E-EBU:	5-ethyl-1-ethoxymethyl-6-benzyluracil
E-EFSeU:	1-(ethoxy-methyl)-(6-phenyl-selenyl)-5-ethyl-uracil
E-EPU:	1-(ethoxy-methyl)-(6-phenyl-thio)-5-ethyluracil
(-)-FTC:	(-)- β -L-2',3'-dideoxy-5-fluoro-3'-thiacytidine
HEPT:	1-[(2-hydroxyethoxy)methyl]6-(phenylthio)-thymine
L'697,593:	5-ethyl-6-methyl-3-(2-phthalimido-ethyl)pyridin-2(1H)-one
L'735,524:	hydroxyaminopentane amide HIV-1 protease inhibitor (Merck)
L'697,661:	3-[[(-4,7-dichloro-1,3-benzoxazol-2-yl)methyl]amino]-5-ethyl-6-methylpyridin-2(1H)-one
L-PDDC:	(-)- β -L-5-fluoro-2',3'-dideoxycytidine
L-FDDC:	(-)- β -L-5-fluoro-dioxolane cytosine
Nevirapine:	11-cyclopropyl-5,11-dihydro-4-methyl-6H-dipyrido[3,2-b:2',3'-e] diazepin-6-one
NNRTI:	nonnucleoside reverse transcriptase inhibitor
P9941:	protease inhibitor (Dupont Merck); [2-pyridylacetyl-IlePheAla- Ψ (CHOH)] ₂
PFA:	phosphonoformate (foscamet)
PMEA:	9-(2-phosphonyl-methoxyethyl)adenine
Ro 31-8959:	hydroxyethylamine derivative HIV-1 protease inhibitor (Roche)
RPI-312:	peptidyl protease inhibitor; 1-[(3s)-3-(N-alpha-benzylloxycarbonyl)-L-asparginyl]amino-2-hydroxy-4-phenylbutyl-N-tert-butyl-L-proline amide
S-2720:	6-chloro-3,3-dimethyl-4-(isopropenylloxycarbonyl)-3,4-dihydroquinoxalin-2(1H)thione
SC-52151	hydroxyethylurea isostere protease inhibitor (Searle)
SC-55389A	hydroxyethylurea isostere protease inhibitor (Searle)
TIBO R82150:	(+)-(5S)-4,5,6,7-tetrahydro-5-methyl-6-(3-methyl-2-butenyl)imidazo[4,5,1-jk][1,4]-benzodiazepin-2(1H)-thione
TIBO 82913:	(+)-(5S)-4,5,6,7-tetrahydro-9-chloro-5-methyl-6-(3-methyl-2-butenyl)imidazo[4,5,1-jk][1,4]benzodiazepin-2(1H)-thione
TSAO-m3T:	[2',3'-bis-O-(tert-butyl)dimethylsilyl]-3'-spiro-5'-(4'-amino-1',2'-oxathiole-2',2'-dioxide)]-b-d-pentofuranosyl-N3-methyl-thymine
U90152:	1-[3-[(1-methylethyl)-amino]-2-pyridinyl]-4-[[5-[(methylsulfonyl)-amino]-1H-indol-2-yl]carbonyl]-piperazine
VB 11,382:	hydroxyethylsulphonamide protease inhibitor (Vertex)
VX-478:	hydroxyethyl-sulphonamide protease inhibitor (Vertex)
XM 323:	cyclic urea protease inhibitor (Dupont Merck)

**Table 2. Combinations of HIV-1 RTase mutations
that confer multi-drug resistance**

RTase genotype	AZT	ddI	Nevirapine	3TC
Single drug resistance				
M41L/T215Y	R	S	S	S
L74V	S	R	Sensitive	S
Y181C or V106A	S	S	Resistant	S
Double drug resistance				
M41L/T215Y/L74V	R	R	S	S
M41L/T215Y/V106A	R	S	R	S
L74V/Y181C	S	R	R	S
M184V	S	R	S	R
Triple drug resistance				
M41L/T215Y/L74V/V106A	R	R	R	S
M41L/T215Y/L74V/V106A/M184V	S	R	R	R

From Larder J Gen Virol 75, 951-957 (1994)

Figure 2. Temporal order of the emergence of mutations conferring AZT resistance

From Larder J Gen Virol 75, 951-957 (1994)

References

- [1] Larder BA & Kemp SD. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (AZT). *Science* 1989; 246: 1155-1158.
- [2] Larder BA, Coates KE & Kemp SD. Zidovudine-resistant human immunodeficiency virus selected by passage in cell culture. *J Virol* 1991; 65: 5232-5236.
- [3] Kellam P, Boucher CA & Larder BA. Fifth mutation in human immunodeficiency virus type 1 reverse transcriptase contributes to the development of high-level resistance to zidovudine. *Proc Natl Acad Sci USA* 1992; 89: 1934-1938.
- [4] St Clair MH, Martin JL, Tudor WG, et al. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. *Science* 1991; 253: 1557-1559.
- [5] Lacey SF & Larder BA. A novel mutation (V75T) in the HIV-1 reverse transcriptase confers resistance to 2',3'-dideohydro-2',3'-dideoxythymidine (D4T) in cell culture. *Antimicrob Agents Chemother* 1994; 38: 1428-1432.
- [6] Gu Z, Gao Q, Li X, Parniak MA & Wainberg MA. Novel mutation in the human immunodeficiency virus type 1 reverse transcriptase gene that encodes cross-resistance to 2',3'-dideoxy-inosine and 2',3'-dideoxycytidine. *J Virol* 1992; 66: 7128-7135.
- [7] Gu Z, Gao Q, Fang H, Salomon H, Parniak MA, Goldberg E, Cameron J & Wainberg MA. Identification of a mutation at codon 65 in the IKKK motif of reverse transcriptase that encodes HIV resistance to 2',3'-dideoxycytidine and 2',3'-dideoxy-3'-thiacytidine. *Antimicrob Agents Chemother* 1994; 38: 275-81.
- [8] Zhang D, Caliendo AM, Eron JJ, Devore KM, Kaplan JC, Hirsch MS & D'Aquila RT. Resistance to 2',3'-dideoxycytidine conferred by a mutation in codon 65 of the HIV-1 reverse transcriptase. *Antimicrob Agents Chemother* 1994; 38: 282-287.
- [9] Fitzgibbon JE, Howell RM, Haberkett CA, Sperber SJ, Gocke DJ & Dubin DT. Human immunodeficiency virus type 1 pol gene mutations which cause decreased susceptibility to 2',3'-dideoxycytidine. *Antimicrob Agents Chemother* 1992; 36: 153-157.
- [10] Slade DE, Vavro CL, Stapleton JT, Swack N & St Clair MH. A cysteine at codon 215 of HIV RT confers resistance to ddC. Second International HIV-1 Drug Resistance Workshop, 1993, Noordwijk, The Netherlands.
- [11] Gu Z, Gao Q, Fang H, Parniak MA, Brenner BG, Wainberg MA. Identification of novel mutations that confer drug resistance in the human immunodeficiency virus polymerase gene. *Leukemia* 1994; 8S1: 5166-5169.
- [12] Schinazi RF, Lloyd RJ, Nguyen M-H, Cannon DL, McMillan, Ilksoy N, Chu CK, Liotta DC, Bazmi HZ & Mellors JW. Characterization of human immunodeficiency viruses resistant to oxathiolane-cytosine nucleosides. *Antimicrob Agents Chemother* 1993; 37: 875-881.
- [13] Tisdale M, Kemp SD, Parry NR & Larder BA. Rapid in vitro selection of human immunodeficiency virus type 1 resistant to 3'-thiacytidine inhibitors due to a mutation in the YMDD region of reverse transcriptase. *Proc Natl Acad Sci USA* 1993; 90: 5653-5656.
- [14] Gao Q, Gu ZX, Parniak MA, Cameron J, Cammack N, Boucher C & Wainberg MA. The same mutation that encodes low-level human immunodeficiency virus type-1 resistance to 2',3'-dideoxyinosine and 2',3'-dideoxycytidine confers high-level resistance to the (-) enantiomer of 2',3'-dideoxy-3'-thiacytidine. *Antimicrob Agents Chemother* 1993; 37: 1390-1392.
- [15] Mellors J, Bazmi H, Weir J, Arnold E, Schinazi R, & Mayers D. Novel mutations in HIV-1 reverse transcriptase confer resistance to foscarnet in laboratory and clinical isolates. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [16] Prasad VR, Lowy I, de los Santos T, Chiang L, & Goff SP. 1991. Isolation and characterization of a dideoxyguanosine triphosphate-resistant mutant of human immunodeficiency virus reverse transcriptase. *Proc Natl Acad Sci USA* 88: 11363-11367.
- [17] Tachedjian G, Gurusinghe A, Hooker D, Deacon N, Mills J, Birch C. In vitro generation and charac-

- terization of foscarnet-resistant HIV-1. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [18] Tisdale M, Parry NR, Cousens D, St. Clair MH, Boone LR. Anti-HIV activity of (1S, 4R)-4-[2-amino-6-cyclopropylamino)-9H-purin-9-yl]-2-cyclopentene-1-methanol (1592U89). Abstracts of the 34th Interscience Conference on Antimicrobial Agents and Chemotherapy, 1994, Orlando, Fla, USA, p92.
- [19] Richman DD, D Havlir, J Corbell, D Looney, C Ignacio, SA Spector, J Sullivan, S Cheeseman, K Barringer, D Pauletti, C-K Shih, M Myers, & J Griffin. Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy. *J Virol* 1994; 68: 1660-1666.
- [20] Richman DD. Resistance of clinical isolates of human immunodeficiency virus to antiretroviral agents. *Antimicrob Agents Chemother* 1993; 37: 1207-1213.
- [21] Larder BA. 3'-Azido-3'-deoxythymidine resistance suppressed by a mutation conferring human immunodeficiency virus type 1 resistance to nonnucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother* 1992; 36: 2664-2669.
- [22] Balzarini J, Karlsson A, Pérez-Pérez M-J, Camarasa MJ, Tarpley WG, De Clercq E. Treatment of human immunodeficiency virus type-1 (HIV-1)-infected cells with combinations of HIV-1-specific inhibitors results in a different resistance pattern than does treatment with single-drug therapy. *J Virol* 1993; 67: 5353-5359.
- [23] Nunberg JH, Schleif WA, Boots EJ, O'Brien JA, Quintero JC, Hoffman JM, Emmini EA & Goldman ME. Viral resistance to human immunodeficiency virus type 1-specific pyridinone reverse transcriptase inhibitors. *J Virol* 1991; 65: 4887-4892.
- [24] Richman D, Shih CK, Lowy I, Rose J, Prodanovich P, Goff S & Griffin J. Human immunodeficiency virus type 1 mutants resistant to nonnucleoside inhibitors of reverse transcriptase arise in tissue culture. *Proc Natl Acad Sci USA* 1991; 88: 11241-11245.
- [25] Mellors JW, Dutschman GE, Im GJ, Tramontano E, Winkler SR & Cheng YC. In vitro selection and molecular characterization of human immunodeficiency virus-1 resistant to non-nucleoside inhibitors of reverse transcriptase. *Mol Pharm* 1992; 41: 446-451.
- [26] Shaw G, Wei X, Johnson V, Taylor M, Decker J, Kilby M, Lifson J, Hahn B, & Saag M. Nucleotide sequence analysis of HIV-1 RNA and DNA from plasma and PBMCs of patients treated with ZDV, ddI, and nevirapine: rapid turnover and resistance development in vivo. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [27] Mellors JW, Im GJ, Tramontano E, Winkler SR, Medina DJ, Dutschman GE, Bazmi HZ, Piras G, Gonzalez CJ & Cheng Y-C. A single conservative amino acid substitution in the reverse transcriptase of human immunodeficiency virus-1 confers resistance to (+)-(5S)-4,5,6,7-tetrahydro-5-methyl-6-(3-methyl-2-butenyl)imidazo[4,5,1-jk][1,4]benzodiazepin-2(1H)-thione (TIBO R82150). *Mol Pharm* 1993; 43: 11-6.
- [28] Balzarini J, Karlsson A, Pérez-Pérez M-J, Vrang J, Walbers J, Zhang H, Öberg B, Vandamme A-M, Camarasa M-J & De Clercq E. HIV-1-specific reverse transcriptase inhibitors show differential activity against HIV-1 mutant strains containing different amino acid substitutions in the reverse transcriptase. *Virol* 1993; 192: 246-253.
- [29] Byrnes V, Blahy O, Condra J, et al. Phenotypic susceptibility of human immunodeficiency virus type 1 RT containing substitutions which engender resistance to nucleoside and nonnucleoside inhibitors. Third Workshop on Viral Resistance, 1993, Gaithersburg, MD, USA.
- [30] Vandamme A-M, Debyser Z, Pauwels R, DeVreese K, Goubau P, Youle M, Gazzard B, Stoffels PA, Cauwenbergh CP, Anne J, Andries K, Janssen PAJ, Desmyter J, De Clercq E. Characterization of HIV-1 strains isolated from patients treated with TIBO R82193. *AIDS Res Human Retro* 1994; 10: 39-46.
- [31] Byrnes VW, Sardana W, Schleif WA, Condra JH, Waterbury JA, Wolfgang JA, Long WJ, Schneider CL, Schlabach AJ, Wolanski BS, Graham DJ, Gotlib L, Rhodes A, Titus DL, Roth E, Blahy OM, Quintero JC, Staszewski S & Emmini EA. Comprehensive mutant enzyme and viral variant assessment

- of human immunodeficiency virus type-1 reverse transcriptase resistance to non-nucleoside inhibitors. *Antimicrob Agents Chemother* 1993; 37: 1576-1579.
- [32] Saag MS, Emimi EA, Laskin OL, et al. A short-term clinical evaluation of L-697,661, a non-nucleoside inhibitor of HIV-1 reverse transcriptase. *N Engl J Med* 1993; 329: 1065-1072.
- [33] Dueweke TJ, Pushkarskaya T, Poppe SM, Swaney SM, Zhao Q, Chen SY, Stevenson M & Tarpley WG. A mutation in reverse transcriptase of bis(heteroaryl)piperazine-resistant human immunodeficiency virus type 1 that confers increased sensitivity to other non-nucleoside inhibitors. *Proc Natl Acad Sci USA* 1993; 90: 4713-4717.
- [34] Demeter L, Resnick L, Nawaz T, Timponi JG Jr., Batts D & Reichman RC. Phenotypic and genotypic analysis of atevirdine (ATV) susceptibility of HIV-1 isolates obtained from patients receiving ATV monotherapy in a phase I clinical trial (ACTG 187): comparison to patients receiving combination therapy with ATV and zidovudine. Third Workshop on Viral Resistance, 1993, Gaithersburg, MD, USA.
- [35] Vasudevachari MB, Battista C, Lane HC, et al. Prevention of the spread of HIV-1 infection with nonnucleoside reverse transcriptase inhibitors. *Virology* 1992; 190: 269-277.
- [36] Balzarini J, Karlsson A, Sardana VV, Emimi EA & De Clercq E. Selection of novel RT mutations upon continued exposure to non-nucleoside reverse transcriptase inhibitors. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [37] Balzarini J, Karlsson A & De Clercq E. Human immunodeficiency virus type 1 drug-resistance patterns with different 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine derivatives. *Mol Pharmacol* 1993; 44: 694-701.
- [38] Nguyen MH, Schinazi RF, Shi C, Goudgaon NM, McKenna PM, & Mellors JW. Resistance of human immunodeficiency virus to acyclic 6-phenylselenenyl- and 6-phenylthiopyrimidines. *Antimicrob Agents Chemother* 1994; 38: 2409-2414.
- [39] de Bethune M-P, Pauwels R, Andries K, et al. AZT resistance reversal by the non-nucleoside reverse transcriptase inhibitor α -APA R18893 in a symptomatic HIV-infected individual. Second International HIV-1 Drug Resistance Workshop, 1993, Noordwijk, The Netherlands.
- [40] Klein J-P, Bender R, Billhardt U-M, Meichsner C, Riess G, Rösner M, Winkler I & Paessens A. Activity of a novel quinoxaline derivative against human immunodeficiency virus type 1 reverse transcriptase and viral replication. *Antimicrob Agents Chemother* 1993; 37: 1659-1664.
- [41] Balzarini J, Velázquez S, San-Félix A, Karlsson A, Pérez-Pérez M-J, Camarasa M-J & De Clercq E. Human immunodeficiency virus type 1-specific [2',5'-bis-O-(tetrabutyl-dimethylsilyl)-beta-D-ribofuranosyl]-3'-spiro-5'-(4'-amino-1',2'-oxathiole-2',2'-dioxide)-purine analogues show a resistance spectrum that is different from that of the human immunodeficiency virus type 1-specific non-nucleoside analogues. *Mol Pharm* 1993; 43: 109-114.
- [42] Balzarini J, Karlsson A, Vandamme AM, Pérez-Pérez M-J, Zhang H, Vrang L, Öberg B, Bäckbro K, Unger T, San-Félix A, Velázquez S, Camarasa M-J & De Clercq E. Human immunodeficiency virus type-1 (HIV-1) strains selected for resistance against the HIV-1-specific 1, 2', 5'-bis-o-(tetrabutyl-dimethylsilyl)-3'-spiro-5'-(4'-amino-1', 2'-oxathiole-2', 2'-dioxide)-1-beta-D-pentofuranosyl (TSAO) nucleoside analogues retain sensitivity to HIV-1-specific non-nucleoside inhibitors. *Proc Natl Acad Sci USA* 1993; 90: 6952-6956.
- [43] Maass G, Immendoerfer U, Koenig B, Leser U, Mueller B, Goody R, & Pfaff E. Viral resistance to the thiazolo-iso-indolinones, a new class of nonnucleoside inhibitors of human immunodeficiency virus type 1 reverse transcriptase. *Antimicrob Agents and Chemother* 1993; 37:2612-7.
- [44] Ho DD, Toyoshima T, Mo H, Kempf DJ, Norbeck D, Chen C-M, Wideburg NE, Burt SK, Erickson JW, Singh MK. Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor. *J Virol* 1994; 68: 2016-2020.
- [45] Kaplan AH, Michael SF, Wehrle RS, Knigge MF, Paul DA, Everitt L, Kempf DJ, Norbeck DW, Erickson JW, R Swanstrom. Selection of multiple human immunodeficiency virus type 1 variants that encode viral proteases with decreased sensitivity to an inhibitor of viral protease. *Proc Natl Acad Sci*

USA 1994; 91: 5597-5601.

- [46] Tisdale M, Myers R, Parry NR, Oliver N, Machera B, & Blair E. Comprehensive analysis of HIV-1 variants individually selected for resistance to six HIV protease inhibitors. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [47] Swanstrom R, Smith T, Petit S, Irlbeck D, Shao W, Wehbie R, Sawhney R, Everitt L, Erickson J. Multiple sequence changes within HIV-1 protease confer reduced sensitivity to a symmetric protease inhibitor. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [48] Maschera B, Blance C, Brown D, Blair ED. Mutations conferring resistance to HIV-1 protease inhibitors can lie outside, as well as within, the enzyme active site and binding pockets. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [49] Kempf D, Markowitz M, Marsh K, Denissen J, Mo H, Bhat T, Park C, Kong X-P, Stewart K, McDonald E, Vasavanonda S, Flentge C, Wideburg N, Robins T, Hsu A, Leonard J, Ho D, Norbeck D. Pharmacokinetic and in vitro selection studies with ABT-538, a potent inhibitor of HIV protease with high oral bioavailability. Abstracts of the 34th Interscience Conference on Antimicrobial Agents and Chemotherapy, 1994, Orlando, Fla, USA.
- [50] Lamarre D, Croteau G, Pilote L, Rousseau P, Doyon L. Molecular characterization of HIV-1 variants resistant to specific viral protease inhibitors. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [51] Rose B, Greytok J, Bechtold C, Alam M, Terry B, Gong Y-F, DeVore K, Patrick A, Colono R, & Lin P-F. Combination therapy with two protease inhibitors as an approach to antiviral therapy. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [52] Emiei E, Schleif WA, Graham D, Deutsch P, Massari F, Teppler H, Squires K, Condra FH. Phenotypic and genotypic characterization of HIV-1 variants selected during treatment with the protease inhibitor L-735, 524. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [53] Otto MJ, Garber S, Winslow DL, et al. In vitro isolation and identification of human immunodeficiency virus (HIV) variants with reduced sensitivity to C-2 symmetrical inhibitors of HIV type-1 protease. *Proc Natl Acad Sci USA* 1993; 90: 7543-7547.
- [54] Jacobsen H, Brun-Vezinet F, Duncan I, Hänggi M, Ott M, Vella S, Weber J, & Mous J. Genotypic characterization of HIV-1 from patients after prolonged treatment with proteinase inhibitor saquinavir. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [55] El-Farrash MA, Kuroda MJ, Kitazaki T, Masuda T, Kato K, Hatanaka M, Harada S. Generation and characterization of a human immunodeficiency virus type 1 (HIV-1) mutant resistant to an HIV-1 protease inhibitor. *J Virol* 1994; 68: 233-239.
- [56] Potts KE, Smidt ML, Stallings WC, Clare M, Pillay D, Richman DD, Bryant ML. In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) variants with decreased sensitivity to hydroxyethylurea isostere containing protease inhibitors. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [57] Pillay D, Smidt ML, Potts KE, Bryant ML, Richman DD. In vitro selection of protease inhibitors resistant human immunodeficiency virus type 1 (HIV-1) strains. Abstracts of the 34th Interscience Conference on Antimicrobial Agents and Chemotherapy, 1993, Orlando, Fla, USA, p7.
- [58] Partaledis JA, Yamaguchi K, Byrn RA. In vitro selection and characterization of HIV-1 viral isolates with reduced sensitivity to inhibitors of HIV protease. Third International Workshop on HIV Drug Resistance, 1994, Kauai, Hawaii, USA.
- [59] King RW, Garber S, Winslow DL, Reid CD, Bacheler LT, Anton E, & Otto MJ. Multiple mutations in the human immunodeficiency virus protease gene are responsible for decreased susceptibility to protease inhibitors. *Antiv Chem Chemother* 1995; 6 (in press).

IV

Related Sequences

This section of the compendium has typically featured HIV related viral sequences—of non-primate lentiviruses and human T-cell lymphotropic viruses I and II. Beginning in 1993, Part IV entries included, with greater emphasis, coding sequences for cellular proteins involved with HIV and SIV pathogenesis. In 1994, Part IV is completely devoted to cellular proteins; however, this does not imply that HIV related viral sequences will not be included in future updates. Kuan-Teh Jeang, one of the database editors, has written a summary for this year's section.

PART IV Related Sequences

Introduction	IV-1
Contents(1994)	2
HIV and Cellular Factors (Kuan-Teh Jeang)	3
CEIF5A (M99499)	13
HIVEP1 (J05011)	14
HSCOUPTF (X16155, X58241)	15
HUMCOUPII (M62760)	16
HUMCYP (Y00052)	17
HUMCYPB (M60857)	18
HUMHIP116 (L34673)	19
HSIL02 (V00564)	21
HUMIL6 (M18403)	22
HUMini1 (U04847)	23
HUMKUP70 (J04611)	24
HUMKUP86 (J04977)	25
HUMLA (J04205, M11108)	27
HUMLCK (M36881)	28
HUMMSS1 (D11094)	29
HUMNB23 (M26697)	30
HNFAT (U08015)	31
HUMNFKB (M58603)	33
HUMP65NFKB (M62399)	35
HUMP68K (M35663)	36
HSZFPBF1 (X51435)	38
HUMSF2P32 (M69039)	41
HUMSP1TF (J03133)	42
HUMTBP1 (M34079)	44
HUMTFIIDA (M34960)	45
HSTGFB1 (X02812, J05114)	46
HUMTRBP (M60801)	48
HSUSF (X55666)	49
HUSFII (S50537)	50
HUMYC1A (L11289)	51
YY1 (M77698)	52
Hcfos (K00650, M16287)	53
HINTAV (U07375)	56
HUMINTBSA (J05633)	57
MMLEF1 (X58636)	59
RNBCOP (X57228)	61
SCSUG1 (X66400)	63

HIV and Cellular Factors

Kuan-Teh Jeang

Chief, Molecular Virology Section

LMM, NIAID, National Institutes of Health, Bethesda, MD 20892-0460

INTRODUCTION

Viruses are obligatory parasites of cells. Thus it is expected that host cell factors contribute importantly to the life-cycle of viruses. In this section, we survey four types of virus-cell interactions involving HIV-1. These four areas include: 1) DNA-binding proteins that recognize target motifs in proviral LTR; 2) RNA-binding proteins that bind HIV-1 RNAs; 3) cellular factors that form protein-protein complexes with HIV-1 regulatory proteins; and 4) cellular genes which are modulated upon viral infection. With rapid increases in knowledge in the area of virus-cell interactions, we anticipate that this initial survey would be expanded extensively in future editions of the data base. Additional discussions on HIV-cell interactions are found elsewhere (Jones and Peterlin, 1994; Jeang and Gagnol, 1994; Garcia and Gaynor, 1994).

I DNA-BINDING PROTEINS

The promoter-enhancers of the human immunodeficiency virus (HIV) are contained in the U3 of the viral long terminal repeat (LTR). HIV-1 U3 is 454 nucleotides long and has binding sites for many transcription factors. Some of these are diagrammed in Figure 1. With the exception of NF- κ B and Sp1 (Ross et al., 1991; Kim et al., 1993; Huang and Jeang, 1993; Huang et al., 1994), direct verification of the importance of the various sequences in viral contexts has not been performed.

Fig. 1. Diagrammatic representation of the positions in U3 and R for some of the DNA-binding proteins that recognize the HIV-1 LTR.

Subgenomic assays in cultured cells indicate that the primary contributors to HIV-promoter activity are the NF- κ B, Sp1, and TATAA (sequences from +1 to -105; Berkhout and Jeang, 1992). Nonetheless, it is likely that other DNA-binding factors also contribute in vivo. Beginning directionally from the 5' end of U3, binding sites for AP-1 (a fos/jun hetero-complex; van Straaten et al., 1983; Hattori et al., 1988) and COUP (a member of the steroid/thyroid receptor superfamily; Cooney et al., 1991) are found between -324 and -357. Both COUP and AP-1 are expressed in human T-cells, and thus could compete with each other for the same proviral DNA sites during HIV-1 infection.

More proximally, between -216 and -254 are binding motifs for nuclear factor from activated T-cells (NFAT; Shaw et al., 1988). NFAT is an intermediating transducer of signals initiated at the

T-cell antigen receptor. Recent evidence suggests that NFAT binding activity is composed of three discrete polypeptides, NFATp (McCaffrey et al., 1993), Fos and Jun (Yaseen et al., 1993).

In the region between -159 and -174 is a binding consensus sequence for USF (Gregor et al., 1990). USF was characterized initially as a positive activator of adenovirus major-late-promoter transcription. In the HIV context, there is conflicting information on whether this factor has moderating (and thus be regarded as a negative regulatory factor; NRF; Lu et al., 1990) or stimulating (Mackewa et al., 1991) effects. Interestingly, USF also binds a second unrelated sequence (-5 to +11) that surrounds the HIV-1 initiator (Hu et al., 1993). USF interaction at the initiator-proximal site strongly activates expression from the TATAA-promoter (Hu et al., 1993). Of note, a factor distinct from USF, which binds the same DNA-sequence, has been cloned and characterized (TFE3; Bechmann and Kadesch, 1991).

LEF-1 is a T-cell specific transcription factor (Waterman and Jones, 1990). Once bound to its cognate site, LEF-1 bends DNA and thereby facilitates the assembly of nucleoprotein complexes at the promoter (Giese et al., 1992; reviewed in Jones and Peterlin, 1994). A high affinity LEF-1 binding site is present at -122 to -143. Two low affinity binding sites exist at -37 to -51 and +17 to +32 (Waterman and Jones, 1990).

NF- κ B (Nabel and Baltimore, 1987) and Sp1 (Jones and Tjian, 1985; Jones et al., 1986) motifs are perhaps the best characterized sequence elements in the HIV-1 LTR. These sequences directly impact viral replication (Ross et al., 1991; Kim et al., 1993; Huang and Jeang, 1993; Huang et al., 1994), viral transcription (Harrich et al., 1990; Berkhout and Jeang, 1992; and references cited therein), and Tat transactivation (reviewed in Jones and Peterlin, 1994; Jeang and Gatignol, 1994). For more extensive discussions of the biochemical and functional properties of NF- κ B (Ghosh et al., 1990; Kieran et al., 1990; Nolan et al., 1991; Liou et al., 1991; and references cited therein) and Sp1 (Dyran and Tjian, 1983; Briggs et al., 1986; Kadonaga et al., 1987; and references cited therein), readers should consult elsewhere.

PRDII-BF1 is a 300 kDa zinc-finger containing protein (Baldwin et al., 1990; Seeler et al., 1994). PRDII-BF1 recognizes and binds the NF- κ B motif; however, it also binds a divergent sequence in R (+27 to +52; Seeler et al., 1994).

Positioned at the junction of U3 and R are sites for LBP (Yoon et al., 1994) and YY1 (Useheva and Shenk, 1994; Seto et al., 1991). While the role for LBP in HIV-1 transcription is not wholly understood (Jones et al., 1988; Kato et al., 1991), binding of YY1 to the LTR has been shown to repress HIV-1 expression and production of virions (Margolis et al., 1994). Besides LBP and YY1, a DNA-mediated activity for the induction of short transcripts (IST) has also been mapped to the same general vicinity (-5 to +26; Sheldon et al., 1993). The cDNA for the cellular factor that mediates IST-activity has not been isolated, and thus the authentic identity of this factor is unknown. HIP 116 is another newly cloned cDNA that binds to the TATA/initiator of the HIV-1 promoter (K.A. Jones, pers. communication).

II. RNA-BINDING PROTEINS

In recent years, it has become evident that RNA-binding proteins play important roles in gene regulation (see reviews, Keene and Query, 1991; Mattaj, 1993; Burd and Dreyfuss, 1994). For HIV-1, cellular proteins that bind viral regulatory RNAs have been studied in detail. In particular, at least eight host cell factors have been described to bind TAR RNA. Similarly, two RRE-binding factors have been characterized. There are also biologically compelling reasons as to why TAR- and RRE-binding proteins are meaningful contributors to the HIV-1 lifecycle.

The HIV-1 leader RNA, TAR, forms a stem-bulge-loop structure of approximately 60+ nucleotides (Muesing et al., 1987; Berkhout and Jeang, 1989). Early, it was reported that many human cellular proteins bound TAR RNA (Gatignol et al., 1989; Gaynor et al., 1989). Since then, some of these factors have been defined further. Tabulating from extant studies, eight proteins associate with either the bulge, loop, or stem of TAR RNA (see Figure 2). TAR loop-binding proteins include p68 (Marciniak et al., 1990), and TRP1/TRP185 (185 kDa; Sheline et al., 1991; Wu et al., 1991), while TRP2 (70-110 kDa; Sheline et al., 1991) binds to TAR-bulge. Proteins that complex with the

double-stranded stem of TAR RNA consist of PI/dsI (newly renamed as PKR; McCormack et al., 1992; Roy et al., 1991), SBP (140 kDa; Rounseville and Kumar, 1992), and TRBP (Gatignol et al., 1991; Gatignol et al., 1993). Finally, two human autoantigens recently have been identified as TAR RNA-binding factors. Lupus antigen Ku (Kaczmarek and Khan, 1993) binds to the loop of TAR, while La (Chang et al., 1994; Svitkin et al., 1994) recognizes U-residues within the overall context of the TAR secondary structure.

RRE-binding proteins have been studied less extensively. We know that TRBP also can bind RRE (Park et al., 1994), and that a 56 kDa factor has been described as a specific RRE-binding protein (Vaishnav et al., 1992). It is expected that other RRE-factors will emerge from future studies.

Fig. 2. Structure of the HIV-1 TAR RNA. Identities of the different proteins that bind to the loop, bulge or stem of the TAR hairpin are indicated.

