BIOLOGY PHYSICS MATHEMATICS ENGINEERING CHEMISTRY

Human Genome news

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

ISSN: 1050-6101

Vol. 5, No. 4, November 1993

U.S. Human Genome Project Updates Goals

New 5-Year Plan Incorporates Research Advances

Unexpected advances in genome research and more-sophisticated understanding of how to achieve long-term objectives have led genome project planners at NIH and DOE to update their initial 5-year goals. The new 5-year plan appeared in the October 1 issue of *Science* in an article coauthored by Francis Collins, Director of the National Center for Human Genome Research, and David Galas, formerly head of the DOE Human Genome Program and Associate Director of the DOE Office of Health and Environmental Research.

New Targets:

- **Technology for Mapping, Gene Identification**
- **Resource Distribution**

The new plan extends research goals in already established categories and adds specific new goals for developing technology for gene identification and mapping. It also provides for outreach programs to distribute genome materials to the scientific community. Although the plan covers the next 5 years of the project (through September 1998), the goals were designed to address both long- and short-term needs.

Obtaining the complete human DNA sequence is still the ultimate goal of the project. Although debate continues over the value of sequencing the whole genome, researchers recognize the importance of DNA sequence information in revealing genes and other biological information that could not be obtained by smaller-scale techniques.

The new goals again assume a funding level for the whole genome program of $200 million annually, adjusted for inflation after 1990. Although this amount was also assumed when the initial goals were developed and implemented, appropriations have never reached that level. U.S. genome project funding for FY 1994 (which began October 1) is about $170 million.

A New Plan Needed

Progress over the last 3 years has put the initial goals well within reach with detailed human genetic maps; improved physical maps of human and model organism genomes; development of DNA sequencing and informatics technology; and identification of major ethical, legal, and social issues (ELSI) concerning the increased availability of genetic information.

Although the first 5-year plan was not due to expire until September 1995, "Advances in genome research have already changed the way research is being done," Collins said. "We need to incorporate these advances into our present research goals to ensure that the program continues to be ambitious and cutting edge."

The genome project has already had a profound impact on biomedical research. In just the past few years, maps generated by project researchers have helped in finding genes associated with dozens of genetic conditions, including Menkes syndrome, the X-linked immune disorder ammaglobulinemia, Huntington's disease, myotonic

In This Issue. . .

Genome News

dystrophy, fragile X syndrome, neurofibromatosis types 1 and 2, and others. In addition to the identification of many more disease genes, other future developments will enable researchers to explore gene mutations and health effects caused by environmental agents.

Developing New Goals

In developing the new goals, an NIH-DOE working group sought advice from scientists, other interested scholars, and public representatives, including many outside the Human Genome Project. (Reports of these meetings are available from HGMIS and the NCHGR Office of Communications; contact information is given on page 12.) The plan was presented to and approved by the NIH National Advisory Council for Human Genome Research and the DOE Health and Environmental Research Advisory Committee.

The following are some general observations underlying specific new goals.

Technology Development. This will continue to be crucial to future program success, particularly in the area of large-scale DNA sequencing. Accomplishments that influence research strategies include new types of genetic markers (i.e., microsatellites) assayable by the polymerase chain reaction (PCR); improved vector systems for cloning large DNA fragments and methods for assembling clones into physical maps; use of sequence tagged sites (STSs) as common physical mapping entities; and improved DNA sequencing technology and automation.

Future Mapping Efforts. These efforts should focus on regions both larger and smaller than a single chromosome, the basic unit of genome analysis to date. (An "average" human chromosome contains about 150 Mb.) Production of whole-genome low-resolution maps is now feasible due to PCR and robotic developments. Increasing attention needs to be paid to fine-detail mapping of smaller DNA regions (one to a few megabases) as well. One million bases is an ambitious dimension for detailed analysis, the plan says, and will provide a "useful bridge" between conventional genetics and larger-scale genomics research as well as a "foundation for innovation" to develop methods with applicability to larger regions. Planners note that progress already achieved allows greater focus on gene information to enrich the maps produced.

Specific goals covering the period between October 1, 1993, and September 30, 1998, appear in the shaded box. More details pertaining to these goals are given below.

Genetic Map. Researchers expect that the genetic map specified in the first 5-year plan will be completed on time, but technological improvements are needed to allow rapid typing of families by nonexperts and simultaneous multimarker testing of

FIVE-YEAR GOALS

Revised 5-Year Research Goals of the U.S. Human Genome Project
October 1, 1993, to September 30, 1998 (FY 1994–98)

Mapping and Sequencing the Human Genome

Genetic Mapping

- Complete the 2- to 5-cM map by 1995.*
- Develop technology for rapid genotyping.
- Develop markers that are easier to use.
- Develop new mapping technologies.

Physical Mapping

- Complete a sequence tagged site (STS) map of the human genome at a resolution of 100 kb.*

DNA Sequencing

- Develop efficient approaches to sequencing one- to several-megabase regions of DNA of high biological interest.
- Develop technology for high-throughput sequencing, focusing on systems integration of all steps from template preparation to data analysis.
- Build up a sequencing capacity to allow sequencing at a collective rate of 50 Mb per year by the end of the period. This rate should result in an aggregate of 80 Mb of DNA sequence completed by the end of FY 1998.
- Sequence selected segments of mouse DNA side by side with corresponding human DNA in areas of high biological interest.

Gene Identification

- Develop efficient methods for identifying genes and for placement of known genes on physical maps or sequenced DNA.

Technology Development

- Substantially expand support of innovative technological developments as well as improvements in current technology for DNA sequencing and for meeting the needs of the Human Genome Project as a whole.

Model Organisms

- Finish an STS map of the mouse genome at a 300-kb resolution.
- Finish the sequence of the *Escherichia coli* and *Saccharomyces cerevisiae* genomes by 1998 or earlier.
- Continue sequencing *Caenorhabditis elegans* and *Drosophila melanogaster* genomes with the aim of bringing *C. elegans* to near completion by 1998.

Informatics

- Continue to create, develop, and operate databases and database tools for easy access to data, including effective tools and standards for data exchange and links among databases.
- Consolidate, distribute, and continue to develop effective software for large-scale genome projects.
- Continue to develop tools for comparing and interpreting genome information.

Ethical, Legal, and Social Implications (ELSI)

- Continue to identify and define issues and develop policy options to address them.
- Develop and disseminate policy options regarding genetic testing services with potential widespread use.
- Foster greater acceptance of human genetic variation.
- Enhance and expand public and professional education that is sensitive to sociocultural and psychological issues.

Training

- Continue to encourage training of scientists in interdisciplinary sciences related to genome research.

Technology Transfer

- Encourage and enhance technology transfer both into and out of centers of genome research.

Outreach

- Cooperate with those who would establish distribution centers for genome materials.
- Share all information and materials within 6 months of their development. This should be accomplished by submission of information to public databases or repositories, or both, where appropriate.

*Goals for map resolution remain unchanged.

large numbers of individuals by researchers studying complex genetic diseases. Also needed are methods for automated polymorphic marker screening and new gene mapping strategies not based on a standard set of polymorphic markers.

Physical Map. An STS-based physical map of the human genome with an average resolution of about 300 kb will be completed within 2 to 3 years. Because this level of detail is not sufficiently useful to either gene mappers or sequencers, the plan calls for markers placed at 100-kb intervals. Such a map would be useful to researchers using conventional methods to isolate genes localized within 100 kb of a mapped marker or in DNA-sequencing preparations.

To facilitate gene finding and DNA sequencing, new approaches are needed for constructing higher-resolution maps and for cloning systems closely tied to development of sequencing technology. The plan also recommends improving clone libraries with regard to stability and chimerism and increasing their accessibility.

DNA Sequencing. Although sequencing costs will meet the original 1996 goal of $0.50/bp, planners estimate that $100 million per year will be needed to develop sequencing technology of sufficient sequencing rate to permit the entire human genome to be sequenced by 2005. Further cost reduction and increased ability to assess sequence accuracy are also critical. The plan recommends expanding the number of groups working on large-scale sequencing, improving conventional gel-based approaches, and developing revolutionary new methods.

Gene Identification. Mapping progress and technological improvements have now enabled project planners to specify the development of gene identification technology as a new goal. Incorporating genes into the rapidly growing body of maps and sequences of both human and model organism genomes will make these resources more useful to researchers exploring their effects on human health.

