

REPORT OF THE SECOND HUMAN GENOME DIVERSITY WORKSHOP
PENNSYLVANIA STATE UNIVERSITY
OCTOBER 29-31, 1992

WRITTEN BY THE ORGANIZING COMMITTEE FOR THE SECOND WORKSHOP
(WEISS, CAVALLI-SFORZA, FELDMAN, KING)

INTRODUCTION

The Second Human Genome Diversity Workshop was successfully held at Penn State University from October 29-31, 1992. The Workshop was essentially organized around 7 groups, each comprising approximately 10 participants, representing the sampling issues in different regions of the world. These groups worked independently, using a common format provided by the organizers; this was adjusted as needed by the individual groups.

The Workshop began with a presentation of the mandate to the participants, and of the procedures to be followed during the workshop. Dr Feldman presented a summary of the results from the First Workshop (the report of which had previously been circulated to the participants). He and the other organizers also presented brief comments giving their perspective on the objectives of the Second Workshop.

Dr Julia Bodmer discussed the study of European genetic diversity, especially in the context of the HLA experience there, and of plans to extend such studies in the coming years. She also discussed surveys of world HLA laboratories in regard to resources related to Human Genome Diversity.

Dr Mark Weiss, who had been an important original administrative sponsor of the workshop series when he was Physical Anthropology director at the U.S. National Science Foundation, presented a few points to the group. These concerned the relevance of nonhuman primate studies for understanding how demographic processes, such as mate exchange between local groups, affected the local dispersion of genetic variation. Primate population geneticists have some relevant experience in interpreting variation at this local level, in particular, with various DNA 'fingerprinting' methods. This experience may be relevant to the Human Genome Diversity Project, in terms of practical and statistical issues.

The Workshop proceeded for most of its 3 days in individual regional groups. At the end, each group presented an overview of its written report, in a plenary session. At that time, various issues of general concern were raised. The current report includes edited, standardized versions of the 7 original regional reports [not included]. The organizers of this meeting, listed above, have attempted to make the enclosed read consistently and clearly, but have attempted not to make any substantive changes to the reports themselves. Participants have been consulted to clarify details unclear in the written reports, and the regional-group leaders have read the edited version to see that it is in fact representative.

Although it cannot be said that there was unanimity among the participants, and there was spirited discussion of various ways to approach the problems of a worldwide sample, and of the populations to be included, there was a remarkable degree of coherence and resolution of difference by the time the final recommendations were made.

Of the organizers of the series of workshops (i.e., those listed on the grant application which funded them), Dr Kidd was unable to participate in the Second Workshop due to illness. Drs Piazza, Pääbo, and (Walter) Bodmer were unable to attend, and did not participate in the organization. Everyone, however, has been given a chance to read and comment on this report. Feedback relevant to clarity and the proceedings has been incorporated.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

DISCLAIMER

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

The participants were limited in number by the cost constraints of the symposium grant. The organizers did their best to invite an unquestionably distinguished worldwide panel of experts. However, funds were limited, and clearly numerous others who would have been qualified could not be invited. To augment the number of expert voices at the meetings, invitations were extended to persons who could support their own costs. A few such persons attended. Also, because they could contribute at little cost to the grant, several members of the Penn State Department of Anthropology also participated.

Before presenting the individual regional reports, we present several overall issues that were raised by the different regional groups. Generally, there was agreement among all the groups on these issues.

GENERAL ISSUES

1. *The need to mix cell-line sampling with larger samples*

It was clear that participants felt that collecting only a limited number of samples for cell-line transformation restricted the scope of the Project in several ways. Many aspects of the Project could be improved by the addition of larger numbers of samples from the studied populations and/or from increasing the number of populations sampled. Increasing the scope by collecting samples for DNA extraction would greatly improve the overall value of the Project. This is consistent with the recommendations of the First (Statistical) Workshop.

2. *The need to address important, but regionally different, disease-related problems*

It is important, from both an academic and a funding point of view, to note that many questions related to disease be considered in this Project, and that the sample relate to such issues. In this way, the Project can contribute to the well-being of many regions of the world in ways that go far beyond the cataloguing of variation and reconstruction of history.

Several disease-related issues were raised specifically by the regional groups. The study of malaria-related genetic variation has long been important in most of tropical Africa and Eurasia. There is a need for continued work on this problem, because new variants are still being discovered and many such variants may exist in populations not yet studied. Similarly, patterns of variation at the HLA loci suggest many aspects of genetic adaptation to disease antigens that are not yet understood. Genetic variation in the industrialized and/or non-tropical parts of the world has been studied extensively, but important problems in the rest of the world have hardly been examined.

There are special problems in regard to leukemia and AIDS viruses. One of the problems is the importance of careful collection of samples so that field and laboratory workers are protected. Equally important, individuals who are sampled and who are found to be HIV-infected must have their confidentiality protected. Generally, no one who is sampled should be at risk of being discriminated against because of any aspect of his disease status (or genotype at any loci). Areas in which rates of infection are discovered to be high should be identified to relevant health authorities, and it would be important generally for the Project to involve health surveys or delivery as part of the studies in many parts of the world.