III. PROTEIN-PROTEIN INTERACTIONS

Protein-protein interactions are well-documented to be important in gene regulation (see reviews, Lewin, 1990; Greenblatt, 1991). Recent studies suggest that many HIV-1 proteins complex with host cell factors. We describe below, in a non-exhaustive fashion, some examples pertaining to Tat, Rev, Gag, and Nef.

Although Tat is best known for transcription, it has other functions (Huang et al., 1994) and has been reported to be a secretable factor that promotes the growth of Kaposi-like cells (Ensoli et al., 1990). There is evidence that Tat can be taken up actively into cells (Frankel and Pabo, 1988) through binding to a cell-surface protein (Weeks et al., 1993) implicated to be $\alpha v \beta 5$ integrin (Vogel et al., 1993). Once inside cells, Tat interacts with multiple partners in activating transcription.

Two factors that bind Tat are themselves critical components of the eucaryotic polII transcription machinery. Genetic evidence supports a critical role for Sp1 in HIV-1 Tat-mediated transactivation (Harrich et al., 1989; Kamine et al., 1991; Southgate and Green, 1991; Berkhout and Jeang, 1992). Interestingly, a direct interaction between Tat and Sp1 (Jeang et al., 1993) has been documented and perhaps reflects observed virus-cell biology. In addition, there is evidence that Tat also contacts TFIID (Kashanchi et al., 1994) and that Tat-Sp1-TFIID could be a multiprotein complex (Huang et al., 1993) inside cells.

A large family of proteins related to the 26S protease from human erythrocytes (Dubiel et al., 1994) are Tat-binding polypeptides. Members of this family include TBP-1 (Nelbock et al., 1990; Ohana et al., 1993), TBP-7 (Ohana et al., 1993; Shaw and Ennis, 1993), MSS1 (Shibuya et al., 1992), and SUG1 (Swaffield et al., 1992). The exact role of this family of proteins in cellular metabolism is not wholly clear; however, the 26S protease seems to regulate the degradation of some cyclins and in this manner has been implicated in modulating the stability of oncoproteins such as c-Mos, c-Myb, c-Myc and p53 (Dubiel et al., 1994).

Two other cellular proteins, a 36 kDa protein (Desai et al., 1991) and a cellular protein kinase (Hermann and Rice, 1993), have been reported to bind Tat. The biological roles of these two factors remain to be defined.

There are at least three cellular factors that bind Rev. Rev localizes to the nucleolus, and work by Laemmli and colleagues (Fankhauser et al., 1991) has demonstrated a tight association between Rev and nucleolar B23 protein. This protein-protein complex likely directs the subcellular localization of Rev.

At the functional level Rev transactivation could be modulated by associations with cellular factors. Two additional proteins have been shown to bind Rev, translation factor eIF-5A (Ruhl et al., 1993), and a p32 protein (YL2; Luo et al., 1994) originally characterized by Krainer and colleagues (Krainer et al., 1991) as a splicing factor-SF2-associated polypeptide. YL2 was identified as a Rev-binding factor through random screening using the yeast-two-hybrid assay; however, in the same assay YL2/p32 also binds HIV-1 and HIV-2 Tat proteins (Trinh and Jeang, unpublished observation; B.R. Cullen, pers. communication).

In addition to Tat and Rev, both HIV Gag and Nef also have cellular partners. p55 and p24 Gag bind cyclophilin A and cyclophilin B proteins (Luban et al., 1993). The cyclophilins are cellular polypeptides originally characterized for their specific binding to cyclosporin A. Nef has been found to bind b-COP, a coat protein from non-clathrin-coated vesicles (R. Benarous, pers. communication). This interaction may be important in promoting the intracellular sequestration of CD4. Another group of investigators has found that Nef binds directly to the lck-kinase in vitro (D. McPhee, pers. communication).

IV. CELLULAR GENES MODULATED BY HIV

Infection of cells by HIV most likely results in activation and repression of many cellular genes. Because of the ambient complexity of gene expression inside cells, it is logistically difficult to dissect those genes that are upregulated from those that are downregulated from those that remain unperturbed. Nevertheless, several examples of genes that respond to HIV infection are known. These include IL-2 (Westendorf et al., 1994), IL6 (Scala et al., 1994), and TGF- β (Buonaguro et al., 1994), among others. The identification of these genes has been based on classical approaches; however, we anticipate that with the advent of mRNA differential display technology (Liang and Pardee, 1992) the exhaustive characterization of cellular genes that respond to infection by HIV should be accomplished soon.

References

REFERENCES

- [1] Baldwin, A. S. J., K. P. LeClair, H. Singh, and P. A. Sharp. 1990. A large protein containing zinc finger domains binds to related sequence elements in the enhancers of the class I major histocompatibility complex and kappa immunoglobulin genes. *Mol. Cell. Biol.* 10:1406-1414.
- [2] Beckmann, H., and T. Kadesch. 1991. The leucine zipper of TEF3 dictates helix-loop-helix dimerization specificity. *Genes Dev.* 5:1057-1066.
- [3] Benichou, S., M. Bomsel, M. Bodeus, H. Durand, M. Doute, F. Letourneur, J. Camonis, and R. Benarous. 1994. HIV-1 nef interacts with B-COP, a coat protein from non-clathrin-coated vesicles. submitted.
- [4] Berkhout, B., and K.-T. Jeang. 1989. trans Activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the trans-acting-responsive hairpin: a quantitative analysis. *J. Virol.* 63:5501-5504.
- [5] Berkhout, B., and K.-T. Jeang. 1992. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the Human Immunodeficiency virus type 1 long terminal repeat. *J. Virol.* 66:139-149.
- [6] Berkhout, B., R. H. Silverman, and K.-T. Jeang. 1989. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. *Cell* 59:273-282.
- [7] Braddock, M., R. Powell, A. D. Blanchard, A. J. Kingsman, and S. M. Kingsman. 1993. HIV-1 TAR RNA-binding proteins control TAT activation of translation in *Xenopus* oocytes. *FASEB J.* 7:214-222.
- [8] Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. 1986. Purification and biochemical characterization of the promoter specific transcription factor, Sp1. *Science* 234:47-52.
- [9] Buonaguro, L., F. M. Buonaguro, G. Giraldo, and B. Ensoli. 1994. The human immunodeficiency virus type 1 Tat protein transactivates tumor necrosis factor beta gene expression through a TAR-like structure. *J. Virol.* 68:2677-2682.
- [10] Burd, C. G., and G. Dreyfuss. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science* 265:615-621.
- [11] Chang, Y., D. J. Kenan, J. Keene, A. Gagnol, and K. T. Jeang. 1994. Direct interactions between the autoantigen La and the human immunodeficiency virus (HIV-1) leader RNA. *J. Virol.* 68:7008-7020.
- [12] Cooney, A. J., S. Y. Tsai, B. W. O'Malley, and M.-J. Tsai. 1991. Chicken ovalbumin upstream promoter transcription factor binds to a negative regulatory region in the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* 65:2853-2860.
- [13] Desai, K., P. M. Lowenstein, and M. Green. 1991. Isolation of a cellular protein that binds to the human immunodeficiency virus Tat protein and can potentiate transactivation of the viral promoter. *Proc. Natl. Acad. Sci. USA* 88:8875-8879.
- [14] Du, H., A. L. Roy, and R. G. Roeder. 1993. Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. *Embo J.* 12:501-511.
- [15] Dubiel, W., K. Ferrell, and M. Rechsteiner. 1994. Tat-binding protein 7 is a subunit of the 26S protease. *Biol. Chem. Hoppe-Seyler* 375:237-240.
- [16] Dynan, W. S., and R. Tjian. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* 32:669-680.
- [17] Ensoli, B., G. Barillari, S. Z. Salahuddin, R. C. Gallo, and F. Wong-Staal. 1990. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* 345:84-86.
- [18] Fankhauser, C., E. Izaurralde, Y. Adachi, P. Wingfield, and U. K. Laemmli. 1991. Specific

complex of human immunodeficiency virus type 1 Rev and nucleolar B23 proteins: dissociation by the Rev response element. *Mol. Cell. Biol.* 11:2567-2575.

- [19] Frankel, A., and C. Pabo. 1988. Cellular uptake of the Tat protein from human immunodeficiency virus. *Cell* 55:1189-1193.
- [20] Garcia, J. A., and R. B. Gaynor. 1994. Regulatory mechanism involved in the control of HIV-1 gene expression. *AIDS* 8:S3-S17.
- [21] Garcia, J. A., D. Harrich, E. Soultanakis, F. Wu, R. Mitsuyasu, and R. B. Gaynor. 1989. Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation. *Embo J.* 8:765-778.
- [22] Garcia, J. A., F. K. Wu, R. Mitsuyasu, and R. B. Gaynor. 1987. Interactions of cellular proteins involved in the transcriptional regulation of the human immunodeficiency virus. *Embo J.* 6:3761-3770.
- [23] Gatignol, A., C. Buckler, and K.-T. Jeang. 1993. Relatedness of an RNA binding motif in HIV-1 TAR RNA-binding protein TRBP to human P1/dsI kinase and *Drosophila* Staufen. *Mol. Cell. Biol.* 13:2193-2202.
- [24] Gatignol, A., A. Buckler-White, B. Berkhout, and K.-T. Jeang. 1991. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* 251:1597-1600.
- [25] Gatignol, A., A. Kumar, A. Rabson, and K.-T. Jeang. 1989. Identification of cellular proteins that bind to the human immunodeficiency virus type 1 trans-activation-responsive TAR element RNA. *Proc. Natl. Acad. Sci. USA* 86:7828-7832.
- [26] Gaynor, R., E. Soultanakis, M. Kuwabara, J. Garcia, and D. S. Sigman. 1989. Specific binding of a HeLa cell nuclear protein to RNA sequences in the human immunodeficiency virus transactivating region. *Proc. Natl. Acad. Sci. U.S.A.* 86:4858-4862.
- [27] Ghosh, S., A. M. Gifford, L. R. Fiviere, P. Tempst, G. P. Nolan, and D. Baltimore. 1990. Cloning for the p50 DNA-binding subunit of NF- κ B: homology to rel and dorsal. *Cell* 62:1019-1029.
- [28] Giese, K., J. Cox, and R. Gosschedl. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* 69:185-195.
- [29] Goyer, C., H. S. Lee, D. Malo, and N. Sonenberg. 1992. Isolation of a yeast gene encoding a protein homologous to the human Tat-binding protein, TBP-1. *DNA Cell Biol.* 13:579-585.
- [30] Greenblatt, J. 1991. Roles of TFIID in transcriptional initiation by RNA polymerase II. *Cell* 66:1067-1070.
- [31] Gregor, P. D., M. Sawadogo, and R. G. Roeder. 1990. The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev.* 4:1730-1740.
- [32] Han, X. M., A. Laras, M. P. Rouseville, A. Kumar, and P. R. Shank. 1992. Human immunodeficiency virus type 1 Tat-mediated trans activation correlates with the phosphorylation state of a cellular TAR RNA stem-binding factor. *J. Virol.* 66:4065-4072.
- [33] Harrich, D., J. Garcia, F. Wu, R. Mitsuyasu, J. Gonzalez, and R. B. Gaynor. 1989. Role of Sp1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. *J. Virol.* 63:2585-2591.
- [34] Hattori, K., P. Angel, M. LeBeau, and M. Karin. 1988. Structure and chromosomal localization of the functional intronless human JUN protooncogene. *Proc. Natl. Acad. Sci. USA* 85:9148-9152.
- [35] Hermann, C. H., and A. P. Rice. 1993. Specific interaction of the human immunodeficiency virus Tat protein with a cellular protein kinase. *Virology* 197:601-608.
- [36] Huang, L. M., and K.-T. Jeang. 1993. Increased spacing between Sp1 and TATAA renders HIV-1 replication defective: Implication for Tat function. *J. Virol.* 67:6937-6944.
- [37] Huang, L. M., A. Joshi, R. Willey, J. Orenstein, and K. T. Jeang. 1994. Human immunodeficiency viruses regulated by alternative trans-activators: genetic evidence for a novel non-transcriptional

- function of Tat in virion infectivity. *Embo J.* 13:2886-2896.
- [38] Jeang, K.-T., R. Chun, N. H. Lin, A. Gatignol, C. G. Glabe, and H. Fan. 1993. In vitro and in vivo binding of human immunodeficiency virus type 1 Tat protein and Sp1 transcription factor. *J. Virol.* 67:6224-6233.
 - [39] Jones, K., P. Luciw, and N. Duchange. 1988. Structural arrangements of transcription control domains within the 5' untranslated leader regions of HIV-1 and HIV-2 promoters. *Genes Dev.* 2:1101-1114.
 - [40] Jones, K. A., J. T. Kadonaga, P. A. Luciw, and R. Tjian. 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science* 232:755-759.
 - [41] Jones, K. A., and B. M. Peterlin. 1994. Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* 63:717-743.
 - [42] Jones, K. A., and R. Tjian. 1985. Sp1 binds to promoter sequences and activates herpes simplex virus "immediate-early" gene transcription in vitro. *Nature* 317:179-182.
 - [43] Kaczmarek, W., and S. A. Khan. 1993. Lupus autoantigen Ku protein binds HIV-1 TAR RNA in vitro. *Biochem. Biophys. Res. Communications* 196:935-942.
 - [44] Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* 51:1079-1090.
 - [45] Kamine, J., T. Subramanian, and G. Chinnadurai. 1991. Sp1-dependent activation of a synthetic promoter by human immunodeficiency virus type I Tat protein. *Proc. Natl. Acad. Sci. USA* 88:8510-8514.
 - [46] Kashanchi, F., G. Piras, M. F. Radonovich, J. F. Duvall, A. Fattaey, C.-M. Chiang, R. G. Roeder, and J. N. Brady. 1994. Direct interaction of human TFIIID with the HIV-1 transactivator Tat. *Nature* 367:295-299.
 - [47] Kato, H., M. Horikoshi, and R. G. Roeder. 1991. Repression of HIV-1 transcription by a cellular protein. *Science* 251:1476-1479.
 - [48] Kenan, D. J., C. C. Query, and J. D. Keene. 1991. RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.* 16:214-220.
 - [49] Kieran, M., V. Blank, F. Logeat, J. Vandekerckhove, F. Lottspeich, O. LeBail, M. B. Urban, P. Kourilsky, P. A. Baeuerle, and A. Israel. 1990. The DNA binding subunit of NF- κ B is identical to factor KBF-1 and homologous to the rel oncogene product. *Cell* 44:261-272.
 - [50] Kim, J., F. Gonzalez-Scarano, S. Zeichner, and J. Alwine. 1993. Replication of type 1 human immunodeficiency viruses containing linker substitution mutations in the -201 to -130 region of the long terminal repeat. *J. Virol.* 67:1658-1662.
 - [51] Krainer, A. R., A. Mayeda, D. Kozak, and G. Binns. 1991. Functional expression of cloned human splicing factor SF2: Homology to RNA-binding proteins, U1, 70k, and drosophila splicing factor. *Cell* 66:383-394.
 - [52] Lee, J.-S., K. M. Galvin, and Y. Shi. 1993. Evidence for physical interaction between the zinc-finger transcription factor YY1 and Sp1. *Proc. Natl. Acad. Sci. USA* 90:6145-6149.
 - [53] Lewin, B. 1990. Commitment and activation at PolII promoters: a tale of protein-protein interactions. *Cell* 61:1161-1164.
 - [54] Li, C., C. Lai, D. S. Sigman, and R. Gaynor. 1991. Cloning of a cellular factor, interleukin binding factor, that binds to NFAT-like motifs in the human immunodeficiency virus long terminal repeat. *Proc. Natl. Acad. Sci. USA* 88:7739-7743.
 - [55] Liang, P., and A. B. Pardee. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-971.
 - [56] Liou, H. C., G. P. Nolan, S. Ghosh, F. Fujita, and D. Baltimore. 1992. The NF- κ B p50 precursor, p105 contains an internal I κ B-like inhibitor that preferentially inhibits p50. *Embo J.* 11:3003-3009.
 - [57] Lu, Y., N. Touzjian, M. Stenzel, T. Dorfman, J. G. Sodroski, and W. A. Haseltine. 1990.

Identification of cis-acting repressive sequences within the negative regulatory element of human immunodeficiency virus type 1. *J. Virol.* 64:5226-5229.

- [58] Luban, J., K. L. Bossolt, E. K. Franke, K. V. Ganjam, and S. P. Goff. 1993. Human immunodeficiency virus type 1 gag protein binds to cyclophilins A and B. *Cell* 73:1067-1078.
- [59] Luo, Y., H. Yu, and B. M. Peterlin. 1994. Cellular protein modulates effects of human immunodeficiency virus type 1. *Rev. J. Virol.* 68:3850-3856.
- [60] Maekawa, T., T. Sudo, M. Kurimoto, and S. Ishii. 1991. USF-related transcription factor, HIV-TF1, stimulates transcription of human immunodeficiency virus-1. *Nucl. Acids. Res.* 19:4689-4694.
- [61] Marciniak, R. A., M. A. Garcia-Blanco, and P. A. Sharp. 1990. Identification and characterization of a HeLa nuclear protein that specifically binds to the trans-activation-response (TAR) element of human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* 87:3642-3646.
- [62] Margolis, D. M., M. Somasundaran, and M. R. Green. 1994. Human transcription factor YY1 represses human immunodeficiency virus type 1 transcription and virion production. *J. Virol.* 68:905-910.
- [63] Mattaj, I. W. 1993. RNA recognition: A family matter? *Cell* 73:837-840.
- [64] McCaffrey, P., C. Luo, T. Kerppola, J. Jain, T. Badalian, A. Ho, Burgeon, E., W. Lane, J. Lambert, T. Curran, G. Verdine, A. Rao, and P. G. Hogan. 1993. Isolation of the cyclosporin-sensitive T cell transcription factor NFATp. *Science* 262:750-754.
- [65] McCormack, S., D. Thomis, and C. Samuel. 1992. Mechanism of interferon action: identification of a RNA-binding domain within the N-terminal region of the human RNA dependent P1/eIF-2 alpha protein kinase. *Virology* 188:47-56.
- [66] Muesing, M., D. Smith, and D. Capon. 1987. Regulation of mRNA accumulation by human immunodeficiency virus trans-activator protein. *Cell* 48:691-701.
- [67] Nabel, G., and D. Baltimore. 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 326:711-713.
- [68] Nelbrock, P., P. Dillon, A. Perkins, and C. A. Rosen. 1990. A cDNA for a protein that interacts with the human immunodeficiency virus Tat transactivator. *Science* 248:1650-1653.
- [69] Nolan, G., S. Ghosh, H. C. Liou, P. Tempst, and D. Baltimore. 1991. DNA-binding and I κ B inhibition of the cloned p65 subunit of NF- κ B, a rel-related polypeptide. *Cell* 64:961-969.
- [70] Ohana, B., P. A. Moore, S. M. Ruben, C. D. Southgate, M. R. Green, and C. A. Rosen. 1993. The type 1 human immunodeficiency virus Tat binding protein is a transcriptional activator belonging to an additional family of evolutionarily conserved genes. *Proc. Natl. Acad. Sci. USA* 90:138-142.
- [71] Park, H., M. V. Davis, J. O. Langland, H. Chang, Y. Nam, J. Tartaglia, E. Paoletti, B. L. Jacobs, R. Kaufman, and S. Venkatesan. 1994. TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proc. Natl. Acad. Sci. USA* 91:4713-4717.
- [72] Pfeifer, K., B. Weiler, D. Ugarkovic, M. Bachmann, H. C. Schroder, and W. E. G. Muller. 1991. Evidence for a direct interaction of Rev protein with nuclear envelope mRNA-translocation system. *Eur. J. Biochem* 199:53-64.
- [73] Ross, E., A. Buckler-White, A. B. Rabson, G. Englund, and M. A. Martin. 1991. Contribution of NF- κ B and Sp1 binding motifs to the replicative capacity of human immunodeficiency virus type 1: distinct patterns of viral growth are determined by T-cell types. *J. Virol.* 65:4350-4358.
- [74] Rounseville, M. P., and A. Kumar. 1992. Binding of a host cell nuclear protein to the stem region of human immunodeficiency virus type 1 trans-activation-responsive RNA. *J. Virol.* 66:1688-1694.
- [75] Roy, S., M. Agy, A. Horvenessian, N. Sonenberg, and M. Katze. 1991. The integrity of the stem structure of human immunodeficiency virus type 1 Tat-responsive sequence RNA is required for interaction with the interferon-induced 68,000-Mr protein kinase. *J. Virol.* 65:632-640.

- [76] Ruhl, M., M. Himmelsbach, G. M. Bahu, F. Mammerschmid, B. Jaksche, H. Wolffe, H. Aschauer, G. K. Farrington, H. Probst, D. Bevec, and J. Hauber. 1993. Eukaryotic initiation factor 5A is a acellular target of the human immunodeficiency virus type 1 rev activation domain mediating trans-activation. *J. Cell Biol.* 123:1309-1320.
- [77] Scala, G., M. R. Ruocco, C. Ambrosino, M. Mallardo, V. Giordano Baldassarre, F., E. Dragonetti, I. Quinto, and S. Venuta. 1994. The expression of the interleukin 6 gene is induced by the human immunodeficiency virus 1 tat protein. *J. Exp. Med.* 179:961-971.
- [78] Seeler, J. S., C. Muchardt, A. Suessle, and R. Gaynor. 1994. Transcription factor PRDII-BF1 activates human immunodeficiency virus type 1 gene expression. *J. Virol.* 68:1002-1009.
- [79] Seto, E., Y. Shi, and T. Shenk. 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. *Nature* 354:241-245.
- [80] Shaw, D., and H. L. Ennis. 1993. Molecular cloning and developmental regulation of dicystostelium discoideum homologues of the human and yeast HIV1 tat-binding protein. *Biochem. Biophys. Res. Communications* 193:1291-1296.
- [81] Shaw, J., P. Utz, D. Durand, J. Tooole, E. Emuel, and G. R. Crabtree. 1988. Identification of a putative regulator of early T cell activation genes. *Science* 244:202-205.
- [82] Sheldon, M., R. Ratnasabapathy, and N. Hernandez. 1993. Characterization of the inducer of short transcripts, a human immunodeficiency virus type 1 transcriptional element that activates the synthesis of short transcripts. *Mol. Cell. Biol.* 13:1251-1263.
- [83] Sheline, C. T., L. H. Milocco, and K. A. Jones. 1991. Two distinct nuclear transcription factors recognize loop and bulge residues of the HIV-1 TAR RNA hairpin. *Genes Dev.* 5:2508-2520.
- [84] Shibuya, H., K. Irie, J. Ninomiya-Tsuji, M. Goehl, T. Taniguchi, and K. Matsumoto. 1992. New human gene encoding a positive modulator of HIV Tat-mediated transactivation. *Nature* 357:700-702.
- [85] Svitkin, Y. V., A. Pause, and N. Sonenberg. 1994. La autoantigen transactivates translation initiation mediated by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *J. Virol.* 68:7001-7007.
- [86] Swaffield, J., J. Bromberg, and S. A. Johnston. 1992. Alterations in a yeast protein resembling HIV Tat-binding protein relieve requirement for an acidic activation in GAL4. *Nature* 357:698-700.
- [87] Usheva, A., and T. Shenk. 1994. TATA-binding protein-independent initiation: YY1, TFIIB, and RNA polymerase II direct basal transcription on supercoiled template DNA. *Cell* 76:1115-1121.
- [88] Vaishnav, Y. N., M. Vaishnav, and F. Wong-Staal. 1991. Identification of a nuclear factor that specifically binds to the Rev-responsive element (RRE) of human immunodeficiency virus type 1. *New Biologist* 3:142-150.
- [89] vanStraaten, F., R. Mueller, T. Curran, C. P. vanBeveren, and I. M. Verma. 1983. Complete nucleotide sequence of a human c-onc gene: deduced amino acid sequence of the human c-fos protein. *Proc. Natl. Acad. Sci. USA* 80:3183-3187.
- [90] Vogel, B., S.-J. Lee, A. Hildebrand, W. Craig, M. Pierschbacher, F. Wong-Staal, and E. Ruoslahti. 1993. A novel integrin specificity exemplified by binding of the $\alpha v \beta 5$ integrin to the basic domain of the HIV tat protein and vitronectin. *J. Cell Biol.* 121:461-468.
- [91] Waterman, M. L., and K. A. Jones. 1990. Purification of TCF-1a, a T-cell-specific transcription factor that activates the T-cell receptor α gene enhancer in a context-dependent manner. *New Biol.* 2:621-636.
- [92] Week, B., K. Desai, P. Loewenstein, M. Klotman, P. Klotman, M. Green, and H. K. Kleinman. 1993. Identification of a novel cell attachment domain in the HIV-1 Tat protein and its 90-kDa cell surface binding protein. *J. Biol. Chem.* 268:5279-5284.
- [93] Westendorp, M. O., M. Li-Weber, R. W. Frank, and P. H. Krammer. 1994. Human immunodeficiency virus type 1 Tat upregulates interleukin-2 secretion in activated T cells. *J. Virol.*

68:4177-4185.

- [94] Wu, F., J. Garcia, D. Sigman, and R. Gaynor. 1991. tat regulates binding of the human immunodeficiency virus trans-activating region RNA loop-binding protein TRP-185. *Genes Dev.* 5:2128-2140.
- [95] Yaseen, N. R., B. Maizel, F. Wang, and S. Sharma. 1993. Comparative analysis of NFAT (nuclear factor of activated T cells) complex in human T and B lymphocytes. *J. Biol. Chem.* 268:14285-14293.
- [96] Yoon, J.-B., G. Li, and R. Roeder. 1994. Characterization of a family of related cellular transcription factors which can modulate human immunodeficiency virus type 1 transcription in vitro. *Mol. Cell. Biol.* 14:1776-1785.
- [97] Zimmermann, K., M. Dobrovnik, C. Ballaun, D. Bevec, J. Hauber, and E. Bohnlein. 1991. trans-Activation of the HIV-1 LTR by the HIV-1 Tat and HTLV-I Tax proteins is mediated by different cis-acting sequences. *Virology* 182:874-878.

LOCUS CEIF5A 1792 bp ss-mRNA VRT 06-APR-1994
 DEFINITION Chicken initiation factor 5A cDNA sequence, complete cds.
 ACCESSION M99499
 SOURCE Gallus gallus embryo whole tissue mRNA.
 REFERENCE 1 (bases 1 to 1792)
 AUTHORS Rinaudo, M.S., Joe, Y.A. and Park, M.H.
 TITLE Cloning and sequencing of a chick embryo cDNA encoding the 20-kDa
 hypusine-containing protein, eIF-5A
 JOURNAL Gene 137, 303-307 (1993)
 COMMENT Translation factor eIF-5A has been shown to bind Rev (Ruhl et al.
 J. Cell. Biol. 123: 1309-1320 1993). At the functional level
 Rev transcription could be modulated by associations with eIF-5A.
 NCBI gi: 289799

BASE COUNT 542 a 347 c 391 g 511 t 1 others
 ORIGIN

```

1 gaattccgga tgcctaaag caagtgcac tctctgcaca cgtctcgccc cgcagccgcc
61 gccATGgccg acgagttgga cttcaagacg ggcgacgccg ggcctctgto caccctatccc
IF cds ->
121 atgcagtgct cggcgctgct taagaacggc ttctgtgtac tgaagggggc cccgtgtaag
181 atcgtggaga tgtccacctc caagacggga aagcacggcc acgccaaagt ccacctgggt
241 ggtattgata tcttcaacgg gaaaaaatac gaagatattt gcccatcaac tctaacatg
301 gatgtcccaa atataaagag gaacgactac cagctgatag ggattcagga tgggtacctc
361 tccctgctga cagaaagtgg cgaagtctgt gaggatctaa agctgccaga aggtgatctg
421 ggaanaagaaa tagaaggaaa gttcaatgcc aatgaagatg tgcagatata cgtcataagt
481 gcaatgaatg aagaatgtgc tgtagccata aagccttgca aaTAAagagg atcgccaggg
                                     <- IF cds end
541 acgggcaact gctcaggtct ggataaacca cagatcttta aatttggttc tgattgtcac
601 caaagctctg gccctcacaa gcaacctcat ttcttttttg aattgttaca aagcagtgtt
661 ggattctgga ttagtggtgc aggctaactg gcagttttca aaatcactga gatttcaatt
721 ccttaattgt aactatttta acattcctgt gggtttcage tatggatctg agaaaccgtt
781 cagggtgtac tagtggtatc attttttatt tgccttgagc aaacagtggt tgtctgtatg
841 aacagacaaa atacagtatt caggggcttt tagataagct agtaggtatc taggattata
901 aagtgaacct gagcacaaat agtatttttc ttgacatgtt tagatagaac tgacaataaa
961 agtagctttt tttttttttt aaagcttaaa tattttgtgg attgggactc ttgcagaact
1021 gtaggtacca atcacaattt ttagaccaag aaactcaagt gatcaaaaat gccagctggc
1081 ttgagcaagc cccagtgagg atgtgcaact taactgtatt acctctgtat tcttttttgc
1141 caacttggtg gcttattgta acgttaaaat ggtgtcaaag cctacctgtg gtttaggcag
1201 gcagttgagt tttagtaga tgagacatag atwggtaag gtgagaatgg atgctaaagc
1261 tatagacttt cactcgaggg ctttgtcaga cttaaaggaa aaaaaaaaaa agaaaaaaaa
1321 aaaacacagt tcaattgctt tctttgttta tatttatctt aaactataaa aggggtgaca
1381 actccgctgg aatctgttgc agtgcttoca gcttaggtt tcaatcttca gtttcagacc
1441 atatttttgt agcacaatga ccacgtgggt tttcaaaaag tgggacaaat tgtagcattg
1501 agcctaaaact aaatgcatgt gtggattgtt gtgttgtttt tttttttttt atagaacaat
1561 tgtctttttc aaatacagac ttgacaaaca attctgaaca accatggctt cactataaaa
1621 ataactagag agttgaaaaa gttagctcct atccacagga gtttggagtc attgaaacac
1681 acatgcacgc ctctaactat ttagtagctt aatggtcaaa cacaaagctt tacagagcta
1741 gtcatggaag gatacctagc ttactcttgt ttatagcact caactagaat tc

```

HIVEP1

LOCUS HIVEP1 1035 bp ss-RNA PRI 28-SEP-1994
 DEFINITION Human immunodeficiency virus enhancer-binding protein (HIV-EP1) sequence, partial cds.
 ACCESSION J05011
 SOURCE Human B-cell lambda gt11 library.
 REFERENCE 1 (bases 1 to 1035)
 AUTHORS Maekawa,T., Sakura,H., Sudo,T. and Ishii,S.
 TITLE Putative metal finger structure of the human immunodeficiency virus type 1 enhancer binding protein HIV-EP1
 JOURNAL J. Biol. Chem. 264, 14591-14593 (1989)
 COMMENT HIV-EP1, also referred to as MBP-1 or PRDII-BF1, is a cellular transcription factor which activates HIV-1 gene expression. The protein contains two zinc finger domains, each of which can bind to either of the two NF-kappa B binding motifs in the HIV-1 enhancer or to the TAR transactivation response element. Gaynor et al. (Genomics 9: 758-61, 1991) have mapped the gene which encodes HIV-EP1 to human chromosome 6p22.3-p24.
 NCBI gi: 326842
 BASE COUNT 328 a 247 c 236 g 224 t
 ORIGIN

```

1 agtgagcaag ataaagaaaa ttcccttaatc aaaagtgaac caagaagaat taaaatattt
HIV-EP1 ->
61 gatggaggat ataagtcasa tgaagagtat gtatatgtcc gaggcagggg aagaggaaaa
121 tacatttgtg aagaatgtgg aatacgttgt aagaaacctc gcatgttaa gaaacacata
181 cgaaccata cagatgtccg cccctaccac tgcacttact gtaacttctc ctttaagact
241 aaaggaaatc tgacaaaaca catgaagtcc aaggcacata gcaagaaatg tgtggattta
301 ggcgtctcag taggtttaat agatgaacag gatcacagaag aatcagatga aaaaacagaga
361 ttcagttatg agcgatctgg atatgatctt gaagaatctg atggcccaga tgaggatgac
421 aatgaaaatg aagacgatga tgaggacagc caggctgaat cagtctctgc agccacaccc
481 tcagtcacag ctagcccgca gcaccttcca tctagaagta gccttcagga cctgtgagt
541 actgacgagg atgtcaggat caccgattgc ttttctgggg tacacacgga cccaatggac
601 gttctgcca gggegtgct caccagaatg actgtcctga gcacagcaca gtctgactac
661 aataggaaga cactctctcc ggggaaggcc aggcagcata ctgcgagaga tgaaaacgac
721 acaattccgt ctgtagcac ttccaggtcc cegtgtctc agatgtctgt ggactaccct
781 gagtcagaag aaattctgag aagttctatg gcaggaaaag ctgttgctat aacacagagc
841 ccacatctg taagaattcc tctgtctgca gctgagcaca gccccagac agcagcgggg
901 atgcttctg tggectcacc acatcttgac cctcaagaac agaagcagca aataactcta
961 cagccgactc caggcttgcc ttctcccac actcatttgt ttagccacct tctttgcat
1021 tcccagcagg aattc
HIV-EP1 ->