Technology Development. Cooperation is encouraged in developing vital new technologies, especially automation and robotics, that are expandable and exportable to basic science laboratories sequencing genomes not being studied in the Human Genome Project.

Model Organisms. Original goals will probably be exceeded for the mouse genetic map, *Drosophila melanogaster* physical map, and DNA sequencing of *Escherichia coli*, *Sacharomyces cerevisiae*, and *Caenorhabditis elegans*. Priorities include completion of the mouse map and sequencing of specified model organisms.

Informatics. Although much progress has been made, further development of accessible, user-friendly tools to collect, organize, and interpret vast amounts of data continues to be crucial to the success of the project. Major future goals are data management, analysis, and distribution.

ELSI. ELSI discussions are tied to both genomic research and use of the data it produces. Initial policy options regarding this use are being developed for four areas identified as having the greatest immediate potential impact on society: privacy, fairness, clinical applications, and professional and public education. Reports on the full range of issues will continue to be presented during the next 2 years.

Policymakers must consider cultural and other social influences as they prepare policies that anticipate the increasing impact on the public of widespread genetic testing for common conditions. Also recommended and encouraged are the active involvement of concerned individuals and groups in developing policy options as well as increased public and professional education at all levels to prevent stigmatization and discrimination.

Training. Because of the increased number of genome centers, more high-quality training programs are expected to be established to meet the need for interdisciplinary training of scientists for genome research.

Technology Transfer. Many new companies have already been established to develop applications of genome research, and collaborations between government-funded genome scientists and the private sector have increased. The plan encourages further

(see Goals, p. 5)

Article adapted from *Science*, **262**, 43–46 (October 1, 1993). Reprints are available from HGMIS and the NCHGR Office of Communications (see addresses, p. 12).

International Cooperation

The new DOE-NIH 5-year plan credits the "spirit of international cooperation and sharing" that has characterized the Human Genome Project and played a major role in its success. The Human Genome Organization was commended for coordinating international research efforts by organizing chromosome workshops to encourage collaboration and expedite chromosome map completion.

Notable international collaborations:

- *Caenorhabditis elegans* sequencing project (United States and United Kingdom).
- Chromosome 16 physical mapping project (Los Alamos National Laboratory and Australia).
- Chromosome 21 high-resolution physical map (Lawrence Livermore National Laboratory and Japan).
- Human genetic map (NIH and Centre d'Etude du Polymorphisme Humain).
- Whole-genome approach to a human physical map (Whitehead Institute and Genethon).◊

Genome News

Informatics Summit Meeting Report

A meeting of genome informatics advisors was held in Baltimore on April 26–27 to assess current OHER bioinformatics efforts and make recommendations on essential planning and coordination of future activities [see *HGN* 5(3), 1–4 (September 1993)]. The report of this meeting is available from HGMIS (see address, p. 12) and through the Johns Hopkins University Computational Biology Gopher under *5. Genome Project/* then *4. DOE Bioinformatics Draft, Version 2.0/.*

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.

DOE Expands LANL Sequence Data Management

The DOE Office of Health and Environmental Research has announced that management of sequence data at Los Alamos National Laboratory (LANL) is now operating independently with an expanded mission and a new name: Genome Sequence DataBase (GSDB). GSDB will function both as a research resource for the specific needs of the Human Genome Project and as a service facility.

For more than 10 years, LANL maintained GenBank®, an electronic database that serves as the national repository for all nucleotide sequence information, through subcontracts and interagency agreements with the National Institute for General Medical Sciences (NIGMS) and the National Center for Biotechnology Information (NCBI), both units of NIH.

Now, operating as GSDB, LANL researchers will continue to accept new direct data submissions and provide update and annotation services for sequences in their care. They will also extend their work in developing new computer tools to improve the value of genetic sequence databases to the international research community. The name GenBank will now be used exclusively by NCBI to describe the nucleotide sequence database services that NCBI will continue to provide to the scientific community.

GSDB Research and Development Activities

- Emphasize database interoperability and remote data access by improving methods for establishing links between sequence and other databases, such as those for genetic maps.
- Produce better data-submission and annotation tools to aid producers of bulk data as well as individual scientists.
- Create data models that better represent biological relationships and facilitate linkage with other databases.
- Expand the concept of electronic data publishing pioneered at LANL.

GSDB Service Facility Data-Management Activities

- Continued electronic processing of direct submissions (normally within 48 hours) at the following e-mail addresses: *datasubs@t10.lanl.gov* for submissions and *update@t10.lanl.gov* for corrections and additions.
- Collaboration with other databases to provide a unified international data-collection entity.
- Increased emphasis on automating the quality-control process, in addition to quality control of submissions by GSDB annotation and review staff.
- Increased emphasis on online submission and maintenance. For the growing portion of the research community on the Internet, GSDB will stress the use of online data-submission tools such as the Annotator's WorkBench (AWB) over batch-submission tools like Authorin.
- Renewed emphasis on remote database access through continued support for relational satellite copies of the LANL database as well as direct, Sybase client-server access for remote queries in standard query language.

GSDB Relationship to Other Databases

GSDB is committed to productive and complementary interactions with other sequence databases such as those at the DNA Data Bank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and NCBI. DOE has been engaged in discussions regarding future relationships between GSDB and other databases to ensure that GSDB activities complement those at other sites to produce improved services for the user community. Data submitted to GSDB are being made available to other sequence databases as soon as processing is complete. Also, data submitted to other databases are incorporated into GSDB as the data become available.

Interoperable Information Resources

For more than a year, DOE has been carrying out an extensive review of its informatics activities in support of genome and structural biology research. Advice and comments from reviewers and the community have emphasized the need for improved, integrated information resources [see *HGN* 5(3), 1–4 (September 1993)]. The report (see side bar) stated that interoperability among crucial databases is essential and noted that current databases are

unable to answer simple queries requiring integration of map, sequence, and other biological data.

Following advice from the report and elsewhere, DOE determined that it must develop and support an integrated information infrastructure for genome and structural biology research. DOE also resolved that major database elements in the integrated infrastructure should emphasize direct access through networked application programming interfaces and allow direct online data submission, annotation, and curation by the research community.

The database component should be both a research project and a production service supporting ongoing biological research, with the research project undertaking development of better data models and direct online tools for data submission and curation and for federated data access.

In the short term, nucleotide data-resource development supported by DOE will take advantage of the specific expertise, facilities, and capacity developed at LANL during its long tenure as a leading U.S. site for nucleotide database development. Over time, DOE nucleotide data resources will undoubtedly evolve in accordance with the developing integrated infrastructure (a "center without walls") and will be subjected to extensive peer review and competitive evaluation.

Historical Role of Los Alamos

In the 1970s, Walter Goad established the Los Alamos Sequence Database, a pioneering effort at LANL that in 1982 evolved into the GenBank project. LANL continued to expand and build the database in collaboration with the firm Bolt, Beranek, and Newman under funding provided by NIGMS and other federal agencies. In 1987 LANL continued to be the site of database design and maintenance, working with IntelliGenetics.

In 1992, NIH transferred its management control for the GenBank project from NIGMS to NCBI at the National Library of Medicine. At that time, DOE and NCBI entered into an Inter-Agency Agreement (IAA) so that LANL could provide assistance in processing direct submissions for NCBI. The IAA noted, "For nine years, LANL has been responsible for the design and management of gene sequence data as part of the GenBank project. . . . In the most recent re-competition, all three proposals which were in the competitive range included LANL as a subcontract for the direct data submission component of the project. Thus, LANL was recognized not only for its past experience in establishing the procedures for collecting and managing biological data, but for its innovative approaches in handling data prior to or independent of the publication process."

Now, NCBI has developed its own capacity for processing direct submissions, freeing LANL to develop new approaches, tools, and services targeted specifically for the genome community.◊

DOE OHER Funds Russian Research

The DOE Office of Health and Environmental Research (OHER) is providing small emergency grants from its Human Genome Program funds to 22 Russian research teams. These teams were identified last December through visits to a dozen institutions and interviews with more than 100 scientific groups in the Moscow, St. Petersburg, and Novosibirsk areas by David Galas (then head of OHER), Elbert Branscomb (Lawrence Livermore National Laboratory), and Raymond Gesteland (University of Utah). The grants are for $5000 a year except for one very large team receiving $30,000. To avoid bureaucratic delays, these funds go directly to investigators, with overhead allowances for their institutions.