It is also now known that mutations related to susceptibility to specific diseases differ among world populations. This statement applies to severe childhood diseases like phenylketonuria and cystic fibrosis, as well as to chronic diseases such as hyperlipidemias, cancers, and perhaps chronic infectious diseases. Little is known about genetic variation in susceptibility to acute infectious or to parasitic diseases (other than malaria), and this Project represents an important opportunity to study such problems.

This can be helpful in two major ways. Specific disease-susceptibility problems can be studied as part of this Project by collecting environmental and family data in selected

parts of the world (examples would be several viral diseases in Papua New Guinea, diabetes in Amerindians and various Pacific islands, various types of cancer in China and arctic populations, hepatitis in various parts of the world). Second, although most genetic diseases may be rare, even small samples from given areas can provide information on the chromosomal haplotypes that exist in those areas. This can be useful in searching for previously unknown mutations (i.e., that do not occur on frequently studied populations such as those of Europe and Japan).

Genetic counseling will not soon be available to many parts of the developing world, but that day will come sometime, and a large fraction of the world's urban populations (e.g., in Africa, Asia, India) have not yet been studied in this way, but could benefit in the near future as genetic diagnostic methods become less expensive and more widely distributed.

3. *The need to see that sampled groups' privacy and well-being are not harmed by this study*

It is standard practice by careful anthropologists not to reveal anything about populations they study that can be used to harm the populations themselves. It is not always possible to prevent exploitation of remote or exotic populations by the outside world, but it is important to do whatever we can to see that the Project does no harm. Indeed, from the beginning it has been hoped that the Project may bring to public attention various issues in which populations are being seriously harmed by the outside world. Such issues are regularly raised by human rights, ecological, and other groups.

In this Project, it is important that the sampled populations understand as fully as practical the implications of the study, and its purposes. Fully informed access to the population, approved by the government of the nation in which they reside, must be provided by the investigator wishing to be sponsored by the Project. The population itself must demonstrably be provided a full level of informed consent. Religious or other cultural concerns must be protected.

This difficult subject will be considered at a future Workshop, and the results will be at the core of Project proposals for funding.

4. *The need to have the actual sampling done as part of full and legitimate anthropological (linguistic, health, social, etc.) studies of the group, by investigators trusted by and familiar to the studied population*

5. *The nature of this report as advisory, but not fixed or proscriptive*

The organizers asked the participants to make specific recommendations of populations that should be sampled, because it was felt that a report that is too general will be difficult to use as the basis for a successful application for Project funds. One reason is that a specific proposal clearly shows the issues and the feasibility of the proposed worldwide study. Another is that a specific agenda is needed in order to draw up a requested budget.

It is clear, however, that it will be impossible to sample *exactly* all the suggested populations on any list drawn up by a general workshop in advance. Some suggested groups may not be accessible. It may be that no qualified anthropologist has access to the group, or that permission to sample it may not be available. At the same time, other investigators may wish to sample different groups, not named in this report, but that are comparable in language, size, degree of isolation or admixture, or relevance to questions of importance to human diversity. Some regions were not adequately represented by the organizers — and for others, different investigators may differ on the recommendations. In addition, some populations may already have been sampled for other reasons by the time the Project is implemented, and this should free up funds for alternative or additional sampling.

The mechanism for sampling that is finally established for the Project must allow for additions or subtractions, or alterations of priority, as knowledge base grows, and to be

consistent with practical constraints. This report is to be viewed as a 'living' document, which identifies major issues and some detailed specific recommendations but which can be changed to meet the needs and practical limits of the Project.

At the same time, it is important that actual sampling NOT be done in a way unrelated to this report or its considerations. The constraint to use this document as a framework for sampling guarantees that the Project will remain an open one, properly determined by sound anthropological expertise. There must be no suspicion that a small number of investigators will use the funding base of this Project to achieve their own private objectives. Anthropologists, linguists, and other qualified persons (including persons concerned with ethical issues) must be included in review panels that judge applications for funding specific populations, to ensure this openness and relevance to the issues raised in the report.

6. *The need for detailed input from specialists in areas not represented by participants at the Second Workshop*

The need to continue to improve the sampling frame as knowledge accumulates, and as input from other investigators is received, has been stated. However, it is important to state clearly that the participants at the Workshop were aware that they represented only a very limited set of views on this subject. Although the participants were among the world's most prominent scholars in their respective areas, there are hundreds of other qualified investigators who could not attend (for practical and budgetary reasons, among others). Finally, many areas were not represented at the meeting. These include the region including Pakistan, the Caribbean, Australia, the Arab-speaking parts of the world, West Africa, and others.

The participants did not wish to represent themselves as having expertise in these areas. The report attempts to identify major issues in these regions, but it will be important that any final Project include a representative series of samples as specified by help provided by experts in the regions.

Generally, the organizers will attempt to continue to publicize the need for assistance and collaboration as widely as possible in world anthropological journals and the like. Recommendations will be changed as such information is received. Finally, wide dissemination of the availability of funding for the Project will be made so that qualified investigators, with access to the recommended populations, can be supported to collect the data.