```

LOCUS HSCOUPTF 1513 bp RNA PRI 29-MAR-1991
 DEFINITION Human chicken ovalbumin upstream promoter transcription factor
 (COUP-TF) sequence, partial cds.
 ACCESSION X16155 X58241
 KEYWORDS DNA-binding protein; steroid receptor; transcription factor.
 SOURCE Human mRNA from cell line Hela s3.
 REFERENCE 1 (bases 1 to 1513)
 AUTHORS Wang, H., Tsai, Y., Cook, G., Beattie, G., Tsai, M.J. and O'Malley, W.
 TITLE COUP transcription factor is a member of the steroid receptor
 superfamily
 JOURNAL Nature 340, 163-166 (1989)
 COMMENT The promoter-enhancers of HIV are contained in the U3 region of
 the viral long terminal repeat. Binding sites for COUP are
 present from -324 to -357 within the U3 region of the LTR of
 HIV-1. A negative role for this element is suggested by mutational
 analysis (Garcia et al. AIDS 8: S3-S17 1994).
 NCBI gi: 30139
 BASE COUNT 319 a 520 c 451 g 223 t
 ORIGIN

```

1 agcagctggc gagatccgca ggaagacgtg gccgggggca accccggcgg ccccaacccc
COUP-TF ->
61 gcagcgcagg cggcccgagg cggcggcgcc ggccggcgcc agcagcagca gcaggcgggc
121 tggggcggcc cgcatacgcc gcagaccccg ggccagcccg gagcgcccg ccccccgcc
181 acggcggggg acaaggccca gggcccgccc ggttcgggac agagccagca gcacatcgag
241 tgcgtggtgt gcggggacaa gtccagcgcc aagcactacg gccaatcac ctgcgagggc
301 tgcaaaagtt tcttcaagag gacgctccgc aggaacttaa cttacacatg ccgtgccaac
361 aggaactgtc ccacgcacca gcaccacccc aaccagtgcc aatactgcc cctcaagaag
421 tgccctcaag tgggcatgag gcgggaagcg gttcagcgag gaagaatgcc tccaaccag
481 cccatccag gccagtaagg actcaccaac ggggaacccc tcaacggcca ctgctacctg
541 tcgggtata tctcgtgtgt gctgcgcgcc gagccctacc ccagtcgag ctacggcagc
601 cagtgcattg agcccaacaa cattatgggc atcgagaaca tctgcgagct ggccgcgcgc
661 ctgctcttca gcgcgctcga gtgggcccgc aacatccctt tcttcccgga tctgcagatc
721 accgaccagg tgcctctgt acgctcacc tggagcgagc tgttcgtgt caacgcggcc
781 cagtgtctta tgcgctgca cgtggcgccg ttgttgccg ccgcccgcct gcctgctctg
841 cccatgtctg ccgaccgct cgtggccttc atggaccaca tccgcatctt ccaggagcag
901 gtggagaagc tcaggcgct acacgtcgac tcagccgagt acagctgctt caaagccatc
961 gtgtgttca cgtcagacgc ctgtggcctg tcggatgcgg cccacatcga gagcctgcag
1021 gagaagtgc agtgcgcact ggaggagtac gtgaggagcc agtaccacca ccagcccgagc
1081 cgttttggca aactgtgtgt ggcactgcc tcgctgcgca ccgtgtcttc ctccgtcatc
1141 gagcagctct tcttcgtcgg ttgtgtagg aaacccccca tcgaaactct catccgcgat
1201 atgttactgt ctgggagcag cttcaactgg ccttacatgt ccacccagtg ctccTAGacc
                                     <-COUP-TF
1261 ttggcgctt cccactgcc ccgtccctt agagactcag aggaaccacc tgggccaagg
1321 actccaaagg cgcggggaca ccgggaagtg cagcgggcca ggcaggctgg gtgggagggg
1381 ggaggggcga gacaggagca gcccacccag cagaaataca atccgagcta caaagcatgg
1441 gaaaaagaga ctcttttagg atcagatctg tggcacgtt ggcgaggaaa acaaaaaaaa
1501 caaaaaaaaaa cgg

```

HUMCOUPII

LOCUS HUMCOUPII 2268 bp ss-mRNA PRI 14-SEP-1992
 DEFINITION Homo sapiens chick ovalbumin upstream promoter transcription factor II (COUP-TFII) cDNA sequence, 5' end.
 ACCESSION M62760
 SOURCE Human female mRNA from HeLa S3 cell line.
 REFERENCE 1 (bases 1 to 2268)
 AUTHORS Wang,L.-H., Ing,N.H., Tsai,S.Y., O'Malley,B.W. and Tsai,M.-J.
 TITLE The COUP-TF compose a family of functionally related transcription factors
 JOURNAL Gene Expression 1, 207-216 (1991)
 COMMENT The promoter-enhancers of HIV are contained in the U3 region of the viral long terminal repeat. Binding sites for COUP are present from -324 to -357 within the U3 region of the LTR of HIV-1. A negative role for this element is suggested by mutational analysis (Garcia et al. AIDS 8: S3-S17 1994).
 BASE COUNT 410 a 833 c 656 g 369 t
 ORIGIN

```

1  cggggccgtac  tgcctttttt  cccctctttc  attctttctc  tcgtcttttt  tctccccctt
61  ctgcgcacga  aggatgtgct  tctagtggtt  gatctgceet  cctctctctc  ttttatacat
121  tctccccctc  gccggcgagt  tgactctttc  cctatgtgtg  tgaggcggcg  gccgcagcag
181  cagcagcagc  ggctccggcg  gccgcagcag  cggcagcagc  gaattcagcg  gccgcggcgg
241  cgctagacgc  agcggctccg  ggcccgcacc  gccggcttcg  gccgcggctc  cggcggcagc
301  gccggcccg  gccggccgca  gggaacggcg  agcggcctcc  acccagcact  gccggcggcg
361  gccggccgg  agagcgaggc  gccgcgcgga  cggccggggc  aggcggcggc  gccggcggcc
421  cagcggcagg  acgacgcgcg  gcagcgcgcg  acgcggaaca  ctttcattgt  gattcccccg
481  gacccgggca  gcctctcggc  cactccggcg  cggcggcctc  cggcccggcc  tgcctggctc
541  cctgggctgc  ccgcaccccg  cgcctcagat  ctctagttcc  tcttgatttc  gatggctttc
601  ctgaatggct  gactgtgggc  tgcctgggac  ttggccccc  gacagtgcgc  tctctctctc
661  ctctacctcc  tcttccacca  ccacctctc  ttctctctcc  tctctctcgc  ccaactctc
721  ggctgcacac  cagctctaag  agcgagagtg  aacgagagag  ggaggagag  agtgagagcg
781  agcgagatct  ttggagagat  ttttttttt  gcctctact  tctgtcttga  agccagacaa
841  tgcacttcag  ctctccctcc  cctccctctt  tctccacgtt  ctgctccccc  tgcctctctt
901  gtcccccttc  cctccctctc  cggcgggaa  ccccccgaaa  ccaacaaagg  tgagccgaga
961  gaaacaaaca  aaacaaacac  accggggccg  acaagccatc  gacaaaactt  tgcaaaagca
1021  aaaacaaaaa  aggaaaaaact  aaccaacctc  aaccaaccag  ccccgagacc  acccggggcg
1081  cctctccgcg  cctctcttga  cctctgcaca  cacaacaggc  gccgcgcgg  agcccgagac
1141  cgggggagtc  gccgcgcgcc  cgcgcgcgcc  cgcagccagg  ggagcaggaa  gtccggacgc
1201  agccccata  gatATGgcaa  tggtagtcag  cagctggcgc  gacccccagg  acgaggtgcc
COUP-TFII cds start ->
1261  cggtcacag  gccagccagg  cctcgccagg  gccgcgcgtg  cccggccccc  cggccggcgc
1321  cccgcacag  ccacagacgc  cgggccaagg  gggcccgacc  agcacgcag  cccagacggc
1381  gcccggtgg  caggcgcgcc  ctggcgccgc  gggtagcgac  aagcagcag  agcagcaaca
1441  catcgagtgc  gtgggtgtcg  gagacaagtc  gagcggcaag  cactacggcc  agttcacgtg
1501  cgagggtgc  aagagcttct  tcaagcgcag  cgtgcggagg  aacctgagct  acacgtgccg
1561  cgccaaccgg  aactgtccca  tcgaccagca  ccacgcgaac  cagtgcagc  actgcgcctt
1621  caaaaagtgc  ctcaaagtgg  gcattgagag  ggaagcggtg  cagaggggca  ggtatgcgcc
1681  gaccacggcg  acccaggggc  agttcgcgct  gaccaacggg  gatccccca  actgcaactc
1741  gtacctgtcc  ggatatattt  cctgtctgtt  gcgcggcgag  ccttatccca  cgtcgcgctt
1801  cggcagccaa  tgcattgcag  ccaacaacat  catgggtatc  gagaacattt  gccgaactgg
1861  cgcgaggatg  ctcttcagcg  ccgtcgagtg  ggcccggaac  atccccctt  tccccgacct
1921  gcagatcacg  gaccaggtgg  cctgtcttcg  cctcacctgg  agcgagctgt  ttgtgttgaa
1981  tgccggcgag  tgcctccatg  ccttcacagt  cggcccgctc  ctggcgcgcg  cggccctgca
2041  tgcctcgccc  atgtccgcgc  accgggtggt  cgcctttatg  gacccataac  ggatcttcca
2101  agagcaagtg  gagaagctca  aggcgctgca  cgttgactca  gccaggtaca  gctgcctcaa
2161  gcccatagtc  ctgttcacct  cagatgcttg  tggctctctt  gatgtagccc  atgtggaaag
2221  cttgcaggaa  aagtctcagt  gtgcttttga  agaatacgtt  aggagcca
COUP-TFII ->

```


LOCUS HUMCYP 723 bp RNA PRI 07-JUN-1987
 DEFINITION Human T-cell cyclophilin cDNA sequence.
 ACCESSION Y00052
 KEYWORDS cyclophilin; cyclosporin A-binding protein.
 SOURCE Human mRNA from Jurkat cell library.
 REFERENCE 1 (bases 1 to 723)
 AUTHORS Haendler,B., Hofer-Warbinek,R. and Hofer,E.
 TITLE Complementary DNA for human T-cell cyclophilin
 JOURNAL EMBO J. 6, 947-950 (1987)
 COMMENT The cyclophilins are cellular polypeptides originally characterized
 for their specific binding to cyclosporin A. Subsequent studies
 have shown that both cyclophilin A and cyclophilin B bind to
 p55 gag and to the p24 capsid proteins (Luban et al. Cell 73:
 1067-78 ,1993).
 NCBI gi: 30308
 BASE COUNT 179 a 165 c 176 g 203 t
 ORIGIN
 1 gtgtactatt agccATGgtc aacccccacg tgtttcttga cattgcogtc gacggogagc
 cyclophilin cds start ->
 61 ccttgggccc cgtctccttt gagctgtttg cagacaaggt cccaaagaca gcagaaaatt
 121 ttctgtctct gagcactgga gagaaaggat ttggttataa gggttcctgc ttccacagaa
 181 ttattccagg gtttatgtgt cagggtggtg acttcacacg ccataatggc actggtggca
 241 agtccatcta tggggagaaa tttagaatg agaacttcat cctaaagcat acgggtcctg
 301 gcattctgtc catggcaaat gctggaccca acacaaatgg ttcccgattt ttcatctgca
 361 ctgccaagac tgagtgggtg gatggcaagc atgtggtgtt tggcaaagtg aaagaaggca
 421 tgaatattgt ggaggccatg gagcgctttg ggtccaggaa tggcaagacc agcaagaaga
 481 tcaccattgc tgactgtgga caactcgaaT AAgtttgact tgtgttttat cttaccacc
 <- cyclophilin cds end
 541 agatcattcc ttctgtagct caggagagca cccctccacc ccatttgcgc gcagtatcct
 601 agaattcttg tgctctcgct gcagttccct ttgggttcca tgttttcctt gttccctccc
 661 atgcctagct ggattgcaga gttaagttaa tgattatgaa ataaaaacta aataacaatt
 721 gtc

HUMCYPB

LOCUS HUMCYPB 851 bp ss-mRNA PRI 27-MAR-1991
 DEFINITION Human cyclophilin B cDNA sequence, complete cds.
 ACCESSION M60857
 KEYWORDS cyclophilin B; cyclosporin A-binding protein;
 peptidyl-prolyl isomerase.
 SOURCE Human mRNA.
 REFERENCE 1 (bases 1 to 851)
 AUTHORS Price,E.R., Zydowsky,L.D., Jin,M., Hunter,C.H., McKeon,F.D. and
 Walsh,C.T.
 TITLE Human cyclophilin B: A second cyclophilin gene encodes a peptidyl
 isomerase with a signal sequence
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 1903-1907 (1991)
 COMMENT The cyclophilins are cellular polypeptides originally characterized
 for their specific binding to cyclosporin A. Subsequent studies
 have shown that both cyclophilin A and cyclophilin B bind to
 p55 gag and to the p24 capsid proteins (Luban et al. Cell 73:
 1067-78 ,1993).
 NCBI gi: 181334
 BASE COUNT 222 a 210 c 240 g 179 t
 ORIGIN
 1 cggaacgca acATGaaggt gtccttgcg gcggccctca tcgaggggtc cgtcttcttc
 cyclophilin B cds ->
 61 ctgctgctgc cgggaccttc tgcggccgat gagaagaaga aggggcccaa agtcaccgtc
 mature peptide ->
 121 aaggtgtatt ttgacctacg aattggagat gaagatgtag gcggggtgat ctttgggtctc
 181 ttoggaaaga ctgttccaaa aacagtggat aattttgttg ccttagctac aggagagaaa
 241 ggatttggct acaaaaacag caaattccat cgtgtaatca aggacttcat gatccagggc
 301 ggagacttca ccaggggaga tggcacagga ggaaagagca tctacggtga gcgcttcccc
 361 gatgagaact tcaaaactgaa gcactacggg cctggctggg tcagcatggc caacgcaggc
 421 aaagacacca acggctccca gttcttcac acgacagtca agacagcctg gctagatggc
 481 aagcatgttg tgtttggcaa agttctagag ggcattggag tggcgaggaa ggtggagagc
 541 accaagacag acagccggga taaacccctg aaggatgtga tcatgcaga ctgcccgaag
 601 atcgaggttg agaagccctt tgccatgcc aaggagTAGg gcacagggac atctttcttt
 <- mature peptide
 <- cyclophilin B cds end
 661 gaggtagcgt ctgtgcaggc cctgtagtec gccacagggc tctgagctgc actggccccg
 721 gtgctggcat ctgggtggagc ggacccactc ccttcacatt ccacagggcc atggactcac
 781 ttttgtaaca aactcctacc aacactgacc aataaaaaaa aatgtggggtt tttttttttt
 841 ttaataaaaa a

LOCUS HUMHIP116 3418 bp ss-mRNA PRI 22-AUG-1994
 DEFINITION Human ATPase, DNA-binding protein (HIP116) cDNA sequence.
 ACCESSION L34673
 KEYWORDS ATPase; DNA-binding protein; SNF2/SWI2-related protein;
 transcription factor.
 SOURCE Homo sapiens mRNA.
 REFERENCE 1 (bases 1 to 3418)
 AUTHORS Sheridan,P.L., Schorpp,M., Voz,M.L. and Jones,K.A.
 TITLE Cloning of a SNF2/SWI2-related protein that binds specifically to
 the HIV-1 promoter and to the SPH motifs of the SV40 enhancer
 JOURNAL Unpublished (1994)
 COMMENT HIP 116 is a newly cloned cDNA whose protein binds to the
 TATA/initiator of the HIV-1 promoter (K.A. Jones, personal
 communication).
 NCBI gi: 531195

BASE COUNT 1135 a 602 c 744 g 937 t
 ORIGIN

```

1 attccccggg tctgaactgga ctccggggcga cttacctttc agtcgtgcgc tctgatccg
61 gcgctcgga tttgtccccg gcttcagggc tgcggggcct ggaaggagcc gtatcgagcc
121 ggctcgaaaa cgatccaggg gagccgaggg gctcctcttg tcatccact cagcgccATG
                                     HIP116 cds start ->
181 tcttggtatg tcaagaggga tccagtttgg aagtacttgc agactgtcca gtatggagtt
241 catggaaatt ttccacgcct ctcatatcca actttctttc cactgtttga attccaagat
301 gttatccctc cagatgactt tctaaotagt gatgaagaag tagattccgt tttatttggg
361 agttttgagag gtcattgtgt tggactacgc tattacacgg gactagttaa taataatgaa
421 atggttgcat tacaacgaga tccataatac ccttatgata agaattgcaat taaagtaaac
481 aatgtgaatg gaaatcaagt tggccattta aagaaagagc ttgcaggtgc ttggccctat
541 atcatggaca acaaatgggc acaaatggaa ggggtagttc cttttggtgc aaacaatgct
601 ttaccatgc ctctgcatat gacttttttg ggaaaagaag aaaaatagaaa agcgggttca
661 gatcagttga agaaccatgg atttaaatgg ggtcctgcac caaaaacttt aggattcaat
721 ttggaagtg gttggggctc tgggaagagc ggaccaagct atagtatgcc agtgcattgt
781 gcagtacaga tgacaactga acagctttaa acagaatttg acaaatgtt tgaagattta
841 aaagaagatg ataaaacca tgaattggaa ccagctgagg ctattgaaac accactgctt
901 ccacatcaaa aacaagctct agcttggatg gtgtcacggg aaaatagcaa agaacttcca
961 ccattctggg aacagcgaaa tgacttatac tataacacaa taacaaattt ttctgagaag
1021 gacggaccag aaatgtcca tggaggaatt ttagctgatg atatgggttt gggtaaaact
1081 cttacagcca ttgcagtaat ccttaccac ttccatgatg gcagacctct tctatttcaa
1141 agagttaaaa agaattctat gaagaaggaa tataatgtta acgatgactc tatgaaactt
1201 ggaggaaaca ataccagtga aaaggcagat ggactaagca aagacgcata tagatttagt
1261 gaacaacca gtatttcaga tatcaaggag aagagtaagt ttgcattgtc agaattgtct
1321 acgtcccgcc ccaaaagaag aaaaactgct gtccagtaca tagaaagcag tgattcagag
1381 gaaattgaaa caagtgaatt gccgcagaaa atgaaaggca aactgaaaaa tgtacagtct
1441 gaaactaag gcagggcgaa agcaggatct tctaaggtta tagaagatgt ggcatttgca
1501 tgtgcattaa ctcatctgt tctacaaca aaaaagaaaa tgttgaaaaa gggagcttgt
1561 gcagtggagg ggtcaaagaa aactgatgtt gaggagagac caagaacaac actgatcatc
1621 tgtccgcttt ctgtgttaag caactggatt gaccagtttg gacaacatat aaaatcagat
1681 gtacacttga atttttatgt ttattatggt cctgatcgta tttagaacc ggcttactt
1741 tcaaaacagg atattgtttt gactacgtat aatattttaa ctcatgacta tggaaactaa
1801 ggagtagtgc cattacatag cataagggtg ctaagagtga tctggatga aggaactgcc
1861 atacgaaatc caaatgctca gcagacaaa gctgtacttg acttagaatc agaaagaaga
1921 tgggttttga caggtaactc aatccagaat tctttaagg acttggtgtc tcttctttcc
1981 tttttaaaac ttaaaccaat tattgataga gaatgggtgg atagaacaat acagcgtcct
2041 gtcacaatgg gagatgaagg aggacttagg cgtttacagt ccttaattaa aaatattaca
2101 cttagaagaa caaagacaag caaaattaaa ggaaaacctg ttttgaggtt accagacgt
2161 aaagtattta ttcagcacat tacactttca gatgaagaga gaaagattta tcagtctgtg
2221 aaaaatgaag gcagagccac tatttgaagg tatttttaat aagggactgt cctggcacat
2281 tatgcagatg tctgtggtct tttgottaga ctggggcaaa tttgttgcca tacttacctt
2341 cttacaaatg cagtgtcttc caatggcccc tcaggaaatg atacacctga agaactgaga
2401 aagaagttaa taagggaagat gaagttaatt ctgagctcag gttcagatga ggaatgtgca

```

HUMHIP116

```

2461 atttgccctgg attctttaa acgttcctgtg ataacacatt gtgcacatgt attttgtaaa
2521 cctgtatatt gccaaagtcac tcagaatgag cagccacatg ctasatgcc tttatgcaga
2581 aatgatatac atgaagataa tttattagaa tgtcctccag aagaattagc acgtgacagt
2641 gagaaaaagt ctgatatgga atggacatcc agttcaaaga ttaatgcgct aatgcacgca
2701 ttgactgact taagaaagaa gaatcccaac ataaaaagtt tggttgtttc tcagtttaca
2761 acattcctgt ctttaataga aataccactt aaagcctctg gatctgtgtt tactcgtttg
2821 gatggttcca tggcccaaaa gaaaagagtt gaatcaattc agtgttttca aaacactgaa
2881 gcaggatctc caactataat gcttctgtcc ttaaaagcag gtggagttgg tttgaatctg
2941 tctgcagctt ctcgagtgtt ttaaatggat ccagcctgga atcctgctgc tgaagatcag
3001 tgctttgaca gatgccatag acttggtcag aagcaagaag ttatcatcac aaaattcatt
3061 gtaaaaggact ctgttgaaga aaatatgctg aaaatacaaa acaaaaagag agaacttgc
3121 gcaggagcct ttggaaactaa aaaaocaaat gctgacgaaa tgaacaagc caaattaat
3181 gaaatcagaa catlaattga cttataAAtt gtgggatttt agtaagaaga ctactatatg
      <- HIP116 cds end
3241 tgagaggcgt gatatctgga tggaaagttg gctggatgat ctccaaagtc gtttcaactc
3301 ttaaagacat cttaatcctg aatgtaaaaca attgttatgt gtttagaate agaatttgat
3361 tttgaacttg agtaattcat ccttacagct atctgtagaa ttagtcattt tttttctt

```

LOCUS HSIL02 801 bp RNA PRI 09-MAR-1993
 DEFINITION Human interleukin-2 (IL-2) cDNA sequence.
 ACCESSION V00564
 KEYWORDS interleukin; signal peptide; lymphocyte regulatory molecule.
 SOURCE Human mRNA.
 REFERENCE 1 (bases 1 to 801)
 AUTHORS Taniguchi,T., Matsui,H., Fujita,T., Takaoka,C., Kashima,N., Yoshimoto,R. and Hamuro,J.
 TITLE Structure and expression of a cloned cDNA for human interleukin-2
 JOURNAL Nature 302, 305-310 (1983)
 REFERENCE 2 (bases 1 to 801)
 AUTHORS Devos,R., Plaetinck,G., Cheroutre,H., Simons,G., Degraeve,W., Tavernier,J., Remaut,E. and Fiers,W.
 TITLE Molecular cloning of human interleukin 2 cDNA and its expression in E. coli
 JOURNAL Nucleic Acids Res. 11, 4307-4323 (1983)
 COMMENT IL-2 is a cytokine which has been shown to induce HIV gene expression (Matsuyama et al. AIDS 5: 1405-1417, 1991). It may mediate this affect by altering the activity of cellular transcription factors (Garcia et al. AIDS 8: S3-S17, 1994).
 NCBI gi: 33780
 BASE COUNT 282 a 147 c 114 g 258 t
 ORIGIN

```

1 atcactctct ttaatcacta ctacacagtaa cctcaactcc tgccacaATG tacaggatgc
      IL-2 cds start ->
61 aactcctgtc ttgcattgca ctaagtcttg cacttgtcac aaacagtgca cctacttcaa
121 gttctacaaa gaaaacacag ctacaactgg agcatttact gctggattta cagatgattt
181 tgaatggaat taataattac aagaatccca aactcaccag gatgctcaca ttttaagttt
241 acatgcccaa gaaggccaca gaactgaaac atcttcagtg tctagaagaa gaactcaaac
301 ctctggagga agtgctaaat ttagctcaaa gcaaaaactt tcacttaaga cccagggact
361 taatcagcaa tatcaacgta atagttcttg aactaaaggg atctgaaaca acattcatgt
421 gtgaatatgc tgatgagaca gcaaccattg tagaatttct gaacagatgg attacotttt
481 gtcaaagcat catctcaaca cttaactTGAa aattaagtgc ttcccactta aaacatatca
      <- IL-2 cds end
541 ggcctttctat ttatttkaat atttaaattt tatatttatt gttgaatgta tgggtttgcta
601 cctattgtaa ctattattct taatttkaaa actataaata tggatctttt atgattcttt
661 ttgtaagccc taggggctct aaaatggttt cacttattta tcccaaaate tttattatta
721 tgttgaatgt taaatatagt atctatgtag attggttagt aaaactattt aataaatttg
781 ataaatataa aaaaaaanaa c

```

HUMIL6

LOCUS HUMIL6 1125 bp ss-mRNA PRI 11-JUN-1993
 DEFINITION Human hybridoma growth factor (interleukin 6) cDNA sequence, complete cds, clone HGF15.
 ACCESSION M18403
 KEYWORDS hybridoma growth factor; interferon-beta-2; interleukin 6.
 SOURCE Human monocyte mRNA.
 REFERENCE 1 (bases 1 to 1125)
 AUTHORS Brakenhoff, J.P., De Groot, E.R., Evers, R.F., Pannekoek, H. and Aarden, L.A.
 TITLE Molecular cloning and expression of hybridoma growth factor in Escherichia coli
 JOURNAL J. Immunol. 139, 4116-4121 (1987)
 COMMENT IL-6 is a cytokine which has been shown to induce HIV gene expression (Matsuyama et al. AIDS 5: 1405-1417, 1991). It may mediate this affect by altering the activity of cellular transcription factors (Garcia et al. AIDS 8: S3-S17, 1994). Draft entry and computer readable copy of sequence [1] kindly provided by J.Brakenhoff 20-SEPT-88.
 NCBI gi: 184631
 BASE COUNT 339 a 246 c 233 g 307 t
 ORIGIN 10 bp upstream of *Ava*I site; Chromosome 7p21.
 1 ttctgcccgc gagcccccgc ggaacgaaag agaagctcta tctcgccctc aggagcccag
 61 ctATGaaact cttctccaca agcgccctcg gtccagttgc cttctccctg gggctgctcc
 IL-6 cds ->
 121 tgggtgttgc tggctgcctc cctgccccag tccccccagg agaagattcc aaagatgtag
 181 ccgccccaca cagacagcca ctcaactctt cagAACgaat tgacaaacaa attcgggtaca
 241 tctctgacgg catctcagcc ctgagaaagg agacatgtaa caagagttaa atgtgtgaaa
 301 gcagcaaaaga ggcactggca gaaaacaacc tgaaccttcc aaagatggtt gaaaaagatg
 361 gatgcttcca atctggattc aatgaggaga cttgcctggt gaaaatcacc actgggtcttt
 421 tggagtttga ggtataccta gactacctcc agaacagatt tgagagttagt gaggaacaag
 481 ccagagctgt gcagatgagt acaaaagtcc tgatccagtt cctgcagaaa aaggcaaaaga
 541 atctagatgc aataaccacc cctgacccaa ccacaaatgc cagcctgctg acgaagctgc
 601 aggcacagaa ccagtggctg caggacatga caactcatct cattctgccc agctttaagg
 661 agttcctgca gtccagcctg agggctcttc ggcaaatgTA Gcatgggcac ctcagattgt
 <- IL-6 cds end
 721 tgggtgttaat gggcatctct tcttctggtc agaaacctgt ccactgggca cagaacttat
 781 gttgtttctt atggagaact aaaagtatga gcgttaggac actattttta ttatttttaa
 841 tttattaata tttaaatatg tgaagctgag ttaattttat taagtcatat ttatatTTTT
 901 aagaagtacc acttgaacaa ttttatgtat tagttttgaa ataataatgg aaagtggcta
 961 tgcagtttga atatcctttg tttcagagcc agatcatttc ttggaaagtg taggcttacc
 1021 tcaataaataa ggttaactta tacatatTTT taaagaataa tttatatTgt atttatataa
 1081 tgtataaatg gtttttatac caataaatgg cattttaaaa aattc

LOCUS HUMIn1 1857 bp mRNA PRI 11-JAN-1994
 DEFINITION Human In1 cDNA sequence, complete cds.
 ACCESSION U04847
 SOURCE Human mRNA from HL60 cell line.
 REFERENCE 1 (bases 1 to 1857)
 AUTHORS Kalpana,G.V. and Goff,S.P.
 TITLE The HIV-1 integrase binds to a novel human protein with sequence similarity to the yeast transcription factor SNF5
 JOURNAL Unpublished
 COMMENT ;
 BASE COUNT 457 a 524 c 476 g 400 t
 ORIGIN

```

1  gccccggccc  cccccagcc  ctctgagcc  ctgcagccc  ggctccggcc  gcccgcctct
61  gccgcgcgca  TGatgatgat  ggcgctgagc  aagaccttgc  ggcagaagcc  cgtgaagttc
In1 cds start ->
121  cagctggagg  acgacggcga  gttctacatg  atcggtcccg  aggtgggaaa  ctacctccgt
181  atgttccgag  gttctctgta  caagagatac  cctcactct  ggaggcgact  agccaactgtg
241  gaagagagga  agaaaatagt  tgcacgtca  catggtaaaa  aaacaaaacc  taacactaag
301  gatcacggat  acacgactct  agccaccagt  gtgacctgt  taaaagcctc  ggaagtggaa
361  gagattctgg  atggcaacga  tgagaagtac  aaggctgtgt  ccacagcac  agagccccc
421  acctacctca  gggaacagaa  ggccaagagg  aacagccagt  gggtaaccac  cctgtccaac
481  agctccacc  acttagatgc  cgtgccatgc  tccacaacca  tcaacaggaa  cgcgatgggc
541  cyagacaaga  agagaacott  ccccttttgc  tttgatgacc  atgaccagc  tgtgatccat
601  gagaacgcac  ctgagcccca  ggtgctggtc  cccatccggc  tggacatgga  gatcgatggg
661  cagaagctgc  gagagccctt  cacotggaac  atgaatgaga  agttgatgac  gcctgagatg
721  ttttcagaaa  tcctctgtga  ccatctggat  ttgaaccggc  tgacgtttgt  gccagccatc
781  gccctctgca  tcagacagca  gatcgagtc  taccocacgg  acagcatcct  ggaggaccag
841  tcagaccagc  gcgtcatcat  caagctgaac  atccatgttg  gaaacatttc  cctgggtggc
901  cagtttgagt  gggacatgtc  agagaaggag  aactcaccag  agaagtttgc  cctgaagctg
961  tgcctggagc  tggggttggg  cggggagttt  gtcaccacca  tcgcatacag  catccgggga
1021  cagctgagct  ggcacagaa  gacctacgcc  ttcagcgaga  accctctgcc  cacagtggag
1081  attgccatcc  ggaacaaggg  ccatgaggac  cagtgggtgc  caotgctgga  gactctgaca
1141  gacgtgaga  tggagaagaa  gatccgcgac  caggacagga  acacgaggcg  gatgaggcgt
1201  cttgccaaca  cgggcccggc  ctggTAacca  gccatcagc  acacggctcc  cacggagcat
      <- In1 cds end
1261  ctgagaagat  tgggcccgt  ctctccatc  ttctggcaag  gacagaggcg  aggggacagc
1321  ccagcgccat  cctgaggatc  gggtaggggt  ggagtggggg  ctccaggtg  gcccttcccg
1381  gtacacattc  catttggtga  gcccagtc  tgccccccac  cccacctcc  ctacctctcc
1441  coagtctctg  gggtcaggaa  gaaaccttat  tttaggttgt  gttttgtttt  tgtataggag
1501  cccagggcag  ggctagtaac  agtttttaaa  taaaaggcaa  caggtcatgt  tcaatttctt
1561  aaatctagtg  tctttatttc  ttctgttaca  atagtgttgc  ttgtgtaagc  aggttagagt
1621  gcacagtgtc  cccaattgtt  cctggcactg  caaaaccaaa  ttaaacaatc  ccacaaagaa
1681  ttctgacatc  aatgtgtttt  cctcagtcag  gtctatttca  agattctaga  agttcctttt
1741  gtaaaacttg  cctttaaacc  tcttctctct  aatgccatca  gatctcttaa  cattggctca
1801  ctgtgggata  tttctcttta  ggttgaattt  ctacgtgaat  atcaaagtgc  ctttttc
  