"These excellent groups are contributing good scientific research at a very low cost," said Marvin Stodolsky, coordinator for the DOE subprogram. "We think is to the benefit of the United States, Russia, and the world that the teams maintain their integrity, and we are making a small contribution in that regard." An extended report on efforts to fund Russian investigators can be found in *Science* **261**, 1382 (September 10, 1993).◊

Integrated Infrastructure To Emphasize Direct, Online Access

Book Features LANL Genome Program

A special issue of *Los Alamos Science* (No. 20, 1992), focusing on the genome program at Los Alamos National Laboratory, has been reprinted and is available for purchase. The 338-page volume contains numerous graphics and includes basic information on the Human Genome Project; principles of classical and molecular genetics; mapping goals and technologies, particularly for chromosome 16; informatics; rapid DNA sequencing; and some of the ethical, legal, and social issues surrounding the project. $38. University Science Books; 20 Edgehill Road; Mill Valley, CA 94941 (415/383-1430, Fax: -3167).◊

Goals *(from p. 3)*

cooperation with industry but cautions that care must be taken to avoid conflicts of interest. Technology transfer from other fields to genome centers must also occur.

Outreach. The private sector is encouraged (with seed funding in some cases) to establish distribution centers for genome materials and respond quickly to the evolving needs of the scientific community. The policy on data and material sharing (within 6 months of creation) has been well accepted.◊

Genome News

Nancy Wexler Receives Lasker Award

Copies of *Genetic Information and Health Insurance*, the insurance task force report mentioned in this article, are available from ELSI Branch, NIH NCHGR; Bldg. 38A, Rm. 617; 9000 Rockville Pike; Bethesda, MD 20892 (301/402-0911, Fax: -1950).

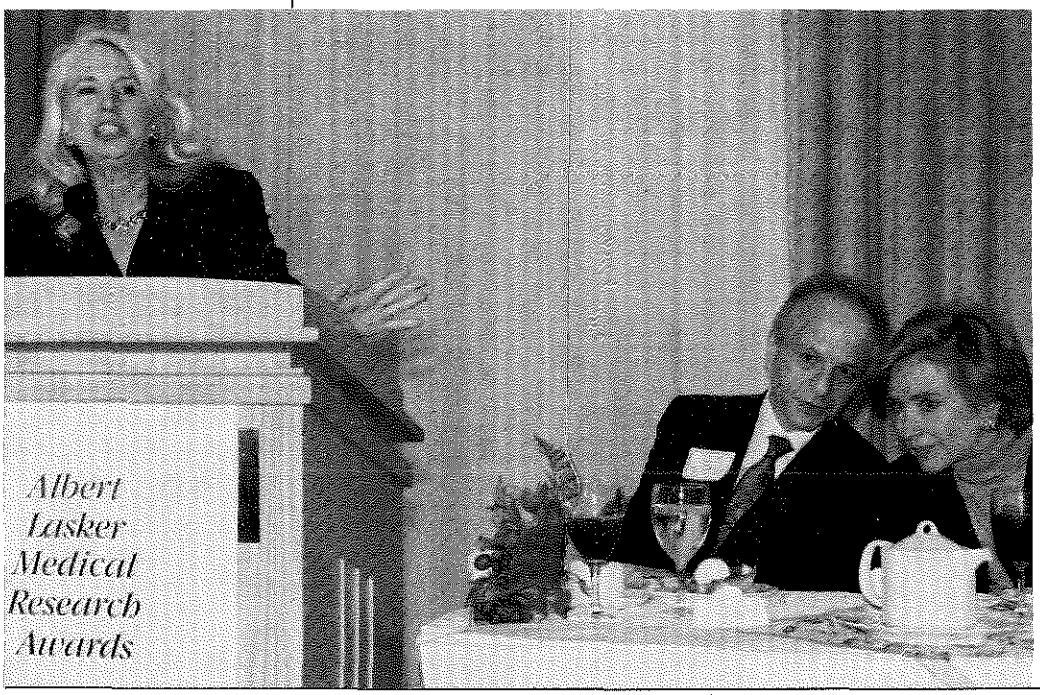
Nancy S. Wexler (Columbia University), chair of the Joint DOE-NIH Human Genome Project Ethical, Legal, and Social Issues (ELSI) Working Group and of the ELSI Committee of the international Human Genome Organization, is the recipient of the Albert Lasker Public Service Award for 1993. The award was presented on September 30 at a luncheon and awards ceremony at which Hillary Rodham Clinton was the keynote speaker. The award cited Wexler's work in mobilizing research, policy development, and scientific advocacy in the worldwide effort to find a cure for Huntington's disease (HD) as well as her service as President of the Hereditary Disease Foundation.

Clinton said of this year's recipients, "These winners join the remarkable group of men and women who have over the decades put health care at the top of our national agenda, men and women whose work has found cures for disease, who have aided the kind of breakthroughs that we've only been able to dream about in the past but now take for granted, men and women who have helped better our health-care system through public awareness and through important legislation."

Clinton also pointed out that personal genetic data generated through Human Genome Project technology might be used as the basis for discrimination by health-insurance carriers unless health-care reform could be achieved. Clinton's health-care reform committee was provided with the report (see sidebar) of the ELSI Working Group task force on genetic information and health insurance. This report, approved by DOE and the Department of Health and Human Services (parent agency of NIH), recommended that access to basic health services should not be conditioned on genetic information.

Wexler, a clinical psychologist, has been involved in the fight against HD since 1968, when she learned that her mother was ill with the genetic condition and that both she and her sister have a 50% chance of developing HD. Her doctoral dissertation on the neuropsychological and emotional consequences of being at risk for HD led to her later appointment as Executive Director of the Commission for the Control of Huntington's Disease and Its Consequences.

Wexler has led medical expeditions to the shores of Lake Maracaibo in Venezuela to work with the largest extended HD family in the world. From this family, Wexler's team has constructed a multigenerational pedigree with more than 13,000 members and collected over 3000 blood samples. These samples have also been used in mapping genes responsible for Alzheimer's disease, kidney cancer, and two types of neurofibromatosis.



Nancy Wexler (Columbia University) accepts the Albert Lasker Public Service Award for 1993. Also on the dais were Jordan Gutterman (University of Texas M. D. Anderson Cancer Center) and Hillary Rodham Clinton, keynote speaker for the awards ceremony.

The Venezuelan samples and pedigrees enabled investigators to pinpoint the gene's precise location, a discovery that many investigators believe will lead to finding a way to prevent or repair HD damage.

Albert Lasker Medical Research awards have been presented continuously for 46 years. Winners receive or share a $25,000 honorarium, a citation highlighting their achievements, and an inscribed statuette of the Winged Victory of Samothrace—the Lasker foundation's traditional symbol of victory over disability, disease, and death. *[Anne Adamson, HGMIS]* ◊

Genome News

NCHGR Scales Up *C. elegans* Sequencing

NIH Awards 5-Year Grant to Washington University Sequencing Center Project

The NIH National Center for Human Genome Research (NCHGR) has awarded a 5-year, $29.7-million grant to Robert H. Waterston (Washington University School of Medicine). This grant will be used to establish the Washington University DNA Sequencing Center, bringing to 18 the number of NCHGR-supported genome and technology centers.

The sequencing center will continue to work closely with collaborators headed by John Sulston at the Sanger Center in Cambridge, England, and its four projects will have the following goals:

- Complete the sequence of the 100-Mb *Caenorhabditis elegans* genome.
- Assist in completing the sequence of the yeast *Saccharomyces cerevesiae* by contributing 3 Mb of sequence in the first 2 years of the project. A small project is included for systematically altering gene expression in the sequenced regions.
- Initiate a 2-year pilot project to evaluate approaches for sequencing human genomic DNA. In the first year, efforts will focus on the 200-kb region of chromosome 16p that harbors the gene for autosomal dominant polycystic kidney disease. In the second and third years, attention will shift to medically important, gene-rich regions of chromosome 7.

Center activities will be supported by four cores, including a development core devoted to implementing robotics and technological improvements. An informatics core will develop software for automated assembly and editing of DNA sequence data and improved interpretation of generated sequence. With these technologies, Waterston hopes to increase the laboratory's annual sequencing capabilities from just over 1 Mb to 10 to 15 Mb.◊

Center Releases Updated Mouse Genetic Map

Researchers Can Use PCR To Detect Mapped SSLPs

The fourth release of the Whitehead Institute/Massachusetts Institute of Technology (MIT) Genome Center Genetic Map of the Mouse became available in October. This map consists primarily of randomly chosen simple-sequence length polymorphisms (SSLPs) (microsatellites) that can be detected using the polymerase chain reaction (PCR), as described by W. Dietrich, et al. [*Genetics* **131**, 423–47 (1992)]. The released map contains 2006 markers that fall into 20 linkage groups spanning about 1470 cM with an average spacing of less than 1 cM.

