Human Genome news

Sponsored by the National Institutes of Health and the U.S. Department of Energy

ISSN: 1050-6101

Vol. 6, No. 4, November 1994

Genome Centers Provide Overview

Goals, Accomplishments, Resources Highlighted

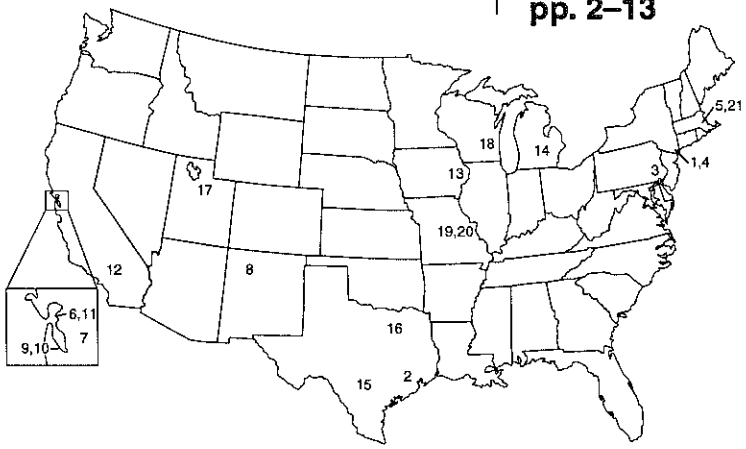
Centers Described, pp. 2–13

Specialized DOE and NIH human genome research centers foster a multidisciplinary approach for addressing major tasks of the U.S. Human Genome Project. Current efforts focus on genetic and physical mapping, DNA sequencing, informatics and technology development, and the societal impact of new genetic tools and information. New genome technology and resources are developed and shared with outside collaborators and the entire research community; many centers offer training opportunities and outreach programs as well. In addition, collaborations with the private sector are encouraged for developing and commercializing new products resulting from genome research. (Contact specific centers for information on technology transfer.)

The DOE Office of Health and Environmental Research (OHER) has established genome centers at three national laboratories; centers are funded annually and peer reviewed through site visits every 2 to 3 years. OHER devoted $29.4 million or 47% of its genome program budget to these laboratories in FY 1994.

The NIH National Center for Human Genome Research supports 18 multidisciplinary Genome Science and Technology Center projects (called GESTECs) in addition to a set of regular research (R01) grants and other funding mechanisms. NCHGR spends around $56.5 million or 53% of its $107-million FY 1994 budget on GESTECs.

Funding levels vary among the centers and GESTECs, reflecting their different goals and scopes.



1. Albert Einstein College of Medicine
2. Baylor College of Medicine
3. Children's Hospital of Philadelphia
4. Columbia University College of Physicians and Surgeons
5. Genome Therapeutics Corporation (Collaborative Research Division) Genome Sequencing Center
6. Lawrence Berkeley Laboratory
7. Lawrence Livermore National Laboratory
8. Los Alamos National Laboratory
9. Stanford Human Genome Center
10. Stanford University DNA Sequence and Technology Center
11. University of California, Berkeley, *Drosophila* Genome Center
12. University of California, Irvine
13. University of Iowa Cooperative Human Linkage Center
14. University of Michigan Medical Center
15. University of Texas Health Science Center at San Antonio
16. University of Texas Southwestern Medical Center at Dallas
17. University Of Utah
18. University of Wisconsin, Madison, *E.coli* Genome Center
19. Washington University School of Medicine
20. Washington University School of Medicine Genome Sequencing Center
21. Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology

Genetic Map Goal Met Ahead of Schedule

Collaborative Efforts Yield Map with Centimorgan Density

An international team of researchers has constructed the most detailed comprehensive human genetic map yet published, with 5840 loci spaced at a mean interval of 0.7 cM. The new high-density map is a compilation of linkage data generated during the past decade by groups led by Jeffrey Murray [Cooperative Human Linkage Center (CHLC)], Jean Weissenbach (Généthon), Ray White (University of Utah), David Ward (Yale University); and over 100 CEPH collaborators.

(see Genetic Maps, p. 14)

In This Issue. . .

Page	
1	Genome Centers Provide Overview
1	Genetic Map Goal Met
15	NIH Explores Cancer-Risk Testing
15	*Arabidopsis* Report Issued
16	Informatics Workshop Held
17	*GDB Forum:* WWW Browser
18	Calendars
19	Funding Guidelines
20	Subscriptions; Center Acronyms

Genome Center Acronym List, p. 20

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
1. ALBERT EINSTEIN COLLEGE OF MEDICINE (AECM) (NIH, established 1993) **RAJU KUCHERLAPATI** *CONTACT:* Kucherlapati (718/430-2069, Fax: -8776, *kucherla@aecom.yu.edu*); AECM; 1300 Morris Park Ave.; Bronx, NY 10461. **OTHER KEY RESEARCHERS** Kenneth Kidd (Yale Univ.) Kenneth Krauter Perry Miller (Yale Univ.) David Ward (Yale Univ.)	Construction of a high-resolution map of chromosome 12, integrating the cytogenetic, linkage, and physical maps; localization of all known chromosome 12 genes onto specific YACs.
2. BAYLOR COLLEGE OF MEDICINE (BCM) (NIH, established 1990) **C. THOMAS CASKEY** David L. Nelson, Associate Director *CONTACTS:* Sandra McMurtry, Administrator (713/798-6524, Fax: -5386, *mcmurtry@bcm.tmc.edu*) or David Nelson (-3122; *nelson@bcm.tmc.edu*); BCM; One Baylor Plaza; Houston, TX 77030-3498. **OTHER KEY RESEARCHERS** Antonio Baldini A. Craig Chinault Robert W. Cottingham, Jr. Richard A. Gibbs Gail E. Herman Charles B. Lawrence Cheng Chi Lee James R. Lupski John M. Shumaker Randall F. Smith Huda Y. Zoghbi	Six core facilities provide support for independently funded physical mapping and positional cloning projects and development of technology, with emphasis on human chromosomes 6, 15, 17, and X and the mouse X chromosome.
3. CHILDREN'S HOSPITAL OF PHILADELPHIA (CHOP) (NIH, established 1991) **BEVERLY S. EMANUEL** Kurt Fischbeck, Associate Director (Univ. of Penna. School of Med.) *CONTACT:* Emanuel (215/590-3856, Fax: -3764, *beverly@mail.med.upenn.edu*); CHOP; 34th St. & Civic Center Blvd.; Philadelphia, PA 19104. **OTHER KEY RESEARCHERS** Callum Bell Jaclyn Biegel Marcia Budarf Kenneth Buetow (Fox Chase Cancer Center) Chris Overton (Univ. of Penna. School of Med.) Eric Rappaport Bruce Roe (Univ. of Okla.) David Searls (Univ. of Penna. School of Med.) Saul Surrey	Construction of high-resolution genetic and physical maps of chromosome 22 in YACs and cosmids. Production of a "sequence-ready" contig of the euchromatic portion of chromosome 22. Sequencing of 25 Mb of chromosome 22 to produce a "sequence-read" map from the "sequence-ready" map.
4. COLUMBIA UNIVERSITY (CU) COLLEGE OF PHYSICIANS AND SURGEONS (NIH, established 1993) **ARGIRIS EFSTRATIADIS** Isidore Edelman, Administrative Director *CONTACTS:* Edelman (212/305-3440, Fax: -1191; *edelman@cuccfa.ccc.columbia.edu*) or Efstratiadis (-5773, Fax: /923-2090, *arg@cuccfa.ccc.columbia.edu*); CU; 701 W. 168th St.; New York, NY 10032. **OTHER KEY RESEARCHERS** Philip Bourne Eftihia Cayanis Stuart G. Fischer Rodney Rothstein James Russo Eric Schon M. Bento Soares Dorothy Warburton Peisen Zhang	Construction of a high-resolution physical map of human chromosome 13, consisting of ordered cosmid contigs aligned to YACs, annotated with STSs at 100-kb average intervals and with cytogenetic assignment of contig members by in situ hybridization. Selection and cytogenetic assignment of cDNAs representing about 20% of the total estimated chromosome 13 genes. Development and integration of gene-based STSs into the physical map.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Characterization of a human chromosome 12 cosmid library and mapping of more than 200 cosmids to specific regions of chromosome 12 by FISH. Generation of monomorphic and polymorphic markers from cosmids mapped to chromosome 12. Construction of YAC contigs covering nearly 40% of the chromosome. Each contig includes polymorphic markers generated by several groups, including Généthon, Univ. of Iowa, CHLC, MMRF, and AECM. Additional markers have been generated from the ends of YACs by AECM. The YAC map also incorporates known chromosome 12 genes. Integration of cytogenetic maps with emerging physical maps by FISH mapping of over 100 individual YACs from contigs to metaphase chromosomes. Development of a Sybase-based relational database for chromosome 12 reagents and maps.	Arrayed chromosome 12 cosmid library for replica plating. CEPH chromosome 12–specific YAC sublibrary. STS collections for mapped positions of chromosome 12 cosmids, YACs, and genes as determined by FISH mapping. URL: *http://paella.med.yale.edu/home.html*
Identification of several new disease loci, including Fragile X syndrome, myotonic dystrophy (DM), Charcot-Marie-Tooth disease, Kallman syndrome, glycerol kinase deficiency, Miller-Dieker lissencephaly, Lowe syndrome, and spinocerebellar ataxia type1. Identification of >2000 YAC clones by PCR screening for target chromosomes. Assembly of contigs in target regions (many contigs >1 Mb from X, 17, 6, and 15). Development of methods for YAC characterization and contig assembly. Development of GDB-Lite and a library screening and clone characterization database capable of periodic automatic submission to GDB. Completion of moderate-scale sequencing projects encompassing over 200 kb of human DNA from the X and other chromosomes (including the genes for Duchenne muscular dystrophy, candidate gene for X-inactivation center, Fragile X locus, and DM). Establishment of several hundred cell lines from patients with chromosome-linked disorders. Assembly of a 35-Mb contig in YACs for Xp21-ter. Development of a Mosaic browser for integrated YAC data scanning.	Total human YAC libraries (CEPH and Wash. Univ.]. Chromosome X YAC libraries (BCM and CHOP). Chromosome 17 and X cosmid libraries (LANL and LLNL.) Screening service for chromosome 17p. Total mouse YAC libraries (St. Mary's and ICRF). DNA pools available for multistep PCR screening (*yaclab@bcm.tcm.edu*). Numerous patient cell lines and chromosome mapping cell panels for 6, 17, and X (hybrids, deletions, etc.). STSs and YACs for many loci on target chromosomes. Database software. NCHGR predoctoral training grant (Chinault). Quarterly newsletter. URL: *http://gc.bcm.tmc.edu:8088*
Construction of a chromosome 22 framework map based on 20 STRPs. Regional assignment of over 300 anchor markers on a regional mapping panel (25 bins). Construction of 15 YAC contigs covering at least 70% of the chromosome. Over 250 STSs screened to identify almost 700 YACs. Contigs contain as many as 78 STSs. Production of clones and markers, including (1) construction of a 1.2x coverage YAC library from a chromosome 22–only somatic cell hybrid; (2) over 200 STSs, 48 CA+TG repeat clones, and 70 *Not* I junction clones; and (3) creation of a normalized human liver cDNA library. Construction of a 100-member radiation hybrid panel (characterized by FISH) and 5 somatic cell hybrids from cell lines with chromosome 22 deletions or translocations. Banking of cell lines with chromosome 22 abnormalities. Development of informatics capabilities, including (1) enhancement of existing Sybase to Quintus Prolog interface to support transparent access of all Sybase data types from Prolog; (2) creation of tool for visualizing DNA sequence features in short sequencing runs; and (3) development of Postscript tool for generating publication-quality depictions of cytogenetic maps.	477 human chromosome 22 STSs from ftp site [*cbil.humgen.upenn.edu* (type: *cdpub/22*)] or WWW server (*http://www.cis.upenn.edu/~cbil/chr22db/chr22dbhome.html*). Somatic cell hybrid regional mapping panel and radiation hybrid panel. Cell lines from patients with chromosome 22 abnormalities. Center-developed software. YAC and cosmid screening resource service for chromosome 22 STSs and probes. Arrayed chromosome 22 cosmid library. Total human YAC libraries (CEPH).
Coverage of an estimated 50% of chromosome 13 with over 500 cosmid contigs aligned to more than 200 YACs; development of 115 STSs. Identification of over 90 cDNAs hybridizing to single-copy sequences of chromosome 13 and development of corresponding STSs. Localization of 265 markers (YACs, cosmids, cDNAs, and half-linking clones) to cytogenetic bands of chromosome 13 by in situ hybridization. Development and genetic mapping of 24 polymorphic STSs spanning chromosome 13. Contribution to Wilson's disease gene identification. Development of relational database and software (integrated mapping package) supporting physical mapping from YAC and cosmid cross-hybridization relationships and STS data.	Total human YAC libraries (CEPH midi- and mega-YACs). Arrayed chromosome 13 midi- and mega-YAC sublibraries (1431 and 827 colonies, respectively; replicate filters and PCR screening pools are available). Arrayed chromosome 13 cosmid library (LANL; 16,996 colonies; replicate filters are available). YAC and cosmid screening service for chromosome 13 probes. Normalized cDNA libraries. Integrated mapping package software. Chromosome 13 somatic cell hybrid mapping panel and deletion cell lines.