```

HUMKUP70

LOCUS HUMKUP70 2086 bp ss-mRNA PRI 15-DEC-1989
 DEFINITION Human lupus p70 (Ku) autoantigen protein cDNA sequence, complete cds.
 ACCESSION J04611
 KEYWORDS Ku; autoantigen; p70 autoantigen.
 SOURCE Human mRNA from hepatoma G2 cell line.
 REFERENCE 1 (bases 1 to 2086)
 AUTHORS Reeves, W.H. and Stoecker, Z.M.
 TITLE Molecular cloning of cDNA encoding the p70 (Ku) lupus autoantigen
 JOURNAL J. Biol. Chem. 264, 5047-5052 (1989)
 COMMENT The HIV-1 leader RNA, TAR, forms a stem-bulge-loop structure of approximately 60+ nucleotides. Since TAR RNA binds many cellular proteins, it is thought to play a role in gene regulation. Recently, it has been reported that the human lupus autoantigen Ku binds to the loop of HIV-1 TAR (Kaczmariski and Khan, Biochem. Biophys. Res. Communications 196:935-942).
 Draft entry and computer-readable sequence [1] kindly submitted by W.H. Reeves, 21-FEB-1989.
 NCBI gi: 178649
 BASE COUNT 589 a 461 c 550 g 486 t
 ORIGIN

```

1  cgettccttg  cgccaaagtg  agcagtagcc  aacATGtcag  ggtgggagtc  atattacaaa
    lupus autoantigen Ku cds start ->
61  accgagggcg  atgaagaagc  agaggaagaa  caagaagaga  accttgaagc  aagtggagac
121  tataaatatt  caggaagaga  tagtttgatt  tttttggttg  atgcctccaa  ggctatgttt
181  gaatctcaga  gtgaagatga  gttgacacct  tttgacatga  gcacccagtg  tatccaaagt
241  gtgtacatca  gtaagatcat  aagcagtgat  cgagatctct  tggctgtggt  gttctatggc
301  acogagaaag  acaaaaatcc  agtgastttt  aaaaatattt  acgtcttaca  ggagctggat
361  aatccagggtg  caaaaacgaat  tctagagctt  gaccagttta  aggggagcaa  gggacaaaaa
421  cgtttccaag  acatgatggg  ccacggatct  gactactcac  tcagtgaagt  gctgtgggtc
481  tgtgccaaac  tcttttagtg  tgtccaatcc  aagatgagtc  ataagaggat  catgctgttc
541  accaatgaag  acaaccccc  tggcaatgac  agtgccaaag  ccagccgggc  caggaccaaa
601  gccgggtgatc  tccgagatac  aggcattctt  cttgacttga  tgcacctgaa  gaaacctggg
661  ggctttgaca  tatccttggt  ctacagagat  atcatcagca  tagcagagga  tggggacctc
721  agggttcact  ttgaggatc  cagcaagcta  gaagacctgt  tgcgggaagg  tggcgccaa
781  gagaccagga  agcagcact  cagcaggtta  aagctgaagc  tcaacaaaga  tatagtgatc
841  tctgtgggca  tttataatct  ggtccagaag  gctctcaagc  ctctctcaat  aaagctctat
901  cgggaaacaa  atgaaccagt  gaaaaccaag  acccggacct  ttaatacaag  tacaggccgt
961  ttgcttctgc  ctacgatac  caagaggtct  cagatctatg  ggagtcgtca  gattatactg
1021  gagaagagg  aaacagaaga  gctaaaacgg  tttgatgata  caggtttgat  gctcatgggt
1081  ctcaagccgt  tggactgtgt  gaagaacac  cattacctga  ggcctcctct  gttcgtgtac
1141  ccagaggagt  cgtgtgtgat  tgggagctca  acctgttca  gtgctctgct  catcaagtgt
1201  ctggagaagg  aggttgacg  attgtgcaga  tacaccccc  gcaggaaacat  ccttccttat
1261  tttgtggctt  tgggtgccca  ggaagaagag  ttggatgacc  agaaaattca  ggtgactcct
1321  ccaggcttcc  agctggtctt  tttacctttt  gctgatgata  aaagggaagat  gccctttact
1381  gaaaaaatca  tggcaactcc  agagcagggt  ggcaagatga  aggcatacgt  tgagaagctt
1441  cgettcacat  acagaagtga  cagctttgag  aaccccgtag  tgcagcagca  cttcaggaa
1501  ctggaggcct  tggccttgga  tttgatggag  ccggaacaag  cagtggacct  gacattgccc
1561  aaggttgaa  caatgaataa  aagactgggc  tccttggtgg  atgagtttaa  ggagcttgtt
1621  taccaccag  attacaatcc  tgaagggaaa  gtaccaaga  gaaaacacga  taatgaagggt
1681  tctggaagca  aaaggcccaa  ggtggagtat  tcagaagagg  agctgaagac  ccacatcagc
1741  aagggtacgc  tgggcaagtt  cactgtgcc  atgctgaaag  aggcctgccc  ggcttaacgg
1801  ctgaagagtg  gtctgaagaa  gcaggagctg  ctggaagccc  tcaccaagca  cttccaggac
1861  TGaccagagg  ccgcgcgtcc  agctgcctt  ccgagtggtg  gccaggctgc  ctggccttgt
    <- Lupus autoantigen Ku cds end
1921  cctcagccag  ttaaaatgtg  tttctcttga  gctaggaaga  gtctacccga  cataagtcga
1981  gggactttat  gtttttgagg  cttctgtgtg  ccattggtgat  ggtgtagccc  tccactttg
2041  ctgttcctta  ctttactgac  tgaataaaga  gccctaagtc  tgtact

```


LOCUS HUMKUP86 3052 bp ss-mRNA PRI 11-JUN-1993
 DEFINITION Human lupus p86 (Ku) autoantigen cDNA sequence, complete cds.
 ACCESSION J04977
 KEYWORDS Ku antigen; nonhistone DNA binding protein; nuclear protein.
 SOURCE Human fetal liver mRNA.
 REFERENCE 1 (bases 1 to 3052)
 AUTHORS Yaneva, M., Wen, J., Ayala, A. and Cook, R.
 TITLE cDNA derived amino acid sequence of the 86kDa subunit of the Ku antigen
 JOURNAL J. Biol. Chem. 264, 13407-13411 (1989)
 COMMENT The HIV-1 leader RNA, TAR, forms a stem-bulge-loop structure of approximately 60+ nucleotides. Since TAR RNA binds many cellular proteins, it is thought to play a role in gene regulation. Recently, it has been reported that the human lupus autoantigen Ku binds to the loop of HIV-1 TAR (Kaczmarzski and Khan, Biochem. Biophys. Res. Communications 196:935-942). Draft entry and computer-readable sequence for (i) kindly submitted by M.Yaneva, 02-JUN-1989.
 NCBI gi: 186791

BASE COUNT 906 a 592 c 708 g 846 t

ORIGIN 58 bp upstream of PvuII site.

```

1 ggccggcgac caaagcgct gaggacggc aacATGgtgc ggtcggggaa taaggcagct
   Lupus autoantigen Ku cds start ->
61 gttgtgctgt gtatggacgt gggtttacc atgagtaact ccattcctgg tatagaatcc
121 ccatttgaac aagcaagaa ggtgataacc atgtttgtac agcgacaggt gtttgcctgag
181 aacaaggatg agattgcttt agtctgttt ggtacagatg gcaatgacaa tcccccttct
241 ggtggggatc agtatcagaa catcacagt cacaacatc tgatgctacc agattttgat
301 ttgctggagg acattgaaag caaatccaa ccaggttctc aacaggctga cttcctggat
361 gcaataatcg tgagcatgga tgtgattcaa catgaacaa taggaagaa gtttgagaag
421 aggcattatt aattattcac tgacctcagc agccgattca gcaaaagtca gctggatatt
481 ataattcata gcttgaagaa atgtgacatc tccctgcaat tcttcttgcc tttctcactt
541 ggcaaggaa atggaagtgg ggacagagga gatggccctt ttcgcttagg tggccatggg
601 ccttcctttc cactaaaagg aattaccgaa cagcaaaaag aaggtcttga gatagtgaag
661 atggtgatga tatctttaga aggtgaagat gggttggatg aaatttatcc attcagtgag
721 agtctgagaa aactgtgctt cttcaagaaa attgagaggg attccattca ctggccctgc
781 cgactgacca ttggctccaa tttgtctata aggattgcag cctataaatc gattctacag
841 gagagagtta aaaagacttg gacagttgtg gatgcaaaa ccttaaaaaa agaagatata
901 caaaaagaaa cagtttattg cttaaatgat gatgatgaaa ctgaagtttt aaaagaggat
961 attattcaag ggttccgcta tgggaagtga atagttcctt tctctaaagt ggaagaggaa
1021 caaatgaat ataatcggg ggggaagtgc tttctgttt tgggattttg taaatcttct
1081 caggttcaga gaagattctt catgggaaat caagttctaa aggtctttgc agcaagagat
1141 gatgaggcag ctgcagttgc actttctctc ctgattcatg ctttggatga cttagacatg
1201 gtggccatag ttgatattgc ttatgacaaa agagctaata ctcaagtcgg cgtggctttt
1261 cctcatatca agcataacta tgagtgttta gtgtatgtgc agctgccttt catggaagac
1321 ttgoggcaat acatgttttc atccttgaaa aacagtaaga aatatgtctc caccgaggca
1381 cagttgaatg ctggtgatgc tttgattgac tccatgagct tggcaagaa agatgagaag
1441 acagacaccc ttgaagactt gtttccaaac accaaatcc caaatcctcg atttcagaga
1501 ttatttctag gtctgtgca cagagtttta catcccggg agcctctacc ccaattcag
1561 cagcatattt ggaatctgct gaatccctcc gctgagggtg caacgaaaag tcaagttcct
1621 ctctctaaaa taaagacctt tttctctctg attgaagcca agaaaaggga tcaagtgact
1681 gctcaggaaa ttttccaaag caacctgaa gatggacct cagctaaaaa attaaagact
1741 gagcaagggg gagccactt cagcgtctcc agtctggctg aaggcagtgat caccctctgtt
1801 ggaagtgtga atcctgctga aaacttccgt gttctagtga aacagaagaa ggcagctttt
1861 gaggaagcga gtaaccagct cataaatcac atcgaaacagt ttttggatac taatgaaaca
1921 ccgtatttta tgaagagcat agactgcac cgagccttcc gggagaagac cattaagttt
1981 tcagaagagc agcgtcttaa caacttctct aaagccttcc aagagaaggt ggaattaaa
2041 caattaaatc atttctggga aattgtgtgc caggatggaa ttactctgat caccaaagag
2101 gaagcctctg gaagttctgt cacagctgag gaagcaaaa agtttctggc ccccaaagac
2161 aaaccaagtg gagacacagc agctgtattt gaagaagggt gtgatgtgga cgatttattg

```

```

2221 gacatgatat AGgtcgtgga tgtatgggga atctaagaga gctgccatcg ctgtgatgct
      <- Lupus autoantigen Ku cds end
2281 gggagttcta acaaaacaag ttggatgcgg ccattcaagg ggagccaaat tctcaagaaa
2341 ttcccagcag gttacctgga gggggatcat ctaattctct gtggaatgaa tacacacata
2401 tatattacaa gggataattt agaccccata caagtttata aagagtcatt gttattttct
2461 ggttgggtga ttattttttc tgtgggtctta ctgactcttg tatattacat acatgctttg
2521 aagtttctgg aaagtagatc ttttcttgac ctagtatata agtgacagtt gcagcctctg
2581 tgatgtgatt agtgtctcat gtggaaccat ggcattggta ttgatgagtt tcttaacctt
2641 ttccagagtc ctctcttccc tgatcctcca acagctgtca cagcttgtgt tgagcaagca
2701 gtagcatttg ctctctcccc acaagcagct gggttaggaa aacctgggt aaggacggac
2761 tcaactctct ttttagttga ggccttctag ttaccacatt actctgcctc tgtatatagg
2821 tggttttctt taagtggggt ggggaagggga gcacaatttc ccttcatact ccttttaagg
2881 agtgagttat ggtgggtggt tcatgaagaa aagacctttt ggcccaatct ctgccatata
2941 agtgaacctt tagaaactca aaaactgaga aatttactac agtagttaga attatatcac
3001 ttcaactgtt tctacttgca agcctcaag agagaaagtt tcgttatatt gg

```

LOCUS HUMLA 1620 bp ss-mRNA PRI 14-MAY-1991
 DEFINITION Human lupus autoantigen La cDNA sequence, clones
 lambda-La[2.1,26.2], complete cds.
 ACCESSION J04205 M11108
 KEYWORDS La antigen; La protein; SS-B antigen; autoantigen.
 SOURCE Human (adult) liver mRNA.
 REFERENCE 1 (bases 223 to 391)
 AUTHORS Chambers,J.C. and Keene,J.D.
 TITLE Isolation and analysis of cDNA clones expressing human lupus La
 antigen
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119 (1985)
 REFERENCE 2 (bases 1 to 1620)
 AUTHORS Chambers,J.C., Kenan,D., Martin,B.J. and Keene,J.D.
 TITLE Genomic structure and amino acid sequence domains of the human La
 autoantigen
 JOURNAL J. Biol. Chem. 263, 18043-18051 (1989)
 COMMENT The HIV-1 leader RNA, TAR, forms a stem-bulge-loop structure
 of approximately 60+ nucleotides. Since TAR RNA binds many
 cellular proteins, it is thought to play a role in gene
 regulation. Recently, it has been reported that the human
 lupus autoantigen La recognizes U-residues within the overall
 context of the TAR secondary structure [Chang et al. J. Virol.
 68, 7008 (1994); Svitkin et al. J. Virol. 68, 7001 (1994)].
 Draft entry and computer-readable sequence for [2] kindly provided
 by J.D.Keene, 28-FEB-1989.
 NCBI gi: 178686

BASE COUNT 620 a 227 c 363 g 410 t

ORIGIN Chromosome 2.

```

1 cccggcggcgc tgggaggtgg agtcgttget gttgctgttt gtgagcctgt ggccgggctt
61 ctgtggggcgc gaaccttaaa gatagccgta ATGctgaaa atggtgataa tgaagaatg
lupus La autoantigen cds start ->
121 gctgccttgg aggcacaaat ctgtcatcaa attgagtatt attttggcga cttcaatttg
181 ccacggggaca agtttctaaa ggaacagata aaactggatg aaggctgggt acctttggag
241 ataatgataa aattcaacag gttgaacgtt ctaacaacag actttaatgt aattgtggaa
301 gcattgagca aatccaaggc agaactcatg gaaatcagtg aagataaaac taaaatcaga
361 aggtctccaa gcaaacccct acctgaagtg actgatgagt ataaaaatga tgtaaaaaac
421 agatctgttt atattaaagg cttcccaact gatgcaactc ttgatgacat aaaagaatgg
481 ttagaagata aaggtcaagt actaaatatt cagatgagaa gaacattgca taaagcattt
541 aagggatcaa tttctgttgt gtttgatagc attgaatctg ctaagaaatt tgtagagacc
601 cctggccaga agtacaaga aacagacctg ctaatacttt tcaaggacga ttactttgcc
661 aaaaaaatg aagaaagaa acaaaataaa gtggaagcta aattaagagc taacacggag
721 caagaagcaa aacaaaagtt agaagaagat gctgaaatga aatctotaga agaaaagatt
781 ggtgcttgc tgaattttc ggttgattta gatgatcaga cctgtagaga agatttacac
841 atacttttct caaatcatgg tgaataaaaa tggatagact tcgtcagagg agcaaaagag
901 gggataattc tatttaaga aaaagccaag gaagcattgg gtaaagccaa agatgcaaat
961 aatggtaacc tacaattaag gaacaaagaa gtgacttggg aagtactaga agggagaggtg
1021 gaaaaagaag cactgaagaa aataatagaa gaccaacaag aatccctaaa caaatggaag
1081 tcaaaaggtc gtagatttaa aggaaaagga aagggttaata aagctgcccc gccctgggtct
1141 ggtaaaggaa aagtacagtt tcagggcaag aaaacgaaat ttgctagtga tgatgaacat
1201 gatgaacatg atgaaaatgg tgcaactgga cctgtgaaaa gagcaagaga agaaacagac
1261 aaagaagaac ctgcatccaa acaacagaaa acagaaaatg gtgctggaga ccagTAGttt
<- lupus
La autoantigen
1321 agtaaaccaa ttttttatcc atttttaata ggttttaaac gacttttgtt tgcggggctt
1381 ttaaaaggaa aaccgaatta ggtccacttc aatgtccacc tgtgagaaag gaaaaatttt
1441 tttgtgtttt aacttgtott ttgtttatgc aaatgagatt tctttgaatg tattgttctg
1501 tttgtgttat ttcagatgat tcaaatatca aaaggaagat ccttccatta aattgccttt
1561 gtaatatgag aatgtattag tacaactaa ctaataaaat atatactata tgaagaagagc

```

HUMLCK

LOCUS HUMLCK 2032 bp ss-mRNA PRI 23-APR-1992
 DEFINITION Human lymphocyte-specific protein tyrosine kinase (lck) cDNA
 sequence, complete cds, clones HK[28.8.2.5].
 ACCESSION M36881
 KEYWORDS lymphocyte-specific protein tyrosine kinase.
 SOURCE Human adult peripheral blood lymphocyte mRNA.
 REFERENCE 1 (bases 1 to 2032)
 AUTHORS Parimutter,R.M., Marth,J.D., Lewis,D.B., Peet,R., Ziegler,S.F. and
 Wilson,C.B.
 TITLE Structure and expression of lck transcripts in human lymphoid cells
 JOURNAL J. Cell. Biochem. 38, 117-126 (1988)
 COMMENT A group of investigators has found that HIV-1 Nef binds directly to
 the lck-kinase in vitro (D. McPhee personal communication).
 NCBI gi: 187033
 BASE COUNT 450 a 576 c 584 g 422 t
 ORIGIN
 1 cgccctggacc atgtgaatgg ggccagaggg ctcccgggct gggcagggac cATGggctgt
 lck kinase cds start ->
 61 ggetgcagct cacacccgga agatgactgg atggaaaaca togatgtgtg tgagaactgc
 121 cattatccca tagtcggact ggatgggaag ggcaggctgc tcatecgaaa tggctctgag
 181 gtgctgggacc cactgggttac ctacgaaggc tccaatccgc cggcttcccc actgcaagac
 241 aacctgggta togtctctga cagctatgag cctctctcag acggagatct gggctttgag
 301 aagggggaac cactcggcat cctggagcag agcggcgagt ggtgggaagg gcagtccttg
 361 accacggggcc aggaaggctt catcccttc aattttgtgg ccaaagcgaa cagcctggag
 421 ccgaacccct ggttcttcaa gaacctgagc cgcaaggacg cggagcggca gctcctggcg
 481 cccgggaaca ctacgggtc ctctctcacc cgggagagcg agagcaccgc cgggtctttt
 541 tcaactgtcg tccgggactt cgaccaaaac caggagagag tgggtgaasca ttacaagatc
 601 cgtaatctgg acaacgggtg ctctctacac tccctctgaa tcacttttcc cggcctgcat
 661 gaactgggtcc gccattacac caatgcttca gatgggctgt gcacacggtt gagccgcccc
 721 tggcagaccc agaagcccca gaagccgtgg tgggaggacg agtgggaggt tcccaggag
 781 acgctgaagc tgggtggagcg gctgggggct gcacagttcg gggaggtgtg gatggggatc
 841 tacaacgggc acacgaagggt ggcggtgaag agcctgaagc agggcagcat gtccccggac
 901 gccttctctg ccgagggcca cctcatgaag cagctgcaac accagcggct ggttcggctc
 961 taagctgtgg tcacccagga gcccatctac atcatcactg aatacatgga gaatgggagt
 1021 ctagtggatt ttctcaagac ccttcaggc atcaagttga ccatcaaca actcctggac
 1081 atggcagccc aaattgcaga aggcattggca ttcattgaag agcggaaatta tattcatcgt
 1141 gaccttctgg ctgccaacat tctggtgtct gacacctga gctgcaagat tgcagacttt
 1201 ggcttagcac gctcattga ggacaaagag tacacagcca gggagggggc caagtttccc
 1261 attaatgtga cagcggcaga agccattaac tacgggacat tcaccatcaa gtcagatgtg
 1321 tgggtctttt ggatcctgct gacggaaatt gtcaccacg gccgcatccc ttaccaggg
 1381 atgaccaacc cggaggtgat tcagaacctg gagcgaggct accgcatggt ggcacctgac
 1441 aactgtccag aggagctgta ccaactcatg aggtgtgct ggaaggagcg cccagaggac
 1501 cggccacact ttgactacct ggcaggtgtg ctggaggact tcttcacggc cacagagggc
 1561 cagtaccagc ctacgctTG Agaggaggcc ttgagaggcc ctggggttct cccctttct
 <- lck kinase cds end
 1621 ctccagcctg acttggggag atggagtctt tgtgcatag tcacatggcc tatgcacata
 1681 tggactctgc acatgaatcc caccacatg tgacacatat gcaccttgtg tctgtacacg
 1741 tgtcctgtag ttgctgggac tctgcacatg tctgtgtcat gtgtagcctg tgcattgtatg
 1801 tcttggacac tgtacaagggt aaccttttct ggtctctcca tttcttgaga ccaccagaga
 1861 gaggggagaa gcttgggatt gacagaagct tctgcccacc taattttctt tctcagatc
 1921 atccagaagt tctgaagggt ccaggacttt atctaatac tctgtgtgct cctccttgg
 1981 gcttggcctg gcacacatca ggagttcaat aatgtctgt tgatgactgc cg

LOCUS HUMMSS1 1478 bp ss-mRNA PRI 24-JUL-1992
 DEFINITION Human MSS1 cDNA sequence, complete cds.
 ACCESSION D11094
 KEYWORDS mammalian suppressor of sgvl; transactivation factor.
 SOURCE Human mRNA from HeLa cell line.
 REFERENCE 1 (bases 1 to 1478)
 AUTHORS Shibuya,H., Irie,K., Ninomiya-Tsuji,J., Goebel,N., Taniguchi,T. and Matsumoto,K.
 TITLE New human gene encoding a positive modulator of HIV Tat-mediated transactivation
 JOURNAL Nature 357, 700-702 (1992)
 COMMENT MSS1 is part of a large family of proteins related to the 26S protease from human erythrocytes. Each of these proteins has Tat-binding activity [1]. MSS1 positively affects gene expression through its interaction with Tat [1].
 NCBI gi: 219930
 BASE COUNT 439 a 287 c 375 g 377 t
 ORIGIN

```

1 ccattgtgct ctaaagggaa ggtgctgtgt aatcattaag gagggagggc ttttggagct
61 gctaaaATGc cggattacct cggtgccgat cagcgaaga ccaaagagga tgagaaggac
MSS1 cds start ->
121 gacaagccca tccgagctct ggatgagggg gatattgcct tgttgaaaac ttatggtcag
181 agcacttact ctaggcagat caagcaagtt gaagatgaca ttcagcaact tctcaagaaa
241 attaatgagc tcaactggtat taaagaatct gacactggcc tggcccccacc agcactctgg
301 gatcttgctg cagataagca gacactccag agtgaacagc ctttacaggt tggcagggtgt
361 acaaagataa tcaatgctga ttcggaggac caaaaataca ttatcaacgt aaagcagttt
421 gccaggtttg tgggtggacct tagtgatcag gtggcaccta ctgacattga agaagggatg
481 agagtgggag tggatagaaa taaatatcaa attcacattc cattgacctc taagattgac
541 ccaacagtta ccatgatgca ggtggasag agactgatg tcaatacag tgatgttggg
601 ggctgtaagg aacagattga gaaactgcga gaagtgttg aaacccatt acttcateca
661 gagaggtttg tgaaccttgg cattgagcct cccaagggcg tgetgctctt tggteccacc
721 ggtacaggca agacactctg tgcgggggca gttgctaate ggactgatgc gtgcttcatt
781 cgagttattg gatctgagct tgtacagaaa taagtcggtg agggggctcg aatgggtcgt
841 gaactctttg aaatggccag aacaaaaaaa gcctgcctta tcttctttga tgaatttgat
901 gctattggag gggctcgttt tgatgatggg gctggagggtg acaatgaagt gcagagaaca
961 atgttgggaa tgatcaatca gcttgatggg tttgatcctc gaggcaatat taaagtgtgt
1021 atggccacta acagacctga tactttggat ccagcactga tgaggccagg gagattggat
1081 agaaaaattg aatttagctt gcccgatcta gagggtcgga cccacatatt taagattcac
1141 gctcgttcaa tgagtgttga aagagatata agatttgaac tgttagcacg actgtgtcca
1201 aatagcactg gtgctgagat tagaagcgtc tgcacagagg ctggtatgtt tgccatcaga
1261 gcacggcgaa aaattgctac cgagaaggat ttottggaag ctgtaataaa ggtcattaag
1321 tottatgcca aattcagtcg tactcctcgt tacatgacat acaacTGAac cctgaaggct
                                     <- MSS1 cds end
1381 ttcaagtga aactttaaat tggaaacctt accttatata gacttggtta taaccaattc
1441 ataaacaaat aaatggcttc aactttagag cacaatgg
  
```

HUMNB23

LOCUS HUMNB23 1293 bp ss-mRNA PRI 26-OCT-1992
 DEFINITION Human nucleolar protein (B23) cDNA sequence, complete cds,
 clone lambda-Zap.
 ACCESSION M26697
 KEYWORDS B23 protein; nucleolar phosphoprotein; nucleolar protein;
 numatrin.
 SOURCE Human amnion mRNA from cell line U.
 REFERENCE 1 (bases 1 to 1293)
 AUTHORS Zhang,X.T., Thomas,D.C. and Samuel,C.E.
 TITLE Isolation and characterization of a molecular cDNA clone of a human
 mRNA from interferon-treated cells encoding nucleolar protein B23,
 numatrin
 JOURNAL Biochem. Biophys. Res. Commun. 164, 176-184 (1989)
 COMMENT Rev localizes to the nucleolus, and work by Laemmli and colleagues
 (Fankhauser et al. Mol. Cell. Biol. 11:2567-2575) has demonstrated
 a tight association between Rev and the nucleolar B23 protein.
 This protein-protein complex likely directs the subcellular
 localization of Rev.
 Draft entry and computer-readable copy of sequence [1] kindly
 submitted by C.E.Samuel, 03-AUG-1989.
 NCBI gi: 189311
 BASE COUNT 422 a 223 c 310 g 337 t 1 others
 ORIGIN
 1 cgggattccg tectgcggcg ttgttctctg gaggcagcgtt cttttatctc cgtccgcctt
 61 ctctctacc taagtgcgtg cggccaccgc ATGgaagatt cgtggacat ggacatgagg
 nucleolar protein B23 cds start ->
 121 cccctgaggg cccagaacta ttttttcggg tgtgaactaa aggcgcgaca agattatcac
 181 ttttaagggtg ataattgatg aaatgagcac cagttatctt taagaacggg cagtttaggg
 241 gctggtgcaa aggatgagtt gcacattgtt gaagcagagg caatgaatta cgaaggcagt
 301 ccaattaaag taacactggc aactttgaaa atgtctgtac agccaacngt ttcccttggg
 361 ggctttgaaa taacaccacc agtggctcta aggttgaagt gtggttcagg gccagtgcat
 421 attagtgga cgcacttagt agctgtggag gaagatgcag agtcagaaga tgaagaggag
 481 gaggatgtga aactcttaag tatactctga aagcgggtct cccctggagg tggtagcaag
 541 gttccacaga aaaaagttaa acttgctgct gatgaagatg atgacgatga tgatgaagag
 601 gatgatgatg aagatgatga tgatgatgat ttgatgatg aggaagctga agaaaaagcg
 661 ccagtgaaga aatctatagc agatactcca gccaaaaatg caaaaaagtc aaatcagaat
 721 ggaaaagact caaaaccatc atcaacacca agatcaaaag gacaagaatc cttcaagaaa
 781 caggaaaaaa ctctaaaaac accaaaagga cctagttctg tagaagacat taaagcaaaa
 841 atgcaagcaa gtatagaaaa aggtggttct cttcccaaaag tgggaagcaa attcatcaat
 901 tatgtgaaga attgcttccg gatgactgac caagaggcta ttcaagatct ctggcagttg
 961 aggaagtctc tttTAgaaaa tagtttaaac aatttggtta aaaattttcc gtcttatttc
 <- nucleolar protein B23 cds end
 1021 atttctgtaa cagttgatat ctggctgtcc tttttataat gcagagtgag aactttccct
 1081 accgtgtttg ataaatgttg tccaggttct attgccaga atgtgttgtc caaaatgccc
 1141 gtttagtttt taaagatgga actccacccc ttgcttgggt ttaagtatgt atggaatggt
 1201 atgataggac atagtagtag cgggtgtcag acatggaaat ggtggggaga caaaaatata
 1261 catgtgaaat aaaactcagt attttaataa agt

LOCUS HNFAT 2743 bp mRNA PRI 08-JUL-1994
 DEFINITION Human NF-AT cDNA sequence, complete cds.
 ACCESSION U08015
 SOURCE Human cDNA from Jurkat cell library.
 REFERENCE 1 (bases 1 to 2743)
 AUTHORS Northrop,J.P., Ho,S.N., Chen,L., Thomas,D.J., Timmerman,L.A.,
 Nolan,G.P., Admon,A. and Crabtree,G.R.
 TITLE NF-AT components define a family of transcription factors targeted
 in T-cell activation
 JOURNAL Nature 369, 497-502 (1994)
 COMMENT NFAT, nuclear factor from activated T cells, is an intermediating
 transducer of signals initiated at the T-cell antigen receptor.
 Within the U3 region of the HIV-1 LTR, there are NFAT binding
 motifs present from -216 to -254. Recent evidence suggests that
 NFAT binding activity results from the interaction of three
 discrete polypeptides. NFATp (McCaffrey et al. Science 262:
 750-754), Fos and Jun (Yaseen et al. J. Biol. Chem. 268:
 14285-14293).
 NCBI gi: 500631