Data Access Information

1. Anonymous ftp to *genome.wi.mit.edu;* login, *anonymous;* password, *user e-mail address.* The release can be found in the directory */distribution/mouse_sslp_release/oct93/.* The file *README* describes the file format and gives other map information.

2. Internet e-mail using a database e-mail server. Locus and assay names of mapped SSLPs, forward and reverse primer sequences, genotypes of loci on the mapping cross, sizes of PCR products on selected standard inbred strains, and other useful information can be obtained. Markers can be requested by name, chromosome, map position, or other criteria such as informativeness. To obtain copies of the most current e-mail query forms with instructions for their use, send a message to *genome_database@genome.wi. mit.edu* with *help* in either the subject line or body text. The completed form should be sent to *genome_database@genome.wi.mit.edu;* answers will be returned automatically.

New markers added to the map will be released near the first of each quarter in 1994. Contact: Lincoln Stein; Whitehead Institute/MIT Genome Center; 9 Cambridge Center; Cambridge, MA 02142 (617/252-1916, Fax: -1902, Internet: *lstein@genome.wi.mit.edu*).◊

¶ HUGO Publishes *Digest*

The European region of the Human Genome Organization (HUGO) began publishing the quarterly newsletter *Genome Digest* in July. Focusing on genome research in Europe, the digest provides information about biological resources, bioinformatics, grants and fellowships, HUGO, meetings, and key global issues and events. To be placed on the mailing list, contact the Editorial Office; HUGO Europe; One Park Square West; London NW1 4LK, United Kingdom (+44/71-935-8085, Fax: -8341).◊

Genome News

Workshop on Educational Resources in Genetics: Organizations and Databases

Educational Resources In Genetics: Organizations and Databases, a five-page listing of names and contact information, is available from Virginia Proud at the address below or through the JHU Computational Biology Gopher system under *5. Genome Project/*, then *6. INFOGEN: Information in Genetics/* [for instructions on accessing Gopher, see *HGN* 5(3), 8 (September 1993)]. Proud wishes to include in future directories and CONTIG activities (see article) the names of individuals and organizations interested in making genetics information available for educational purposes. Contact Proud at Laboratory of Medical Genetics; University of Alabama at Birmingham; 908 20th Street South, Room 323; Birmingham, AL 35294-2050 (205/934-4973, Fax: /975-6389, Internet: *gene003@uabdpo.dpo.uab.edu*).◊

A multidisciplinary group of educators, computer specialists, clinical geneticists, counselors, librarians, and consumers representing 15 constituent organizations met September 26–27 at Johns Hopkins University (JHU) for the Workshop on Educational Resources in Genetics: Organizations and Databases. The meeting was organized by the education committee of the Council of Regional Networks of Genetic Services (CORN) and Genome Data Base and sponsored in part by the Maternal and Child Health Bureau. Chaired by Virginia Proud (University of Alabama and CORN Consumer Database Subcommittee), the workshop was designed to develop a system for easy access to basic medical genetics information and to coordinate existing and potential databases.

Representatives from constituent groups (including Alliance for Genetic Support Groups, March of Dimes Birth Defects Foundation, and National Organization of Rare Disorders) expressed the need for easily accessible, accurate information for the thousands of callers who contact their organizations each month. As the impact of the Human Genome Project begins to be realized, questions and requests for information will increase, they said. Valuable information and educational resources developed through the Human Genome Project, including educational programs and curricula, videos, and print publications, also need to be made available to the genetics and education communities and to consumers.

Attendees presented material already developed—such as data sets, bibliographies, and online access—that, when compiled and integrated, could avoid duplication of effort and prove useful to other groups and individuals. Under the direction of Dan Jacobson and David Kingsbury (both of JHU), participants also accessed Online *Mendelian Inheritance in Man* (OMIM) and other genetic databases worldwide through the JHU Computational Biology Gopher. They concluded that some good resources on medical genetics and clinical medical conditions are now available; files already in computer-readable format or in defined databases could quickly be made accessible to the entire genetics community through Gopher.

Attendees were enthusiastic about developing these resources into a consumer-oriented database called INFOGEN. Although it had been proposed as a bibliography of educational resources, those present felt that a full-text database, possibly with graphics, would be more useful; information could also be distributed in hard copy, CD-ROM, or floppy disk. Teachers of genetics expressed interest in serving as reviewers and resource developers for the database, and an advisory board and related working committees are being set up.

To continue the developing collaboration, the Consortium of Teachers in Genetics (CONTIG) was proposed at this meeting. CONTIG would bring together collaborators to develop educational systems for helping to bridge the gap between (1) scientific data generated through basic genetics research and the Human Genome Project and (2) useful applications of this data for the public and clinicians. "Teachers" in CONTIG include but are not limited to various educators who are also geneticists, counselors, librarians, medical informatics specialists, consumers, and representatives from the Human Genome Project.

[Virginia Proud, University of Alabama] ◊

Some Collaborating Groups

- Alliance of Genetic Support Groups
- American Society of Human Genetics (ASHG)
- Biological Sciences Curriculum Study (BSCS)
- Birth Defects Information Services, Inc. (BDIS)
- California Genetics Education Resource Center
- Council of Regional Networks for Genetics Services (CORN)
- DOE
- DOE HGMIS
- NIH National Center for Human Genome Research
- Gallaudet University Genetic Services Center
- Genetic Services Branch, Maternal and Child Health Bureau
- Genetics Education Programs for Teachers
- GDB and OMIM
- Great Plains Genetics Services Network
- Helix: A National Directory of DNA Diagnostic Laboratories
- Institute of Medicine
- International Society of Nurses in Genetics (ISONG)
- March of Dimes Birth Defects Foundation (MOD)
- Michigan Human Genetics Videotape Lending Library
- National Research Register for Hereditary Hearing Loss
- National Science Foundation
- National Organization for Rare Disorders, Inc. (NORD)
- National Society of Genetic Counselors
- National Information Center for Children and Youth with Disabilities (NICHCY)
- National Library of Medicine (NLM)
- National Center for Education in Maternal and Child Health (NCEMCH)
- New York State Library
- Rare Disorder Network
- Sensory Genetics/ Neuro-Development Vision Screening Project
- Shodair Hospital Family Resource Library
- Texas Genetic Network (TEXGENE)
- University of Michigan Human Genome Center Education Center

IOM Issues Report on Genetic Testing

On November 4, the Committee on Assessing Genetic Risks of the Institute of Medicine (IOM) of the National Academy of Sciences released its report on a study of a variety of issues raised by the rapid proliferation of genetic tests. Designed to produce policy options on urgent issues, the report was commissioned in 1990 as one of several proactive initiatives of the NIH-DOE ELSI component of the U.S. Human Genome Project. The 24-month study, cosponsored by the L. Markey Trust and chaired by Arno Motulsky (University of Washington, Seattle), was carried out by a panel of experts in human genetics, law, health education, economics, ethics, medicine, insurance, psychology, and regulation. Jane Fullarton (now at Tascon) was staff director of the project.

The report, designed to produce policy options on urgent issues, was commissioned in 1990 as one of several proactive initiatives of the NIH-DOE ELSI component of the U.S. Human Genome Project.

In the report, the committee attempted to set down broad guidelines and policy principles to guide the development and use of genetic tests. They pointed out that genetic testing is already used or may be used in newborn screening, prenatal testing, genetic carrier testing, medical diagnosis, and presymptomatic or predictive testing. Predictive testing, which is likely to have major impact, raises new problems because of the ability to identify individuals at risk long before they develop symptoms. The committee expressed concern about commercial pressures urging premature introduction of presymptomatic tests and recommended that pilot studies be carried out before certain types of tests are used routinely.

The committee recommended that genetic testing should be voluntary and that the individual should decide what to do with the genetic information, all forms of which should be confidential and not disclosed without the person's consent. All organizations that generate or maintain genetic information or samples should have established procedures for protecting confidentiality.