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
5. GENOME THERAPEUTICS CORPORATION (COLLABORATIVE RESEARCH DIVISION) GENOME SEQUENCING CENTER (NIH, established 1994) **JEN-I MAO** *CONTACT:* Mao (617/893-5007 x242, Fax: /642-0310, *mao@cric.com);* 100 Beaver St.; Waltham, MA 02154. **OTHER KEY RESEARCHERS** George Church (Harvard Medical School) Ronald Lundstrom Peter Richterich Hershel Safer Douglas R. Smith	Improvement of the overall throughput and cost-efficiency of multiplex sequencing by 10 times. Sequencing of *Mycobacterium leprae* and *Mycobacterium tuberculosis* genomes and large regions of the human genome. Automation and improvement of front-end multiplex sequencing protocols. Development of a high-throughput integrated system for automated hybridization and infrared fluorescence detection. Development and implementation of software tools to support sequence-production needs including base calling, sequence assembly, contig editing, and sequence analysis.
6. LAWRENCE BERKELEY LABORATORY (LBL) (DOE, established 1988) **MOHANDAS NARLA** *CONTACTS:* Narla (510/486-4251, Fax: -6746, *mohandas_narla@macmail.lbl.gov*) or Jennifer Knox, Executive Assistant to Center Director (*JLKnox@lbl.gov*); LBL; Human Genome Center; 74-157, 1 Cyclotron Road; Berkeley, CA 94720. **OTHER KEY RESEARCHERS** Jan-Fang Cheng, John McCarthy Joseph Jaklevic, Michael Palazzolo William Kimmerly, Martin Pollard William Kolbe, Edward Rubin Victor Markowitz, Ed Theil Christopher Martin, Manfred Zorn	Development and implementation of directed methodologies, automation, and informatics tools for cost-effective and accurate high-throughput human DNA sequencing. Data management and distribution through collaborations among biologists, the automation group, and computer scientists. Assembly of physical maps of complete chromosome arms rooted in templates appropriate for sequencing and development of capacity to produce genomic sequence of multiple megabases per year. One target is the chromosome 5q31 region. Discovery of novel human genes in targeted sequencing regions by analyzing mouse phenotypes created through alteration of genetic content of syntenic regions.
7. LAWRENCE LIVERMORE NATIONAL LABORATORY (LLNL) (DOE, established 1990) **ANTHONY V. CARRANO** *CONTACTS:* Linda Ashworth, Assistant to Center Director (510/422-5665, Fax: -2282, *ashworth1@llnl.gov*) or Carrano (-5698, *carrano1@llnl.gov*); LLNL; Human Genome Center; 7000 East Ave., L-452; P.O. Box 808; Livermore, CA 94551. **OTHER KEY RESEARCHERS** Joe Balch, Jane Lamerdin Mark Batzer, Richard Langlois Brigitte Brandriff, Greg Lennon Elbert Branscomb, Ray Mariella Emilio Garcia, Harvey Mohrenweiser Jeff Garnes, Anne Olsen Jeff Gingrich, Tom Slezak	Development of new cloning, mapping, and sequencing technologies focused on the assembly, closure, and characterization of a high-resolution ordered clone map of human chromosome 19. The final high-resolution map will consist of cosmid contigs with YACs, BACs, and PACs and an *Eco*R1 restriction map for the minimal spanning set of cosmids; the map will be aligned with genetic maps of chromosome 19. Isolation, mapping, and sequencing of chromosome 19 cDNAs with emphasis on full-length clones. Physical mapping and sequencing of targeted gene families in human and other relevant species. Comparative sequencing of selected mouse and human genomic regions. Development of software solutions for data management and integrated map display for physical mapping. Construction of NLGLP chromosome-specific lambda and cosmid libraries (with LANL) for distribution. Development of instrumentation in support of physical mapping and DNA sequencing. Transfer of technology and knowledge to industry, universities, and the public.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Completion of 1-Mb mycobacterial and human genomic sequences using software REPLICA and GTAC developed by Church, Leon Mintz, and Gary Gryn (HHMI, Harvard Medical School). About 900 kb are in GenBank. Development of improved internal sequencing standard and a set of 20-plex phagemid vectors for chemical and dideoxy multiplex sequencing. Demonstration of highly sensitive infrared detection from membranes of multiplex sequencing. Construction of prototype for automated hybridization of multiplex membranes. Development of software modules and programs for automated primer picking and assembly verification. Development of improved software for lane definition and band location on multiplex internal standards and near completion of improved base-calling software. Development of automated management tools for maintenance of individual sequencing projects. Development of program to assemble cosmids containing long repeated regions. When available, the mate information from double-ended sequencing is incorporated.	Mycobacterial and human sequence data. Improved multiplex phagemid vectors. Test data set for sequence assembly. X-gel, multiplex sequencing software for semi-automated lane finding and image analysis. Human chromosome 10 physical mapping data and mapped cosmid and YAC clones. URL: *http://www.cric.com*
Completion of 1.5 Mb of sequence, with more than 1 Mb already deposited in public databases. Directed strategy is roughly comparable for both invertebrate and mammalian templates, with 600 to 700 kb/yr generated. Completion of automation modules, including an image station that captures and analyzes mapping information from agarose gels, a colony picker, a robotic library replicator, and a modified Biomek that sets up PCR assays and sequencing reactions. A water-based thermocycler and a 12-channel oligonucleotide synthesizer are now in testing stages. Development of software, including mapping tools for generating distance-orientation-gene-size resolution (DOG) tag and transposon maps, database management tools, enhancements and new data models for ACEDB, software tools for sequencing and marker submission, and automatic logging and report generators for project management. Construction of 1.2 Mb P1 map in the interleukin gene cluster region of human chromosome 5q. Construction of 3-Mb P1 map in the Down syndrome region of human chromosome 21. Construction of in vivo transgenic mice library containing 2 Mb of human sequences from the Down syndrome region. Development of transgenic mice containing nested deletions in the syntenic region of human chromosome 5q31.	Automated colony picker and imaging station with supporting hardware for gel casting (M. Pollard, 510/486-4561). Databases for representing and displaying genomic data (E. Theil, -7501). Biological resources: mapped YACs, P1s, and cDNAs in the Down syndrome region of chromosome 21; mapped YACs and P1s in the interleukin gene cluster region of chromosome 5; in vivo transgenic mice library containing 2 Mb of human DNA from the Down syndrome region (for transgenic mice library: E. Rubin, -5072, Cheng, -6590). URL: *http://genome.lbl.gov/GenomeHome.html*
Coverage of an estimated 95% of chromosome 19 in cosmid contigs. Over 85% of the chromosome is spanned by ordered islands of cosmid, YAC, BAC, PAC, and P1 phage clones, with distances determined between >200 selected cosmids spaced an average 260 kb apart. FISH localization of over 400 cosmids to chromosome bands, including an ordered set of 190 cosmids spaced an average 260 kb apart. Localization of over 400 genes, genetic markers, and other loci at the cosmid level on the chromosome 19 map. Fine-scale mapping in selected regions, including the cytochrome P450, pregnancy-specific glycoprotein, fucosyltransferase, and zinc finger genes or gene families. Construction of *Eco*R1 complete digest restriction maps spanning 32 Mb of chromosome 19. Establishment of Mosaic access to current chromosome 19 physical map. Identification of novel *Alu* repeat sequences for studying human genetic variation. Sequence determination and comparison of over 250 kb from chromosome 19 regions, including DNA repair genes and their rodent homologs. Development of instrumentation supporting high-density filter production, DNA sequencing, and flow cytometry. Construction of NLGLP chromosome-specific cosmid libraries from sorted chromosomes for chromosomes 1, 2, 3, 7, 9, 12, 18, 19, 21, 22, X, and Y. Development of integrated mapping analysis software with interactive graphical display and linkage to local and national databases.	NLGLP large-insert lambda and cosmid chromosome-specific libraries. Cosmid filters, FISH mapping, and cDNA and YAC, BAC, PAC screening for chromosome 19 probes. Assistance in database development, systems management, networking, and contig assembly by fingerprinting and automated restriction mapping. Graduate and postdoctoral research training through the Institute of Genetics and Genomics at LLNL (Mohrenweiser, 510/423-0534). URL: *http://www-bio.llnl.gov/bbrp/genome.html*

¶ Publication

Handbook of Human Genetic Linkage, by Joseph Terwilliger (University of Oxford) and Jurg Ott (Columbia University), contains detailed instructions for carrying out linkage analysis, guiding the reader step by step through each process. Chapters are grouped into three main sections: two-point linkage analysis; multipoint linkage analysis; and such advanced topics as mutation rates, gene frequencies, linkage disequilibrium, complex diseases, and computer simulation. Each chapter ends with exercises that can be performed on an IBM PC–compatible computer, and solutions are supplied at the end of each section. [Johns Hopkins University Press; Hampden Station; Baltimore, MD 21211-2190 (800/537-5487 or 410/516-6956, Fax: -6998).] ◊