BASE COUNT 529 a 959 c 788 g 467 t
 ORIGIN

```

1 gaattccgca gggcgcgggc accggggggc gggcagggtc cggagccacc ggcgaggtcc
61 tagggcgcg ggcgggcccc gccacggcg caccagcccc tcgatgactt tctccgggg
121 cgcggcgcg tgagccgggg gcgagggtc tcttccggga gaccggcccc cggcagcgcg
181 gggcgggcac ttctctgtg cctccggcg ctgtccact ccccgccgc gccgagcgga
                                     NFAT cds start ->
241 TGccaagcac cagctttcca gtcccttcca agtttccact tggccctgcg gctgcggtct
301 tggggagagg agaaactttg gggcccgcg cggcgcgcg cggcaccatg aagtcagcgg
361 aggaagaaca ctatggctat gcctcttcca acgtcagccc cggcctgccc ctccccacgg
421 cgcactccac cctgcggggc cgttgcacac accttcagac ctccacacgg ggcacatccc
481 cgcggcgcg tcacccctcg gggtagcgag cagcttttga cgggtgggccc gggggctact
541 tctctctctc cggccacacc aggcctgatg gggcccttgc cctggagagt cctgcctcgg
601 agataacctc gtgcttgggc ctgtaccaca acaataacca gtttttccac gatgtggagg
661 tgggaagacgt cctcccttag tccaaacggt ccccttccac ggccacgctg agtctgccc
721 gcttgagggc ctacagagac cctctgtgac tgagcccgcc cagcagcctg tctcccgga
781 gctgcaactc agaggcctcc tctacagagt ccaactactc gtaaccgtac gcttcccccc
841 agacgtcgcc atggcagctc cctgcgtgt ctcccaagac cagggacccc gaggagggct
901 ttccccggcg gctggggggc tgcacactgc tgggttcccc gcagcactcc cctccacct
961 cgcggcgcg cagcgtcact gaggagagct ggctgggtgc cggctcctcc agaccgggt
1021 ccccttgcaa caagagggaag tacagcctca acggccggca gccccttacc tcacccccc
1081 actgccccac gccgtccccc caaggctccc cgggggtcag cgtgacccac gactcgtggt
1141 tgggcaaacac caccagctac accagctcgg ccatcgtggc cggcatcaac gcgctgacca
1201 ccgacagcag cctggacctg ggagatggcg tccctgtcaa gtcccgcaag accacctgg
1261 agcagcgccc ctccagtgcg ctcaagtgag agcccgtcgg ggaggacctg ggcagcccc
1321 cgcggcgcg cgaacttcgg cccgaagact actcctcttt ccagcacatc aggaaggcg
1381 gcttctgaga ccagtaacctg ggggtgcgc agcaacccta ccagtggcg aagcccaagc
1441 cctgtcccc taegtctac atgagccga cctgcgcgc cctggactgg cagctgccgt
1501 cccactcagg cccgtatgag ctccggattg aggtgcagcc caagtcccac caccgagccc
1561 actacgagac ggagggcagc cggggggcg tgaaggcgtc gggcgaggga cccccatcg
1621 tgcagctgca tggctacttg gagaatgagc cgtgatgct gcagcttttc attgggacgg
1681 cggacgaccg cctgctgcgc ccgcacgct tctaccaggt gcaccgcatc acagggaaga
1741 ccgtgtccac caccagccac gaggetatcc tctccaacac caaagtctg gagatccac
1801 tctgcggga gaacagcatg cagagccgtc ttgactgtgc cggatcctg aaactcagaa
1861 actccgacat tgaacttcgg aaaggagaga cggacatcgg gagggaaga acacgggtac
1921 ggtcgtgtgt cgcgttcac gtcccgcaac ccagcgccg cagctgtcc ctgcaggtgg
1981 cctcccaacc catcgaatgc tcccagcgt cagctcagga gctgcctctg gtggagaagc
2041 agagcagcga cagctatccg gtctggggcg ggaagaagat ggtcctgtct ggccacaact
2101 tctgcagga ctccaaggtc attttcgtgg agaaagcccc agatggccac catgtctggg
2161 agatgggaag gaaaactgac cgggacctgt gcaagccgaa ttctctggtg gttgagatcc
  
```

HNFAT

```

2221 cgcacatttcg gaatcagagg ataaccagcc cggttcacgt cagttttctac gtctgcaacg
2281 ggaagagaaa gcgaagccag taccagcgtt tcacctacct tcccgccaac ggtaacgcca
2341 tctttctaac cgttaagccgt gaacatgagc gcgtgggggtg ctttttcTAA agacgcagaa
                                     <- NFAT cds end
2401 acgacgtcgc cgtaaagcag cgtggcgtgt tgcacattta actgtgtgat gtcccgttag
2461 tgagaccgag ccatcgatgc cctgaaaagg aaaggaaaag ggaagcttcg gatgcatttt
2521 ccttgatccc tgttgggggt ggggggcggg ggttgcatatc tcagatagtc acggttattt
2581 tgcttcttgc gaatgtataa cagccaaggg gaaaacatgg ctcttctgct ccaaaaaact
2641 gaggggggtcc tgggtgtgcat ttgcacctta aagctgctta cggtgaaaag gcaaataggt
2701 atagctattt tgcaggcacc tttaggaata aactttgctt tta

```


LOCUS HUMNFKB 3625 bp ss-mRNA PRI 20-FEB-1991
 DEFINITION Human nuclear factor kappa-B (NF-kappa-B) DNA binding subunit
 sequence, complete cds.
 ACCESSION M58603
 KEYWORDS nuclear factor kappa-B DNA binding subunit.
 SOURCE Human premyeloid mRNA.
 REFERENCE 1 (bases 1 to 3625)
 AUTHORS Meyer,R., Hatada,E.N., Hohmann,K.-P., Haiker,M., Bartsch,C.,
 Roethlisberger,U., Lehm,H.-W., Schlaeger,E.J., van Loon,A.P. and
 Scheidereit,C.
 TITLE Cloning of the DNA-binding subunit of human nuclear factor kappa-B:
 The level of its mRNA is strongly regulated by phorbol ester or
 tumor necrosis factor alpha
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 966-970 (1991)
 COMMENT NF-kappa-beta motifs are perhaps the best characterized sequence
 elements in the HIV-1 LTR (Nabel and Baltimore Nature 326:711-
 713). NF-kappa-beta directly impacts viral replication (Ross et
 al. J. Virol. 65:4350-4358; Kim et al. J. Virol. 67:1658-1662;
 Huang and Jeang J. Virol. 67:6937-6944; Huang et al. EMBO J. 13:
 2886-2896), viral transcription (Harrich et al. J. Virol. 63:
 2585-2591; Barkhout and Jeang J. Virol. 63:5501-5504), and Tat
 transactivation (reviewed in Jones and Peterlin Annu. Rev. Biochem.
 63:717-743). For more extensive discussions of the biochemical and
 functional properties of NF-kappa-beta (Ghosh et al. Cell 62:
 1019-1029; Kieran et al. Cell 44: 261-272; Nolan et al. Cell 64:
 961-969; Liou et al. EMBO J. 11: 3003- 3009).
 NCBI gi: 189177

BASE COUNT 933 a 937 c 956 g 799 t

ORIGIN

```

1  ggccacogga ggggccoggc gacgatcgct gacagcttcc cctgcccctc ccgtcggtcg
61  ggccgccagc cgcgcagcc ctgggcctgc acgcagccac cggccccgct cccggagccc
121 agcgccgccc aggcgcagc cgcgcggcca gtaaggcggc gcgcgccgcy gccacogcgg
181 gcctgcctt tccctcgcc gcgctgcgcc atggcgcggc gctgactggc ctggcccggc
241 ccgcgcggc tcccgctgc ccgacccgc actcgggccc gcccgggctc cgccctgcgc
301 ccgcctcttc attctccagc cggcagggcc cgcgccttag gaggggagag ccaccgcgc
361 caggagggcg aacgcggact cggcaccgcy cttcagaATG gcagaagatg atccatattt
NF-kappa-beta DNA binding subunit cds start ->
421 ggggaaggcc gaacaaatgt ttcatttggc tccttctttg actcatacaa tatttaatcc
481 agaagtattt caaccacaga tggcactgcc aacagatggc ccataccttc aaatattaga
541 gcaacotaaa cagagaggat ttctgttccg ttatgtatgt gaaggeccat cccatggtgg
601 actacctggt gctcttagtg aaaagaacaa gaagtcttac cctcaggtca aaatctgcaa
661 ctatgtggga ccagcaaagg ttattgttca gttggtcaca aatggaaaaa atatccacct
721 gcatgcccac agcctggtgg gaaaacactg tgaggatggg atctgcactg taactgctgg
781 acccaaggac atggtggtcg gcttcgcaaa cctgggtata cttcatgtga caaagaaaaa
841 agtatattga acactggaag cagcaatgac agaggcgtgt ataaggggct ataactctgg
901 actcttggtg caccctgacc ttgcctattt gcaagcagaa ggtggagggg accggcagct
961 gggagatcgg gaaaaagagc taatccgcca agcagctctg cagcagacca aggagatgga
1021 cctcagcgty gtgcggctca tgtttacagc tttcttccg gatagcactg gcagcttcac
1081 aaggcgcccty gaacccgtgg tatcagagcy catctatgac agtaaagccc ccaatgcac
1141 caacttgaaa attgtaagaa tggacaggac agctggatgt gtgactggag gggaggaaat
1201 ttatcttctt tgtgacaaag ttcagaaaga tgacatccag attcgatttt atgaagagga
1261 agaaaatggt ggagttctggg aaggatttgg agatttttcc cccacagatg ttcatagaca
1321 atttgccatt gtcttcaaaa ctccaaagta taaagatatt aatattacaa aaccagcctc
1381 tgggtttgtc cagcttcgga ggaatctga cttggaaact agtgaaccaa aacctttcct
1441 ctactatcct gaaatcaaa ataaagaaga agtcagaggg aacgcgcaga agctcatgcc
1501 caatttttct gatagtttct gcggtggtag tgggtgcgga gctggaggcg gaggeatgtt
1561 tggtagtgcc ggtggaggag ggggcaactg aagtaacaggt ccagggtata gcttccaca
1621 ctatggattt cctacttatg gtgggattac tttccatcct ggaactacta aatctaattg
1681 tgggatgaag catggaacca tggacactga atctaaaaag gaccctgaag gttgtgacaa

```

```

1741 aagtgatgac aaaaacactg taaacctctt tgggaaagtt attgaaacca cagagcaaga
1801 tcaggagccc agcgaggcca cegtgtggaa tggtagggtc actctaactg atgcaacagg
1861 aacaaaagaa gagagtgtg gagttcagga taacctcttt ctagagaagg ctatgcagct
1921 tgcaaaagagg catgccaatg cctttttcga ctacgagggtg acaggagacg tgaagatgct
1981 gctggccgtc cagcgccatc tcactgtgtg gcaggatgag aatggggaca gtgtottaca
2041 cttagcaatc atccaccttc atttccaact tgtgagggat ctactagaag tcacatctgg
2101 tttgatttct gatgacatta tcaacatgag aatgatctg taccagacgc ccttgcactt
2161 ggcagtgate actaagcagg aagatgttgt ggaggatttg ctgagggctg gggccgacct
2221 gagcctttctg gaccgcttgg gtaactctgt ttgcaacctg gctgccaaag aaggacatga
2281 taaagtcttc agtatcttac tcaagcaca aaaggcagca ctacttcttg accaccccaa
2341 cggggacggg ctgaatgcca ttcatctage catgatgagc aatagcctgc catgtttgct
2401 gctgctggtg gcgctgggg ctgacgtcaa tgcctaggag cagaagtccg ggcgcacagc
2461 actgcacctg gctgtggagc acgacaacct ctcatggca ggctgcctgc tcttggaggg
2521 tgatgccat gtggacagta ctacctacga tggaaacaca cctctgcata tagcagctgg
2581 gagagggctc accaggctgg cagctcttct caaagcagca ggagcagatc ccttgggtga
2641 gaactttgag cctctctatg acctggatga ctcttgggaa aatgcaggag aggatgaagg
2701 agttgtgctt ggaaccacgc ctctagatat ggccaccagc tggcagggtat ttgacatatt
2761 aatgggaaa ccatatgagc cagagtctac atctgatgat ttactagcac aaggagacat
2821 gaaacagctg gctgaagatg tgaagctgca gctgtataag ttactagaaa ttctgatcc
2881 agacaaaaac tgggtactc tggcgagaa attaggtctg gggatactta ataatgcctt
2941 cgggctgagt cctgctctt ccaaaacact tatggacaac tatgaggtct ctgggggtac
3001 agtcagagag ctggtggagg cctgagaca aatgggtctc accgaagcaa ttgaagtgat
3061 ccaggcagcc tccagcccag tgaagaccac ctctcaggcc cactcgctgc ctctctgcc
3121 tgcttcaca aggcagcaaa tagacgagct ccgagacagt gacagtgtct gcgacacggg
3181 cgtggagaca tcttccgca aactcagctt taccgagtct ctgaccagtg gtgctcact
3241 gctaaactct aacaaaatgc cccatgatta tgggcaggaa ggacctctag aaggcaaat
3301 tTAGcctgct gacaatttcc cacacgtgt aaaccaaagc cctaaaattc cactgcgttg
    <- NF-kappa-beta DNA binding subunit cds end
3361 tccacaagac agaagctgaa gtgcaccaa aggtgctcag agagccggcc cgctgaatc
3421 attctcgatt taactcgaga ccttttcaac ttggcttctt ttcttggttc ataatgaat
3481 tttagtgttg ttcacttaca gatagtatct agcaatcaca acactggctg agcggatgca
3541 tctggggatg aggttgctta ctaagctttg ccagctgctg ctggatcaca gctgctttct
3601 gttgtcattg ctgttgcctc tctgc

```

LOCUS HUMP65NFkB 1767 bp ss-mRNA PRI 02-DEC-1991
 DEFINITION Human NF-kappa-B transcription factor p65 DNA binding subunit cDNA
 sequence, complete cds.
 ACCESSION M62399
 KEYWORDS NF-kappa-B transcription factor.
 SOURCE Human mRNA.
 REFERENCE 1 (bases 1 to 1767)
 AUTHORS Ruben,S.M., Dillon,P.J., Schreck,R., Henkel,T., Chen,C.-H.,
 Mahar,M., Baeuerle,P.A. and Rosen,C.A.
 TITLE A novel rel-related human cDNA that potentially encodes the 65 kDa
 subunit of NF-kappaB
 JOURNAL Science 251, 1490-1493 (1991)
 COMMENT NF-kappa-beta motifs are perhaps the best characterized sequence
 elements in the HIV-1 LTR (Nabel and Baltimore Nature 326:711-
 713). NF-kappa-beta directly impacts viral replication (Ross et
 al. J. Virol. 65:4350-4358; Kim et al. J. Virol. 67:1658-1662;
 Huang and Jeang J. Virol. 67:6937-6944; Huang et al. EMBO J. 13:
 2886-2896), viral transcription (Harrich et al. J. Virol. 63:
 2585-2591; Berkout and Jeang J. Virol. 63:5501-5504), and Tat
 transactivation (reviewed in Jones and Peterlin Annu. Rev. Biochem.
 63:717-743). For more extensive discussions of the biochemical and
 functional properties of NF-kappa-beta (Ghosh et al. Cell 62:
 1019-1029; Kieran et al. Cell 44: 261-272; Nolan et al. Cell 64:
 961-969; Liou et al. EMBO J. 11: 3003- 3009).
 NCBI gi: 189503
 BASE COUNT 354 a 626 c 470 g 317 t
 ORIGIN

```

1 gaattccggc gaattggctcgt tctgtagtgc acgcgcgggg cccagctggc accccggccc
61 cccccccggg accccggccc TGgacgaact gttccccctc atcttccggg cagagccagc
NFkB p65 subunit cds start ->
121 ccaggcctct ggcccttatg tggagatcat tgagcagccc aagcagcggg gcattgcgctt
181 ccgtctacaag tgcgagggggc gctccggggg cagcatccca ggcgagagga gcacagatac
241 caccagagac caccacacca tcaagatcaa tggctacaca ggaccagggg cagtgcgcat
301 ctccctgggc accaaggacc ctctcaccg gctcaccccc caggagcttg taggaaagga
361 ctgccgggat ggcttctatg aggtcagct ctgccgggac cgtcgcctcc acagtctcca
421 gaacctggga atccagtgtg tgaagaagcg ggaacctggg caggtctatca gtcagcgcct
481 ccagaccacac aacaacccct tccaggttcc tatagaagag cagcgtgggg actacgacct
541 gaatgctgtg cggctctgct tccaggtgac agtgcggggc ccacagggca ggccctccg
601 cctgcgcctt gtcttctctc atccatctt tgacaatcgt gcccacaaca ctgcgagct
661 caagatctgc cgagtgaacc gaaactctgg cagctgcctc ggtggggatg agatcttctt
721 actgtgtgac aaggtgcaga aagaggacat tgaggtgtat ttacagggac caggtcggga
781 ggcccaggc tcttcttcgc aagctgatgt gcaccgacaa gtggccattg tgttccggac
841 cctccctac gcagacccca gctgcaggc tctgtgtcgt gtctccatgc agctgcggcg
901 gcttccgac cgggagctca gtgagcccat ggaattccag tacctgccag atacagaaga
961 tcttcacggg attgaggaga aacgtaaaag gacatatgag accttcaaga gcatcatgaa
1021 gaagagtctt ttacagcggc ccacagacc cgggcttcca cctcgaacca ttgtgtgccc
1081 ttcccgagc tcagcttctg tccccagcc agcaccacag cctatccct ttactcctc
1141 cctgagcacc atcaactatg atgagtttcc caccatgggt ttctctctg ggagatcag
1201 ccaggcctcg gcttggccc cggccctctc ccaagtctg cccaggtct cagccctgc
1261 cctgtctcca gccatgggat cagctctggc ccaggcccca gcccctgtcc cagtctagc
1321 ccaggccct cctcaggtcg tggcccccac tgcccccagg cccacccagg ctggggaagg
1381 aacgtgttca gaggccctgc tgcagctgca gtttgatgat gaagacctgg gggccttgc
1441 tggcaacagc acagaccag ctgtgttcac agacctggca tccgtcgaca actccgagtt
1501 tcagcagctg ctgaaccagg gcatacctgt ggccccccac acaactgagc ccattgctgt
1561 ggagtaacct gaggtataaa ctgcctagc gacagggggc cagagggccc ccgaccagc
1621 tctgtctcca ctggggggcc cggggtctcc caatggctc ctttcaggag atgaagactt
1681 ctctccattt gggacatgg acttctcagc cctgctgagt cagatcagct cCTAAGgggg
<- NFkB p65
1741 tgacgcctgc cctccccaga gcactgg

```

HUMP68K

LOCUS HUMP68K 2562 bp ss-mRNA PRI 29-JAN-1991

DEFINITION Human p68 kinase complete cDNA sequence.

ACCESSION M35663

KEYWORDS p68 kinase.

SOURCE Human mRNA from interferon-treated Daudi cell.

REFERENCE 1 (bases 1 to 2562)

AUTHORS Meurs,E., Chong,K., Galabru,J., Thomas,N.S., Kerr,I.M., Williams,B.R. and Hovanessian,A.G.

TITLE Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon

JOURNAL Cell 62, 379-390 (1990)

COMMENT Human p68 kinase, also referred to as P1/dsI, DAI, and PKR, complexes with the double-stranded stem of TAR RNA (McCormack et al. Virology 188:47-56; Roy et al. J. Virol. 65:632-640). It is thought that TAR activates this kinase and results in the phosphorylation and down-regulation of eIF-2 (Garcia and Gaynor AIDS 8:S3-S17).

Draft entry and computer-readable sequence for [1] kindly submitted by E.Meurs, 22-JUN-1990.

NCBI gi: 189505

BASE COUNT 843 a 478 c 500 g 741 t

ORIGIN

```

1 cagtttcttg agcaaatcca gtttgccctc ctggatttgt aaattgtaat gacctcaaaa
61 ctttagcagt tcttccatct gactcaggtt tgcttctctg gcggtcttca gaatcaacat
121 ccacacctcc gtgattatct gcgtgcattt tggacaaagc ttccaaccag gatacgggaa
181 gaagaaATGg ctggtgatct ttcagcaggt ttcttcatgg aggaacttaa tacataccgt
p68 kinase cds ->
241 cagaagcagg gagtagtact taaatatcaa gaactgcta attcaggacc tccacatgat
301 aggaggttta catttcaagt tataatagat ggaagagaat ttccagaagg tgaaggtaga
361 tcaaagaagg aagcaaaaaa tgcgcagccc aaattagctg ttgagatact taataaggaa
421 aagaaggcag ttagtccctt attattgaca acaacgaatt cttcagaagg attatccatg
481 ggaatttaca taggccttat caatagaatt gccagaaga aaagactaac tgtaaattat
541 gaacagtgtg catcgggggt gcatgggcca gaaggatttc attataaatg caaatgggga
601 cagaagaagt atagtattgg tacaggttct actaacacag aagcaaaaca attggcgcgt
661 aaacttgcac atcttcagat attatcagaa gaaacctcag tgaactctga ctacctgtcc
721 tctggttctt ttgctactac gtgtgagtc caaagcaact ctttagtgac cagcacactc
781 gcttctgaat catcatctga aggtgacttc tcagcagata catcagagat aaattctaac
841 agtgacagtt taaacagttc ttctgtgctt atgaatggtc tcagaataaa tcaagggaag
901 gcaaaaagat ctttggcacc cagatttgac ctctctgaca tgaaggaaac aaagtatact
961 gtggacaaga ggtttggcat ggatttttaa gaaatagaat taattggctc aggtggattt
1021 ggccaagttt tcaaagcaaa acacagaatt gacggaaga cttacgttat taaacgtgtt
1081 aatatataata acggaaggcc ggagcgtgaa gtaaaagcat tggcaaaact tgatcatgta
1141 aatattgttc actacaatgg ctgttgggat ggatttgatt atgacctga gaccagtgat
1201 gattctcttg agagcagtga ttatgatcct gagaacagca aaaatagttc aaggtcaaag
1261 actaagtgcc ttttcatcca aatggaattc tgtgataaag ggaccttggg acaatggatt
1321 gaaaaaagaa gaggcgagaa actagacaaa gttttggctt tggaaactct tgaacaaata
1381 acaaaagggg tggattatat acattcaaaa aaattaatc atagagatct taagccaagt
1441 aatatattct tagtagatac aaacaaagta aagattggag actttggact tgaacatct
1501 ctgaaaaatg atggaagcgc aacaaggagt aagggaactt tgcgatatac gagcccagaa
1561 cagatttctt cgcaagacta tggaaaggaa gtggacctct acgctttggg gctaattctt
1621 gctgaacttc ttcatgtatg tgacactgct ttgaaacat caaagttttt cacagacctt
1681 cgggatggca tcatctcaga tatatttgat aaaaaagaaa aaactctctt acagaattta
1741 ctctcaaaga aacctgagga tcgacctaac acatctgaaa tactaaggac cttgactgtg
1801 tggagaaaaa gccagagaaa aatgaacga cacacatgtT AGagcccttc tgaaaaagta
<- p68 kinase cds end
1861 tctgcttctt gatatgcagt ttctcttaaa ttatctaaaa tctgctaggg aatatcaata
1921 gatatttacc ttttatttta atgttctctt taatttttta ctatttttac taatctttct
1981 gcagaaacag aaagggtttt tcttttttgc ttcaaaaaca ttcttacatt ttactttttc
2041 ctggctcatc tctttatttt tttttttttt ttttaagac agagctctgc tctgttgccc

```

```

2101 aggcctggagt gcaatgacac agtccttggct cactgcaact tctgcctctt gggttcaagt
2161 gattctcctg cctcagcctc ctgagtagct ggattacagg catgtgccac ccacccaact
2221 aattttttgtg tttttaataa agacaggggtt tcaccatgtt ggccaggctg gtctcaaact
2281 cctgacotca agtaatccac ctgcctcggc ctcocaaagt gctgggatta cagggatgag
2341 ccaccgcgcc cagcctcatc tctttgttct aaagatggaa aaaccacccc caaatTTTTt
2401 tttatacta ttaatgaatc aatcaattca tatctattta ttaaatttct accgctttta
2461 ggccaaaaaa atgtaagatc gttctctgcc tcacatagct tacaagocag ctggagaaat
2521 atggtactca ttaaaaaaaa aaaaaaaaag tgatgtacaa cc

```

HSZFPEBF1

LOCUS HSZFPEBF1 9020 bp RNA PRI 14-MAR-1990
 DEFINITION Human PRDII-BF1 sequence, clone PRDII-BF1.
 ACCESSION X51435
 SOURCE Human lambda gt11 library.
 REFERENCE 1 (bases 1 to 9020)
 AUTHORS Fan, C.M. and Maniatis, T.
 TITLE A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence
 JOURNAL Genes Dev. 4, 29-42 (1990)
 COMMENT Human PRDII-BF1, also referred to as HIV-EP1 or MBP-1, is a cellular transcription factor which activates HIV-1 gene expression. The protein contains two zinc finger domains, each of which can bind to either of the two NF-kappa B binding motifs in the HIV-1 enhancer or to the TAR transactivation response element. Gaynor et al. (Genomics 9: 758-61, 1991) have mapped the gene which encodes HPRDII-BF1 to human chromosome 6p22.3-p24.
 NCBI gi: 38017

BASE COUNT 2745 a 2164 c 1930 g 2181 t
 ORIGIN

```

1  cgctgctggg  cggcggcagg  gtggcggagc  gagcggggga  cgggggagcg  gcgggcgcga
61  ggagggttatg  tttgtgtttg  gggttgtcaa  gtgaaggagg  gatcccaggc  gccgcgcgcg
121  ccgcgcgcggg  ggtcgcgaga  tcccgagccg  cggccgcgcg  catcagcagc  gcagtcaggg
181  gccggctgca  gccgcagctc  cgcggggcgt  cctggcagca  gcacatggat  taattgatgt
241  atgttgagtt  tatggagctg  ccttttggtg  gcttgcttta  tctgcagttt  ttaagaagaa
301  aaagaaggcc  ctgagtcaaa  gaagATGcct  cgaactaacc  aaattcatcc  cagaaatcta
      PRDII-BF1 cds start ->
361  agagacaaaa  ttgaagaagc  acaaaaagaa  cttaatgggg  cagaagtttc  aaaaaaagaa
421  atcttacagg  ctgggtgtta  aggaacttcg  gaatccctta  aaggtgtgaa  acgcaaaaag
481  atcgtagctg  agaatacact  gaaaaaata  ccaaaatccc  cactgagaaa  tctctctcag
541  gcaaaacata  aacaaaatc  agaagagtc  tctttgcgcg  ttcttcatag  tgcctcggag
601  tctcacaga  aacagaatta  tattcctgta  aaaaatggga  agcagtttac  caaacaatat
661  ggagaaacac  ctggaataat  tgctgaagcc  tcaaaatctg  aagaatctgt  ctccccaaag
721  aagcccttgt  ttctgcagca  accatctgaa  ctgcgtagat  ggagatccga  aggcgtgat
781  cctgccaat  tcagtgcct  cgatgaacaa  tgtgactcaa  gttccttgtc  aagtaaaacc
841  aggactgaca  atagcgaatg  catctcttct  cattgtggca  ctacgtcccc  ctctataca
901  aacactgcat  tcgatgtctt  actgaaagca  atggagccag  aactgagcac  cttgtcacia
961  aagggctcac  cttgtgcaat  taagacagaa  aaactgaggg  caaataaaac  tgcacgttcc
1021  cctcccaaat  taaaaaacag  ttcaatggat  gcccacaatc  agacttcaca  ggaattgggt
1081  gctgaateac  agtcttcttg  tacctcatat  acagtccata  tgtctgtctg  tcagaagaat
1141  gagcaagggg  caatgcagtc  agcttctcat  ttgtatcatc  aacatgaaca  ctttgttccc
1201  aaatccaacc  aacataatca  acagcttcg  ggggtgttcag  gtttcacagg  atcactgaca
1261  aatctgcaaa  atcaagagaa  tgccaaactt  gaacagggtt  ataataatgc  agtgacatca
1321  tctgtaggcc  taacttcacc  ttccagtata  tctcagggtt  ctctcaaaa  ccagcaaatg
1381  gattctgctt  cacttttgtc  aataagtcg  gctaattcta  cacagtcgcc  ccccatgcca
1441  atctataatt  caactcatgt  tgctctgttt  gttaatcaaa  gcgtagagca  aatgtgcaat
1501  cttctctctg  aagatcagaa  gccaaaaaaa  caaggaaaat  atatttgtga  gtattgcaat
1561  agagcatgtg  caaagcctag  tgtgctttta  aagcatatcc  gctccacac  tggagagcga
1621  cctatccct  gtgtgacttg  tggattttca  ttttaagacta  aaagtaatct  gtataagcac
1681  aagaaatccc  acgcacatac  tatcaaaactg  ggtcttgtct  tgcaaccaga  tgcgtgtggc
1741  ttgttcttgt  cccacgagtc  ccccaaagca  cttagtattc  attcagacgt  agaagacagt
1801  ggggagagcg  aggaggaagg  cgcactgat  gagagacagc  atgacctggg  cgcctgggag
1861  ctgcagaatg  tgcacataat  aaagaggatg  tcaaatgctg  aaactttact  aaaatcaage
1921  ttcactccaa  gcagtcacga  aaatgtgata  ggtgactttt  tgtacagga  cagactgtga
1981  gaatcacaa  ctgtgacaga  gttaccgaaa  gttgtggtcc  accatgtcac  tgtgtcccc
2041  ttaagaactg  acagtcacaa  ggccatggat  cccaagcctg  aactttctag  tgcacaaaag
2101  cagaaggacc  ttcagggtgac  aaacgtacag  ccactttcag  ccaacatgtc  ccagggtgga
2161  gtctccaggt  tggagactaa  tgagaattcc  caccagaaag  gcgacatgaa  tccactggaa
2221  ggaaagcaag  actctcacgt  aggaacggta  caagccagc  tacaaggga  gcagggtacc
2281  gattactccc  aagagcagca  aggaagctc  ctgagtcctc  gaagtttagg  aagtaaggat