The committee strongly supported nondirective counseling and felt that appropriate medical options should be given only after a full discussion of the medical and social consequences of genetic information, including its possible impact on employment and health insurance. The panel also favored legislation to prohibit possible discrimination in health insurance and employment. Specific recommendations covered a wide variety of circumstances, including the following:

- Informed consent is not merely signing a form but a process of education and the opportunity to have questions answered.
- As more tests for different diseases are bundled and performed together, multiple genetic traits can be detected in an individual. Such multiplex tests should be grouped into categories that raise similar problems of informed consent, for example, treatable vs untreatable and early-onset vs late-onset conditions. Each group of tests should be administered separately and at different times of life.
- More public education and much more intensive coverage of genetics in schools at all levels are critical.
- Genetic tests must be accurate, effective, and interpreted with close to "zero-error" tolerance. The highest standards must be applied to new genetic tests and to laboratories performing tests.
- Private and public health plans, geneticists, and consumers must work together to develop guidelines for reimbursement for appropriate genetic services, which should also be incorporated into health-care reform plans.
- A broadly representative national advisory committee, aided by an expert working group on genetic testing, should be established to set standards for genetic tests and policies for the use of new tests in medical practice.
- Significant gaps exist in data on current genetic testing and screening. More information is needed to make appropriate decisions in this area.◊

Prepublication copies of the report *Assessing Genetic Risks: Implications for Health and Social Policy* are available from the National Academy Press at 2101 Constitution Avenue, NW,; Washington, DC 20418 (202/334-3313 or 800/624-6242).

¶ **Genome-Related Publication:** ***Computers & Chemistry*** is devoting a series of special issues [edited by Andrzej Konopka (BioLingua Research)] to state-of-the-art computational molecular biology. Topics for these issues were selected during Workshops on Open Problems in Computational Molecular Biology, held annually in Telluride, Colorado [see *HGN* **3**(5), 11–12 (January 1992) and **5**(1), 8 (May 1993)]. Peer-reviewed papers pertaining to the 1991 and 1992 workshops were published in *Computers & Chemistry* **16**(2) (1992) and **17**(2) (1993), respectively, and a third special issue is planned for 1994. [David Claridge; Pergamon Press, Ltd.; Chemistry Department; Headington Hill Hall; Oxford OX3 OBW, U.K. (Fax: + 44/865-60285).] ◊

GDB Forum

SURVEY OF COMPUTER USAGE

The development of computer software to support the Human Genome Project—from analytical tools for the individual investigator to public databases such as GDB, GenBank®, and SWISS-PROT—is hampered by the lack of reliable statistics on computer usage by geneticists, molecular biologists, physicians, and other data users.

This survey is a modest attempt to collect this information, *NOT ONLY FOR GDB's USE* but for sharing with all groups who are working to develop better tools for gathering and analyzing biological data. The development of more useful software to aid research will be greatly facilitated by users who fill out this form and return it to GDB by fax, direct mail, or e-mail. Each response will be very much appreciated.

Survey Return Information
GDB User Support
Johns Hopkins University
2024 E. Monument Street
Baltimore, MD 21205
410/955-7058, Fax: /614-0434
Internet: *help@gdb.org*
Electronic survey form: *survey.txt* is available via ftp from *ftp.gdb.org* and via Gopher (*gopher.gdb.org*) in the Genome Project/GDB section.

1. How would you classify your organization? (circle one)
   A. University B. Genome Center C. Industry D. Hospital/Patient care E. Other _____
2. What is your major responsibility in your organization? (circle one)
   A. Genetics research B. Administrative/Management C. Computer/Systems D. Education E. Patient care F. Other _____
3. What kind of computers do you use? (rank usage with 1 = most frequent; specify model)
   A. __ Mac ( ) B. __ PC ( ) C. __ Sun ( ) D. __ DEC Alpha ( ) E. __ Other UNIX workstation ( )
   F. __ ASCII terminal ( ) G. ___ Other ( )
4. Which computers do you use to access GDB/OMIM? (circle all that apply)
   A. Mac B. PC C. Sun D. DEC Alpha E. Other Unix workstation F. ASCII terminal G. Other _____
5. Which of your computers has Internet access? (circle all that apply)
   A. Mac B. PC C. Sun D. DEC Alpha E. Other Unix workstation F. Other _____
6. Which of your computers has a CD-ROM? (circle all that apply)
   A. Mac B. PC C. Sun D. DEC Alpha E. Other Unix workstation F. Other _____
7. Which of your computers has a modem? (circle all that apply; specify speed)
   A. Mac ( ) B. PC ( ) C. Sun ( ) D. DEC Alpha ( ) E. Other Unix workstation ( ) F. Other _____ ( )
8. Which Internet services do you use? (circle all that apply)
   A. E-mail B. FTP C. Gopher D. WAIS E. USENET news F. World Wide Web G. Other _____
9. If you have problems using hardware/software, where do you go most often for help? (circle one)
   A. Local technical support B. Knowledgeable coworker C. Manufacturer/Distributor D. Manual
10. Which biological databases do you use? (circle all that apply)
    A. GDB B. OMIM C. GenBank D. Medline E. PIR (Protein Identification Resource)
    F. SWISS-PROT G. PDB (Protein Data Bank) H. Entrez I. Other _____
11. Which methods do you use to access GDB data? (circle all that apply)
    A. GDB application (APT forms) B. GDB/Accessor C. WAIS D. Gopher E. FTP F. HGM book (HGM11, CCM92)
12. Which methods do you use to access OMIM data? (circle all that apply)
    A. OMIM application (IRX) B. GDB/Accessor C. WAIS D. Gopher E. FTP F. MIM book
13. Do you presently have data to submit to GDB? (circle one)
    A. No B. Yes, waiting for complete data C. Yes, waiting to find out how to submit it
14. Would a software package analogous to GenBank's Authorin increase your willingness to submit data to GDB?
    A. Yes B. No
15. Please add any comments you feel would help improve GDB services.

(Optional) Name_____ Organization_____

GDB 5.2 Enhances Output

In Genome Data Base (GDB) Version 5.2 to be released this winter, output report generation has been revised to "look and feel" like the data managers. The new Output Manager includes the following capabilities:

Report Contents: (1) generate reports that include data updates only; (2) specify order of related data; (3) specify inclusion of page breaks; (4) specify inclusion of SQL query; and (5) output Field Values, Online Help, and User Accounts.

Report Definition, Saving, and Generation: (1) Save and use session defaults for Table/Detail format and order of related data, (2) define and modify reports without having to generate them, and (3) cancel scheduled reports before they are generated.

A detailed description of how to use the Output Manager and other new features will be available online in "Release Notes" under "News."◊

GDB/Accessor Available for Macintosh

GDB/Accessor is a Macintosh program for accessing GDB and related genetic databases on the Internet. The software was developed and is maintained by Corprew Reed and Tom Marr (both at Cold Spring Harbor Laboratory) in cooperation with GDB.

The program, which has a typical Macintosh "look and feel," uses multiple windows and allows the user to carry out the following tasks:

- Define GDB searches for locus, map, and probe information.
- Choose loci, maps, or probes for inclusion in a report generated directly on the Macintosh.
- Design reports by selecting data content and format.
- Review reports on the screen.
- Manipulate text (save, print, cut, paste).
- Define direct searches and use cross references to retrieve GDB-related data from sequence, disease, and mouse databases on the Internet.

The software is designed to run without loss of functionality under System 6 or 7. No database logins or e-mail addresses are needed. Minimum requirements are stated below.

For Either System

- MacTCP Version 1.0 or above (MacTCP Version 1.1.1 recommended),
- About 600 K to store the program, documentation, and configuration files,
- About 1.5 MB of free memory, and
- At least 640 by 480 screen resolution (13-in. monitor or larger).

For System 6

- System software 6.0.4 or above (Version 6.0.8 recommended) and
- At least 2 MB of RAM to run under Finder and 2.5 MB under MultiFinder.

For System 7

- System software 7.0 or above and
- At least 4 MB of RAM.

GDB/Accessor is available at no cost via anonymous ftp:

- SERVER NAME: *ftp.gdb.org*,
- DIRECTORY: *pub/mac/accessor*, and
- FILES: *accessor.sit.hqx* (program in stuffed, BinHex format) and *accessor.readme* (converting BinHex file into Macintosh application).

Users with appropriate hardware but without ftp access should contact GDB User Support for alternate means of obtaining software (see address above). Questions should also be sent to GDB User Support.◊

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. To obtain a user's local SprintNet (Telenet) number for locations within the United States: 800/736-1130.