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
8. LOS ALAMOS NATIONAL LABORATORY (LANL) (DOE, established 1988) **ROBERT K. MOYZIS** Larry L. Deaven, Deputy Director *CONTACT:* Lynn Clark, Technical Coordinator (505/667-9376, Fax: -2891; *clark@telomere.lanl.gov*); LANL; Center for Human Genome Studies, MS M886; Los Alamos, NM 87545. **OTHER KEY RESEARCHERS** Michael Altherr, Norman Doggett, Joe Gatewood, Deborah Grady, Jim Jett, Dick Keller, Jon Longmire, Pat Medvick, Julie Meyne, Rob Pecherer, David Torney, Michael Yesley	Assembly of complete, high-resolution (<10 kb), sequence-ready chromosome 16 map; low-resolution (0.5 Mb) map of chromosome 5; and high-resolution, sequence-ready map of Cri du Chat region on chromosome arm 5p. Determination of molecular basis of chromosome structure and function and isolation of selected disease genes on chromosomes 5 and 16. Short-term development and support for large-scale physical mapping and sequencing projects and long-term development of tools for storage, manipulation, and analysis of genome data. Development and application of new methods for physical mapping and sequencing; use of robotics in handling and storing DNA fragments; construction of DNA libraries from flow-sorted chromosomes; and rapid, inexpensive, large-scale sequencing. Studies of ethical, legal, and social issues arising from the increased availability of genome data. Transfer of technologies and medically important information to industry and the medical community. Establishment of Internet access via WWW and Mosaic to physical mapping data.
9. STANFORD HUMAN GENOME CENTER (SHGC) (NIH, established 1990) **RICHARD M. MYERS** David R. Cox and Douglas Vollrath, Co-Directors *CONTACTS:* Myers (415/725-9687, Fax: -9689; *rmyers@camis.stanford.edu*), Cox (-8042, Fax: -8058, *cox@camis.stanford.edu*), Vollrath (/723-3290, Fax: -7016, *vollrath@genome.stanford.edu*), or Cristina Estébanez, Administration (/812-1915, Fax: -1916, *cxe@camis.stanford.edu*); SHGC; Dept. of Genetics; 855 California Ave.; Palo Alto, CA 94304. **OTHER KEY RESEARCHERS** Lane Conn, Sid Cowles, Jian-Bing Fan, Richard Goold, Cynthia Keleher, Kathleen McKusick, Christopher Mader, John Quackenbush, Elizabeth Stewart, Laura Stuve	Generation of 1000 STSs and high-resolution radiation hybrid and YAC STS-content maps for human chromosome 4. Generation of 30,000 STSs throughout the human genome. Construction of two radiation hybrid maps of the entire human genome with average resolutions of 500 kb (6500 STSs) and 100 kb (30,000 STSs) by 1995 and 1998, respectively. Development and application of technologies for rapid sequencing of human genomic DNA. Development of informatics resources for analyzing and managing human genome mapping data and for sequencing data generated in the center. Development of an integrated education program that includes teacher education, outreach efforts, and genome science curriculum development.
10. STANFORD UNIVERSITY DNA SEQUENCE AND TECHNOLOGY CENTER (SUDSTC) (NIH, established 1993) **RONALD W. DAVIS** David Botstein, Co-Director *CONTACT:* Jackie Couture (415/812-1968, Fax: -1975, *couture@genome.stanford.edu*); SUDSTC; 855 California Ave.; Palo Alto, CA 94304. **OTHER KEY RESEARCHERS** Pat Brown, Mike Cherry, Fred Dietrich, Richard Hyman, Rick Norgren, Peter Oefner, Victoria Smith	Development of a robust high-throughput sequencing methodology. Center efforts are divided into three parts: (1) automation, (2) sequencing, and (3) investigation of genome-wide approaches to analyzing biological function of uncharacterized DNA sequence. Current sequencing efforts focus on sequencing all of chromosome V and half of chromosome IV from *Saccharomyces cerevisiae* as part of the international effort to sequence the entire yeast genome by 1996.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Construction of integrated chromosome 16 physical-genetic-cytogenetic map, including: • low-resolution YAC contig map providing nearly complete coverage of euchromatin; map consists of 350 STSs, 600 CEPH mega-YACs and 220 flow-sorted chromosome 16–specific YACs that are localized to and ordered within a somatic cell hybrid breakpoint map (1-Mb average resolution); and • a high-resolution, sequence-ready, fingerprinted cosmid contig map covering 60% of chromosome 16 and anchored to YAC and breakpoint maps via STSs developed from cosmid contigs and by hybridizations between YACs and cosmids. Construction of human chromosome 5 framework STS map consisting of 306 markers; of these, 60 were assigned regionally on the p arm at a 1-Mb resolution [Joan Overhauser (Thomas Jefferson University)], and 100 were assigned regionally on the q arm at a 1.5-Mb resolution [John Wasmuth (UC Irvine)]. Biological developments, including (1) identification and cloning of the human telomere; (2) determination of unusual 3-D structure of telomeric DNA [with Alex Rich (MIT)]; (3) identification and cloning of highly conserved centromeric repetitive DNA regions, likely human centromere components; and (4) FISH technique enabling physical orientation of probes. Development of technology, including (1) NLGLP (with LLNL) chromosome-specific libraries: over 2600 DNA libraries sent to research and production laboratories worldwide, including complete digest libraries for each human chromosome; partial-digest phage and cosmid libraries for human chromosomes 4, 5, 6, 8, 10, 11, 13, 14, 15, 16, 17, 20, X, and Y; and complete digest low-chimeric YAC libraries for human chromosomes 5, 9, 16, and 21; (2) flow-cytometry techniques for DNA fragment sizing and detecting single DNA molecules, resulting in a CRADA with LTI for codevelopment of rapid DNA-sequencing technology; (3) robot for high-density cosmid-YAC array replication and distribution; and (4) parallel primer walking directly off cosmid contig clones.	Phage and cosmid libraries developed within NLGLP (see accomplishments). CEPH Mark II to VII YAC libraries. STS collections for chromosomes 16 (320) and 5 (306). SIGMA, a graphical map editor. cDNA Inform database and software for comparison of sequences. Modified Biomek robot and LANL robot for high-density array construction. Graduate and postdoctoral research training through LANL and DOE Human Genome Distinguished Postdoctoral Fellowship (LANL: -5919; Oak Ridge Institute for Science and Education: 615/576-9975). URL: *http://www.lanl.gov*
Generation of 1300 STSs from human chromosome 4, including 170 meiotically mapped SSRs and 50 known genes. PCR used to localize most STSs within 9 bins. Generation of an additional 3500 STSs distributed throughout the human genome and determination of their chromosomal assignment using a somatic cell hybrid panel. Sources for STS sequences include Généthon and Univ. of Iowa CHLC microsatellite repeat markers. Isolation of YACs from CEPH Mark I and mega-YAC libraries for 50 Mb of chromosome 4p (600 YACs with 389 STSs; 650-kb average YAC insert size) and 150 Mb of chromosome 4q (2700 YACs with 900 STSs; 800-kb average YAC insert size). Generation of an STS-content map of 4p (155-kb average resolution) from these STSs and YACs; 318 of 389 STSs were uniquely ordered. Construction of a radiation hybrid map of chromosome 4 that includes 689 STSs, providing a comprehensive map at a resolution of 0.5 Mb and a framework map at a resolution of 1.0 Mb. Construction of a panel of 86 radiation hybrid cell lines for whole-genome mapping and construction from these hybrids of a map containing the same STS map order as that produced with chromosome 4–only set of radiation hybrids. Establishment of a relational database, computational tools for automatic radiation hybrid map construction and YAC STS-content map generation, and an ftp server containing much of the center's data (*shgc.stanford.edu*).	Oligonucleotide sequences and complete STS sequences with PCR conditions for over 1300 chromosome 4 STSs. Oligonucleotide sequences, complete STS sequences with PCR conditions, and chromosomal assignments for 3500 STSs distributed around the rest of the human genome. YAC clones mapped by STS content for 1200 chromosome 4 markers (Research Genetics, Inc., 800/533-4363). Set of 86 radiation hybrids (containing portions of the entire human genome) that allows markers to be mapped at about 500-kb resolution (Research Genetics). Educational resources, including tours of the center; lectures on genome science and human genetics for teachers, students, and lay groups; and information on genome science curricula development.
Shotgun sequencing of about 540 kb (>95%) of yeast chromosome V and 300 kb of chromosome IV (average redundancy of 10). Efficient DNA sequencing using automation and a 96-well format; two ABI Catalyst robots, one Beckman Biomek 1000 robot with side loader, and five ABI 373 sequencing robots used. Development of software (Guile) for extending Biomek 1000 robot. Development of the *Saccharomyces* Genome Database. Construction of a prototype robotics system for automatic plaque and colony picking. Design and construction of a rapid gel-pouring assembly and construction of a random-access microtiter plate server that can integrate with new instrumentation. Completion of an automated multiplex oligonucleotide synthesizer that has produced about 7000 oligonucleotides. Development of a new DNA-shearing instrument that yields uniform, cloning-efficient fragments. Use of transposon mutagenesis in gene function studies. --- **AVAILABLE RESOURCES** Sequence data for most of yeast chromosome V and a large region of chromosome IV [Fred Dietrich (*dietrich@genome.stanford.edu*)]. Assistance in database development [Fabien Petel (*fabien@genome.stanford.edu*) and M. Cherry (*cherry@genome.stanford.edu*)]. Yeast Genome Information Server (*genome.stanford.edu*). ➠	(*continued from column at left*) *Saccharomyces* Genome Database (415/725-8956, Fax: /723-7016; *yeast-curator@genome.stanford.edu*). Information for Biomek 1000 software (*ssmith@genome.stanford.edu*), ABI Catalyst software (*jdc@genome.stanford.edu*), oligonucleotide synthesizer (*brennan@genome.stanford.edu*), DNA shearing (*oefner@genome.stanford.edu*), image analysis and robotics (*rick@genome.stanford.edu*), base-calling software (*abemo@genome.stanford.edu*), gel slider pouring apparatus (*lashkari@genome.stanford.edu*), preparation of M13 libraries from poorly growing cosmids and lambda vectors (*hyman@genome.stanford.edu*). URL: *http://genome-gopher.stanford.edu/staff-tech.html*

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
11. UNIVERSITY OF CALIFORNIA, BERKELEY (UCB), *DROSOPHILA* GENOME CENTER (NIH, established 1992) **GERALD M. RUBIN** *CONTACT:* Rubin (510/643-9945, Fax: -9947); 539 LSA Bldg.; UCB; Berkeley, CA 94720-3200. **OTHER KEY RESEARCHERS** Daniel L. Hartl (Harvard Univ.) William Kimmerly (LBL) Suzanna Lewis Christopher H. Martin (LBL) Michael J. Palazzolo (LBL) Allan C. Spradling (Carnegie Institute of Washington, HHMI)	Generation of a physical map of the *Drosophila melanogaster* genome by STS-content mapping of a P1 library to serve as templates for DNA sequencing. Integration of sites of high biological interest into the physical map, including known genes, cDNAs, and lethal P-element insertion sites. Sequencing of the 120-Mb euchromatic portion of the *Drosophila* genome.
12. UNIVERSITY OF CALIFORNIA, IRVINE (UCI) (NIH, established 1993) **JOHN J. WASMUTH** John D. McPherson, Co-Director *CONTACTS:* Wasmuth (714/856-7067), McPherson (-8242), or Judy Brown, Administrative Assistant (-7067); (Center Fax: /725-3403); UCI; Natl. Hum. Gen. Res. Center; Coll. of Med.; Irvine, CA 92717.	Construction of 150-kb-resolution radiation hybrid map of human chromosome 5 including all highly polymorphic STRP markers and every known gene on the chromosome. Establishment of a radiation hybrid contig (from single 5- to 7-Mb fragments of chromosome 5) that spans the entire chromosome in minimally overlapping segments. Isolation of about 100 cosmids from each of 30 bins defined by the final radiation hybrid contig. Isolation and sequencing of about 30 exons from each of the same 30 bins via exon amplification of pooled cosmids.
13. UNIVERSITY OF IOWA (UI) COOPERATIVE HUMAN LINKAGE CENTER (CHLC) (NIH, established 1992) **JEFFREY C. MURRAY** *CONTACTS:* Nancy Newkirk (319/335-6899) or Murray (-6946, *jeff-murray@uiowa.edu*); (Center Fax: -6970); CHLC; UI; Iowa City, IA 52242. **OTHER KEY RESEARCHERS** Kenneth Buetow (Fox Chase Cancer Center) Pui Kwok (Washington Univ.) Val C. Sheffield James L. Weber (Marshfield Medical Research Foundation) Robert F. Weir	Expansion of the linkage map of the entire human genome to an average resolution of 1 cM based on high-heterozygosity STRP markers. Evaluation of human subject protection and provision of training for social and behavioral scientists on issues of human and molecular genetics Provision of computer resources for linkage mapping and a database of linkage marker and map data. Evaluation of strategies for mapping complex genetic disorders.
14. UNIVERSITY OF MICHIGAN MEDICAL CENTER (UMMC) (NIH, established 1990) **MIRIAM H. MEISLER, Director** *CONTACT:* Meisler (313/763-5546, Fax: -9691; *miriam.meisler@med.umich.edu*); UMMC; Ann Arbor, MI 48109-0618. **OTHER KEY RESEARCHERS** Diane Baker Mike Boehnke David Burke Sally Camper Jeff Chamberlain Thomas Gelehrter David Ginsburg Thomas Glover Jerry Gorski Sun-Wei Guo David Law Dorene Markel Jerry Slightom (Upjohn Company) Spencer Thomas	Study of genetic diseases through development and use of novel technologies. Advancement and development of genomic technology, especially development of microsatellite markers, FISH mapping, radiation hybrid mapping, genetic mapping, YAC technology, and DNA sequencing. Comparative human-mouse mapping and gene identification with focus on human 17q and mouse chromosome 11. High-resolution genetic and physical mapping of both species.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Preparation of robotically arrayed 10,000-member bacteriophage P1 library for automated screening and contig assembly. Construction of a map by in situ hybridization against polytene chromosomes for over 2500 P1 clones (about 1.5 hits) that are being used as starting points for contig assembly. Mapping of over 1500 STSs to the P1 library. Mapping by in situ hybridization of over 2000 P-transposable-element insertion sites. Completion of over 1 Mb of *Drosophila* genomic sequence.	Mapped P1 clones that are directly available from 16 different sites. *Drosophila* strains with mapped single P-element insertions (Indiana *Drosophila* Stock Center). Adaptation of *C. elegans* database (ACEDB) to manage and graphically display *Drosophila* data. Primer information and complete STS sequence (average, 300 bp) with PCR conditions of over 1500 STSs of known chromosomal position.
Placement of 1200 loci on chromosome 5 radiation hybrid map. Order of 850 loci was established with odds >10,000:1. Identification of 7 single-fragment radiation hybrids with different, defined segments of 5q; total coverage of 5q is about 30%. Establishment of complete cosmid contig spanning about 3 Mb in 5q31-32. Isolation of the hyperekplexia and achondroplasia genes.	Natural deletion hybrids of chromosome 5. Selected radiation hybrids with defined 5 to 10 Mb of 5q. (Entire set will become available when the complete RH "contig" is complete.) Set of about 300 cosmids for FISH, distributed relatively evenly along the chromosome. Placement of markers (STRPs or genes) from other laboratories on the RH map with data being maintained confidentially for 3 months.
Development of libraries highly enriched for STRPs. Incorporation of CEPH genotypes on new STRP microsatellite markers into preexisting multipoint linkage maps. Construction of multipoint linkage maps using the CEPH database (version 6.0) and Univ. of Iowa CHLC marker data. Provision of a public computer interface for data analysis and acquisition. **AVAILABLE RESOURCES** Laboratories at MMRF and CHLC available for genotyping to establish disease linkages or conduct mutation detection searches for specific gene or genes with CEPH maps and STRP resources. Fluorescein-labeled human linkage screening set based on CHLC version 4a autosomal screening set (Research Genetics, 800/533-4363). Training for secondary science teachers to explore technology and social and ethical issues related to the Human Genome Project. ➠	*(continued from column at left)* Biomedical ethics visiting associates (319/335-9631, Fax: -8515, *jay-horton@uiowa.edu*). CHCL newsletter (for hard copy, contact CHLC Administration). Directories and files via anonymous ftp (*ftp.chlc.org*) and Gopher (*gopher.chlc.org*). CHLC Information: *help@chlc.org* or *info-server@chlc.org* Server for genetically mapping submitted markers using CHLC data sets: *linkage-server@chlc.org* *URL: http://www.chlc.org*
Identification of families and individuals with genetic diseases, resulting in collection of 550 blood samples, immortalization of 500 cell lines, and investigation of over 200 families with breast cancer. Generation of new microsatellite repeat polymorphisms, including 20 dinucleotide and more than 80 tetranucleotide markers. The latter are from chromosome 17, and many map near the breast cancer locus on 17q. Development of algorithms and software for radiation hybrid mapping, FISH mapping, and estimation of allele frequency based on pedigree data. Construction of radiation hybrid panels and characterization of four somatic cell hybrid lines for the long arm of chromosome 17; development of large-scale FISH mapping of YAC and cosmid clones. Direct cycled sequencing of PCR products and direct sequence analysis of CA repeats. Screening of 5 to 10 STS markers against 4 YAC libraries each month to complete the chromosome 17 physical map. Production of 5 different cDNA libraries. Development of methodology in which microdissected material is labeled by PCR and mapped back to metaphase chromosomes.	cDNA libraries from human retinal pigment epithelium, retina, kidney, fetal brain, and whole fetus. Microdissection technology and protocols. Chromosome 17 radiation hybrid panel. Four chromosome 17q somatic cell hybrids. Tetranucleotide and dinucleotide markers from chromosome 17. Chromosome 17q YACs. (The center serves as a community screening resource.) Human and mouse P1 libraries. Mouse interspecific backcross; 550 animals typed for chromosome 11 markers. Radiation hybrid mapping software (RHMAP). Educational resources for clinicians, genetic counselors, journalists, and high school teachers. URL: *http://mendel.hgp.med.umich.edu/Home.html*

