M. I. Simon

Construction of genome-wide physical BAC contigs using mapped
cDNA as probes: Toward an integrated BAC library resource for
genome sequencing and analysis

DOE/ER/60891--77

DE-FG03 89ER60891

Steve C. Mitchell, Diana Bocskai, Yicheng Cao, Robert Xuequn Xu,
Mei Wang, Troy Moore*, So Hee Dho, Enrique Colayco, Christie Gomez,
Gabriella Rodriguez, Annabel Echeverria, Melvin I. Simon,
and Ung-Jin Kim

Progress Report July 1995 - Jan. 1997

*Annual Report
1996*

The goal of human genome project is to characterize and sequence entire genomes of human and several model organisms, thus providing complete sets of information on the entire structure of transcribed, regulatory and other functional regions for these organisms. In the past years, a number of useful genetic and physical markers on human and mouse genomes have been made available along with the advent of BAC library resources for these organisms. The advances in technology and resource development made it feasible to efficiently construct genome-wide physical BAC contigs for human and other genomes. Currently, over 30,000 mapped STSs and 27,000 mapped Unigenes are available for human genome mapping. ESTs and cDNAs are excellent resources for building contig maps for two reasons. Firstly, they exist in two alternative forms - as both sequence information for PCR primer pairs, and cDNA genomic libraries efficiently for large number of DNA probes by combining over 100 cDNA probes in each hybridization. Second, the linkage and order of genes are rather conserved among human, mouse and other model organisms. Therefore, gene markers have advantages over random anonymous STSs in building maps for comparative genomic studies.

As a preliminary work for the ongoing "BAC-EST" project, we are currently screening our human BAC libraries with thousands of cDNA probes. We have thus far used over 3,000 Unigene probes and the number will increase to 7,000 in this year. Our goal is to screen the library with at least 27,000 markers, most of which are in the form of cDNA probes. This represents at least 1 marker per every 100 kb of euchromatin regions.

We plan to deconvolute the positive BACs to each marker by sorting the library into groups of BACs that are positive to specific pools, arraying each group on small hybridization filters, then hybridizing the filters with individual probeto precisely align these mapped BAC clones against any known sequence

contigs by means of sequence match. Putative contigs or clone overlaps identified by markers or sequence match are verified via restriction fingerprint analysis. The BAC clone resources integrated with physical mapping information will be useful for building sequence-ready contigs on any chromosomal region.

- Research Genetics, Huntsville, AL

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

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
electronic image products. Images are
produced from the best available original
document.**