GDB, OMIM Training Schedule

Contact U.S. GDB User Support Office (below). General User Classes will be held in Baltimore on February 14–15, April 18–19, and June 13–14, 1994.

User Support Offices

UNITED STATES
GDB User Support
Genome Data Base
Johns Hopkins University
2024 E. Monument Street
Baltimore, MD 21205-2100
410/955-7058
Fax: /614-0434
Internet: *help@gdb.org*

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: *dbk261@cvx12.dkfz-heidelberg.de*

UNITED KINGDOM
Christine Bates
Human Gene Mapping Program Resource Center
CRC, Watford Road
Harrow, Middx HA1, 3UJ
United Kingdom
+ 44/81-869-3446
Fax: -3807
Internet: *cbates@uk.ac.crc*

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*

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*

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

GDB Adds EST Data

GDB now includes 8977 locus and cloned probe records of over 9000 expressed sequence tags (ESTs) downloaded from the dbEST database of the National Center for Biotechnology Information (NCBI). Both types of records contain references to GenBank® numbers, and locus records are assigned D-numbers that have been returned to NCBI to update corresponding records in dbEST. The easiest way to access this data in GDB is first to retrieve the GDB identification number G00-043-725 in the Citation Manager, then call either Probe or Locus Manager to retrieve the EST data.◊

Resources

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 Correspondence Address
Betty K. Mansfield
ORNL
P.O. Box 2008
Oak Ridge,TN 37831-6050
615/576-6669
Fax:/574-9888
BITNET: *bkq@ornlstc*
Internet: *bkq@ornl.gov*

Sponsor Contacts

Daniel W. Drell
DOE Program Office
Germantown, MD 20545
301/903-6488, Fax: -8521
Internet: *daniel.drell%er@mailgw.er.doe.gov*

Leslie Fink
NIH National Center for Human Genome Research
Bethesda, MD 20892
301/402-0911, Fax: -4570
Internet: *lsf@cu.nih.gov*

National Center for Human Genome Research

☛ Software, Services, Electronic Data Access

This material sent to HGMIS is not reviewed or evaluated and is not a comprehensive list.

ACEDB is a Windows-based *Caenorhabditis elegans* database containing genomic data that includes strain and sequence, physical and genetic maps, and a bibliography. Contact: Richard Durbin; Medical Research Council Laboratory of Molecular Biology; Cambridge CB2 2QH, England (+44/223-248011, Internet: *rd@mac-lmb.cam.ac.uk*).◊

GeneScape is a relational database of *Escherichia coli* genomic map data for Macintosh computers. It allows access to information being collected in datasets of DNA sequences (EcoSeq), maps (EcoMap), and genetic data (EcoGene). Available by anonymous ftp to *ncbi.nlm.nih.gov*.◊

Yeast Genome Database is under development. The Yeast Genome Information Server can be accessed by Gopher at *genome.stanford.edu*.◊

FlyBase is a *Drosophila melanogaster* database being built and maintained by Michael Ashburner (University of Cambridge, England); Bill Gelbart (Harvard University); Thom Kaufman and Kathy Matthews (both at Indiana University); and John Merriam (University of California, Los Angeles). This database will continue the tradition of the RedBook but in an electronically accessible and searchable form. Preliminary and working parts of FlyBase are available at the Indiana University Bio archive, which can be accessed by ftp from *ftp.bio.indiana.edu* or by Gopher at *fly.bio.indiana.edu*.◊

Screening Service for Chromosome 21 Yeast Artificial Chromosomes (YACs) is available, and a P1 library screening service is being started. Contact: David Patterson; Eleanor Roosevelt Institute for Cancer Research (303/333-4515, Fax: -8423, Internet: *davepatt@druid.hsc. colorado.edu*).◊

Chromosome 17 YAC Screening Service is available. Contact: Craig Chinault (713/798-6075, Internet: *chinault@bcm.tmc.edu*). For electronic access to information on chromosome 17 YACs, send e-mail to *yaclab@ bcm.tmc.edu*.◊

Chromosome 9 Anonymous FTP Server has been established by John Attwood (University College, London; *john@mrc-hbgu.ucl.ac.uk*). The address *diamond.gene. ucl.ac.uk (128.40.82.1)* may be used to download chromosome 9 workshop abstracts, figures, and reports (to be published in *Cytogenetics and Cell Genetics*). After connecting to the server, users should log in as *anonymous* and give their electronic mail address as a password. The workshop files are in the directory */pub/c9workshop* in both plain text and postscript format. An electronic mailing list has also been arranged for the chromosome 9 community. To be added to the list, send e-mail to Attwood at the above address.◊

GRAM V1.4 (Genomic Restriction Map AsseMbly), a software tool for assembling restriction maps from single-digest data, is being distributed by the Center for Human Genome Studies at Los Alamos National Laboratory (LANL). The principal author, designer, and developer of GRAM is Carol Soderlund, a 1991 DOE Human Genome Distinguished Postdoctoral Fellow.

Using single-digest restriction fragments as input, GRAM outputs one or more plausible restriction maps with clone alignment. The exact restriction map can rarely be generated due to the complexity of the problem and the uncertainty in the data. Therefore, the GRAM algorithm uses a set of heuristics and stochastic algorithms to approximate the solution. With GRAM's interactive graphics, the user can query the GRAM solution and correct outstanding errors. Example execution times for the GRAM algorithms are 3 seconds on an input of 14 clones with a restriction map of 24 fragments and 18 seconds on 47 clones with a restriction map of 72 fragments. GRAM is available from LANL without charge, and the following can be obtained through anonymous ftp: the GRAM source code, which is written in C and uses Athena widgets; a postscript file of the GRAM manual; a set of sample input files; and the executable software that was compiled on a Sun Sparcstation and statically linked with the X11R4 libraries. For more information on GRAM, contact Soderlund [LANL; T-10 MS K710; Los Alamos, NM 87545 (Internet: *cari@t10.lanl.gov*)].◊

GDB-Lite Tool Developed for Macintosh. GDB-Lite is a user-friendly browsing and data-entry tool developed for the Macintosh by Randall Smith and Joanna Power of Baylor College of Medicine. With the point-and-click interface of the 4th Dimension (4D) relational database management system, investigators can use the automatic join features to incorporate data from several Genome Data Base (GDB) tables. They can also enter their own probe and locus information directly into a compatible format for electronic submission to GDB. User documentation and the 4D structure and data files for GDB-Lite may be obtained by Internet anonymous ftp from *gc.bcm.tmc.edu* (subdirectory *GDB-Lite_0.3*). GDB-Lite requires a Macintosh with a hard disk, monochrome or high-resolution monitor of at least 12 in., and the 4D Runtime Version 3.0 from the software manufacturer ACIUS (Cupertino, CA). For more information, e-mail to *gdb-lite@bcm.tmc.edu* or contact Power (713/798-4689, Fax: -5386).◊

PROSEARCH: Fast Searching of Protein Sequences. ProSearch is a computer program that allows rapid identification of biologically relevant motifs in protein sequences and generation of a report describing the patterns and their locations. Although user-developed patterns can be added to any search, ProSearch currently uses the 690 different patterns in the PROSITE database of Amos Bairoch (University of Geneva), which consists of protein motifs characteristic of biochemical functions or protein families in the SWISS-PROT database. ProSearch, written in the AWK computer language that can run on all platforms commonly found in laboratories, is available by anonymous ftp from *receptor.mgh.harvard.edu*. The program was developed by Lee F. Kolakowski, Jr. [Massachusetts General Hospital and Harvard Medical School (MGH HMS)]; Jack Leunissen (University of Nijmegen, Netherlands); and Joseph Smith (University of Pennsylvania) [*BioTechniques* **13** (6), 919–21 (1992)]. Contact: Lee Kolakowski; MGH HMS; 149 13th St.; Charlestown, MA 02129 (Fax: 617/726-5669; Internet: *kolakowski@helix.mgh.harvard.edu*). ◊

☛ Free Resources?