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
15. UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT SAN ANTONIO (UTHSCSA) (NIH, established 1992) **SUSAN L. NAYLOR** *CONTACT:* Naylor (210/567-3842, Fax: -6781, *naylor@uthscsa.edu*); UTHSCSA; 7703 Floyd Curl Drive; San Antonio, TX 78284-7762. **OTHER KEY RESEARCHERS** Peter Cartwright (Univ. of Utah) David Housman (MIT) Robin Leach Peter O'Connell Stephanie Sherman (Emory Univ.) Brad Windle	Construction of radiation hybrid, genetic linkage, and contig maps for human chromosome 3. Development of 2500 STSs at 100-kb intervals on chromosome 3. Development of a chromosome 3–specific database (with the Univ. of Utah genome center).
16. UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER AT DALLAS (UTSMCD) (NIH, established 1990 at Salk Institute; relocated 1994) **GLEN A. EVANS** Harold R. Garner, Associate Center Director *CONTACTS:* Evans (214/648-1660, *gevans@swmed.edu*), Suzie Hayes, Administrative Services Officer (-1636), or Garner (-1661, *garner@smed.edu*); (Center Fax: -1666); UTSMCD; 6000 Harry Hines Blvd.; Dallas, TX 75235-8591. **OTHER KEY RESEARCHERS** Anne Bowcock David Burbee Kim Jackson Michael Lovett Shane Probst Lori Romberg Roger Schultz Ron Scott Sylvia Thomas	Development of a complete, contiguous, error-free physical map of human chromosome 11 at 100- to 200-kb resolution, based on STS-content mapping and nonchimeric YAC contigs. Development of a high-resolution (1- to 5-kb) physical map of human chromosome 11 and selected other chromosomes based on genomic sequence sampling and contig ordering using direct-visualization FISH. Determination of the one-pass DNA sequence of 30% of chromosome 11. Development of tools, instrumentation, and infrastructure for high-throughput automated DNA sequencing of human chromosome 11 and selected other regions of the human genome by parallel primer walking. Development of informatics support, biocomputational tools, and infrastructure for high-throughput automated DNA sequencing of the human genome. Development of additional new methods, technologies, and instrumentation for automated mapping and sequencing, including nanovolume sample processing and handling on DNA chips. Production of resources necessary for rapid identification and functional characterization of biologically and medically important genes, including cDNA isolation and genotyping.
17. UNIVERSITY OF UTAH (UU) (NIH, established 1990) **RAYMOND F. GESTELAND** and **ROBERT B. WEISS** *CONTACT:* Gesteland (801/581-5190, Fax: /585-3910, *rayg@gene1.med.utah.edu*); UU; 6160 Eccles Genetics Bldg.; Salt Lake City, UT 84112. **OTHER KEY RESEARCHERS** Jeff Botkin Peter Cartwright Mark Leppert Harold Swerdlow	Development of resources and technologies for sequencing and mapping, including (1) organized front-end strategies for large sequencing projects; (2) DNA sequencing technology, including capillary and multiplex approaches; (3) informatics tools for automation of large-scale sequencing and for distributive database searches; (4) pilot large-scale sequencing; and (5) automated genotyping technology.
18. UNIVERSITY OF WISCONSIN, MADISON (UWM) *E.COLI* GENOME CENTER (NIH, established 1991) **FREDERICK R. BLATTNER** *CONTACT:* Blattner (608/262-2534, Fax: /263-7459, *ecoli@genetics.wisc.edu*); UWM; 445 Henry Mall; Madison, WI 53706. **OTHER KEY RESEARCHERS** David Argentar Valerie D. Burland Guy Plunkett III Debra Rose Pat Wathen	Determination of the complete genome sequence of *E. coli* K12 strain MG1655 (4.7 Mb). Determination of the sequences of selected *E. coli* phages. Identification and annotation of genes and features of *E. coli* and its phages. Comparison and confirmation of all data with available published sequence to provide a coherent, consistent view. Distribution of results to research community. Development of improvements in sequencing technology and costs. Cooperation and coordination with other *E. coli* genome projects.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Production of framework hybrids that divide chromosome 3 into 23 regions. Construction of genetic linkage maps (one with 120 SSRs and 52 markers and another with 198 dinucleotide repeats) and a radiation hybrid map with 150 loci. YAC contigs covering over 80% of chromosome 3. Database contains 20,000 YACxSTS and YACx probe records. Characterization of 450 simple repeat polymorphisms for chromosome 3. Design and assembly of PCR primers for 81 genes and RFLPs. Development of a semlautomated system for PCR reactions and automated gel loading robotics for both CEPH screening and radiation hybrids.	Framework hybrids for chromosome 3. PCR primers for polymorphic dinucleotide repeats and genes mapping to chromosome 3. Scheme for automating CEPH typing. Chromosome 3–specific database. Pre- and postdoctoral training in human genome research. YAC clones, chromosome 3 cosmid library, P1 screening. URL: *http://mars.uthscsa.edu*
Preparation of a collection of over 700 standardized STSs, including random sequences mapped by somatic cell hybrid analysis or FISH. Isolation of over 900 YAC clones (>1 Mb) and over 1000 chromosome 11–specific, nonchimeric YAC clones regionally mapped to chromosome 11. Development of a YAC clone contig map of chromosome 11 with continuous coverage of 80% of the chromosome. Development and application of directed end-cloning strategies to produce additional STSs from the contig ends for gap filling. Development of a novel and potentially rapid strategy for producing high-resolution physical maps and sequencing templates using parallel sequencing and clone fingerprinting (denoted genomic sequence sampling). Development of two useful robotic systems, Prepper and GAS (Genome Automation System) for high-throughput sample processing. Began development of advanced technology for nanovolume sample processing (Garner) and advanced sequencing using DNA microdevices and chips [in collaboration with Michael Heller (Nanogen, Inc., San Diego)]. Identification of the site of a novel suppressor oncogene active in cervical carcinoma at the distal end of chromosome 11.	Cosmids: Arrayed chromosome 11–specific libraries cSRL, 16,000 clones; c11q (11q13-11qter), 1200 clones. cCLM *Giardia* cosmid libraries, 12,000 clones, 2 hosts. YACs: Total genome libraries (St. Louis, CEPH Mark I, CEPH Mark VI to VII mega-base, CEPH/Généthon "mega-YAC" subset). Chromosome 11–specific library (T. Shows, RPMI, Buffalo, NY). STSs: Over 700 STSs produced by this lab and >1000 STS for chromosome 11 (available online via Internet). Cell hybrids: 15 cell hybrid chromosome 11 mapping panels; monochromosomal hybrids for all human chromosomes. Instrumentation: Prepper, GAS, GIST (Genome Informatics System on Transputers), PCR system, Hyb system (Garner).
Development of a sensitive fluorogenic assay suitable for DNA detection on membranes; an automated chamber for unattended, sequential probing of multiplexed DNA sequencing patterns, including on-board image acquisition and base calling; magnetic bead technology for DNA preparation; transposon system for ordered sequencing of large fragments in the multiplex system; efficient multiplex method for large-scale genotyping; pilot-scale sequencing of human and mouse genomic regions; informatics tools for automating sequencing systems, transposon mapping, and database management searching; 12 mapping projects in the visitors' laboratory; and ELSI seminars and projects.	Transposon front-end reagents for sequencing. Mapping reagents and software. Multiplex sequencing protocols. URL: *http://www-genetics.med.utah.edu*
Establishment of a production sequencing facility capable of producing 1 Mb/yr, with accuracy in the range of less than one error in 10^5 nucleotides. Placement in GenBank of 1,061,766 bp of complete, contiguous, fully annotated *E. coli* sequence. (Largest contiguous segment of *E. coli* data in GenBank is now 1,244,699 bp, which includes Wisconsin and Japanese submissions.) Sequencing of 116,000 bp of phages phi80 and 933W in progress; provisional data available on ftp site. Development and application of technical innovations in robotics and sequencing strategies. Establishment of an ftp site (*ecoliftp.genetics.wisc.edu*).	Full-coverage ordered clone bank of *E. coli* MG1655 in lambda vectors; Janus M13 "flipping" vector. Sequencing production laboratory (5000 sq. ft.) equipped with ABI, LI-Cor and radioactive sequencing technology. High-speed network of UNIX workstations and Macintosh and IBM-compatible computers. Condor parallel processing system allows multiple workstations to solve a single problem. Extensive custom and commercial software for sequence acquisition and analysis, including assembly, gene detection, annotation, and Informix-based DNA sequence production management.