The intention of the organizers is that the actual data collection will be done not by them, but by investigators (anthropologists, health workers, and so on) who have regular working relationships with the populations. The organizers' role will be to make the availability of the Project widely known so that such persons will apply for funding and will collect the data.

7. *The need for joint sponsorship of the Project by many nations and funders*

It is clear that U.S. funding organizations cannot and will not be able to provide the entire amount of funding for this Project. The Project cannot succeed unless it is truly a worldwide effort. Those nations that have funding to provide to scientific projects must help to support this Project. Agencies such as the United Nations and the science or health ministries of wealthy or industrial nations worldwide must be involved. This will not only make it possible to pay the costs of such a project, but will spread the control of the Project around the world and will help to guarantee that the legitimate interests of the different regions of the world are represented fully.

8. *The need for the Project to make permanent, rather than temporary, technology transfer to laboratories in areas of the world that will be involved in sample processing but that do not have such capability currently*

The organizers have made it clear, from the original application for funding, that an objective of the Human Genome Diversity Project would be to help establish or upgrade

laboratories in many parts of the world, so that those laboratories could do modern genetics work and could participate in the Project. An immediate objective will be to enable laboratories near to Project field collection sites. Many laboratories around the world may be able, or nearly able, to do cell transformation or other sample-preservation procedures. Those that need it should be provided with the equipment and the training for this Project. One series of labs that would be optimal for this would be regional HLA laboratories. The Project intends to establish a close collaborative relationship with the HLA establishment worldwide for this purpose.

In addition, it was pointed out clearly that it will not be sufficient merely to train and support local laboratories solely for a through-put processing role in the Project. The Project cannot be a worldwide success, of benefit to our species as a whole, unless permanent resources are established in all its different regions. The suggestion was made that when the Project is planned, it will be important to consider how regional laboratories, established or funded by the Project, can retain control of samples specific to their region, to serve distribution, data banking, and analytic functions. One possibility is that such laboratories retain stocks of all cell lines processed by them, as well as retaining the distribution function for all DNA samples that they process. If the latter is limited in amount, regional studies would have to be done by investigators working with the regional labs. A condition for establishing such labs, however, would have to be that they cooperate on an open basis with investigators interested in their region.

9. *The need to develop a standardized "questionnaire" for ascertaining demographic, ecologic, and other information about individuals and populations included in the Project sample*

The next Workshop, or some other effort, should develop a standard set of information that would always be collected in association with studies funded by the Project. Other studies around the world that might contribute to the Project, even if not funded directly by it, should be asked to collect and provide the same kind of ancillary data. Age, sex, social-group affiliation, parental and grandparental birthplace, genealogical relationships, language spoken, geographic location and ecotype, climate, and other variables should be collected. Dr Baer submitted some suggested formats, which are retained by the organizers for future use; a copy is available to any interested party who requests it.

10. *Sensitivities must be recognized to the maximum possible extent*

Many issues of sensitivity to various points of view need to be considered. For example, the establishment of permanent cell lines needs to be explained in terms that are understandable, but that do not mislead subjects in any population. English terms such as 'immortalization' of cell lines can be badly misunderstood, and have been avoided in this report. Similarly, there is no fully acceptable way to refer to populations that are in danger of physical extinction or of disruption as integral genetic units (gene pools); some existing terms such as 'endangered' populations can have various connotations. Many populations around the world, especially isolates living traditional lifestyles, will soon disappear as independent units, because of disease, economic or physical deprivation, genetic admixture, or cultural assimilation. In this report, we refer to such groups as "Isolates of Historic Interest" (IHI's), because they represent groups that should be sampled before they disappear as integral units so that their role in human history can be preserved.

The organizers have attempted to use terminology in this report that is as sensitive as possible in these regards. Undoubtedly, errors have been made. There is no attempt to make judgments about the inherent values of any people or populations, cultures, languages, and the like. The ethics review mechanisms of the Project will be designed to protect, to the extent that it is possible, against any such insensitivities.

=====

THINGS ADDED FROM OTHER FILES AND NEEDING TO BE FOLDED IN

=====

The elucidation of genetic variation in underserved populations, including most African populations, will provide important baseline data, which will be of value in studying these genetic diseases. The markers known to be linked to disease loci have been almost exclusively defined in Caucasoid peoples and they may not, therefore, be informative in the peoples of Africa.

The role of genetic factors in conferring on the individual protection against infectious (including parasitic) diseases is likely to be best understood by studying individuals living in ecological areas where these diseases are still prevalent, that is, in regions of developing countries. This has been strikingly true for the malaria-protective traits, like sickle cell haemoglobin, the thalassaemias, and G6PD deficiency, and may well be the case for many other genes.

The study of the human genome, including elucidating its diversity, should not detract, in any way, from the need to address the health problems of the Third World, the bulk of which could be solved by the wide-scale application of knowledge already available; what is needed is the will to do so and the commitment of adequate resources.

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.