***HGN* would like to be informed about freely available informatics and educational resources being developed for use by genomics researchers.**

LBL Chromosome 21 Physical Mapping Database

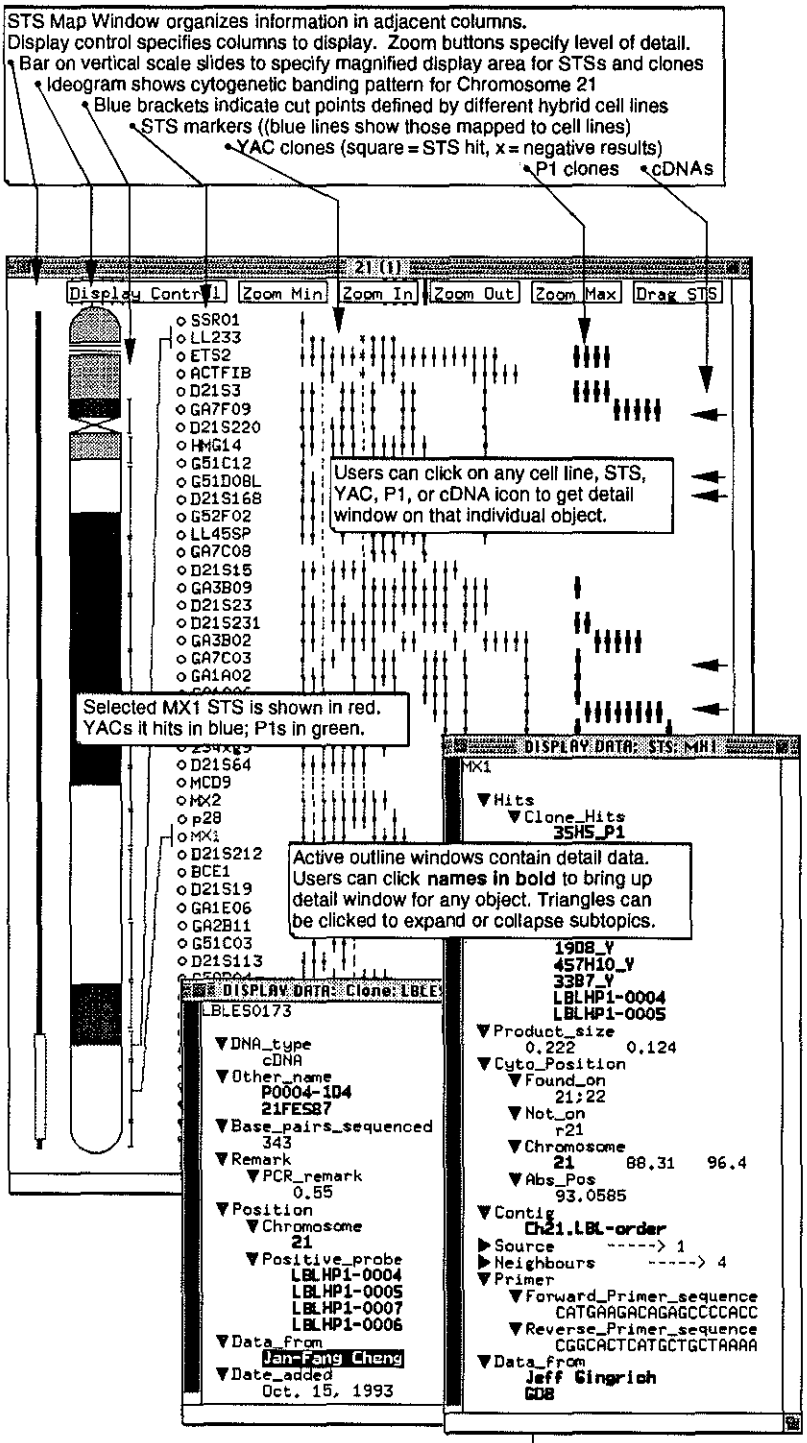
In keeping with the DOE policy of providing timely access to genome data, the Human Genome Center at Lawrence Berkeley Laboratory (LBL) is making data available to other members of the research community from its Chromosome 21 Physical Mapping Database (21Bdb). 21Bdb was developed by the genome center and its Genome Computing Group.

This initial release (1.1) of 21Bdb covers several different types of data and their relationships, including 250 sequence tagged site (STS) markers, 102 P1 clones, 816 yeast artificial chromosome (YAC) clones, 217 cDNA clones, and 160 DNA sequences. New LBL data and revisions will be added weekly. Users are encouraged to add questions and comments.

21Bdb extends ACEDB, which was originally developed for *Caenorhabditis elegans* and allows all parts of the database to be cross-referenced, resulting in a rich web of interconnections. The ACEDB user interface permits exploration of data via "point and click" with the workstation mouse or with a versatile query facility developed by LBL. LBL staff are collaborating with the original ACEDB authors, Richard Durbin (Medical Research Council, U.K.) and Jean Thierry-Mieg (CNRS, France), to continue its development. 21Bdb is one of several databases that LBL is helping to develop for different genomes, including those of *Drosophila melanogaster*, soybean, wheat, and forest trees.

The 21Bdb STS window displays columns of ordered physical map markers, including STSs from Genethon and LBL, YACs, P1s and cDNAs localized to the STSs, and radiation hybrid cell-line breakpoints. The map is displayed graphically with zoom capability, and users can "click" individual biological elements to bring up details on any object, such as STS primer sequence and product size and cDNA sequence. Authorized users can rearrange map markers by "drag and drop" or by respecifying position. Other ACEDB modules allow users to display and search DNA sequences for open reading frames, genes, and other features and compare multiple maps that share common markers.

Remote access to the full Unix version of 21Bdb is available to anyone on the Internet running X-windows software, including Unix workstations and Macintosh and IBM-PC computers. Remote users run 21Bdb at LBL with the graphic interface display on their own local computers. Those familiar with Unix and ACEDB may copy the entire 21Bdb database and ACEDB software. For more detailed information, use anonymous ftp to access *genome.lbl.gov* and see the files in */pub/21Bdb*, or send e-mail to *21Bdb@genome.lbl.gov*.◊



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

December

16. †Nicholas Dracopoli: Genetic Anal. of Melanoma; Bethesda, MD [NCHGR Lecture Series, E. Feingold, 301/496-7531, Fax: /480-2770]

January

15–22. *Gene Therapy; Silverthorne, CO (abs. deadline: Sept. 1) [Keystone Symposia, 303/262-1230, ext. 110, Fax: -1525]

15–22. *Mol. Biol. of Human Genetic Disease; Silverthorne, CO (abs. deadline: Sept. 1) [see Keystone contact above]

20. †Ellen Wright-Clayton: All in the Family—Special Problems Posed by Genetic Res. in Human Subjects; Bethesda, MD [see contact: Dec. 16]

21–28. *Tranposition and Site-Specific Recombination—Mechanism & Biol.; Park City, UT (abs. deadline: Sept. 22) [see Keystone contact: Jan. 15–22]

24–25. *NIH Natl. Adv. Council for Human Genome Res.; Bethesda, MD [J. Ades, 301/402-2205, Fax: -2218]

26–27. BioEast '94; Washington, DC [BioConferences Intl., 301/652-3072, Fax: -4951]

24–27. Plant Genome II; San Diego [Scherago Intl., 212/643-1750, Fax: -1758]

31–Feb. 5. Mol. Genetics of Tumor Progression and Metastasis; Big Sky, MT [AACR, 215/440-9300, Fax: -9313]

February

5–10. Advances in Gene Tech.: Mol. Biol. and Human Genetic Disease at the 1994 Miami Bio/Technol. Winter Symposia; Fort Lauderdale, FL [S. Black, 305/547-3597, Fax: /324-5665]

7–11. Intl. Conf. on Comparative Gene Mapping in Terrestrial and Aquatic Vertebrates; Oslo [H. Lewin, 217/333-5998, Fax: /244-5617 or Ø. Lie, +47-22/96-47-82, Fax: -86]

13–20. *Tumor Suppressor Genes; Taos, NM (abs. deadline: Oct. 27 [see Keystone contact: Jan. 15–22]

14–16. Mol. Biol. and Ecol. of Gene Transfer and Propagation Promoted by Plasmids; Madrid [CIMB, A. Gonzalez, +34-1/435-4240, Fax: /576-3420]

17. †Neil Risch: Susceptibility Gene Localization and Exclusion for Complex Diseases; Bethesda, MD [see contact: Dec. 16]

18–20. 2nd Intl. Workshop on Human Chromosome 15 Mapping; Oxford, UK [T. Donlon, 808/978-8350, Fax: -8053, Internet: *donlon@uhunix.uhcc.hawaii.edu*]