```

2341 tctgggttact tttcacgttc tgaaagtgc gatcaaacag tgagtccacc aactcccttt
 2401 gccagaaggt tcccagcaca gaacaagact ctgggaaggag taaggagccc tctgcagetc
 2461 ttgtcaccac gtcaacaccc tctgctttgc cacagggaaa aggcatttget tttaccaggt
 2521 cagatgcgcc caacctttggc cacaanaaca cttgaggagc ggatatcgaa gcttatctca
 2581 gacaatgaag ctttggtaga tgacaagcaa ctggatagtg tgaagccgcg gagaacctca
 2641 ctgtcaagac gaggaagcat tgattccccc aatcataca tatttaaga ttctttccag
 2701 tttgatttaa aaccagtggg acggagaaca agttcaagct ctgatatacc gaagtcacct
 2761 ttaccccta ctgaanaac aaagcaagt tttctctgt ctgtaocttc acttgactgt
 2821 ttacctatca caagaagtaa ttccatgcgc accacaggtt attcagcagt acctgcaaat
 2881 ataatacttc ctccatcatc actaagagga agtcagtcac ttgatgacaa aattggcgct
 2941 ttctatgatg atgtctttgt atcgggacct aacgctcttg tgccccagag tgggcatccc
 3001 cgtacacttg tgagacaagc agccatagaa gactcttcag caaatgaaag tcatgttctt
 3061 ggtactggac agtccctgga tgagagccac caaggatgac atgtctgttg tgaagccatg
 3121 tcaagtggga gcaaggcact ggcacaagc ccacatatag aaaaaaaga gtctcatcaa
 3181 gggcgaggga caatgtttga gtgtgaaact tgtagaanaa ggtataggaa actggaaaaat
 3241 tttgaanaac ataagaatt ttactgttct gaggtaacat gacaaaaaac aaaggtagcc
 3301 atgagagaa ctgagcacag cctgtgccc ggcggtctgc aacctcagat tctacactac
 3361 agagtgcctg ggtccctcgg catctgggaa cagacgcccc agataagaaa aaggaggaaa
 3421 atgaaaagt tggggatga tgaagaact cagcaaatg aaagtggaa atctccaaa
 3481 agttctgaag gccttcagtt tcagaatgct ctgggtctga atccagttt gcctaaacat
 3541 agtgttacca taagaagtga ccagcagcat aaaaatatac agttgcaaaa ctcccatatt
 3601 caocttgttg ccaggggccc tgagcagacc atggatcccc agctgtcgac catcatggaa
 3661 caacagataa gttcagcagc ccaggacaag atagaactgc agagacacgg aactggaatc
 3721 tctgtcatcc agcacacca ctccctgagc agcccaact cattedgaaa gectgagcct
 3781 tttgaagag cctccocagt ttctttccag gactgaata gaacggggaa ttccgggtct
 3841 ctanaagtg taggaatctc ccaagaggaa agtcacccct ctgggagcgg gtctcatcct
 3901 caacagcttg cactatcaga cgtctcaga ggagaacttc aggaagctc cagaagaggt
 3961 caagtgaac gacatgtgtt aggacagccc tcaagactta tccggcagca caacatccaa
 4021 gttccagaga ttttggcac agaagaacca gatcgagacc tggaaagctc atgccatgat
 4081 caagaaaagt cagagaagtt cagttggccc cagcgtagtg aaacctgtc aaaattgcca
 4141 acagagaaac tgccacccaa aaagaaaagg ctccgtctgg ctgagataga acattcctca
 4201 acagaatcga gctttgatcc cactctctcc aggagtctaa gtaggagag cagtttatct
 4261 cacactcaa gtttctcagc ctcttttagc atagaggagc ttctaaaaac ggaggcttcc
 4321 cccaaaatcg attttctaaa taaagccgag ttctttatga ttccagctgg cttgaatact
 4381 ctgaatgttc ctggatgtca ccgggaatg aggcgtactg catcagaaca gattaattgc
 4441 accgaaacgt caatggaggt ctctgatctc agaagcaaat cattogattg tggagcctc
 4501 accccacccc agacaacacc acttactgaa ttgcagcctc catcttcacc ttctcgagt
 4561 ggagtgtctg ggcattgtgc tctcttagaa agaaggagag gccactggt accggcaata
 4621 tctttgggca tagcccccag tagtcatctg tctctgtac acccaacatc ttccaaaat
 4681 actgctcttc ccagtgtgaa tgcagtcca tatcagggc ctccagctcac tagtacatct
 4741 ttagctgagt tttctgcaa tactttgcac tctcagactc aggttaagga tctgcaggca
 4801 gaacatcaa actccagctc taccacagtt ttctctgttc aacagctctg tgatatcaat
 4861 ttgttaaatc aaatccatgc accgcctagc caccagagca cadagctatc tctgcaagt
 4921 tctacgcagg gtagcaagcc agataanaat tctgttttat ctgggtcttc taaaagttag
 4981 gattgctttg ctcccaata ccaattgcac tctcaggttt tcaactcagg cccatcttgc
 5041 tcttctaate ctgtgcattc tttgccaaat caagttattt cagatccagt tggacagat
 5101 cattgtgtga catcagcaac attaccaacc aaattaattg acagcatgtc taattcgcac
 5161 cctctgtctc caccagagct caggcccttt ggaagtccag tgcaagaagt gccatcatca
 5221 ttcatgctgc ccatacgcct gcagagttag gtctctgctt actgttttgc taaactcaca
 5281 tccctgccac aaatactagt gacccaagat ctgccccatc agccaatttg ccagactaat
 5341 catagtgtag tgccaatcag tgaagaacaa aattctgtgc caacattaca aaaaggtcat
 5401 cagaatgctt tgccaaaccc agagaaggaa tttctatgtg aaatgtttt ttcagagatg
 5461 agccaaaatt ctctctctac agaatecttg ccataaactc agaaaatata tgttggtcga
 5521 ctttccctc aacaagaatc ttacgtctcg agtaaaagga tgctttcccc agcaaatagt
 5581 ttagacattg ccattggaaa gcaccagaag cgggccaag atgaaaatgg agctgtttgt
 5641 gcaacagacg tgagaccttt agaggctttg agttcgagag ttaatgaagc tagtaaacag
 5701 aagaagccta ttttagtgag acaggtttgt actacagagc cctggagcgg tbtgatgttg
 5761 aaaaaggatg tttttctca acctgaaact agtaaatgag ctgttaattt gacaaatgtt
 5821 ttaccagctg ataattcatc aacggatgc tctaaatttg tcttataga acctataagt

```

5881 gaattgcagg aatttgaaaa catcaagtc tccacatcat taactottac agttcgaagt
5941 tcaactgctc ettcagaaaa tactcatott tctcctttga aatgtacaga caataacca
6001 gaaaggaagt ctcagggggt taaaaatcaa ggtgacaaag tgaacatcca agagcaaagt
6061 caacggccag tcacttctct ttcatgtttt aacatcaagg acaccagca gctgggtttc
6121 cctagcctga aaactacaac caactttaca tgggtgttatc tcttaaggca gaagtcgttg
6181 cttttgcctc agaaggacca gaaaacttca gcttatactg attggacagt aagcgccagt
6241 aatccaaatc cactcggttt gccacaaaaa gttgcaattg ctctccttaa tccaaaacag
6301 aacactggaa aatcactata ctgtcaagca ataactacc attccaagtc agacttattg
6361 gtctattcaa gcaagtggaa aagcagctta agcaagagag cattaggtta tcaaaagtcc
6421 acagttagtg aattcagcaa taagatgcc tctgaaatta acagtggca agataaagaa
6481 aattccttaa tcaaaagtga accaagaaga attaaaatat ttgatggagg atataagtc
6541 aatgaagagt atgtatatat ccgaggcagg ggaagaggaa aatacatttg tgaagaatgt
6601 ggaatacgtt gtaagaaacc tagcatgtta aagaaacaca taagaaacca tacagatgtc
6661 cgccctacc actgcactta ctgtaacttc tcttttaaga ctaaaggaaa tctgacaaaa
6721 cacatgaagt ccaaggcaca tagcaagaaa tgtgtggatt taggcctctc agtaggttta
6781 atagatgaac aggatacaga agaatcagat gaaaaacaga gattcagtta tgagcgatct
6841 ggatgatgac ttgaagaatc tgatggccca gatgaggatg acaatgaaaa tgaagacgat
6901 gatgaggaca gccaggctga atcagtcctg tcagccacac cctcagtcac agctagcccc
6961 cagcaccttc catctagaag tagccttcag gacctgtga gtactgacga ggaatgtcagg
7021 atcaccgatt gctttctctg ggtacacacg gacccaatgg acgttctgcc caggggcctg
7081 ctccacagaa tgactgtcct gaggcacagca cagtctgact acaataggaa gacactctct
7141 ccgggggaagg ccaggcagcg tgctgcgaga gatgaaaacg acacaattcc gtctgtagac
7201 acttccaggt cccgtgtca tcagatgtct gtggactacc ctgagtcaga agaaattctg
7261 agaagttcta tggcaggaaa agctgttgot ataacacaga gccatcctc tgtaagactt
7321 cctcctgctg cagctgagca cagcccccag acagcagcgg ggaatgcctc tgtggcctca
7381 ccacatcctg accctcaaga acagaagcag caaataactc tacagccgac tccaggettg
7441 ccttctcccc acactcattt gtttagccac ctctctttgc attcccagca gcaatcgagg
7501 acacottata atatgggttc agttgggggg atccatgtgg taacctgctg cctcacatac
7561 tccaggtttg tgcccttca ggtggacca gtgcagctca cgaaccctgc tgtcagtgte
7621 gttcacagaa ctttgggtac tcataggaat acggtcacag aagtgtctgg cactacaaac
7681 cctgctggag tggctgaatt aagcagtggt gtgccatgta ttctatctgg ccaaatccgc
7741 gtgccaggcc ttcagaacct aagtacccca ggcttgcaat cactccctc gttaaagcatg
7801 gaaaccgtca atattgtagg cctagccaat acaaatatgg cccacaaagt ccatccacca
7861 ggaactggctc tgaatgtgtt cggactgcag gttctgactg caaaccttc atcacaaagc
7921 agccccgccc ctcaggcaca cattccaggt ctccagatct tgaacatagc attgcccacc
7981 ttaatcccc cagtcagtea agtagccgtt gatgcacagg gagctccaga aatgccagct
8041 tcccaaagca aagcatgcga gacacaacc aagcagactt ctgtagccag cgcaaacag
8101 gtcagcagga ccgagtctcc tcagggggta cctacagtcc agcgggaaaa tgcaaaaaaa
8161 gttctgaatc caccctgccc tgcagggtgac catgcaagga ttgatggcct gagtaaaatg
8221 gacacagaga aggtgtcctc ggcaaatcac gtgaagccca agcctgaact cacttccata
8281 cagggccaac cagcgtccac gtcacaaact ctgctgaagg cacattctga agtttttaca
8341 aagccctcag gccagcagac tctctctcca gacagacagg ttcccaggcc cacaggacta
8401 ccgaggaggc agcccactgt gcacttcagc gacgtgagca gcgatgatga cgaggacagg
8461 cttgtgatag caaccTGAlg gatttttatt ttattttgct ttttttttat ataacttta
<- PRDII-BF1 cds end
8521 aagggtttctt tgaaaaacct cctttcotta aagcacattt ttctgacata aactcatgac
8581 taatctttgt gcaatcatga acttttgacc aataattgtt gttttgtgtc agctccagcc
8641 atttttgtac atgttgtata gacaattgtg ccttttagga gctttatgtt tagaaactgt
8701 acagattgtt gaatatctat atacataaaa atatatata tatgtatatg aaaaccaggt
8761 agttatctgt gtttagtaag gaaaaacctt caataaatac aatgattaa attatatgtt
8821 ccactgttga atataaattt tatggctatg gggcagagtt tctgtgtata aattagtatg
8881 taactccat atttatgtat tcataattagt ctttgaatat gggctctgtc tcttctgtga
8941 agacagtaac tttaacttcc agacagattt tctgtgttat gaaatgttcc agtaaaatat
9001 tgtttactga cttttaaaaa

```


LOCUS HUMSF2P32 1240 bp ss-mRNA PRI 13-JAN-1993
 DEFINITION Human pre-mRNA splicing factor SF2p32 complete sequence.
 ACCESSION M69039
 KEYWORDS SF2p32.
 SOURCE Homo sapiens Female cDNA.
 REFERENCE 1 (bases 1 to 1240)
 AUTHORS Krainer,A.R., Mayada,A., Kozak,D. and Binns,G.
 TITLE Functional expression of cloned human splicing factor sf2: Homology to RNA-binding proteins, U1 70K, and drosophila splicing regulators
 JOURNAL Cell 66, 383-394 (1991)
 COMMENT SF2p32 is a polypeptide from HeLa cells that copurifies with the essential pre-mRNA splicing factor SF2. It may be a subunit of SF2, although its function is presently unknown.
 SF2p32 is 92% identical to the mouse YL2 protein and is thought to be the human homolog. Recently, YL2 has been shown to bind to the basic domain of the HIV-1 Rev protein and through this binding helps mediate its role (Luo et al. J. Virol. 68:3850-3856). It is hypothesized that Rev either inhibits RNA splicing or directly transports singly and unspliced HIV-1 transcripts out of the nucleus.
 Purification of this protein from HeLa cells is described in Krainer, A.R., Conway, G.C., and Kozak, D. (1990) Purification and Characterization of SF2, a Human Pre-mRNA Splicing Factor. Genes Dev. 4, 1158-1171.
 NCBI gi: 338042

BASE COUNT 326 a 297 c 299 g 318 t

ORIGIN

```

1 ctgcgctgcg tgcctcgtgt gctggggtcc tccgtgcgcg gcttcgcgcg tgcgcgcgcc
SF2p32 ->
61 gctcgccttt tccggcagct cctgcagcgc gcaccccggc tgtgcaccgc gcccttcggg
121 ctgctcagcg tgcgcgcagc ttcgagcgcg cggccgggccc tccctgcggcc tcgaggaccc
181 tgcgcctgtg gctgtggctg cgctcgcctg cacaccgacg gagacaaagc ttttgttgat
241 ttcttgagtg atgaaattaa ggaggaaaaga asaattcaga agcataaaac cctccctaag
301 atgtctggag gttgggagct ggaactgaat gggacagaag cgaaattagt gcggaaagtt
361 gccggggaaa aaatcacggt cactttcaac attacaaca gcctccacc aacatttgat
421 ggtgaggagg aacctcgca agggcagaag gttgaagaac aggagcctga actgacatca
481 actcccaatt tcgtggttga agttataaag aatgatgatg gcaagaaggc ccttgtgttg
541 gactgtcatt atccagagga tgaggttga caagaagacg aggcctgagag tgacatcttc
601 totatcaggg aagttagctt tcagtcactt ggcgagtcct aatggaagga tactaattat
661 aactcaaca cagattcctt ggaactgggc ttatatgacc acctaatgga tttccttgcc
721 gaccgagggg tggacaacac ttttgcagat gagctggtgg agctcagcac agccctggag
781 caccaggagt acattacttt tctgaagac ctcaagagtt ttgtcaagag ccagTAGagc
                                     <- SF2p32
841 agacagatgc tgaaagccat agtttcatgg caggcttttg ccagtgaaca aatcctactc
901 tgaagctaga catgtgcttt gaaatgatta tcctcctaatt atcatggggg aaaaaatacc
961 aaatttaaat tatatgtttt gtgttctcat ttattatcat tttttctgt acaaatctat
1021 tattctctaga tttttgtata acatgataga cataaaattg gtttatctcc tgetgagttc
1081 tcttagaagt acctgctggg ttctcagttc cagttcccat cctttgattg atcactctcc
1141 ctgacatcca cctgtatgac tttgtcacca aatgctgtct tctctttctc caatcaagaa
1201 ataataatcc ctcgagttta caaaacaaaa aaaaaaaaaa

```

HUMSP1TF

LOCUS HUMSP1TF 2655 bp ss-mRNA PRI 30-SEP-1988
 DEFINITION Human transcription factor Sp1 cDNA sequence, 3' end, clone Sp1-1.
 ACCESSION J03133
 KEYWORDS transcription factor SP1.
 SOURCE Human mRNA from HeLa cell line (library of C.Hauser).
 REFERENCE 1 (bases 1 to 2655)
 AUTHORS Kadonaga,J.T., Carner,K.R., Masiarz,F.R. and Tjian,R.
 TITLE Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain
 JOURNAL Cell 51, 1079-1090 (1987)
 COMMENT The Sp1 motifs are perhaps the best characterized sequence elements in the HIV-1 LTR. These sequences directly impact viral replication (Ross et al. J. Virol. 65:4350-4358; Kim et al. J. Virol. 67:1658-1662; Huang and Jeang J. Virol. 67:6937-6944; Huang et al. EMBO J. 13:2886-2896), viral transcription (Harrich et al. J. Virol. 63:2585-2591; Berkhout and Jeang J. Virol. 63:5501-5504), and Tat transactivation (reviewed in Jones and Peterlin Annu. Rev. Biochem. 63:717-743). For more extensive discussions of the biochemical and functional properties of Sp1 (Dyran and Tjian Cell 32:669-680; Briggs et al. Science 234:47-52; Kadonaga et al. Cell 51:1079-1090).
 Draft entry and computer-readable sequence for [1] kindly provided by J.T.Kadonaga, 17-NOV-1987.
 NCBI gi: 339517

BASE COUNT 717 a 726 c 630 g 582 t

ORIGIN 540 bp upstream of HincII site.

1 acaggtgagc ttgacctcac agccacacaa ctttcacagg gtgccaatgg ctggcagatc
 Sp-1 cds ->

```

61 atctcttctt cctctggggc tacccttacc tcaaagggaac agagtggcag cagtaccaat
121 ggcagcaatg gcagtgagtc ttccaagaat cgcacagtc ctggtgggca gtatgtgtg
181 gctgccgctc ccaacttaca gaaccagcaa gttctgacag gactacctgg agtgatgcct
241 aatattcagt atcaagtaat cccacagttc cagaccgttg atgggcaaca gctgcagttt
301 gctgccactg gggcccaagt gcagcaggat ggttctggtc aaatacagat cataccaggc
361 gcaaaccaac agattatcac aaatcgagga agtggaggca acatcattgc tgctatgcc
421 aacctactcc agcaggctgt cccctcccaa ggctctggta ataattgact ctacaggacg
481 actcagtatg tgaccaatgt accagtggcc ctgaatggga acatcacctt gctacctgtc
541 aacagcgttt ctgcagctac cttgactccc agctctcagg cagtcacgat cagcagctct
601 gggctccagg agagtggctc acagcctgtc acctcaggga ctaccatcag ttctgccagc
661 ttggtatcat cacaagccag ttccagctcc tttttacca atgccaatag ctactcaact
721 actactacca ccagcaacat ggggaattatg aactttacta ccagtggatc atcagggacc
781 aactctcaag gccagacacc ccagagggtc agtgggctac aggggtctga tgctctgaac
841 atccagcaaa accagacatc tggaggctca ttgcaagcag gccagcaaaa agaaggagag
901 caaaaccagg agacacagca gcaacaaatt cttatccagc ctacagctagt tcaaggggga
961 caggccctcc aggcctcca agcagcacca ttgtcagggc agacctttac aactcaagcc
1021 atctcccagg aaaccttcca gaacctccag cttcaggctg ttccaaactc tggctccatc
1081 atcatccgga caccaacagt ggggcccatt ggacaggtca gttggcagac tctacagctg
1141 cagaacctcc aagttcagaa ccacaagcc caaacaatca ccttagcccc aatgcagggt
1201 gtttctctgg ggcagaccag cagcagcaac accactctca caccattgc ctacgtctgt
1261 tccattctct ctggcacagt cactgtgaat gctgtcaac tctctccat gccaggctc
1321 cagaccatta acctcagtgc attgggtact tcaggaatcc aggtgcacc aattcaagge
1381 ctgccgttgg ctatagcaaa tgcctccagg gatcatggag ctacagcttg tctccatggg
1441 gctggcgttg atggaatata tgatgacaca gcaggtggag aggaaggaga aaacagccca
1501 gatgcccaac cccaagccgg tcggaggacc cggcgggaag catgcacctg cccctactgt
1561 aaagacagtg aaggaagggg ctgggggat cctggcaaaa agaaacagca tatttgcac
1621 atccaaggct ggggaaagt gtatggcaag acctctcacc tgggggcaac cttgcgctgg
1681 catacaggcg agaggccatt tatgtgtacc tggctcact gtgggaaacg cttcacacgt
1741 tgggtgagc tacagaggca caaacgtaca cacacaggtg agaagaaatt tgccctgcc
1801 cagtgctcta agcgttcat gaggagtgc cactgtcaa aacatatcaa gaccaccag
1861 aataagaagg gagggccagg tgtagctctg agtgtgggca ctttgcctcc ggacagtggg

```

```

1921 gcagggttcag aaggcagtg cactgccact ccttcagccc ttattaccac caatatggta
1981 gccatggagg ccctctgtcc agagggcatt gcccgctctg ccaacagtgg catcaacgtc
2041 atgcagggtgg cagatctgca gtccattaat atcagtggca atggcttctg Agatcaggca
                                     <- Sp-1 ods end
2101 cccgggggcca gagacatatg ggccataccc cttaaccccg ggatgcaagg tagcatgggt
2161 ccaagagaca tggagagag agccatgaag cattaaaatg catgggtgtg agaagaatca
2221 ggagagggat acaagagagg agatggggtc cgggcacca tctgtatcat cagtgcctct
2281 ttgaagggtg gaaacattag tgaaaattct gttgggtgcca cgtcttgatg agcatttggt
2341 tgaccccaagt ttcttcttac acttcttacc ccagccctacc ctctctgcat ttctctctct
2401 agctcttcca tgatggattc ccccccttt cctaaagcca tcatgccttg ataaatatat
2461 atgatcattg aaatactttt taacaaaaaa cagattctat attattatat atatatatat
2521 atataaagat atatagagat gcattcacag ggttggctgg agaggaagac cattctgtga
2581 ccaaaatacc ttggtcattt tttttatatt gctttatttc cctatggctg agccttgttg
2641 tgacacatca agctt

```

HUMTBPI

LOCUS HUMTBPI 1341 bp ss-mRNA PRI 14-MAY-1990
 DEFINITION Human immunodeficiency virus tat transactivator binding protein-1 (tbp-1) cDNA sequence, complete cds.
 ACCESSION M34079
 KEYWORDS Tat-binding protein.
 SOURCE Human mRNA from Jurkat T-cell line.
 REFERENCE 1 (bases 1 to 1341)
 AUTHORS Nelbock, P., Dillion, P.J., Perkins, A. and Rosen, C.A.
 TITLE A cDNA corresponding to a protein that interacts with the human immunodeficiency virus Tat transactivator
 JOURNAL Science 248, 1650-1653 (1990)
 COMMENT TBP1 is part of a large family of proteins related to the 26S protease from human erythrocytes. Each of these proteins has Tat-binding activity [1]. The exact role of this family of proteins in cellular metabolism is not wholly clear; however, the 26S protease seems to regulate the degradation of some cyclins and in this manner has been implicated in modulating the stability of oncoproteins such as c-Mos, c-Myb, c-Myc, and p53 (Dubiel et al. Biol. Chem. Hoppe-Seyler 375: 237-240). Draft entry and computer-readable [or printed] sequence for (1) kindly submitted by C.A.Rosen, 08-MAY-1990.
 NCBI gi: 338699
 BASE COUNT 350 a 337 c 412 g 242 t
 ORIGIN

```

1 gaattccggc gacogtgtgg gatgaggcgg agcaagatgg aattggggag gaggtgctca
61 agATGtccac ggaggagatc atccagcgca caoggtctgt ggacagtgg atcaagatca
TBP-1 cds ->
121 tgaagagtga agtgttgaga gtcacccatg agctccaagc catgaaggac aagataaaag
181 agaacagtga gaaaatcaaa gtgaacaaga ccctgccgta ccttgtctcc aacgtcatcg
241 agctcctgga tgttgatcct aatgaccaag aggaggatgg tgccaatatt gacctggact
301 cccagaggaa gggcaagtgt gctgtgatca aaacctctac acgacagacg tacttccctc
361 ctgtgattgg gttggtggat gctgaaaagc taagagcagg agacctggtg ggtgtgaaca
421 aagactccta tctgatacct gagacgctgc ccacagagta tgactcgccg gtgaaggcca
481 tggaggtaga cgagaggccc acggagcaat acagtgaacat tgggggtttg gacaagcaga
541 tccaggagct ggtggaggcc attgtcttgc caatgaacca caaggagaag tttgagaact
601 tggggatcca acctccaaaa ggggtgctga tglatgggcc ccaggggacg gggaagaccc
661 tcttggcccg ggctgtgccc gcacagacta agggccacct cctaaagctg gctggccccc
721 agctggtgca gatgttcatt ggagatgggt ccaagctagt ccgggatgcc tttgccctgg
781 ccaaggagaa agcgccctct atcatcttca ttgatgagtt ggatgccatc ggcaccaagc
841 gctttgacag tgagaaggct ggggacccgg aggtgcagag gacaatgctg gagcttctga
901 accagctgga tggcttccag cccaacaccc aagttaaggt aattgcagcc acaaacaggg
961 tggacatcct ggaccccgcc ctctccgct cgggcgcgct tgaccgcaag atagagttcc
1021 cgatgcccac tgaggaggcc cgggccagaa tcatgcagat ccactccga aagatgaatg
1081 tcagtctctg cgtgaactac gaggagctgg ccgctgcac agatgacttc aatggggccc
1141 agtgcaaggg tgtgtgtgtg gagggcgggc tgatgcact gggcaggggt gccacggagc
1201 tcacccacga ggactacatg gaaggcatcc tggaggtgca ggccaagaag aaagccaacc
1261 tacaatacta cgccTAGgca cacaggccag cccagtcctc acggctgaag tgcgcaataa
<- TBP-1 cds end
1321 aagatggttt agggggaatt c

```

LOCUS HUMTFIIDA 1035 bp ss-mRNA PRI 15-SEP-1990
 DEFINITION Human transcription factor IID cDNA sequence, complete cds.
 ACCESSION M34960
 KEYWORDS transcription factor IID.
 SOURCE Human mRNA from Hela fibroblast cell line.
 REFERENCE 1 (bases 1 to 1035)
 AUTHORS Kao,C., Lieberman,P., Schmidt,M., Zhou,Q., Pei,R. and Berk,A.J.
 TITLE Cloning of the human TATA binding factor: Expression of a transcriptionally active TFIID protein
 JOURNAL Science 248, 1626-1626 (1990)
 COMMENT There is evidence that Tat contacts TFIID (Kashanchi et al. Nature 367:295-299) and that Tat-Spl-TFIID could be a multiprotein complex inside cells (Huang et al. EMBO J. 13:2886-2896).
 Draft entry and computer-readable sequence kindly submitted by C.C.Kao, 05-JUN-1990.
 NCBI gi: 339493

BASE COUNT 304 a 275 c 236 g 220 t
 ORIGIN

```

1 ATGgattcaga acaacagcct gccaccttac gtcagggtt tggcctcccc tcagggtgcc
TFIID cds ->
61 atgactcccg gaatccctat ctttagtcca atgatgcctt atggcactgg actgacocca
121 cagcctattc agaacaccaa tagtctgtct attttgsaag agcaacaaag gcagcagcag
181 caacaacaac agcagcagca gcagcagcag cagcagcaac agcaacagca gcagcagcag
241 cagcagcagc agcagcagca gcagcagcag cagcagcagc aacaggcagt ggcagctgca
301 gccgttcagc agtcaacgtc ccagcaggca acacagggaa cctcaggcca ggcaccacag
361 ctcttccact cacagactct cacaactgca cccttgccgg gcaccactcc actgtatccc
421 tcccccatga ctcccatgac ccccatcact cctgccacgc cagcttcgga gagttctggg
481 attgtacggc agctgcaaaa tattgtatcc acagtgaatc ttgggtgtaa acttgacctt
541 aagaccattg cacttcgtcg ccgaaacgcc gaataaate ccaagcgggtt tgctgcggta
601 atcatgagga taagagagcc acgaaccacg gcactgattt tcagttcttg gaaaatgggtg
661 tgcacaggag ccaagagtga agaacagtcc agactggcag caagaaaata tgctagagtt
721 gtacagaagt tgggttttcc agctaagttc ttggacttca agattcagaa catggttggg
781 agctgtgatg tgaagtttcc tataaggtta gaaggccttg tgctaccca ccaacaattt
841 agtagttatg agccagagtt atttctgggt ttaatttaca gaatgatcaa acccagaatg
901 atcaaaccca gaattgttct ccttattttt gtttctggaa aagttgtatt aacaggtgct
961 aaagtcagag cagaatttta tgaagcattt gaaaacatct acctattct aaagggattc
1021 aggaagacga cgTAA

```

<- TFIID cds end

HSTGFB1

LOCUS HSTGFB1 2745 bp RNA PRI 06-JUL-1989
 DEFINITION Human transforming growth factor-beta (TGF-beta) sequence.
 ACCESSION X02812 J05114
 KEYWORDS signal peptide; transforming growth factor-beta.
 SOURCE Human.
 REFERENCE 1 (bases 1 to 2745)
 AUTHORS Derynck,R., Jarrett,J.A., Chen,E.Y., Eaton,D.H., Bell,J.R.,
 Assoian,R.K., Roberts,A.B., Sporn,M.B. and Goeddel,D.V.
 TITLE Human transforming growth factor-beta complementary DNA sequence
 and expression in normal and transformed cells
 JOURNAL Nature 316, 701-705 (1985)
 COMMENT TGF-beta is a cellular cytokine which has been shown to inhibit
 HIV-1 gene expression (Garcia and Gaynor AIDS 8:S3-S17).
 The sequence downstream from the translational stop codon
 (2015-2017) was taken from a genomic clone.
 NCBI gi: 37092
 BASE COUNT 527 a 938 c 801 g 479 t
 ORIGIN
 1 acctccctcc gccgagcagc cagacagcga gggccccggc cgggggcagg ggggacgcc
 61 cgtccggggc acccccccg gctctgagcc gcccgcgggg cgggectegg cccggagcgg
 121 aggaaggagt cgcgaggag cagcctgagg cccagagtc tgagacgagc cgcgcgcgc
 181 cccgccactg cgggaggag ggggaggagg agcgggagga gggacgagct ggtccggaga
 241 agaggaaaaa aacttttgag acttttccgt tgcgcctggg agccggaggc ggggggacct
 301 cttggcgoga cgtgccccc caggaggcca ggacttgggg accccagacc gcctcccttt
 361 gccgccccgg acgcttgcct cctccctgcc cctacacagg cgtccctcag gcgccccat
 421 tcgggaccag cctcggggag tcgcgaccc ggctccccc aaagactttt ccccagacct
 481 cgggcgcacc cctgcacgc cgccttcact cccggcctgt cctctgagcc cccgcgcac
 541 ctgaccctt tctcctccag gagacggatc tctctccgac ctgccacaga tccctattc
 601 aagaccaccc accttctggg accagatcgc gccatctag gttatttccg tgggatactg
 661 agacaccccc ggtccaagcc tccctccac cactgcgcc ttctccctga ggagcctcag
 721 ctttccctcg aggcctcct accttttggc gggagacccc cagccctcgc agggggcggg
 781 cctcccccac acaccagccc tgttcggcgt ctgggcagtg cggggggggc ccgcctcccc
 841 cATGccgccc tcggggctgc ggtgctgcc gctgctgcta ccgctgctgt ggtactggt
 TGF-beta ->
 901 gctgacgcct ggcccgccgg ccgcgggact atccacctgc aagactatcg acATGgagct
 alternative cds start ->
 961 ggtgaaggcg aagcgcacg aggcacatcg cggccagatc ctgtccaagc tggggctcgc
 1021 cagccccccg agccaggggg aggtgccgcc cggcccgctg cccgaggccg tgcctgcct
 1081 gtacaacagc accgcgcacc gggctggccg ggagagtga gaaccggagc ccgagcctga
 1141 ggccgactac tacgccaagg aggtcacccg cgtgctaatt gtggaaacct acaacgaat
 1201 ctatgacaag ttcaagcaga gtacacacag catatatatg ttcttcaaca catcagagct
 1261 ccgagaagcg gtacctgaac ccgtgttgcct ctcgccggca gagctgogtc tgcctgaggag
 1321 gctcaagtta aaagtggagc agcacgtgga gctgtaccag aaatacagca acaattcctg
 1381 gcgatacctc agcaaccggc tgcctggcacc cagcgactcg ccagagtggc tatcttttga
 1441 tgtcacccga gttgtgcggc agtggtttgag ccgtggaggg gaabttgagg gctttcgcct
 1501 tagcgcacac tgcctctgtg acagcaggga taacacactg caagtggaca tcaacggggt
 1561 cactaccggc cgcgaggtg acctggccac cattcatggc atgaaccggc ctttctcgt
 1621 tctcatggcc acccgcctgg agaggggcca gcattctgaa agctcccgcc accgcgagc
 TGF-beta mature cds ->
 1681 cctggacacc aactattgct tcagctccac ggagaagaac tgctgctgc gccagctgta
 1741 cattgacttc cgcaggacc tcggtggaa gtggatccac gagcccaagg gctaccatgc
 1801 caacttctgc ctggggcctt gccctacat ttggagcctg gacacgcagt acagcaagg
 1861 cctggccctg tacaaccagc ataaccggg cgcctggcg ggcctgctt ggcgtgcgca
 1921 ggcgtggag ccgctgccc tcgtgtacta cgtgggcgc aagcccaagg tggagcagct
 1981 gtccaacatg atcgtgcct cctgcaagtg cagcTGAggt cccgccccgc cccgccccgc
 <- TGF-beta cds end
 2041 ccgggcaggc ccggccccac ccgccccgc cccgctgcc ttgccatgg gggctgTATT
 signal ->
 2101 TAaggacacc gtgccccag cccacctggg gccccattaa agatggagag aggactggg

```

2161 atctctgtgt cattgggggc ctgectgggg tatccatccc tgacgttccc ccactcccac
2221 tccctctctc tccctctctg cctcctcctg cctgtctgca ctattccttt gcccggcac
2281 aaggcacagg ggaccagtgg ggaacactac tgtagttaga tctatttatt gagcaccttg
2341 ggcactgttg aagtgcctta cattaatgaa ctcatlcagt caccatagca acactctgag
2401 atggcagggg ctctgataac acccatttta aaggttgagg aaacaagccc agagaggtta
2461 agggaggagt tctgcccac caggaacctg ctttagtggg ggatagtga gaagacAATA
                                     signal ->
2521 AAagatagta gttcaggcca ggggggggtgc tcacgcctgt aatcctagca cttttgggag
2581 gcagagatgg gaggatactt gaatccaggc atttgagacc agcctgggta acatagttag
2641 accctatctc tacaaaacac ttttaaaaaa tgtacacctg tgggtcccag tactctggag
2701 gctaagggtg gaggatcact tgatcctggg aggtcaaggc tgcag