GENOME CENTERS and GESTECs Directors, Other Key Researchers*	MAJOR GOALS
19. WASHINGTON UNIVERSITY SCHOOL OF MEDICINE (WUSM) (NIH, established 1990) **DAVID SCHLESSINGER** Bernard H. Brownstein, Assistant Director *CONTACT:* Brownstein (314/362-3613 or -1199, Fax: -3203, *buddy@sequencer.wustl.edu*); Center for Genetics in Medicine; 4566 Scott Ave., Box 8232; St. Louis, MO 63110. **OTHER KEY RESEARCHERS** Frank Burough Ellson Chen (Applied Biosystems Division, Perkin-Elmer) Terry Featherstone Pui Kwok Ramalah Nagaraja Volker Nowotny John Rice David States	Construction of STS-based integrated physical and genetic maps of the X chromosome. Implementation of megabase-level sequencing of selected X chromosome regions.
20. WASHINGTON UNIVERSITY SCHOOL OF MEDICINE (WUSM) GENOME SEQUENCING CENTER (NIH, established 1993) **ROBERT H. WATERSTON** *CONTACT:* Paula Kassos, Administrator (314/286-1802, Fax: -1810, *pkassos@watson.wustl.edu*); WUSM, Box 8501; 4444 Forest Park Blvd.; St. Louis, MO 63108. **OTHER KEY RESEARCHERS** Warren R. Gish — David J. States LaDeana Hillier — Mark Vaudin H. Mark Johnston — Richard K. Wilson Elaine R. Mardis	Complete determination by 1998 of genomic sequence (100 Mb) of the nematode *C. elegans* (with the Sanger Centre). Contribution to the rapid completion of the yeast *S. cerevisiae* genome sequence with production of 2 to 3 Mb in 2 years. Development of technology and software for efficient large-scale genome sequencing. ***C. elegans* Sequencing Collaborators:** Sanger Centre (Medical Research Council, U.K.), John Sulston, Director. *Contact:* Jane Rogers, Administrator (+44-223/834938, Fax: /494919, *jrh@sanger.ac.uk*); Sanger Centre; Hinxton Park; Hinxton, Cambridge; UK CB10 1RQ. **Other Key Researchers (Sanger Centre):** Mary Berks — Simon Dear — Karen Thomas Alan Coulson — Richard Durbin **Other Collaborators:** Philip Green (Univ. of Wash., Seattle) Jean Thierry-Mieg [Centre National de la Recherche Scientifique (Montpellier, France)] Roger Staden [MRC Laboratory of Molecular Biology (Cambridge, U.K.)]
21. WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH and MASSACHUSETTS INSTITUTE OF TECHNOLOGY (MIT) (NIH, established 1990) **ERIC S. LANDER** David Page and Nat Goodman, Associate Directors *CONTACT:* Lander (617/258-5192, Fax: -6505; *lander@mitwibr.bitnet*); Whitehead Institute for Biomedical Research; Nine Cambridge Center; Cambridge, MA 02142. **OTHER KEY RESEARCHERS** Daniel Cohen (CEPH) Joe Nadeau (Jackson Laboratory) Shirley Tilghman (Princeton Univ., HHMI)	Construction in 3 years of a low-resolution physical map of the human genome based on 8500 STSs. Construction of genetic and physical maps of the mouse genome, including a high-resolution genetic map consisting of 6000 SSLPs, integrating 25% of these markers with the Copeland-Jenkins cross showing gene location. Production within 5 years of a low-resolution physical contig map (average size, 10 to 20 Mb) of the mouse genome based on 10,000 STSs.

*All located at centers unless otherwise noted.

MAJOR ACCOMPLISHMENTS	AVAILABLE RESOURCES
Completion (with collaborators) of YAC coverage of a number of loci, including 2 Mb in the Huntington's disease region and 4 Mb in the major histocompatibility complex. Formulation, testing, and implementation of STS-content mapping. Assembly of contigs >65% of chromosome 7 from materials that include >147 STSs [of which 100 are highly polymorphic linkage probes obtained originally from Jean Weissenbach (Institut Pasteur) and Ray White (Univ. of Utah)] and a set of >5400 chromosome 7–specific YACs that are low in cocloning. Assembly of over 85% of the X chromosome in contigs using 1400 STSs and 5000 X-specific YACs. The contigs range up to 21 Mb in length, with fully rationalized maps available for Xq26-qter. Large contigs are now being aligned in Xp, Xq13, and Xq24-q26; 200 highly polymorphic gene-specific markers are being placed along the chromosome with cognate YACs. STS and YAC information has been submitted to GDB. Wide distribution of Wash. Univ. human YAC library. Submission to ATCC for worldwide distribution of all X chromosome–specific YACs, including an X-specific YAC library. Assembly of robot-assisted workstation to screen YAC libraries; includes capacity for up to 1800 PCR reactions/d. Development of algorithms and software for analysis of STS-content and radiation hybrid mapping data and for choosing PCR primers. Design and construction of a database for STS-content mapping data. Implementation of SEGMAP for visual map representation and for testing STS-based DNA segment contigs.	Assorted genomic and X-specific YAC libraries (made at Wash. Univ. and elsewhere); all configured for PCR-screening YACs with STSs. Implemented rapid PCR screening of X-specific cosmid libraries. High-throughput PCR machine. Information on PCR conditions and end-cloning protocols. UNIX-based software to run a Zymark side-loader arm in conjunction with a Biomek 1000 robot for robot-assisted screening, and information for preparing large libraries for PCR screening. SEGMAP (YAC contig assembly software). Reliable procedures to recover YAC insert ends and sequence them automatically. Software for STS development and data storage. URL: *http://ibc.wustl.edu* (see *Center for Genetics* under *Institute Resources and Related Organizations*).
Completion of 8 Mb of nematode genomic sequence, containing about 1500 predicted genes (with the Sanger Centre). Completion of 1 Mb of yeast genomic sequence, including the complete sequence of chromosome VIII. Adaptation of methods to human genomic sequencing and application to human 16p21 cosmids. Development and implementation of the object-oriented database ACEDB to provide the sequence and physical map in the context of other information available for the nematode (with Durbin and Thierry-Mieg). Improvements in the Staden assembly-editing package (xbap and xgap); scripts to automate data transfer, handling, and analysis with improved data tracking; improved base calling and development of alternative assembly methods (with Green and Thierry-Mieg). Development of GENEFINDER, a program that accurately and rapidly predicts genes in nematode genomic DNA.	Nearly complete *C. elegans* clonal physical map with cosmid and YAC clones. Nearly complete *S. cerevisiae* clonal physical map, with cosmid and lambda clones. 1500 sequence-tagged and mapped cDNA clones. Software packages, including ACEDB, GENEFINDER, xbap/xgap, oligoselection, data tracking, and processing suites.
Construction of a mouse genetic map having 5250 SSLPs and human STS-content data with ≥3700 STSs. Construction of a mouse genome YAC library with 700-kb average inserts. Implementation of an efficient approach for screening entire YAC libraries. Implementation of a center-pioneered object-oriented database for genomic data. Introduction of YACs into mouse embryonic stem cell lines and mouse germline. Development of new technologies to increase automation ("waffle iron" thermocycler, which can handle 16 microtitre dishes at once), STS screening robot.	Mouse genome YAC library (Research Genetics, 800/533-4363; Genome Systems, 800/248-7609). Primers for mouse SSLPs (Research Genetics). Software for choosing primers and genetic mapping. "Waffle Iron" thermocycler prototype (Intelligent Automation Systems, Cambridge, MA). Data are available by (1) anonymous ftp (*genome.wi.mit.edu*; login, *anonymous*; password, *user e-mail address*); (2) internet e-mail using a database e-mail server (for copies of query forms, send message to *genome_database@genome.mit.edu* with *help* in either subject line or body text); and (3) WWW at *http://www-genome.wi.mit.edu* Newsletter available via WWW URL and e-mail (*newsletter@genome.mit.edu*).

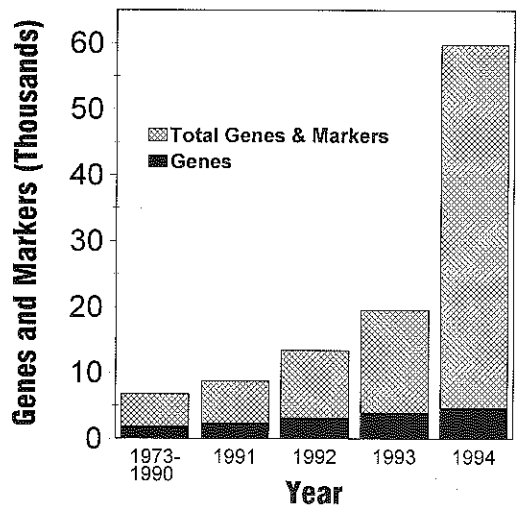
Genome News

STRPs Spark Progress in International Mapping Effort

Genetic Maps *(from p. 1)*

This achievement represents the fulfillment of a Human Genome Project goal, first established at the project's inception in 1990 and restated in last year's revised goals, to achieve a 2- to 5-cM linkage map by 1995 [*HGN* 5(4), 1–3, 5 (November 1993)]. An article describing the maps appears in the special genome issue of *Science* [**265**, 2049–54 (September 30, 1994)].

The rapid saturation of the genetic map was propelled by the discovery of a new type of DNA-based marker discovered in the late 1980s. The new markers belong to a class of short (1- to 5-bp) DNA sequences first reported by Jim Weber of the Marshfield Medical Research Foundation for the dinucleotide CA. These variants are repeated up to thousands of times throughout the genome. In a few cases, repeats of trinucleotides can hyperexpand, leading to such inherited disorders as Fragile X syndrome, Huntington's disease, or myotonic dystrophy. Because the number of repeats can vary greatly among individuals, these short tandem repeat polymorphisms (STRPs) have become one of the most useful tools for genetic mappers.



The figure depicts the rapid growth of mapping data from 1973, when 64 genes were reported, through 1994 (as of 10/28/94), when over 40,000 genes and markers were submitted to the Genome Data Base (GDB). Nearly 60,000 genes and markers have been collected in all. Types of markers include D-segments—arbitrary DNA segments of unknown function—such as RFLPs and STRPs. Before 1991, all mapping data were gathered at Human Gene Mapping Workshops; data from 1991–94 were submitted to GDB. (Graph data provided by GDB.)

Chart Available Electronically

A chart containing a subset of information accompanies the *Science* article; more detailed presentations of the map and data sets are available electronically via e-mail, ftp, and WWW from the Genome Data Base (*help@gdb.org*, *ftp.gdb.org*, and *http://gdbwww.gdb.org/*) and from CHLC (*info-server@chlc.org*, *ftp.chlc.org*, and *http://www.chlc.org/*).

The abundance and even distribution of STRPs throughout the genome and their ability to be assayed by PCR have enabled researchers to generate efficiently and rapidly the markers needed to saturate the genetic map. The new maps incorporate over 3600 STRPs, in addition to over 400 genes and nearly 1800 other markers such as RFLPs and anonymous DNA segments. These maps describe human genetic diversity at a mean resolution of 0.7 cM.

Of the three largest groups collaborating on the new linkage map, the Généthon group contributed markers containing dinucleotide repeats, the most common type of STRP [see *HGN* 6(3), 5 (September 1994) for announcement of map based on these markers]. CHLC and the Utah group generated primarily tri- and tetranucleotide repeats, which are easier to genotype but less frequent. The Yale University group provided cytogenetic anchor points for a subset of the linkage markers (using FISH of YACs).

Using Genetic Maps

Genetic maps are used for starting gene hunts, and their usefulness increases with marker density and quality. A variety of strategies can be applied using polymorphic markers, including linkage, association, allele-sharing, and loss-of-heterozygosity studies. In *Science*, Lender and Schork discuss which strategy to apply in a particular situation [**265**, 2037–48 (Sept. 30, 1994)]. Once the gene is localized to a particular area, researchers turn to physical maps (ordered clone sets) to retrieve flanking DNA segments for further detailed study. In 1986, the gene for the immune disorder chronic granulomatous disease was the first to be isolated by using linkage maps with a procedure now known as positional cloning. Genetic maps have been used since to help localize about 40 genes, including those for cystic fibrosis, Fragile X syndrome, myotonic dystrophy, and types of colon and breast cancer.

Other uses for the new high-density maps are to study candidate genes by substituting an informative STRP as a surrogate for a less-informative gene and to explore complex, multifactorial, or polygenic disorders. Because their construction is based on STSs, high-resolution genetic linkage maps featuring PCR-based markers also serve as a framework for constructing physical maps and integrating genetic, cytogenetic, and physical maps.

(see Maps, p. 15)

What Are Genetic Maps?