```

HUMTRBP

LOCUS HUMTRBP 1368 bp ss-mRNA PRI 03-MAY-1991
 DEFINITION Human TAR RNA binding protein (TRBP) cDNA sequence, complete cds.
 ACCESSION M60801
 KEYWORDS TAR RNA binding protein.
 SOURCE Human mRNA from HeLa D98/AH-2 cell line.
 REFERENCE 1 (bases 1 to 1368)
 AUTHORS Gatignol,A., Buckler-White,A.J., Berkhout,B. and Jeang,K.-T.
 TITLE Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR
 JOURNAL Science 251, 1597-1600 (1991)
 COMMENT Human TRBP complexes with the double-stranded stem of TAR RNA (Gatignol et al. Science 251:1597-1600; Gatignol et al. Mol. Cell. Biol. 13:2193-2202).
 NCBI gi: 339908
 BASE COUNT 279 a 405 c 395 g 289 t
 ORIGIN

```

1 gctcttgggt tctgtagttt tctcgcgatc caaaaggctc cgtgcccaaa gcaaATGctg
TRBP cds start ->
61 gccgccaacc caggcaagac ccgcgtcagc cttctgcagg agtatgggac cagaataggg
121 aagaagcctg tgtacgacct tctcaaagcc gagggccaag cccaccagcc taatttcacc
181 ttccgggtca ccgttgggga caccagctgc actggtcagg gccccagcaa gaaggcagcc
241 aagcacaagg cagctgaggt ggccctcaaa cactcaaaag gggggagcat gctggagccg
301 gccctggagg acagcagttc tttttctccc ctgactctt cactgcctga ggacattccg
361 gtttttactg ctgcagcagc tgcataccca gttccatctg tagtccaac caggagcccc
421 gccatggaac tgcagccccc tgtctccctt cagcagtcct agtgcaaccc cgttgggtgt
481 ctgcaggagc tgggtggtgca gaaaggctgg cggttgcccg agtacacagt gaccaggag
541 tctggggccag ccacccgcaa agaattcacc atgacctgtc gagtggagcg ttccattgag
601 attgggagtg gcacttccaa aaaattggca aagcgggaatg cggcggccaa aatgctgctt
661 cgagtgcaca cggtgccctt ggatgcccgg gatggcaatg aggtggagcc tgatgatgac
721 cacttctcca ttggtgtggg ctccgccttg gatggtcttc gaaaccgggg ccagggttgc
781 acctgggatt ctctacgaaa ttcagtagga gagaagatcc tgtccctccg cagttgctcc
841 ctgggctccc tgggtgccc ggcccttgc tgcctccgtg tcttcagtga gctctctgag
901 gagcaggcct ttcacgtcag ctacctggat attgaggagc tgagcctgag tggactctgc
961 cagtgcctgg tggaaactgt caccagcccg gccactgtgt gtcattggct tgcaccacc
1021 agggaggcag ccggtggtga ggctgcccgc cgtgcccctg agtaacctca gatcatggca
1081 ggcagcaagT GAagccccag ctggactcat ggatgtgcac cctttgctcc ctgctcttcc
<- TRBP cds end
1141 tgcctctggg ctcatgtatc tgcgcagctc tggtaacctc tgtgggtgcc atctctacct
1201 ctgacacaga ctgcctgcct tgaagctgag aaggcacagg gcaaggagcc aaggaccaca
1261 gagcctcagc cagcccagga tccgtctcca ttttattggt gatgatgaat gggaatgaaa
1321 tcaggggggt gtctactaga gcttgggaata aatatgctgc tttgtgga

```


LOCUS HSUSF 1739 bp RNA PRI 15-FEB-1991
 DEFINITION Human usf sequence for late upstream transcription factor,
 clone dI2.
 ACCESSION X55666
 KEYWORDS DNA-binding protein; helix-loop-helix domain;
 transcription factor; upstream stimulatory factor.
 SOURCE Human mRNA from lambda-Zap library.
 REFERENCE 1 (bases 1 to 1739)
 AUTHORS Gregor, P.D., Sawadogo, M. and Roeder, R.G.
 TITLE The adenovirus major late transcription factor USF is a member of
 the helix-loop-helix group of regulatory proteins and binds to DNA
 as a dimer
 JOURNAL Genes Dev. 4, 1730-1740 (1990)
 COMMENT USF was characterized initially as a positive activator of the
 adenovirus major-late-promoter transcription. Within the LTR of
 HIV-1, a consensus binding sequence for USF is present from -159 to
 -174. However, there is conflicting information on whether this
 factor has moderating (and thus be regarded as a negative
 regulatory factor; Lu et al. J. Virol. 64: 5226-5229 1990) or
 stimulating (Maekawa et al. Nucl. Acids Res. 19: 4689-4694)
 effects. Interestingly, USF binds a second sequence near the HIV-1
 initiator (-5 to +11). USF interaction at the initiator-proximal
 site strongly activates HIV-1 expression from the TATAA-promoter.
 Of note, a factor distinct from USF binds the same DNA sequence
 (TFE3; Bechmann and Kadesch Genes Dev 5:1057-1066).
 NCBI gi: 37614

BASE COUNT 426 a 435 c 491 g 387 t
 ORIGIN

```

1 ggaattcctt gaaaattttc ctgggatagg aaaggactta gcaactcagc ctgtgaatct
61 aggagataca aagacctcca aaaaaggacc agttctctgg atgtgcccc tcacagagag
121 ATGaaggggc agcagaaaac agctgaaacg gaagagggga cagtgcagat tcaggaaggt
USF cds ->
181 gcagtggtta ctgggggaaga cccaaccagt gtggctattg ccagcatcca gtcagctgoc
241 accttccttg accccaacgt caagtaactg ttccgaactg agaatggggg ccaggtgatg
301 tacagggtga tccaggtgtc tgaggggcag ctggatggcc aaactgaggg aactggcgcc
361 atcagtggtc accctggcac tcaatccatg acccaggcgg tgatccaggg tgccttcacc
421 agtgatgatg cagttgacac ggaggggaca gctgctgaga cgcactatac ttacttccc
481 agcacggcag tgggagatgg ggcagggggg accacatcgg ggagtacagc tgctgttgtt
541 actaccaggg gctcagaggg actgctgggg caggcgaccc ctctggcac tggctcaattc
601 tttgtgatga tgtcaccaca agaagtactg caggagaggaa gccagcgctc aattgccct
661 aggaactcac cttattcccc gaagtcagaa gctccccgga cgaactcgga tgagaaacgc
721 agggctcagc ataataagt ggagcgtcgc cgccgagaca agatcaacaa ctggatcgtg
781 cagctctcca agataatccc agactgctct atggagagca ccaagtctgg ccagagttaa
841 ggtgggattc tatccaaagc ttgtgattat atccaggagc ttgggcagag taaccaccgc
901 ttgtctgaag aactgcaggg acttgaccaa ctgcagctgg acaatgacgt gcttcgacaa
961 caggtggaag atcttaaaaa caagaatctg ctgcttcgag ctcaattgag gcaccacgga
1021 ttagaggtcg tcatcaagaa tgacagcaac TAActatggg gattcagggg ctttgggccc
<- USF cds end
1081 aagaactgca gatagcccag gagcaacagc ctaatccctg gccctttcc ttcaactgcc
1141 cacttctggc atgggacagg gggaagttca gaaggtgtgt ccttgaactg aggcctgtg
1201 atatggcgcc ctgcagtggt gtgaagcaca caatgtggaa cgtgcactga cagccttgc
1261 cccccccacc atgcagcccc tgggccttgt gctcctctcg cacaatgcat gtgctgtctc
1321 catgctggat actggacaca ctaaactctg gggcttgtcc tgtgcttgc tagagtgcc
1381 agcagaggtt tgctgacagg tgatgctctg gcttgcccc gaactctggc acttccattg
1441 gttcttccct tccctggagc tgaggttttag atgtgcaacc tgtggctcag gggagcaagc
1501 ttacacaaga agtgagggaa ggtgttttag cagtggtctg tgcccatgaa gaggagattg
1561 gccagtgaga agctgaggcc tatgcagaca tctctggagc cagagagaa acacaggcag
1621 ggcccaactg gggccttccc ccttgtgggg ggtgttttt ttttttttt ttttttttt
1681 ttttttttt tttttttttt aagataaat tgttcaagc caaaaaaaa aggaattcc

```

HUSFII

LOCUS HUSFII 1299 bp mRNA PRI 11-MAR-1993
 DEFINITION Human upstream stimulatory factor II cDNA sequence.
 ACCESSION S50537
 SOURCE Human mRNA from Namalwa B cell line.
 REFERENCE 1 (bases 1 to 1299)
 AUTHORS Sirtio,M., Walker,S., Lin,Q., Kozlowski,M.T., Klein,W.H. and Sawadogo,M.
 TITLE Members of the USF family of helix-loop-helix proteins bind DNA as homo- as well as heterodimers
 JOURNAL Gene Expr. 2, 231-240 (1992)
 COMMENT USF was characterized initially as a positive activator of the adenovirus major-late-promoter transcription. Within the LTR of HIV-1, a consensus binding sequence for USF is present from -159 to -174. However, there is conflicting information on whether this factor has moderating (and thus be regarded as a negative regulatory factor; Lu et al. J. Virol. 64: 5226-5229 1990) or stimulating (Maekawa et al. Nucl. Acids Res. 19: 4689-4694) effects. Interestingly, USF binds a second sequence near the HIV-1 initiator (-5 to +11). USF interaction at the initiator-proximal site strongly activates HIV-1 expression from the TATAA-promoter. Of note, a factor distinct from USF binds the same DNA sequence (TFE3; Bechmann and Kadesch, Genes Dev 5:1057-1066).
 NCBI gi: 261070
 BASE COUNT 339 a 369 c 367 g 224 t
 ORIGIN
 1 ggccggccggca cgagcgggtgg ccgagggcagc gggatccaga cccgggtcca gcatgtccag
 USFII ->
 61 taccagttec gcacagagac aaatggagga caggctgtga tccaaaatcc cttcagcaat
 121 ggtggcagtc cggcggccga ggctgtcagc ggggaggcac gatttgcta ttcccagcg
 181 tccagtgtgg gagatactac ggctgtgtcc gtacagacca cagaccagag cttgcaggt
 241 ggaggccagt tctacgtcat gatgacgcc caggatgtgc ttcagacagg aacacagagg
 301 acgatggccc ccgggacaca cccttactct ccaaaaattg atggaaccag aacacccga
 361 gatgagagga gaagagccca gcacaacgaa gtggagcggg ggccgaggga caagatcaac
 421 aactggatcg tccagcttcc gaaaatcatt ccagactgta acgcagacaa cagcaagacg
 481 ggagcgagta aaggagggat cctgtccaag gcctgogatt acatcgggga gttgogccag
 541 accaaccagc gcatgcagga gaccttcaaa gaggccgagc ggctgcagat ggacaacgag
 601 ctcttgaggc agcagatcga ggagctgaag aatgagaacg ccctgcttcg agcccagctg
 661 cagcagcaca acctggagat ggtgggcgag ggcacccggc agTGAagccc gccaccacca
 <- USFII cds end
 721 cgcagccgac gccgcccacg ccggcctctg ctgccccctt cccagccct tagcacagag
 781 agggacacat gcccctcccc cagctgcgtt tttttatagt agatttttaa caaaaaacgg
 841 ggagaaataa tgcatttctg tggatacagt gccacccgac ctctccact tggaaacggt
 901 atcctccctg cccatccgtc tgtctgtcgc ccttctccg gccctcgcta agccccggca
 961 cttctagtgg tctacctgg caagagggag ggtacagagc ctctgccaac gtcccgctgg
 1021 tgctcctgc tctctggagg tactgagaca ggggtgctgat gggagggagg ggagcctttg
 1081 gggggggcac cgggggctg gacctatgca gggaggccac gtcccacccc acctcttgtt
 1141 tctgggtccc tgctccctt tgggggtgtg gtgtgtgttt taattttctt tatggaaaaa
 1201 ttgacaaaaa aaaaatagag agagaggtat ttaactgcaa taaactggcc ccattgtggaa
 1261 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

LOCUS HUMYC1A 1520 bp ss-mRNA PRI 17-MAR-1994
 DEFINITION Human YC1 cDNA sequence.
 ACCESSION L11289
 SOURCE Homo sapiens mRNA from lambda gt11 library.
 REFERENCE 1 (bases 1 to 1520)
 AUTHORS Lu,Y., Ueberla,K., Touzjian,N. and Haseltine,W.A.
 TITLE YC1: A human repressor of HIV-1 and IL2R alpha promoter transcription
 JOURNAL Unpublished (1993)
 COMMENT :
 BASE COUNT 459 a 382 c 311 g 368 t
 ORIGIN
 1 gaattcgggc cgagacttgg aaaccccbaa gtgtccgga cccctgacgg cagctccctt
 61 ccagcttcAT Gggcaagtg tggaacacgc agatgtaccc tcagtacgcc acctactatt
 YC1 cds start ->
 121 acccccagta tctgcaagcc aagcagtttc tgggtcccagc ccaccccatt gccccctccc
 181 gtcccagcac caccagcagt aataacaaca gtagcagcag tagcaactca ggatgggagc
 241 agctcagcaa aacgaacctc tatatccgag gactgcctcc ccacaccacc gaccaggacc
 301 tgggtgaagct ctgtcaacca tatgggaaaa tagtctccac aaaggcaatt ttggataaga
 361 caacgaacaa atgcaaaagg tatggttttg tgcactttga cagccctgca gcagctcaaa
 421 aagctgtgtc tgcctgaag gccagtgggg ttcaagctca aatggcaaa caacagggaac
 481 aagatcctac caacctctac atttctaatt tgcactctc catggatgag caagaactag
 541 aaaatatgct caaaccattt ggacaagtta tttctacaag gatactacgt gattccagtc
 601 ctacaagtcg tgggtgtggc tttgctagga tggaaatcac agaaaaatgt gaagctgtta
 661 ttggtcatit taatggaaaa tttattaaga caccaccagg agtttctgcc cccacagaac
 721 ctttatgtg taagtttgct gatggaggac agaaaaagag acagaaccca aacaataaca
 781 tccctaattg aagaccatgg catagagaag gagaggtag acttgetgga atgacactta
 841 cttacgaccc aactacagct gctatacaga acggatttta tccctcacca tacagtattg
 901 ctacaaacgg aatgatcact caaacttcta ttacacccta tattgcatt cctgtatctg
 961 cctaccagg gcaagtcct tegtggatgc aacctcaacc atatattcta cagaccctg
 1021 gtgccgtgtt aactccctca atggagcaca ccatgtcact acagcccgca tcaatgatca
 1081 gccctctggc ccagcagatg agtcactctg cactaggcag caccggaaca tacatgectg
 1141 caacgtcagc tatgcaagga gectacttgc cacagtatgc acatatgcag acgacagcgg
 1201 ttctgttga ggaggcaagt ggtcaacage aggtggctgt cgagacgtct aatgaccatt
 1261 ctccatatac ctttcaacct aataagTAAc tgtgagatgt acagaaagggt gttcttacat
 <- YC1 cds end
 1321 gaagaagggtg tgaaggctga acaatcatgg atttttctga tcaattgtgc tttaggaaat
 1381 tattgacagt ttgtcacagg ttcttgaaaa cgttatttat aatgaaatca actaaaacta
 1441 tttttgctat aagttctata aggtgcataa aacctttaa ttcacttagt agctgttccc
 1501 ccgaacagggt ttatttttagt

LOCUS YY1 2353 bp ss-mRNA PRI 25-MAY-1993
 DEFINITION Homo sapiens GLI-Krupple related protein (YY1) cDNA sequence, complete cds.
 ACCESSION M77698
 KEYWORDS GLI-Krupple related protein.
 SOURCE Homo sapiens mRNA.
 REFERENCE 1 (bases 1 to 2353)
 AUTHORS Shi,Y., Seto,E., Chang,L.-S. and Shenk,T.
 TITLE Transcriptional repression by YY1, a human GLI-Krupple related protein, and relief of repression by adenovirus Ela protein
 JOURNAL Cell 67, 377-388 (1991)
 COMMENT Positioned at the junction of U3 and R in the HIV-1 genome, is a binding site for YY1 (Useheva and Shenk Cell 76:1115-1121; Seto et al. Nature 354:241-245). Binding of YY1 to the LTR has been shown to repress HIV-1 expression and production of virions (Margolis et al. J. Virol. 68:905-910).
 BASE COUNT 625 a 598 c 608 g 522 t
 ORIGIN
 1 cgcgcgagacg agcagcgggcc gagcgcagcgc gggcgcggggc gcaccgagggc gaggggaggcg
 61 gggaagcccc gcgcgcgcgc cccgcgcgcgc ccttcccccc gcgcgcgcgc cctctctccc
 121 cgcgcgcgcgc gcgcgcgcgc cctctctccc cctctctccc acggcgcggcc gccctctcgc
 181 cgcgcgcgcgc gcgcgcgcgc agcgcgcgcgc gcgcgcgcgc tggcgcgcgga gccctcagcc
 241 ATGgcctcgg gcgcacccct ctacatcgcc acggacgggt cggagatgcc ggccgagatc
 YY1 cds ->
 301 gtggagctgc acgagatcga ggtggagacc atcccggtgg agaccatcga gaccacagtg
 361 gtggcgaggg agggaggagg ggcgcgcgcgc gcgcgcgcgc gcgcgcgcgc cgcacacggc
 421 ggcgggggcg gccacgggga cgcgcgcgcgc caccaccacc accatcacca ccaccaccac
 481 cgcgcgcgcgc tgcctctgca gcgcgcgcgc accgcgcgcgc gcgcgcgcgc gcacccaggc
 541 caggaggtga tcttggtgca ggcgcgcgcgc gaggtggtgg gcgcgcgcgc ctgcgcgcgc
 601 ctgcgcgcgc aggcgcgcgc cgcgcgcgcgc attctcctcc cggcgccgcgc gccgcgcgcgc
 661 ggcgagcagc actacattga acaaacgcgc gtcacgcgcgc cgcgcgcgcgc caagagcgcgc
 721 ggcggcgcgc cgtcgcgcgc ggcgcgcgcgc cgcgcgcgcgc aggcgcgcgc caagaagagc
 781 ggcaagaaga gttacctcag cgcgcgcgcgc ggcgcgcgcgc gcgcgcgcgc cgcgcgcgcgc
 841 ggcaacaaga agtgggagca gaagcaggtg cagatcaaga ccttgagggg cgcgcgcgcgc
 901 gtcaccatgt gtcctcaga tcaaaaaaaa gatattgacc atgagacagt ggttgaagaa
 961 cagatcattg gagagaactc acctcctgat tattcagaat atatgcagc aaagaaactt
 1021 cctcctggag gaatacctgg cattgcacct tcagatccca acaaacctgg agaatttgct
 1081 agaatgaagc caagaaaaat taaagaagat gatgctccaa gaacaatagc ttgcctcat
 1141 aaaggctgca caaagatggt cagggataac tggccatgca gaaaacatct gcacacccac
 1201 ggtcccgagc tccacgtctg tgcagaatgt ggcaaaagct ttgttgagag tcaaaaacta
 1261 aaacgcacac aactgggttc tactggagag aagccctttc agtgcacgtt cgaaggctgt
 1321 gggaaaacgt ttctactgga ctccaatttg cgcacacatg tgcgaatcca taccggagac
 1381 aggccttatg tgtgcccctt cgatgggtgt aataagaagt ttgctcagtc aactaacctg
 1441 aaatctcaca tcttaacaca tgcataaggcc aaaaacaacc agTGAaaga agagagaaga
 <- YY1 cds end
 1501 ccttctctga ccacgggaag catcttccag aagtgtgatt gggaataaat atgcctctcc
 1561 tttgtatatt atttctagga agaattttaa aatgaatcc tacacaccta agggacatgt
 1621 tttgataaag tagtaaaaat taaaaaaaat aaactttact aagatgacat tgctaagatg
 1681 ctctatcttg ctctgtaate tctgttcaaa aacacagttt tttgttaaag tctggtccca
 1741 acaggaggac aattcatgaa ctctgcacac aaagacaatt ctttatacaa cagtgtctaa
 1801 aatgggaact cttttcacat ttttataaat atgaagctca cctgttgctt acaatttttt
 1861 taatttttga ttttccagc gtgcataatt tacacttttt tggggatagc cttagtaatt
 1921 ctacgtgtga ttttcttggc ggttgataac ttgtcttgc gtagattttt tttaaaagaa
 1981 tgggcagtta catgcatact tcaaaagtat ttctctgtaa aaaaaaaaaa agttatatag
 2041 gttttgtttg ctatcttaat ttgtgttga ttctttgatg ttaacacatt ttgtataatt
 2101 gtatcgtata gctgtattga atcatgtagt atcaaatatt agatgtgatt taatagtgtt
 2161 aatcaattta aaccattttt agtcactttt tttttccaaa aaataactgc cagatgctga
 2221 tgttcagtgt aatttctttg cctgttctag tacagaaagt ggtgctcagt tgtagaatgt
 2281 attgtacctt ttaacacctg atgtgtacat cccatgtaac agaaagggca acaataaaat
 2341 agcaatccca aag

LOCUS Hcfos 6210 bp ds-DNA PRI 15-MAR-1990
 DEFINITION Human fos proto-oncogene (c-fos) sequence, complete cds.
 clone pc-fos(human)-1.
 ACCESSION K00650 M16287
 KEYWORDS c-myc proto-oncogene; fos oncogene; proto-oncogene.
 SOURCE Human DNA from lymphoblastic cell line (CCRF-CM).
 REFERENCE 1 (bases 1 to 4165)
 AUTHORS van Straaten,F., Mueller,R., Curran,T., van Beveren,C.P. and Verma,I.M.
 TITLE Complete nucleotide sequence of a human c-onc gene: Deduced amino acid sequence of the human c-fos protein
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80, 3183-3187 (1983)
 REFERENCE 2 (sites)
 AUTHORS Treisman,R.
 TITLE Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5' element and c-fos 3' sequences
 JOURNAL Cell 42, 889-902 (1985)
 REFERENCE 3 (bases 4166 to 6210)
 AUTHORS Verma,I.M., Deschamps,J., van Beveren,C.P. and Sassone-Corsi,P.
 TITLE Human fos gene
 JOURNAL Cold Spring Harb. Symp. Quant. Biol. 51, 949-958 (1986)
 COMMENT Between -216 and -254 of the HIV-1 LTR are binding sites for NFAT, nuclear factor from activated T cells (Shaw et al. Science 244:202-205). Recent evidence suggests that NFAT binding activity results from the interaction of three discrete polypeptides, NFATp (McCaffrey et al. Science 262: 750-754), Fos and Jun (Yaswen et al. J. Biol. Chem. 268:14285-14293).
 BASE COUNT 1497 a 1571 c 1619 g 1523 t
 ORIGIN Chromosome 14q24.3-q31.
 1 gcaggaacag tgctagtatt gctcagagccc gagggctgga ggtaggggga tgaaggctcg
 61 cttocaogct ttgcactgaa ttaggggctag aattggggat gggggtaggg ggcatttct
 121 tggggagccg aggtttaagt cctcggggtc ctgtactega tgcggtttct cctatctctg
 181 agcctcagaa ctgtcttcag tttccgtaca aggttaaaaa ggcgtctctt gccccatccc
 241 ccccgacctc ggggaacaagg gtccgcattg aaccagggtgc gaatgttctc tctcattctg
 301 cgcggttccc gcttcccttc cccagagcgc ggcccccgcc tcccccgca ctgcacctc
 361 ggtgttggtt gcagccccgc agcagttccc gtcaatccct ccccccttac acaggatgtc
 -> transactivator region
 421 catattagga catctgcgtc agcaggtttc cagggecttt cctgtagcc ctggggggag
 481 ccatccccga aacccctcat cttggggggc ccacagagacc tctgagacag gaactgcgaa
 541 atgttcacga gattaggaca cgcgcccaagg cgggggcagg gagctgcgag cgtcggggac
 601 gcagccgggc ggcgcagaa ggcgccaggc ccgcgcgcca cccctctggc gccaccgtg
 661 ttgagcccggt gacgtttaca ctcatccta aaacgcttgt tataaaagca gtggtgogg
 721 cgcctcgtac tccaaccgca tctgcagcga gcaactgaga agccaagact gagccggogg
 781 ccgcgggcga ggaacagagc agtgacogtg ctcctaccca gctctgcttc acagcgccca
 841 cctgtctccg cccctcggcc cctcggccgg ctttgcttaa ccgccacgAT Gatgttctcg
 cfos exon 1 start ->
 901 ggcttcaacg cagactacga gggtcatcc tcccgtgca gcagcggcgc ccgggcggg
 961 gatagcctct ctactacca ctcaaccgca gactccttct ccagcatggg ctgcgctgtc
 1021 aacgcgcagg taaggctggc ttcccgtcgc cgcggggccg ggggcttggg gtcgcggagg
 <- cfos exon 1 end
 1081 aggagacacc gggcgggacg ctccagtaga tgagttaggg gctcccttgt gcttgagggg
 1141 aggtctgcgt ggcgggagcg gtgcggctc gggggctcgg gaattgctct gaggcagcgc
 1201 acgcttgcca tagtaagaat tggttcccc ttccggaggc aggttcgttc tgagcaacct
 1261 ctggtctgca ctccaggacg gatctctgac attagctgga gcagacgtgt cccaagcaca
 1321 aactcgctaa cttaggcctg gcttcttcgg ggaggtggca gaaagcggca atccccctc
 1381 ccccggcagc ctggagcagc gaggagggat gagggaggag ggtgcagcgg gggggtgtgt
 1441 aaggcagttt cattgataaa aagcgagttc attctggaga ctccggagcg gcgcctgcgt
 1501 cagcgcagac gtcagggata ttataacaa aaccccttcc aagcaagtga tgctgaaggg

```

1561 ataacgggaa cgcagcggca ggatggaaga gacaggcact gcgctgcgga atgcctggga
1621 ggaaaagggg gagacctttc atccaggatg agggacattt aagatgaat gtccgtggca
1681 ggatcgtttc tcttcactgc tgcctgcggc actgggaact cgcaccacct gtgtccgga
1741 cctgctcgct cagctcggtt tcccccttct gttttgttct aggaattctg caggacctg

      cfos exon 2 start ->
1801 gccgtctcca gtgccaactt cattccdacg gtdactgcca tctcgaccag tccggacctg
1861 cagtggctgg tgcagcccg cctcgtctcc tctgtggccc catcgagac cagagccctt
1921 caacctttcg gagtccccgc cccctccgct ggggcttact ccagggtggt cgttgtgaag
1981 accatgacag gaggcgcgag gcagagcatt ggcaggaggg gcaagggtga acagggtagg
      <- cfos exon
      2 end
2041 aactctagcg tactcttctt gggaaatgtg gggctgggtg ggaagcagcc cgggagatgc
2101 aggagcccg tacagaggat gaagccactg atggggctgg ctgcacatcc gtaactggga
2161 gccctggctc caagcccatc ccaccccaac tcagactctg agtctcacc taagaagtac
2221 tctcatagtt tcttccctaa gtttcttacc gcattgcttc agactgggtt cttctttgtt
2281 ctcttgctga ggcatttatt ttaaatgcaa gtcacacctt ttctgcaact gcaggtcaga
2341 aatgggttca cagtgggtg ccaggaagca gggagctgc aggagccagt tctactgggg
2401 tgggtgaatg gaggtgatgg cagacacttt tactgaatgt cgggtctttt ttgtgattat
2461 tctagttatc tccagaagaa gaagagaaaa ggagaatccg aagggaagg aataagatgg

cfos exon 3 ->
start
2521 ctgcagccaa atgcgcgaac cggaggaggg agetgactga tacactccaa ggggtaggta
      <- cfos exon
      3 end
2581 ctctgtgggt tgctcctttt taaaacttaa gggaaagtgt gagattgagc ataaggcccc
2641 ttgagtaaga ctgtgtctta tgccttctct tatcctctgt tacaaggag acagaccaac

      cfos exon 4 start ->
2701 tagaagatga gaagtctgct ttgcagaccg agattgccaa cctgctgaag gagaaggaaa
2761 aactagagtt catcctggca gctcaccgac ctgctgcaa gatcctgat gacctgggtt
2821 tcccagaaga gatgtctgtg gcttcccttg atctgactgg gggcctgcca gaggttgcca
2881 ccccgagtc tgaggaggcc ttcaccctgc ctctctcaa tgacctgag cccaagccct
2941 cagtggaaac tgtcaagagc atcagcagca tggagctgaa gaccgagccc ttgatgact
3001 tccgttccc agcatcatcc aggcacagtg gctctgagac agcccgctcc gtgcagaca
3061 tggacctatc tgggtccttc tatgcagcag actgggagcc tctgcacagt ggctccctgg
3121 ggaaggggcc catggccaca gagctggagc cctgtgcac tccggtggtc acctgtactc
3181 ccagctgcac tgcctacacg tcttcccttg tcttcccta cccgaggct gactcttcc
3241 ccagctgtgc agctgcacc cgcagggcaa gcagcagcaa tgagccttcc tctgactcgc
3301 tcagctcacc cagctgctg gccctgTGAG ggggcaggga aggggaggca gccggcacc

      <- cfos cds end
3361 acaagtgcc ctgcccagc tgggtgcatta cagagaggag aaacacatct tccctagagg
3421 gttcctgtag acctaggag gaccttatct gtgcgtgaaa cacaccaggc tgtgggcttc
3481 aaggacttga aagcatccat gtgtggactc aagtccctac ctcttccgga gatgtagcaa
3541 aacgcattga gtgtgtattg tcccagtgca cacttcagag agctggtagt tagtagcatg
3601 ttgagccagg cctgggtctg tgtctcttct ctcttctctc ttagtcttct catagcatta
3661 actaatctat tgggttcatt attggaatta acctgggtgt gcatattttt aaattgtatc
3721 tagtgagctt gattttaaca ataactactg tgttccctgg aatagtgtgt tctgattaga
3781 aatgaccaat attatactaa gaaaagatac gactttattt tctggtagat agaaataaat
3841 agetatatcc atgtactgta gttttctctt aacatcaatg ttcattgtaa tgttactgat
3901 catgcattgt tgagggtggc tgaatgttct gacattaaca gttttccatg aaaacgtttt
3961 attgtgtttt taattttatt attaagatgg attctcagat atttatattt ttattttatt
4021 ttttctacc ttgaggtctt ttgacatgtg gaaagtgaat ttgaatgaaa aatttaagca
4081 ttgtttgctt attgttccaa gacattgtca ataaaagcat ttaagttaa tgcgaccaac
4141 cttgtgctct tttcattctg gaagtcttct aagtctctga aaggcattat tggagaccag
4201 tttgtcaaga aggttagctg ctggaggggg acacaccctc tgtctgatcc ctatcaaaag
4261 aggacaagga aactatagag ctgatttttag aatattttac aaatacatgc cttccattgg
4321 aatgctaaga ttttctactg ctctctggga cgggaaaccg ctgtgtaaca gcttttgttg
4381 gaactaattt tttctgtttc agtactcgca gggggaataa tttaaatttt gttgtgctaa
4441 tattaaattc agatgttttg atcttaagg atccctttaa gcaaacagaa cctagctttg
4501 tacagactat ttttaacttt tattctcaca aaatcacgtg gaggttatt ctacttcaa

```

```

4561 gatgagcaaa ttgaagaatg gttagaataa acaactttct tgatatcccg ttatcggcat
4621 tagaatcttc ctgctcgtta tegtatccag caggctgaac tgcctcttga tacttggtta
4681 aaaaaaattt tcagggcggg ogcgggtggc catgcctgta atcctagcac ttggggaggc
4741 cgaggcaggc ggatcacctg aggtcgggag ttcgagacca gcctgaccac catggagaaa
4801 ccccgctctt actaaaaata caaaattagc ctgggtgtgtt ggtgcatgcc tgtaatccta
4861 gctacttgag aggcctgagc agggaaatca ctgaaactcg ggaggcggat gttgcagcga
4921 actgagattg cgcatttgc ctcagcctg ggcaacaaga ttgaaactct gtttaaaaaa
4981 aaaagtcttc actaatgtgt acattttttt gtactctttt attctcgaaa ggggaaggagg
5041 gctatttccc tatcccttat taataaatgc attgtgtgtt ctggtttctc taataccata
5101 tgcctctcat tcagtttata gtgggcggaa gtgggggaga aaaagtgtct cagaatcaa
5161 aagatatctc aaacagcaca aataatggct gatcgttctg caaacaaaaa gttacataat
5221 agctcaagaa ggagaagtca acatgactct gaacaagctt taacttagaa actttatcat
5281 ctttaaggag aacgtgacct ttgtccagga cgtctctggt aatggggcac ttacacacac
5341 atgcacacgt acaaaccaca gggaaaggag accgccttc tgcctctgct cgcgagtac
5401 acgcaggcac catgcactat gttttcacac acactgggtg gaagaagagc ttcagcgaca
5461 gtcttctaata gcttttgtga taatgaaaat cactgggtgc ttatgggggtg tcataattcaa
5521 tcgagttaaa agttttaatt caaaatgaca gttttactga ggttgatgtt ctcgctctatg
5581 atatctctgc cctcccata aaaaatggaca tttaaagca acttaccgct ctttagatca
5641 ctcttatatc acacaccact tgggtgtctg tttctgctag acttgatgag acagtggcct
5701 taggatecct gtttgctgtt caaagggcaa atattttata gcctttaaat atacctaac
5761 taaatacaga attaatataa ctaacaaaca cctggctctga aataacaagg tgatctacc
5821 tggaaggaa cagctgtgtg ggccaggagc ggtggctcac acctgtatt ccagcacttt
5881 gggaggctga gacaggagga tcaactggag ccaggagttt gagaccagcc tgggcascac
5941 ggcaaaaccc agtgtgtctc tgttgtccca gctacactac tcaggaggtt gaggcaggag
6001 tatgacttga gcctgggagg gggaggttgc agagaactga tattgcacca ccactgcact
6061 ccagcctggg tgacacagca aaacctatc tcaaaaaaaa aaaaaaaaaa aagggaacca
6121 gctggttctt gtaggtgtgc aataataaca accagaggaa gaaaaggag acgatttccc
6181 agatgaagaa gggcagctgg acctcggac

```

HINTAV

LOCUS HINTAV 1074 bp DNA PRI 30-MAR-1994
 DEFINITION Human integrin alpha v sequence, promoter region and partial cds.
 ACCESSION U07375
 SOURCE Human.
 REFERENCE 1 (bases 1 to 1074)
 AUTHORS Donahue,J.P., Sugg,N. and Hawiger,J.
 TITLE The integrin alphav gene: Identification and characterization of the promoter region
 JOURNAL Unpublished
 COMMENT There is evidence that HIV-1 Tat can be taken up actively into cells through binding to a cell-surface protein (Weeks et al. J. Biol. Chem. 268:5279-5284) implicated to be alpha-v-beta-5 integrin (Vogel et al. J. Cell. Biol. 121:461-468). Once inside cells, Tat interacts with multiple partners in activating transcription.
 NCBI gi: 466371
 BASE COUNT 143 a 394 c 328 g 209 t
 ORIGIN
 1 agctttctctc atttttaaac aacagtcgca cggaggttcc cggcgggaca agggaaacgtg
 61 ggtgcccttg ctactcccggt ggacgcgggt agattgggac gctggaccgt atctccccgc
 121 cccgcgcccc acgcctctctc aggtgctcag cctgaggcct tegtccagga gcgctgcgcg
 181 tgacccaggc tcaggagctg ggggccccctg cacagacgct caggtctcgg gacaggcggt
 241 gactgcactc acggaagtac gctgagctct cccctgtaga agggcgctc tctccccca
 301 ctctctctctc cagctccaca gcagctccc gggcgggctc ctctctctc caggtctct
 361 ccagtgcgcg cgcggctctc aggcctgagg tgcggcgctc acccggcag tccccagct
 421 cagacgctgc gtggagcggc ggagcgggag ggaagcaaag gaccgtctgc gctgctgtc
 481 ccgccccggg cgtctctgdc cctctgtccc tggcggtcgc tcgaagctc agcctcttg
 541 cctgccccgg agctgtcccg ggtagccga gaagagagcg gccggcaagt ttgggcgcg
 601 gcaggcgggg ggcgcggggc actgggcgct tgcgtggggt ggggggaggt ggtaccgct
 661 cccggcttgg cgtcccggc gcacttcggc gATGgcttct cgcgcgggc gacggctggt
 vitronectin alpha subunit cds start ->
 721 cctcggtccc cggggctcc cgtttcttct ctggggactc ctgtacctc tgtgcgcgc
 781 cttaaaccta gacgtggaca gtcttgccga gtactctggc ccgagggaa gttacttcgg
 841 cttegcgctg gattttcttg tgcacagcgc gtcttcgtta gtggcgcac ttggaactgg
 901 agctggcccc ctccccacc gcgcgcacc acccagcgtt tctccattgg gattgatttc
 961 cattgggatt gatttccgag agatcccggc gtctgtctgt ctacccatc ctactctca
 1021 gagtgggtta tatctgggt ggaggaata ttgggggaa tgaactccct ttgg
 vitronectin alpha subunit <-

LOCUS HUMINTB5A 3117 bp ss-mRNA PRI 30-JUL-1991
 DEFINITION Human integrin beta-5 subunit cDNA sequence, complete cds, clone 10.
 ACCESSION J05633
 KEYWORDS integrin beta-5 subunit.
 SOURCE Human epithelial lung adenocarcinoma mRNA.
 REFERENCE 1 (bases 1 to 3117)
 AUTHORS McLean, J.W., Vestal, D.J., Cheresch, D.A. and Bodary, S.C.
 TITLE cDNA sequence of the human integrin beta-5 subunit
 JOURNAL J. Biol. Chem. 265, 17126-17131 (1990)
 COMMENT There is evidence that HIV-1 Tat can be taken up actively into cells through binding to a cell-surface protein (Weeks et al. J. Biol. Chem. 268:5279-5284) implicated to be alpha-v-beta-5 integrin (Vogel et al. J. Cell. Biol. 121:461-468). Once inside cells, Tat interacts with multiple partners in activating transcription.
 Draft entry and computer-readable sequence for [J. Biol. Chem. (1990) in press] kindly submitted by J.W.McLean, 20-JUL-1990.
 NCBI gi: 186504

BASE COUNT 718 a 819 c 892 g 688 t

ORIGIN

```

1 cccgggagtg cagcgaccgc gccgcgcgtg agggaggcgc cccaccATGc cgggggcccc
   integrin beta-5 subunit cds start ->
61 ggcgccgcgtg taogccctgcc tccctggggct ctgcgcgcctc ctgccccggc tcgcaggtct
   integrin beta-5 subunit mature peptide ->
121 caacatatgc actagtggaa gtgccacctc atgtgaagaa tgtctgctaa tccacccaaa
181 atgtgccttg tctccaaag aggaacttcg aagcccacgg tccatcacct ctgggtgtga
241 tctgagggca aaccttgtca aaaatggctg tggagggtgag atagagagcc cagccagcag
301 cttccatgtc ctgaggagcc tgccccctcg cagcaagggt tgggctctg caggctggga
361 cgtccattcg atgacaccac agggagattgc cgtgaacctc cggccccgtg acaagaccac
421 cttccagcta caggttcgcc aggtggagga ctatcctgtg gacctgtact acctgatgga
481 cctctccctg tccatgaagg atgacttggc caatatccgg agcctgggca ccaactcgc
541 ggaggagatg aggaagctca ccagcaactt cgggttggga ttgggtctt ttgttgataa
601 ggacatctct cctttctctt acacggcacc gaggtaccag accaatccgt gcattggtta
661 caagttgttt ccaaattgcg tccctcctt tgggttccgc catctgctgc ctctcacaga
721 cagagtggac agcttcaatg aggaagtctg gaaacagagg gtgtccccga accgagatgc
781 cctgagggg ggctttgatg cagtactcca ggcagccgtc tgcaaggaga agattggctg
841 gcgaaaggat gcactgcatt tgctggtgtg cacaacagat gatgtgcccc acatgcatt
901 ggcaggaana ttgggagggc ttgtgcagcc acacgatggc cagtgcacc tgaacgaggc
961 caacgagtac actgcattca accagatgga ctatccatcc cttgccttgc ttggagagaa
1021 attggcagag aacaacatca acctcatctt tgcagtgaca aaaaaccatt atatgctgta
1081 caagaatttt acagccctga tccctggaac aacggtggag attttagatg gagactccaa
1141 aaatattatt caactgatta ttaatgcata caatagtatc cggctctaaag tggagtgtgc
1201 agtctgggat cagcctgagg atcttaactc cttctttact gctacctgcc aagatggggg
1261 atcctatcct ggtcagagga agtgtgaggg tctgaagatt ggggacacgg catcttttga
1321 agtatcattg gaggcgcgaa gctgtcccag cagacacacg gagcatgtgt ttgccctgcg
1381 gccggtggga ttccgggaca gcctggaggt gggggtcacc tacaactgca cgtgcggctg
1441 cagcgtgggg ctggaaccca acagcccgag gtgcaacggg agcgggacct atgtctgcgg
1501 cctgtgtgag tgcagcccg gctacctggg caccaggtgc gagtgcacgg atggggagaa
1561 ccagagcgtg taccagaacc tgtgccggga ggcagagggc aagccactgt gcagcggcg
1621 tggggactgc agctgcaacc agtgtctctg cttcgagagc gaggttggca agatctatgg
1681 gcctttctgt gagtgcgaca acttctctg tgcagggaac aaggaggtcc tctgctcagg
1741 coatggcgag tgcactgcg ggggaatgca gtgccatgca ggttacatcg gggacaactg
1801 taactgctcg acagacatca gcacatgcg gggcagagat ggccagatct gcagcgagcg
1861 tgggcactgt ctctgtgggc agtgccaatg cagcgagccg ggggcctttg gggagatgtg
1921 tgagaagtgc cccacctgcc cggatgcacg cagcaccacg agagattgcg tcgagtcccc
1981 gctgctccac tctgggaacc ctgacaacca gacctgccac agcctatgca gggatgaggt
2041 gatcacatgg gtggacacca tctgaaaga tgaccaggag gctgtgctat gtttctacaa

```

HUMINTB5A

```

2101 aaccgcccaag gactggcgtca tgatgttcac ctatgtggag ctccccagtg ggaagtccaa
2161 cctgacogtc ctgagggagc cagagtgtgg aaacaccccc aacgccatga ccctcctcct
2221 ggctgtggtc ggtagcatcc tcttgtttgg gcttgcactc ctggctatct ggaagctgct
2281 tgtcaccatc cagcaccgga gggagtttgc aaagtttcag agcgagcgat ccagggccccg
2341 ctatgaaatg gcttcaaatc cattatacag aaagcctatc tccacgcaca ctgtggactt
2401 caccttcaac aagttcaaca aatcctacaa tggcactgtg gacTGAtgtt tcttctcog
                                     <- integrin beta-5
2461 aggggctgga ggggggatct gatgaaaagg tcagactgaa acgccttgca cggctgctcg
2521 gcttgatcac agctccctag gtaggcacca cagagaagac cttctagtga gcctgggcca
2581 ggagcccaca gtgcctgtac aggaagggtgc ctggccatgt cactggctg ctaggccaga
2641 gccatgccag gctgcgtccc tccgagcttg ggataaagca aggggacctt ggcgctctca
2701 gctttccctg ccacatccag cttgttgtcc caatgaaata ctgagatgct gggctgtctc
2761 tcccttccag gaatgctggg cccccagcct ggccagacaa gaagactgtc aggaagggtc
2821 ggagtctgta aaaccagcat acagtttggc tttttcaca ttgatcattt ttatatgaaa
2881 taaaaagatc ctgcatttat ggtgtagttc tgagtccctg gacttttctg cgtgatggct
2941 atgccttgca cacaggtgtt ggtgatggg ctgttgagat gcctgttgaa ggtacatcgt
3001 ttgcaaatgt cagtttcttc tctgtccgt gtttgtttag tacttttata atgaaaagaa
3061 acaagattgt ttgggattgg aagtaaagat taaaaccaaa agaatttgtg tttgtca

```

LOCUS MMLEF1 2460 bp RNA ROD 16-DEC-1993
DEFINITION Mouse lymphoid enhancer binding factor 1 LEF1 sequence.
ACCESSION X58636
SOURCE Mus musculus RNA.
REFERENCE 1 (bases 1 to 2460)
AUTHORS Travis,A., Amsterdam,A., Belanger,C. and Grosschedl,R.
TITLE LEF-1, a gene encoding a lymphoid-specific DNA-binding protein with an HMG domain, regulates T cell receptor alpha enhancer function
JOURNAL Genes Dev. 5, 880-894 (1991)
COMMENT LEF-1, lymphoid enhancer-binding factor, is a T-cell specific transcription factor. It is capable of binding the minor groove of a DNA helix through its high mobility group (HMG) domain, a domain characteristic of other eukaryotic regulatory proteins. This interaction results in a bending of the DNA helix and thereby facilitates the assembly of nucleoprotein complexes at the promoter (Giese et al. Cell 69: 185-195 1992, Jones and Peterlin Annu. Rev. Biochem. 63: 717-743 1994). Within the LTR of HIV-1, one high affinity LEF-1 binding site exists at -122 to -143 and two low affinity binding sites are present; one located from -37 to -51 and another from +17 to +32 (Waterman and Jones New Biol. 2: 621-636 1990).
 NCBI gi: 52887
BASE COUNT 603 a 768 c 629 g 460 t
ORIGIN

```

1  eggaggagcc gagggggccac cggcagctct ttgctttgac cgaggggagg cggggagaga
61  ggcgagccag gcagaaactc gagccgggaa caaagaggga tcgggctcga gtgtgtgtgt
121 gtgtgtgtgt gtgtgtgtgt gtaagtgtgt gtatcggccc gagctgcggg
181 ctgggacatt tgggcccga gctcctgctg tgactcccca agactccggc gtgccagcca
241 ccgcagatcc ccagcgctca tcatcaca aaattattctt ggcaaaacttc tctttttctc
301 cctcccccct ccagcagatt aaatgctcct ccagaaggaa aaccgaagcg aaaggggaagg
361 aaagaagctc taacggcgac gtctgcagcc cggkggctct ttattgttta ctctgaagga
421 agtggacttt tcgggtatttt ctgattcttc tcttacctcc gctggggcaa agggagcctc
481 ttggccagct ctctctctct caaaaacaaa acaaaaacaaa accccaatc acogaaatca
541 gcatccgagg aatccgtgga gaggttttgc acggtgacgc accacgggac cctcgtttgc
601 actttggcaa gtctgtctcc ctccgggaacc ggtgtctctg aagcacagac ccagcaaacg
661 agtgcggcgt ggggtggcca ccagccctag gcagcctagc ctagtgcacg cggagcgctg
721 agacgctctc agagcgcccg gcgtgcccgc ctgcagctct ggtgcccttc gccccaggcc
781 gggggcgccg ogctctctta agtgggaaag cggggcaatc gcagaggctc ctgcagctgc
841 tccagcgttg ggccctctgc ttgccacccc ggttccatt cacagtctc ccgcactcag
901 tottccaaact ccccccctc acatcgaccc tccctctctg ctgtctcgga gtgtctccgc
961 agcggagcgg agattacaca gcgcggggA TGccccaaact ttcgggagga ggggcgggg

```

LEF1 cds start ->

```

1021 gggaccggga actctgggcc accgatgaga tgatcccctt caaggacgaa ggcatcccc
1081 agaaggagaa gatcttggcc gagatcagtc atcccgaa gaaggggagac ttagccgaca
1141 tcaagtcatc tttgggttaac gagtcgaaa tcatccagc cagcaacggg catgaggtgg
1201 tcagacaagc ccggtctctc caggagcctt accacgaaa ggccagagaa caccctgatg
1261 aaggaaaagca tccagacgga ggcctgtaca acaagggaacc ctctactcc agttactctg
1321 gctacataat gatgcccatt atgaacagcg acccgtacat gtcaaatggg tccctttctc
1381 caccatcccc gaggaatca aataaagtgc ccgtggtgca gccctctcac ggggtccacc
1441 cgtcaccccc cctcatcacc tacagcgacg agcacttttc tccgggatcc caccgctcac
1501 acatcccgte agatgtcaac tccaagcaag gcatgtccag acaccctcca gctctgaaa
1561 tccccacctt ctaccccttg tctccggggc gcgttggaca gatccccc cccattggct
1621 ggcaagggtca gctctgttat cccatcacgg gtggattcag gcaacctac ccatctcac
1681 tgtcaggcga cacttccatg tccaggtttt cccatcatat gattcctggg ccccttgcc
1741 cccacacaac tggcatccct catccagcta ttgtaacacc tcaggtcaa caggagcacc
1801 cccacacgga cagtgaacta atgcaagtga agcctcaaca cgaacagaga aaggagcagg
1861 agcccaaaag acctcatatt aagaagctc tgaatgctt catgttatat atgaagaaa
1921 tgagagcgaa tgtcgtagct gagtgcacgc taaaggagag tgcagctatc aaccagatcc

```

MMLEF1

```

1981 tgggcagaag atggcacgcc ctctcccggg aagagcaggc caataactat gaactagcac
2041 ggaaagagag acagctacac atgcagcttt atccaggctg gtcagcgcca gacaattatg
2101 gcaagaagaa gaagaggaag agagagaagc tacaggagtc gacttcaggt acagggtcca
2161 gaatgacagc tgctacatc TGAaacatgg tggtaagaga agctccttc caacgtgcaa
      <- LEF1 cds end
2221 agccaaggca ggcaccccag gccctcttct ggagatggaa gcttggtgaa accccagact
2281 gtctccacag cttgcccggc tgaccccaag gaacactgac agcaacctta cctgaggtc
2341 actgctagcg ctgacccgaa gacacagtca ctgccacctc ttccttctgt cgtctactgc
2401 aagcgccgac ttccaaaaag aaagccgaaa acggttggtt cggaaaaaaa aaaaaaaaaa

```

LOCUS RNBCOP 3073 bp RNA ROD 07-FEB-1991
 DEFINITION Rat beta COP sequence.
 ACCESSION X57228
 SOURCE Rat mRNA.
 REFERENCE 1 (bases 1 to 3073)
 AUTHORS Duden,R., Griffiths,G., Frank,R., Argos,P. and Kreis,T.E.
 TITLE beta COP, a 110kD protein associated with non-clathrin coated vesicles and the Golgi complex shows homology to beta adaptin
 JOURNAL Cell 64, 649-665 (1991)
 COMMENT HIV-1 nef has been found to bind beta-COP, a coat protein from non-clathrin-coated vesicles. This interaction may be important in promoting the intracellular sequestration of CD4.
 beta COP shows significant sequence homology to beta adaptin (clathrin-coated protein).
 NCBI gi: 55818

BASE COUNT 893 a 643 c 691 g 846 t

ORIGIN

```

1 cttcccgacc cccaccccca cggagccaga tttaccctggc acctttctgto cttttcatct
61 tgtttttgagt tggttataaa caggaaccAT Gaccgcagct gagaacgtgt gctatacgtt
    beta COP cds start ->
121 aattaacgtg ccaatggact cagaaccccc ttctgaaatc agcttaaaaa atgaacctgga
181 gaaaggcgat gtgaagtcaa agactgaagc tttgaagaaa gtgatcatta tgattctgaa
241 tggggaaaag cttcctggac tctgatgac gatcattcgt ttctgtctgc ctcttcagga
301 tcacaccatc aaaaagtgtc ttctggctct ctgggaaatt gttccgaaaa caactccaga
361 tgggaggttc ttacatgaaa tgattcttgt gtgtgatgca tacagaaagg atctccagca
421 tctaatgag tttattctgt gatctactct tegtttcctt tgcaaatga aggaagcaga
481 attgctggaa cctctgatgc ctgctatccg tgcctgtttg gaacatcgtc acagctatgt
541 taggagaaac gctgttttgg ccactctacac catctacaga aattttgaaa atcttatacc
601 tgatgctcct gagctgatac atgattttct ggtaaacgag aaggatgcaa gctgcaaaag
661 aaatgcattt atgatgctca ttcatgcaga tcaggatcga gctttggatt atttaagtac
721 atgtattgat caagttcaga catttggaga cattctacag ttggttattg ttgaactaat
781 ttataaggtc tgcctatgta atccatcaga gagggcccggt ttcatctgct gtatctataa
841 cttactgcag tcactctagtc ctgctgtaaa atacgaagct gctgggacac tggtgacact
901 gtcaagtgtc ccaactgcaa taaaggctgc tgcctcagtg tacattgatt taattattaa
961 ggaaagtgtat aacaatgtaa agctcattgt cctggaccga ttggtagaac taaaagagca
1021 tctgtctcac gagecaggtcc tccaagatct ggtcatggac atctacagag tactgagtao
1081 accagaccta gaagtggcca agaaaacgct gcagtttagca ctggatcttg tctcatctag
1141 gaatgttgaa gagttggtta ttgtcttgaa gaaggaagta attaaaacaa ataactgtgc
1201 tgagcatgaa gacactgaca aataccgaca gcttctctgt cgaacactgc attcctgtct
1261 tgtccgattt ccagatatgg ctgcaaatgt tattctgtgt ctaatggaat ttctcagtga
1321 cagtaatgaa gcagcagctg ctgatgtctt ggagtttgtg cgtgaagcca ttcagcgtct
1381 tgacaacctg agaatgctga ttgttgagaa gatgctggaa gtcttccatg ctattaagtc
1441 tgtcaagatt tatcgaggag cactgtggat cctgggagag tactgcagta ctaaggaaga
1501 cattcagagt gtgatgactg aagtcgcgag gtcccttggg gagatcccaa tcttagaatc
1561 agaaataaag aaagaagctg gcgagttaaa accagaagaa gaataaactg ttgggcccggt
1621 tcagaaattg gttacagaga tgggcaccta tgcctactcag agtgcactca gcagctctag
1681 acctaccaag aaagaagagg acagaccacc cttgagagga ttctctctgg atggagattt
1741 ctttgttgct gcctcccttg ccacaactct gaccaagatt gcattacgct atgtagcatt
1801 ggttcaggag aagaaaaagc aaaactcttt tgttctgag gctatgttgc ttatggcaac
1861 catcttcat ttgggaaaat cctctcttcc taagaagcca atcactgat atgatgtaga
1921 ccgaatttcc ctgtgcttca aggtcttata tgaatgctca cctttaatga atgacatttt
1981 taataaggag tgcagacagt ctctttccca aatgttgtct gccaaactcg aagaagagaa
2041 actatcccaa aagaagaagt ctgaaaagag gaatgtaaca gtacagcctg atgaccccat
2101 ttctttcatg caactaactg ctaagaatga gatgaactgc aaggaggatc agttccagct
2161 gagtttgttg gcagcgatgg gtaacactca gaggaagag gcagcagacc ccctggcgtc
2221 caaactcaac aaggtcactc agttgacagg tttctctgat ccagtgtatg cagaagctta
2281 tgttcatgtc aatcagtatg atattgtcct ggatgttctt gttgtaaacc aaaccagtga
2341 tactttgcag aactgcacat tagagttagc tactctaggg gatctgaaac tctgtgaaaa
2401 accatctcct ttgactcttg ctctcatga ctttgcgaat attaaagcca atgtcaaaat

```

RNBCOP

```

2461 agcatcaaca gaaaatggaa taattttcgg caatatagtt tatgatgtct ctggagcagc
2521 aagtgcacaga aactgtgtgg tcttcagtgat catccacatt gacatcatgg actatatcca
2581 gcttgcgaact tgcactgatg ctgagttcgg tcagatgtgg gccgaatttg aatgggaaaa
2641 taaagtgaca gttaacacca acgtgactga cctgaatgat tatttacagc acatcctcaa
2701 gtccaccaac atgaagtgcc tgactccaga gaaggccctt tctggctact gtggcttcat
2761 ggcagccaac ctatatgtc gctctatatt tggagaggat gcacttgcaa atgtcagcat
2821 tgagaagcca gttcaccagg gaccagatgc tgcgtttact ggccatataa gaattcgtgc
2881 aaaaagtcag ggaatggcct tgagccttgg agataaaatc aacctgtctc aaaagaagac
2941 tagtctcTAA gaatgaacaa acggtcctcg aaagtgtaa gttaaaattt aggtgtgggc
      <- beta-COP cds end
3001 tgattggact ccgacgtttt tgtactcttt catgctttat gttttagaaa tttgagttca
3061 tgcagaattc att

```

LOCUS SCSUG1 1671 bp DNA PLN 09-JUL-1992
 DEFINITION S.cerevisiae SUG1 sequence, complete cds.
 ACCESSION X56400
 KEYWORDS ATPase; sug1 gene.
 SOURCE S.cerevisiae DNA.
 REFERENCE 1 (bases 1 to 1671)
 AUTHORS Swaffield,J.C., Bromberg,J.F. and Johnston,S.A.
 TITLE Alterations in a yeast protein resembling HIV Tat-binding protein
 relieve requirement for an acidic activation domain in GAL4
 JOURNAL Nature 357, 698-700 (1992)
 COMMENT SUG1 is part of a large family of proteins related to the 26S
 protease from human erythrocytes. Each of these proteins has
 Tat-binding activity [1]. The exact role of this family of
 proteins in cellular metabolism is not wholly clear; however,
 the 26S protease seems to regulate the degradation of some
 cyclins and in this manner has been implicated in modulating
 the stability of oncoproteins such as c-Mos, c-Myb, c-Myc,
 and p53 (Dubiel et al. Biol. Chem. Hoppe-Seyler 375: 237-240).
 NCBI gi: 4590
 BASE COUNT 529 a 289 c 366 g 487 t
 ORIGIN
 1 gaattogett tggaaatggc acogttgccc cacacgtaac aaaaagagcc ttttcttcaa
 61 taatacccat tgtttgaaag agctggettc tgactcgttt gtttaacctc tttttgcgac
 121 tgaaaacttg tttttctatt tccctttttc ttcgggtggg ttttttaata tgaatcata
 181 ttttgcatac ttctctaccc ttattcttgt tcgggaattt togtatttta aaactgtcaa
 241 aatgcgcagc aggtggcaaa agttgtaata tcagtaacag ttaggttagg ttttatgtaa
 301 acggaacaac agtatagtgg gatactatca cgacaATGac agctgctgta acatcctcaa
 yeast SUG1 cds start ->
 361 atatatgtatt agaaaccac gaaagtggta tcaaaccata ttttgaacag aagatccaag
 421 aaacagaatt aaagatccgc tccaaaacag aaatgggtcg cagactggaa gctcaaagga
 481 atgcattgaa tgacaaagta cgttttatca aggacgaact ggtctatata caggaacctg
 541 gatcttatgt ggggtgaagt ataaagatcg tgtctgacaa aaaagttctc gttaaagtgc
 601 aacctgaggg aaaatacatc gtggatgttg caaaagatat aaacgtgaag gacttaagg
 661 catctcaaag agtttgtcta aggagtgaact ctatatgtt gcataaagtt cttgagaata
 721 aggetgaccc actagtctcg ttgatgatgg tggaaaaagt tctgattcc acatacgata
 781 tggttggtag tttgacaaag caaataaagg agattaaaga agttattgaa ttgcccgtaa
 841 aacatcctga actttttgaa agtttgggta ttgcgcaacc aaaggggtgc atcttatatg
 901 gccccctgg tacagggaaa accttattgg caagagctgt cgcacatcac actgattgta
 961 aattcatcag ggtcagtggt gcggaactgg tacaaaagta tatcggcgaa ggttcccgta
 1021 tggtcctgta gctgtttgtg atggctagag aacatgctcc ctcaattate tttatggatg
 1081 aaatcgattc cattggctct actcgtgtag aaggttctgg tgggtggtgat tcagaagttc
 1141 aaagaacaat gttagaacta ctaaaccaat tggacgggtt tgaacctctc aaaaatatca
 1201 agatcataat ggccacgaat agactagata ttctagatcc agcacttttg agaccgggta
 1261 gaatagatag gaagattgaa tttccacctc caagtgtcgc agctagagct gaaattttaa
 1321 gaatocatto cagaaaaatg aatctaactc gtggtatcaa cttgagaaag gttgctgaaa
 1381 agatgaacgg ttgttctggt gcogatgtca aaggtgtctg tacagaagct ggtatgtatg
 1441 ctttaagaga aagaagaata cacgttactc aagaagactt cgaactcgtc gtgggtaagg
 1501 ttatgaacaa gaaccaagaa acggccattt ctgtcgccaa gctgttcaag TGAtgtattc
 <- yeast SUG1
 1561 gtttcacttg gtcattgta tatgttactt agtgtatgtg tataatatca tagtttagca
 1621 ttaactaaga ggcaaaatta taccgctgga tatacaggat ataacggta c

Communications

The enclosed diskettes comprise the electronic portion of Part V, providing HIV and SIV nucleic acid and protein sequences in ASCII characters. As a general rule, American and non-European users receive nucleic acid files in GenBank format with protein translations of coding sequences given in the features tables. European users receive nucleic acid files in EMBL format and protein files in Swiss-Prot format. Both PC-DOS and Apple Macintosh formatted diskettes are available; which format you receive depends upon what you have requested. If you wish to change the diskette format or the file formats, simply contact us at 505-665-1356 or fax a message to us at 505-665-3493. If you wish to contact us via the postal service, our address is: HIV Sequence Database, T-10, MS K710, LANL, Los Alamos, NM, 87545, U.S.A. The diskettes contain READ.ME files to provide you with the most recent information regarding their contents. No alignments are included on the diskettes in this release of the Compendium.

Sequences and alignments are also available electronically from a server here at Los Alamos called "atlas." In the following instructions for accessing atlas, what you actually type at your terminal is shown in **boldface** type [with an explanation shown in square brackets]. What the server responds with is shown indented in *italic* type.

ftp atlas.lanl.gov [this ftp command connects your computer to atlas]

Name (atlas.lanl.gov): [atlas asks for your user name]

anonymous [for offsite users anonymous is the only name that works]

Guest login ok, send ident as password.

Password:

[enter your e-mail address or your name as ident]

Guest login ok. access restrictions apply.

cd pub [change to the directory called pub]

CWD command successful

cd aids-db

ls [this lists the files and directories inside of aids-db]

AIDS.A.TAR.Z, AIDS.E.TAR.Z, NUC.A, PRO.A, ALIGN, NUC.E, PRO.E, PROGS, README.TXT, ROADMAP

pwd [to show your location in the file hierarchy]

"/pub/aids-db" is current directory.

cd NUC.A/HIV_1 [move down 2 directories]

ls [all the GenBank format files are listed]

get hivrls4 [this copies the file hivrls4 to your home computer]

mget *.nf [this copies all files in the current directory with the suffix .nf to your home computer]

bye [this disconnects you from the server]

A map of the aids-db directories and files on atlas is reproduced on the following page. Many of the directories contain ReadMe files that describe the contents.

PARTIAL MAP OF HIV DATABASE ON "ATLAS"

KEY

symbol	meaning
BOLD TYPE	Directories
<i>italic type</i>	Files
→	Files and directories analogous to the one shown expanded exist below here
gray boxes	Directories of a similar type and with similar contents