Genes and other DNA fragments located on the same chromosome are said to be linked. Genetic linkage maps (also called genetic maps or linkage maps) show the relative order of and approximate spacing between specific, identifiable DNA regions that researchers use as signposts along the chromosomes. These landmarks or markers include detectable genes or other DNA stretches. Distances between markers, measured in centimorgans (cM), are not actual physical distances but rather describe the frequency with which the markers are coinherited. Two markers are said to be 1 cM apart if they are separated (not coinherited) by recombination 1% of the time. A genetic distance of 1 cM is roughly equivalent to a physical distance of 1 million base pairs.

Human genetic mapping began in the 1930s when researchers studying family inheritance patterns noted that color blindness and hemophilia were linked to the X chromosome. Mapping proceeded slowly and was limited to genes linked to visible physical traits until the late 1970s, when a new type of marker, based on detectable DNA sequence differences, was discovered. Individual variations in base sequence cause restriction enzymes to cut DNA at specific sites, producing fragments of different lengths (restriction fragment length polymorphisms or RFLPs). Although the discovery of RFLP markers spurred genetic mapping, researchers found them to be relatively rare and unevenly dispersed in the human genome, difficult to analyze (genotype), and not variable or "informative" enough. The newer STRP markers have several technological advantages over RFLPs, including their ability to be amplified by PCR.◊

NIH Grants Explore Issues Related to Genetic Testing for Cancer Risk

The recent isolation of genes that increase a person's likelihood of developing breast, ovarian, or colon cancers brings with it the technological potential for testing large numbers of people to see if they carry the predisposing genes. DNA testing for susceptibility to some cancers may offer the opportunity for early preventive interventions before invasive cancer develops.

At the moment, little is known about the prevalence of predisposing mutations in large populations; correlation between mutations and the development of cancer; ability of tests to predict risk accurately; and social, psychological, and economic costs of being tested. Similarly, investigators do not know how the general population views the use of genetics in health care or whether the medical profession could provide adequate counseling before and after testing.

To help answer these questions, the NIH National Center for Human Genome Research (NCHGR), National Cancer Institute, National Institute of Nursing Research, and National Institute of Mental Health have jointly awarded more than $2.5 million in research grants (see box at right). The 3-year grants will support 11 research projects in a consortium coordinated by the NCHGR Ethical, Legal, and Social Implications Branch. The consortium format will allow researchers to compare findings on common issues, reduce duplicated research efforts, and promote information sharing on informed consent issues and DNA test quality assurance.

NCHGR Director Francis Collins noted that recent cloning of the BRCA1 gene brings a particular urgency to address questions that surround DNA testing for cancer susceptibility. With these awards, he continued, "A highly qualified group of investigators will be funded to tackle these questions. We're glad to be able to do it in such a timely way."◊

Consortium Principal Investigators and Research Plans

Wylie Burke *(Fred Hutchinson Cancer Research Center and University of Washington, Seattle)*
- Provide genetic counseling and DNA testing for the BRCA1 mutation in women from families with a risk for breast cancer. Examine alternative forms of counseling by genetic specialists and primary care providers.
- Gather information about reactions to genetic testing for breast cancer from women receiving "routine" health care and from providers of genetic services and primary care.

Mary B. Daly *(Fox Chase Cancer Center, Philadelphia)*
- Employ oncology nurses as primary patient counselors involved in DNA testing for BRCA1 in an ethnically diverse population.

Bonnie Flick *(University of Utah, Salt Lake City)*
- Study how adolescent girls are affected by BRCA1 testing of their parents.

Judy Garber *(Dana-Farber Cancer Institute, Boston)*
- Prepare teams of genetic counselors and nurses to educate and counsel families who have the BRCA1 gene; examine the impact of BRCA1 testing.

Karen Glanz *(University of Hawaii, Honolulu)*
- Study an array of factors that inhibit or motivate multiethnic Hawaiian residents to seek DNA testing for colon cancer susceptibility.

Gail Geller *[Johns Hopkins University (JHU), Baltimore, Md.]*
- Explore informed consent and develop a model protocol for use in BRCA1 testing.

Ellen Gritz (M.D. Anderson Cancer Center, University of Texas, Houston)
- Characterize the psychosocial and behavioral impact of DNA testing for hereditary nonpolyposis colon cancer.

Caryn Lerman *(Georgetown University, Washington, D.C.)*
- Study methods for educating and counseling women who seek DNA testing for cancer predisposition.

Gloria Petersen *(JHU)*
- Study social and psychological factors that underlie decisions about genetic testing among people with a family history of colon cancer.

Kathryn Taylor *(Princess Margaret Hospital, Toronto, Canada)*
- Focus on health-care providers who deliver information to patients seeking DNA testing; develop guidelines for disseminating health-risk information.

Arabidopsis Report Issued

An ad hoc committee representing the community of *Arabidopsis* researchers met June 8–9 in Arlington, Virginia, to explore the feasibility of a large-scale, federally funded U.S. *Arabidopsis thaliana* genome project. The committee agreed that a project should begin as soon as possible and recommended continuing collaborations with participants in the Multinational Coordinated *Arabidopsis* Genome Research project, especially the European Community *Arabidopsis* genome sequencing program.

The group urged that funds be provided for the following:
- Completion of the *Arabidopsis* physical-genetic map and creation of sequence-ready clone collections by 1997.
- Pilot sequencing and technology-development projects with the goal of sequencing 10 Mb of *Arabidopsis* by 1999.
- Pilot project scale-up and complete sequencing of the 100-Mb genome by 2004.

The committee's report has been posted on the *Arabidopsis* newsgroup (*arabidopsis@net.bio.net*), with hard copies available from Machi Dilworth; National Science Foundation; 4201 Wilson Blvd., Rm. 685; Arlington, VA 22230 (703/306-1422, Fax: -0349, Internet: *mdilwort@nsf.gov*).◊

Maps *(from p. 14)*

New Goals

Although the short-term genetic mapping goal has been achieved, the maps still have gaps and lack anchor points at chromosomal telomeres and centromeres. Investigators believe that maximally useful genetic maps will require increased marker density, next at the level of 1 marker per 100 kb. A complete description of human genetic varation might require a marker every 100 to 1000 bp. *[Denise Casey, HGMIS]*◊

This newsletter is prepared at the request of the DOE Office of Health and Environmental Research and the NIH National Center for Human Genome Research by the Biomedical and Environmental Information Analysis Section of the Health Sciences Research Division at Oak Ridge National Laboratory, which is managed by Martin Marietta Energy Systems, Inc., for the U.S. Department of Energy, under Contract DE-AC05-84OR21400.◊

Human Genome news

This newsletter is intended to facilitate communication among genome researchers and to inform persons interested in genome research. Suggestions are invited.

Human Genome Management Information System

Managing Editor
Betty K. Mansfield

Editors/Writers
Anne E. Adamson
Denise K. Casey
Kathleen H. Mavoumin

Production Manager/Editor
Judy M. Wyrick

Production Assistants
K. Alicia Davidson
Larry W. Davis
Sheryl A. Martin
Laura N. Yust

HGMIS Address
Betty K. Mansfield
Oak Ridge National Laboratory
1060 Commerce Park
MS 6480
Oak Ridge,TN 37830
615/576-6669
Fax: /574-9888
Internet: *bkq@ornl.gov*

Sponsor Contacts

Daniel W. Drell
DOE Program Office
Health Effects and Life Sciences Division
Germantown, MD
301/903-6488, Fax: -8521
Internet: *daniel.drell@mailgw.er.doe.gov*

Leslie Fink
NIH National Center for Human Genome Research
Bethesda, MD
301/402-0911, Fax: -4570
Internet: *lsf@cu.nih.gov*

National Center for Human Genome Research

Informatics Workshop Held at CSHL Meeting

As part of the annual Cold Spring Harbor Laboratory (CSHL) Genome Mapping and Sequencing meeting in May, David States (Washington University, St. Louis) convened a workshop on modular software development and data interchange in genome research. Molecular biologists routinely exchange protocols and reagents, but difficulties in technology exchange have led to duplication of effort in genome informatics. The informatics workshop focused on ways to improve the modularity of software supporting genome mapping and sequencing and to enhance reliable data communication among centers and other researchers.

Workshop speakers were Nat Goodman (Whitehead Institute for Biomedical Research); Jean Thierry-Mieg (CNRS, Montpellier, France); James Ostell (National Center for Biotechnology Information); Tom Slezak (Lawrence Livermore National Laboratory); Ken Fasman [Genome Data Base (GDB)]; and States, who served as moderator.

In presenting a modular view of software requirements for large-scale sequencing laboratories, States defined independent roles for sequence base calling, shotgun map assembly, consensus sequence generation, multiple sequence alignment, editing, and data storage. He urged that developers concentrate on specific tasks rather than build an entire system to test a novel idea that impacts only one activity.

Goodman outlined database technology applicable to genome sequencing and pointed out that a critical issue in modular software development is agreement on information content at data interfaces. Once content agreement has been achieved, technical issues of data interchange can be addressed. In particular, Ostel and Thierry-Mieg concurred that translators between ASN.1 and .ace data formats could be written easily and that work in this area should be encouraged.

Slezak pointed out that centers' different biological approaches have led to a variety of map representations. Fasman added that implementation of various anonymous ftp servers, Gopher sites, and World Wide Web pages has resulted in very heterogeneous collections of information resources. In addition, sites may update their own servers on variable and unreliable schedules and have different data formats, definitions, and semantics, all of which add to difficulties in using and maintaining up-to-date data collections. Fasman reminded the group that centralized data repositories such as GDB were developed to address these issues and need community support to implement solutions.

States and Goodman are planning a similar workshop for the next CSHL mapping and sequencing meeting. Ideas and suggestions should be sent to States (314/362-2135, Fax: -0234, Internet: *states@ibc.wustl.edu)* or Goodman (617/252-1904, Fax: -1902, Internet: *nat@genome.wi.mit.edu*).◊

INFORMATICS DISCUSSION GROUP: A mailing list and discussion group has been set up on WWW (URL *http://www-genome.wi.mit.edu/informatics/sharing_archive/index.html*). To join the discussion group, *lstein@genome.wi.mit.edu*; to send mail to the group, *sharing@genome.wi.mit.edu*).]

¶ Publications

Assessing Genetic Risks: Implications for Health and Social Policy reports on a state-of-the-art study, conducted by a multidisciplinary panel of experts, on the current status and future implications of genetic testing [*HGN* 5(4), 9 (November 1993)]. This study, commissioned in 1990 by the NIH-DOE Ethical, Legal, and Social Implications component of the U.S. Human Genome Project and undertaken by the Institute of Medicine, explored scientific aspects of genetic risk assessment and the many societal problems surrounding such testing. The book was edited by Lori Andrews (American Bar Foundation); Jane Fullarton (Staff Study Director, now at Tascon); Neil Holtzman (Johns Hopkins University Hospital); and Arno Motulsky (University of Washington, Seattle), committee chair. Hard cover, 338 pp. 1994. [National Academy Press; 2101 Constitution Ave., NW; Washington, DC 20418 (202/334-3313 or 800/624-6242).] ◊

Intractable Neurological Disorders, Human Genome Research and Society: Proceedings of the Third International Bioethics Seminar in Fukui, 19–21 November, 1993 was edited by Norio Fujiki (Fukui Medical School, Japan) and Darryl R.J. Macer (Eubios Ethics Institute). The book includes about 80 papers on bioethics; eugenics; ethical, legal, and social aspects of the Human Genome Project; medical genetics; patenting of life forms; public perceptions of disease; genetic technology; and genetic counseling in many countries around the world. English, 320 pp.; Japanese, 340 pp.; 1994. [Eubios Ethics Institute; P.O. Box 125; Tsukuba Science City; Ibaraki 305, Japan (+81-298/53-4662, Fax: -6614, Internet: *macer@sakura.tsukuba.ac.jp*) or 31 Colwyn Street; Christchurch 5, New Zealand.] ◊

☛ Mapping Panel Resource

The second version of Mapping Panel #2 is available as cell cultures or DNA from the Human Genetic Mutant Cell Repository of the National Institute of General Medical Sciences (NIGMS). Version 2 consists of 24 human-rodent somatic cell hybrids, each retaining a single intact human chromosome. In the new monochromosomal hybrid for chromosome 1, GM/NA13139, 96% of the cells retain the human chromosome; in the new hybrid, GM/NA13140, 68% of the cells retain chromosome 20.

The panel has been characterized by the following tests, at minimum: (1) G-banded chromosome analysis, (2) in situ hybridization using biotinylated total human DNA, (3) Southern blot hybridization, and (4) analysis by the polymerase chain reaction.

For information or a repository catalog, contact the NIGMS Human Genetic Mutant Cell Repository; Coriell Cell Repositories; Coriell Institute for Medical Research; 401 Haddon Avenue; Camden, NJ 08103 (800/752-3805 or 609/757-4848, Fax: -9737).◊

GDB WWW Browser Links to GSDB Sequence Data

GDB and the Genome Sequence Data Base (GSDB), hosted at the National Center for Genome Resources in Santa Fe, New Mexico, have developed improved WWW connections between human gene mapping and sequence data. Detailed information about loci and probes retrieved from GDB's WWW server now includes links to GSDB DNA sequence entries (URL: *http://gdbwww.gdb.org*).

Selecting the DNA sequence accession number or numbers associated with a GDB locus or probe retrieves a tabular summary of all related GSDB sequences. The summary shows the entry name, length, accession number, and a brief sequence description. Summary lines are linked to detailed GSDB flatfile sequence entries, which in turn are linked to GDB through gene symbol and map location. For example, following the GSDB map location link retrieves all GDB loci mapped to the same cytogenetic region as the sequenced locus. Not all GDB loci and probes have related GSDB sequences, and not all GSDB entries have map locations.◊

Computer Survey Results

A GDB survey of computer usage was included in the November 1993 and January 1994 issues of *HGN*. Although GDB had planned to publish the results, the number of responses was not sufficient to draw any meaningful conclusions. As always, GDB welcomes input from the user community.◊

Home Pages Describe Genome Programs

Three new WWW home pages offer information about genome programs:

- DOE Office of Health and Environmental Research (OHER) Home Page for the DOE Human Genome Program (*http://www.er.doe.gov/production/oher/hug_top.html*) includes pointers to other programs within OHER (*http://www.er.doe.gov/production/oher/oher_top.html*) and the Office of Energy Research (*http://www.er.doe.gov*). Links are made to additional biological and environmental information and to the Human Genome Management Information System (HGMIS), GDB, and other sites.
- HGMIS Home Page at Oak Ridge National Laboratory (*http://www.ornl.gov/TechResources/Human_Genome/home.html*). Includes text on the DOE Human Genome Program; funding and fellowships; and selected genome resources. The *1993 DOE Human Genome Program Report*, two latest issues of *HGN*, calendars of genome meetings and training events, and a list of chromosome editors are also accessible.
- GDB Home Page extension, Human Genome Project Resources and Meetings (*http//gdbwww.gdb.org/gdbdoc/genomic_links.html*), includes links to human genome centers, human chromosome-specific servers, model organism genome projects, and GDB and OMIM related pages. The page can also be accessed through the URL of the GDB Home Page (*http://gdbwww.gdb.org*).

These WWW sites point to each other and are coordinating their efforts with other WWW sites to provide Internet access to genome information.◊

GDB USER SUPPORT, REGISTRATION

To become a registered user of GDB and OMIM, contact one of the User Support offices listed below (a user may register to access both Baltimore and a remote node). Questions, problems, or user-registration requests may be sent by telephone, fax, or e-mail. User-registration requests should include name, institutional affiliation, and title (if applicable), street address (no P.O. box numbers), telephone and fax numbers, and e-mail address.

The Help Line in Baltimore is staffed from 9 a.m. to 5 p.m. EST for information on accounts and training courses, technical support, and data questions. Calls received after hours will be forwarded to the appropriate voice mail and returned as soon as possible.

GDB, OMIM Training Schedule

"GDB/OMIM and Genomic Data on the Internet" classes will be held in Baltimore on March 20–21, 1995, and June 5–6, 1995. These courses offer thorough coverage of the structure, content, and roles of GDB and OMIM; discuss the strengths and weaknesses of various interfaces for searching the data; and explore related genomic resources available worldwide on the Internet. In addition to using GDB and OMIM application software, participants will learn how to retrieve phenotype, mapping, and sequence data with tools such as ftp, e-mail, Gopher, and the WWW hypertext browser NCSA Mosaic. Contact the U.S. GDB User Support Office.

User Support Offices

UNITED STATES
GDB User Support
Genome Data Base
Johns Hopkins University
2024 E. Monument Street
Baltimore, MD 21205-2100
410/955-9705
Fax: /614-0434
Internet: *help@gdb.org*

AUSTRALIA
Alex Reisner
ANGIS
Electrical Eng. Bldg. J03
University of Sydney
Sydney, N.S.W. 2006
Australia
+ 61/2-692-2948
Fax: -3847
Internet: *reisner@angis.su.oz.au*

FRANCE
Philippe Dessen
Service de Bioinformatique
CNRS-INSERM
7 rue Guy Moquet - BP8
94801 Villejuif Cedex
France
+33/14559-5241
Fax: -5250
Internet: *gdb@genome.vjf.inserm.fr*

GERMANY
Otto Ritter
Molecular Biophysics Dept.
German Cancer Research Center
Im Neuenheimer Feld 280
D-6900 Heidelberg
Germany
+ 49/6221-42-2372
Fax: -2333
Internet: *dok261@cvx12.dkfz-heidelberg.de*

ISRAEL
Jaime Prilusky
Bioinformatics Unit
Weizmann Institute of Science
76100 Rehovot, Israel
+972/8-343456
Fax: -344113
Internet: *lsprilus@weizmann.weizmann.ac.il*

JAPAN
Mika Hirakawa
JICST GDB Center
Numajiri Sangyo Building
783-12, Enokido
Tsukuba City, Ibaraki 305,
Japan
+81/298-38-2965
Fax: -2956
Internet: *mika@gdb.gdbnet.ad.jp*

NETHERLANDS
GDB User Support
CAOS/CAMM Center
Faculty of Science
University of Nijmegen
P.O. Box 9010
6500 GL NIJMEGEN
Netherlands
+ 31/80-653391
Fax: -652977
Internet: *post@caos.caos.kun.nl*

SWEDEN
GDB User Support
Biomedical Center
Box 570
S-751 23 Uppsala
Sweden
+ 46/18-174057
Fax: -524869
Internet: *help@gdb.embnet.se*

UNITED KINGDOM
Administration
HGMP Resource Centre
Hinxton, Cambridge
CB10 1RQ
United Kingdom
+ 44/1223-494511
Fax: -494512
Internet: *admin@hgmp.mrc.ac.uk*

¶ Electronic Data for Chromosome 11

Data from a paper on the radiation hybrid (RH) map of human chromosome 11 that appeared in *Nature Genetics* [8, 70–76 (September 1994)] is accessible via anonymous ftp (*ftp.well.ox.ac.uk* in the directory *pub/genetic/rh11map*). The data, contributed by a group led by Michael James (The Wellcome Trust Centre for Human Genetics, U.K.) includes a file of over 600 STSs with their primer sequences and the RH map. These files will be updated regularly as new STSs are placed on the RH map. [Contact: Michael James (Fax: +44-865/742-196 or -187, Internet: *michael.james@well.ox.ac.uk*).]◊

Calendar of Genome-Related Events* (acronym list, p. 16)

January 1995

3–6. Biotechnol. Comput. Track HICSS-28; Maui, HI (paper deadline: June 15) [L. Hunter, 301/496-9300, Fax: -0673, *hunter@work.nlm.nih.gov* or T. Takagi, +81-3/5449-5614, Fax: -5434, *workshop@ims.u-tokyo.ac.jp*]

5–11. Oncogenes: 20 Yrs. Later; Keystone, CO (reg. deadline: Sept. 7) [Keystone Symp., 303/262-1230, Fax: -1525]

6–12. Bacterial Chromosomes; Santa Fe, NM (reg. deadline: Sept. 7) [see contact: Jan. 5–11]

9–10. Intl. Cong. on Cellular Therapy & Tissue Engin.; Washington, DC [BioConf. Intl., W. Small, 301/652-3072, Fax: -4951]

9–12. BioEast '95; Washington, DC [see contact: Jan. 9–10]

9–15. Mol. Toxicology; Copper Mountain, CO (reg. deadline: Sept. 7) [see contact: Jan. 5–11]

10. 2nd Natl. Biotechnol. Summit; Washington, DC [see contact: Jan. 9–10]

11. Intl. Conf. on Hum. Gene Therapy; Washington, DC [see contact: Jan. 9–10]

12. Russell F. Doolittle: DNA, Genet., and Biotechnol.; Gaithersburg, MD [TIGR/NIST Distinguished Speaker Ser., D. Hawkins, 301/869-9056, Fax: -9423]

15–19. Plant Genome III; San Diego (abs. deadline: Nov. 1) (PG-I and PG-II abs. available on WWW, *http://probe.nalusda.gov:8000/plant/index.html)* [Scherago Intl., 212/643-1750, Fax: -1758, *scherago@biotechnet.com*]

19. Michael Lovett: Fishing for Complements—Finding Genes by Direct Select.; Bethesda, MD [NCHGR Lect. Ser., E. Feingold, 301/496-7531, Fax: /480-2770, *fey@cu.nih.gov*]

20. DIMACS Sp. Yr.: Distinguished Lect. Ser.—Joachim Messing [For specific location, contact M. Farach, 908/445-4580, Fax: -5932, *special@dimac.rutgers.edu*]

20–26. Genet. Networks; Santa Fe, NM (reg. deadline: Sept. 7) [see contact: Jan. 5–11]

23–25. Exploiting Mol. Diversity: Small Mol. Libraries for Drug Discovery; San Diego [CHI, B. Keddy, 617/487-7989, Fax: -7937]

26–27. Exploiting Biol. Diversity: Libraries & Mutations for Res. & Drug Dev.; CHI, La Jolla, CA [see contact: Jan. 23–25, above]

February 1995

3. DIMACS Sp. Yr. Distinguished Lect. Ser.—Richard Karp [see contact: Jan. 20]

4–9. Miami 1995 Bio/Technol. Winter Symp. on the Adv. in Gene Technol.:Protein Engin. and Structural Biol.; Fort Lauderdale, FL (poster deadline: Nov. 1) [MBWS, 800/642-4363, Fax: 305/324-5665, *mbws@mednet.med.miami.edu*]

6–8. **DIMACS Sp. Yr.: Phylogeny Workshop; Princeton, NJ [see contact: Jan. 20]

6–9. 4th Pacific Rim Biotechnol. Conf.; Melbourne [ABA, I. Prince, +61-3/905-3449, Fax: -5686, *ian.prince@eng.monash.edu.au*]

9–10. Commercializing Oligonucleotide-Based Therapeutics: Latest Clin. Trials & Appl.; Coronado, CA [IBC, 508/481-6400, Fax: -7911]

10 or 13. DIMACS Sp. Yr.: Distinguished Lect. Ser.—Eugene Lawler [For specific location, see contact: Jan. 20]

12–16. 2nd Intl. Conf. on Antisense Nucleic Acids: Biol., Pharmacology, Therapy; Garmisch-Partenkirchen, FRG [K.-H. Schlingensiepen, +49-551/201-616, Fax: -499]

16. Marty Rosenberg: DNA, Genet., and Biotechnol.; Gaithersburg, MD [see contact: Jan. 12]

16. Maynard Olson: Proposed Technical Scenario for Sequencing the Hum. Genome; Bethesda, MD [see contact: Jan. 19]

16–21. 1995 AAAS Annu. Meet. and Sci. Innovation Exposition: Unity in Diversity; Atlanta (reg. deadline: Jan. 31) [AAAS, 703/671-1400, Fax: -7695]

18–21. Mol. Approaches to Lab. Diagn.; San Francisco (abs. deadline: Nov. 25) [Office of Continuing Med. Edu., 415/476-5808, Fax: -0318]

20–23. BioExpo 95; Nice, FR [C. Metais, +33-1/47-56-21-05, Fax: -20]

March 1995

6. **DIMACS Sp. Yr.: Gene Finding and Gene Prediction Mini-Workshop; Philadelphia [see contact: Jan. 20]

6–7. Intl. Symp. on Fluorescent Proteins and Appli.; Palo Alto, CA [J. Larrick or D. Youvan, 415/694-1420, Fax: -7717, *paimmfp@aol.com*]

6–8. 2nd Annu. HGP: Commercial Implications; CHI, San Francisco [see contact: Jan. 23–25]

6–9. 2nd Joint Meet. of March of Dimes and Am. Coll. of Med. Genet.; Los Angeles [M. Ryan, 301/571-1825, Fax: -1895]

9–10. Adv. in Genet. Screening and Diagn. of Hum. Dis.; CHI, San Francisco [see contact: Jan. 23–25]

9–15. Discovery of Therapeutic Agents; (reg. deadline: Sept. 28) Lake Tahoe, CA [see contact: Jan. 5–11]

10. DIMACS Sp. Yr.: Distinguished Lect. Ser.—Eric Lander; [see contact: Jan. 20]

16. Lori Andrews: Emerging Issues in Genet. Testing; Bethesda, MD [see contact: Jan. 19]

16–17. Human Genet. Towards 2000; Melbourne [R. Cotton, +61-3/345-5045, Fax: /348-1391]

17–23. Toward the Genet. Manipulation of Insects; Tamarron, CO [see contact: Jan. 5–11]

18–22. 86th Annu. Meet. of AACR; Toronto [AACR, 215/440-9300, Fax: -9313]

19–22. Electrophoresis '95; Rockville, MD [D. Forio, 913/843-1221, Fax: -1274]

20–21. **DIMACS Sp. Yr.: Global Minimization of Nonconvex Energy Functions Mini-Workshop; New Brunswick, NJ [see contact: Jan. 20]

22. **DIMACS Sp. Yr.: Antibody Struct. and Sequence Mini-Workshop; New Brunswick, NJ [see contact: Jan. 20]

23. M.R.C. Greenwood: DNA, Genet., and Biotechnol.; Gaithersburg, MD [see contact: Jan. 12]

23–29. Repair and Processing of DNA Damage; Taos, NM [see contact: Jan. 5–11]

24. **DIMACS Sp. Yr.: Sequence-Based Methods for Protein Folding Mini-Workshop; New Brunswick, NJ [see contact: Jan. 20]

26–30. Baculovirus and Insect Cell Gene Expression Conf.; Pinehurst, NC [Glaxo, Inc., C. Hoffman, 919/990-6441, Fax: -6470, *crh11316@glaxo.com*]

26–Apr. 1. Gene Therapy and Mol. Med.; Steamboat Springs, CO [see contact: Jan. 5–11]

April 1995

2–8. Sci. and Engin. of Immunoprotected Cell Transplants; Frisco, CO (reg. deadline: Nov. 2) [see contact: Jan. 5–11]

3–9. Frontiers of NMR in Mol. Biol.-IV; Keystone, CO (reg. deadline: Nov. 2) [see contact: Jan. 5–11]

4–10. Nuclear Matrix: Involvement in Replication, Transcription, Gene Splicing, and Cellular Regul.; Hilton Head, SC (reg. deadline: Nov. 2) [see contact: Jan. 5–11]

4–10. Epigenet. Regul. of Transcription; Hilton Head, SC (reg. deadline: Nov. 2) [see contact: Jan. 5–11]

8–9. 2nd Intl. Workshop on Chromosome 20; Woods Hole, MA [C. Smith, 617/353-8500, Fax: -8501, *cls@buenga.bu.edu*]

10–11. Dev. of Small Mol. Mimetic Drugs; San Francisco [CHI, 617/487-7989, Fax: -7937]

17. DIMACS Sp. Yr.: Distinguished Lect. Ser.—Charles Cantor (also offered Apr. 18) [For specific location, see contact: Jan. 20] ◊

An extended list of genome events is available from HGMIS. See p. 16 for contact information.

*Dates and meeting status may change; courses may also be offered at other times and places; check with contact person.
**Attendance is either limited or restricted.

For Your Information

Training Calendar*

January 1995

3. Understanding PCR; Columbia, MD [Exon-Intron, Inc., 800/407-6546, Fax: -3983]

3–7. Recombinant DNA Methodol.; Washington, DC [CATCMB/CUA, 202/319-6161, Fax: -4467, *millerm@cua.edu*]

5. RNA Isolation Strategies; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

9. Intro. to PCR; Houston [BTP, S. Chance, 800/821-4861, Fax: 603/267-1993]

9–13. **Adv. Linkage Course; New York (appl. deadline: Nov. 10) [K. Montague, 212/960-2507, Fax: /568-2750, *jurg.ott@columbia.edu*]

9–13. Basic Cell and Tissue Culture; CATCMB/CUA, Washington, DC [see contact: Jan. 3–7]

9–13. Recombinant DNA Methodol.; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

10–11. Quantitative RNA-PCR; BTP, Houston [see contact: Jan. 9]

12–13. Basic Cloning & Hybridization Techniques; BTP, Houston [see contact: Jan. 9]

16–20. PCR Tech.; Germantown, MD [LTI, L. Kerwin, 800/952-9166, Fax: 301/258-8212]

17. rDNA: An Overview; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

19. Principles of Chemiluminescence; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

23. Differential Display PCR/Long PCR; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

23–27. Anal. of Gene Expression; LTI, Germantown, MD [see contact: Jan. 16–20]

28–30. PCR Tech.; CATCMB/CUA, Washington, DC [see contact: Jan. 3–7]

February 1995

6–10. RNA Isolation & Gene Expression; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

6–11. cDNA Library Tech.; LTI, Germantown, MD [see contact: Jan. 16–20]

13–17. Recombinant DNA Tech. I; LTI, Germantown, MD [see contact: Jan. 16–20]

20–24. PCR Methodol.; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

26–Mar. 11. Intro. to Nucleic Acids Tech.; Chapel Hill, NC [Carolina Workshop, W. Litaker, 919/962-8920, Fax: /966-6821]

27–Mar. 3. Recombinant DNA Methodol.; Exon-Intron, Inc., Columbia, MD [see contact: Jan. 3]

March 1995

6–10. In Situ Hybridization Tech.; LTI, Germantown, MD [see contact: Jan. 16–20]

13–17. Receptor Binding Tech.; CATCMB/CUA, Washington, DC [see contact: Jan. 3–7]

13–17. Recombinant Baculovirus Tech.; LTI, Germantown, MD [see contact: Jan. 16–20]

April 1995

3–7. Cell Culture Tech.; LTI, Germantown, MD [see contact: Jan. 16–20] ◊

U.S. Genome Research Funding Guidelines

Note: Investigators wishing to apply for funding are urged to discuss their projects with appropriate agency staff before submitting proposals.

NIH National Center for Human Genome Research (NCHGR)

Application receipt dates:

- R01, P01, R21, R29, P30, P50, K01,* and R13 grants – February 1, June 1, and October 1.
- Individual postdoctoral fellowships – April 5, August 5, and December 5.
- Institutional training grants – January 10, May 10, and September 10.
- Small Business Innovation Research Grants (SBIR: firms with 500 or fewer employees) – April 15, August 15, and December 15.
- Research supplements for underrepresented minorities – applications are accepted on a continuing basis.
- Requests for Applications (RFAs) – receipt dates are independent of the above dates. Notices will appear in *HGN* and other publications.

*Expedited review possible. Check with NCHGR during application development phases.

Program announcements are listed in the weekly *NIH Guide for Grants and Contracts*,* which is available electronically through one of the following methods.

- Gopher (*gopher.nih.gov*).
- Institutional Hubs. A designee receives automatic updates and distributes them locally to researchers. Send a message naming the responsible person to BITNET: *q2c@nihcu* or Internet: *q2c@cu.nih.gov*.
- NIH Grant Line (also known as DRGLINE): Electronic bulletin board updated weekly. Connection is through a modem (301/402-2221), and files can be transmitted rapidly via BITNET or Internet. The Grant Line is also accessible by Telnet to *wylbur.cu.nih.gov*. When connection is open, type *VT100*. At the INITIALS prompt, type *BB5* and at the ACCOUNT prompt, type *CCS2*. For more information, contact John James (301/594-7270, Fax: -7384).

Full text of RFAs listed in the NIH grants guide may also be obtained from NIH NCHGR in Bethesda, Maryland (301/496-0844).

DOE Human Genome Program

For funding information or general inquiries, contact the program office via

- 301/903-6488 or Internet (*genome@er.doe.gov*). Relevant documents are available by ftp (*oerhp01.er.doe.gov* in directory */genome*).

DOE Human Genome Distinguished Postdoctoral Fellowships

Next deadline: February 1, 1995. For further information, contact

- Linda Holmes, Oak Ridge Institute for Science and Education: 615/576-9934, Fax: /241-5219.

SBIR Grants

DOE and NIH invite small business firms to submit grant applications addressing the human genome topic of SBIR programs, which are designed to strengthen innovative firms in research and development and contribute to the growth and strength of the nation's economy. For more information on human genome SBIR grants, contact

- Kay Etzler; c/o SBIR Program Manager, ER-16; DOE; Washington, DC 20585 (301/903-5867, Fax: -5488). DOE SBIR applications are due March 1, 1995.
- Bettie Graham; Bldg. 38A, Rm. 610; NIH; 9000 Rockville Pike; Bethesda, MD 20892 (301/496-7531, Fax: /480-2770).

National SBIR conference: Chicago, IL (April 26–28, 1995). Conference Hotline: 407/791-0720.◊

SELECTED CALENDAR ACRONYMS

AAAS Am. Assoc. for the Advancement of Sci.
AACR Am. Assoc. for Cancer Res.
ABA Australian Biotechnol. Assoc.
ATCC Am. *Type Culture* Collection
BTP Biotechnol. Train. Programs
CATCMB/CUA Ctr. for Adv. Train. in Cell and Mol. Biol./Cathol. Univ. Am.
CEPH Centre d'Etude du Polymorphisme Humain
CHI Cambridge Healthtech Inst.
DIMACS Discrete Math. & Comp. Sci.
GDB/OMIM Genome Data Base/Online *Mendelian Inheritance in Man*
HGP Hum. Genome Proj.
HICSS Hawaii Intl. Conf. on Sus. Sci.
IBC Intl. Bus. Comm.
MBWS Miami Bio/Technol. Winter Symp.
NCHGR Natl. Ctr. for Human Genome Res.
TIGR/NIST The Inst. for Genomic Res./Natl. Inst. of Standards and Technol.
WWW World Wide Web

Human Genome Management Information System Subscription/Document Request* (Vol. 6, No. 4)

Name _____
(First) (MI) (Last)

Affiliation _____

Department/Division _____

Street/P.O. Box/Building _____

City/State/Zip Code _____

Country _____ **Area of Interest** _____

Phone _____ **Fax** _____

E-Mail Address _____

1. ___ *Human Genome News* ___New Subscriber ___Change of Name/Affiliation/Address (circle all that apply) ___Drop Subscription
2. ___ Reprint of "A New Five-Year Plan for the U.S. Human Genome Project" *(Science*, October 1, 1993) by Francis Collins and David Galas
3. ___ DOE *Human Genome 1993 Program Report* ___DOE *Primer on Molecular Genetics*
4. ___ *Meeting Report: DOE Informatics Summit*—DRAFT (April 26–27, 1993, Baltimore, Maryland)

*Please type, print carefully, or enclose a business card to ensure efficient shipping. To change name/address/affiliation or drop your subscription to *Human Genome News*, enclose your current *HGN* address label. Send to HGMIS address shown below and on p. 12.

Acronym List for Centers Article

ABI Applied Biosystems Inc.

ATCC American *Type Culture* Collection

BAC bacterial artificial chromosome

CEPH Centre d'Etude du Polymorphisme Humain

CRADA Cooperative Research and Development Agreement

ELSI ethical, legal, and social issues

FISH fluorescence in situ hybridization

GDB Genome Data Base

GESTEC Genome Science and Technology Center

HHMI Howard Hughes Medical Institute

ICRF Imperial Cancer Research Fund

LTI Life Technologies, Inc.

MIT Massachusetts Institute of Technology

MMRF Marshfield Medical Research Foundation

NCHGR National Center for Human Genome Research

NLGLP National Laboratory Gene Library Project

PAC P1 artificial chromosome

PCR polymerase chain reaction

SSR simple sequence repeat

STRP short tandem repeat polymorphism

STS sequence tagged site

YAC yeast artificial chromosome

OAK RIDGE NATIONAL LABORATORY • MANAGED BY MARTIN MARIETTA ENERGY SYSTEMS, INC. • FOR THE U.S. DEPARTMENT OF ENERGY

Betty K. Mansfield
HGMIS
Oak Ridge National Laboratory
1060 Commerce Park, MS 6480
Oak Ridge, TN 37830, U.S.A.

Postmaster: Do Not Forward.
Address Correction Requested.
Return Postage Guaranteed.

BULK RATE
U.S. POSTAGE
PAID
OAK RIDGE, TN
PERMIT NO. 3

RECYCLED ♻
RECYCLABLE