19–24. Cancer: Perturbations in Cell Cycle Control and Genomic Integrity; AACR, Banff, Alberta [see contact: Jan. 31–Feb. 5]

28. Mol. Bioinformatics IEE 94 Colloquium; London (paper deadline: Oct. 31) [S. Schulze-Kremer, +49-30/463-3040, Fax: /464-4097, Internet: *steffen@kristall.chemie.fu-berlin.de*]

28–Mar. 2. HGP: Commercial Implications; San Francisco [CHI, 617/487-7989, Fax: -7937]

March

3–4. Gene Transcription-Based Therapeutics; CHI, San Francisco [see contact: Feb. 28–Mar. 2]

4–10. *Mol. Basis of Cancer Therapy; Tamarron, CO (abs. deadline: Oct. 27) [see Keystone contact: Jan. 15–22]

5–11. Growth Factors, Development, and Cancer; AACR, Interlaken, Switzerland [see contact: Jan. 31–Feb. 5]

6–8. 3rd Intl. Chromosome 5 Workshop; Laguna Beach, CA [J. Wasmuth, 714/856-8242, Fax: /725-3403]

17. †Jeffrey Trent: Genomic Applications of Chromosome Microdissection; Bethesda, MD [see contact: Dec. 16]

April

7–8. 2nd Intl. *Nature Genetics* Conf.: Mouse Genetics, Transgenics, and Polygenics; Toronto [D. Berger, 212/477-9699, Fax: /505-1364]

9–11. 3rd Intl. Workshop on Human Chromosome 9; Cambridge, U.K. [S. Povey, +44-713/807-410, Fax: /873-496, Internet: *mpovey@mrc-crc.ac.uk*]

10–13. 85th Ann. Mtg. of AACR; San Francisco (abs. deadline: Oct. 25) [see contact: Jan. 31–Feb. 5]

21. †Ronald Davis: Technol. Development for High-Throughput DNA Sequencing; Bethesda, MD [see contact: Dec. 16]

23–26. 4th European Workshop on Cytogenetics and Mol. Genetics of Human Solid Tumors; Noordwijkerhout, The Netherlands (abs. deadline: Nov. 15) [J. van Dam, +31-20/566-4801, Fax: /696-3228]

24–28. 1st World Cong. on Computational Med., Public Health, and Biotech.; Austin, TX (abs. deadline: Dec. 31) [L. Bockoven, 512/471-2472, Fax: -2445, Internet: *compmed94@chpc.utexas.edu*]

25–27. Gene Therapy: New Technologies & Applications; CHI, Bethesda, MD [see contact: Feb. 28–Mar. 2]

26–28. *Natl. SBIR Conf.; Houston [Hotline, 407/791-0720, Fax: -0098]

27–29. Modulation of Signal Transduction & Gene Expression; CHI, Bethesda, MD [see contact: Feb. 28–Mar. 2]

Training Calendar**

January

3–7. Recombinant DNA Methodology; Washington, DC [CATCMB/CUA, M. Miller, 202/319-6161, Fax: -4467]

10. Intro. to PCR; Gainesville, FL [BTP, 800/821-4861, Fax: 603/659-4708]

10–14. †*Advanced Linkage Course; New York (application deadline: Nov. 24) [K. Montague, 212/960-2507, Fax: /568-2750]

10–14. PCR Techniques; Germantown, MD [LTI, 800/952-9166, Fax: 301/258-8212]

11–12. Quantitative RNA PCR; BTP, Gainesville, FL [see contact: Jan. 10]

13–14. Clin. Applications of PCR; BTP, Gainesville, FL [see contact: Jan. 10]

13–14. Intro. to Mol. Cytogenetics: Metaphase and Interphase Chromosome Anal.; Gaithersburg, MD [Oncor, Inc., 800/776-6267, Fax: 301/926-6129]

15–17. PCR Techniques; CUA/CATCMB, Washington, DC [see contact: Jan. 3–7]

17–22. cDNA Library Techniques; LTI, Germantown, MD [see contact: Jan. 10–14]

February

3–4. Intro. to Mol. Cytogenetics: Tissue In Situ Hybridization; Oncor, Inc., Gaithersburg, MD [see contact: Jan. 13–14]

➡

*Attendance at meetings listed with asterisk is either limited or restricted. Dates may change; check with contact person.
**Dates and course status may change, and courses may also be offered at other times and places; check with contact person.
†NCHGR-funded event.
‡DOE-funded event.

For Your Information

3–4. Intro. to Mol. Cytogenetics: Tissue in Situ Hybridization; Oncor, Inc., Gaithersburg, MD [see contact: Jan. 13–14]

3–4. Managing the Commercialization and Transfer of Technol.; Los Angeles [UCLA Short Course Program Office, 310/825-3344, Fax: /206-2815]

6–9. *Genomic Information: Ethical Implications; Seattle (application deadline: Nov. 30) [B. Brownfield, 206/543-5447, Fax: /685-7515]

14–15. GDB/OMIM Training Courses [see schedule, p. 11]

14–18. DNA-Protein Interactions; LTI, Germantown, MD [see contact: Jan. 10–14]

16–18. Advanced Mol. Cytogenetics: Probe Labeling; Oncor, Inc., Gaithersburg, MD [see contact: Jan. 13–14]

21–25. Recombinant DNA Techniques I; LTI, Germantown, MD [see contact: Jan. 10–14]

28–Mar. 4. PCR Methodology; Columbia, MD [Exon-Intron, 410/730-3984, Fax: -3983]

28–Mar. 5. Recombinant DNA Techniques II: Anal. of Gene Expression; LTI, Germantown, MD [see contact: Jan. 10–14]

March

7–11. In Situ Hybridization and rDNA Technology; Exon-Intron, Columbia, MD [see contact: Feb. 28–Mar. 4]

7–11. In Situ Hybridization Techniques; LTI, Germantown, MD [see contact: Jan. 10–14]

7–11. Recombinant DNA Methodology; CATCMB/CUA, Washington, DC [see contact: Jan. 3–7]

14–18. Receptor Binding Techniques; CATCMB/CUA, Washington, DC [see contact: Jan. 3–7]

20–25. Psychoneurogenetics; Ventura, CA [GRC, 401/783-4011, Fax: -7644]

21–25. Recombinant DNA Methdology; Exon-Intron, Columbia, MD [see contact: Feb. 28–Mar. 4]

April

18–22. Baculovirus Techniques; LTI, Germantown, MD [see contact: Jan. 10–14]

25–29. RNA Isolation & Characterization; Exon-Intron, Columbia, MD [see contact: Feb. 28–Mar. 4]

May

16–20. Recombinant DNA: Techniques & Applications; Rockville, MD [ATCC, 301/231-5566, Fax: /770-1805]

23–27. Biotech. for Business; Durham, NC [M. Pirrung, 919/660-1579, Fax: -1591]

24–27. PCR Applications/Cycle; ATCC, Rockville, MD [see contact: May 16–20]

June

12–15. †*Genomic Information: Ethical Implications; Seattle [see contact: Feb. 6–9]

20–25. Workshop for Secondary Sci. Teachers; Kansas City, KS [D. Collins, 913/588-6043, Fax: -3995]

23–24. Advanced Mol. Cytogenetics: Customized Res. Assistance; Oncor, Inc., Gaithersburg, MD [see contact: Jan. 13–14]

U.S. Genome Research Funding Guidelines

Note: Investigators wishing to apply for NIH and DOE funding are urged to discuss their projects with 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 and 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 under-represented 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 through

- Hard-copy subscription: call 301/496-7441.
- Electronic version (E-Guide): Access through one of the following methods.
  1. Institutional Hubs. A designee receives automatic updates and distributes them locally to researchers. To use this NIH-preferred method, send a message naming the responsible person to Rebecca Duvall (BITNET: *q2c@nihcu*, Internet: *q2c@cu.nih.gov*).
  2. NIH Grant Line (also known as DRGLINE). User reads electronic bulletin board for weekly updates. Connection is through a modem, and files can be transmitted rapidly via BITNET or Internet. For more information, contact John James (301/496-7554 or BITNET: *zns@nihcu*).

*Full text of RFAs listed in the NIH grants guide may be obtained from either of the two electronic sources or from NIH NCHGR in Bethesda, Maryland (301/496-0844).

DOE Human Genome Program

Solicitations for proposals are announced in the *Federal Register*, *Science*, and other publications. Proposals for FY 1995 will be due in summer 1994.

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

- 301/903-6488, Fax: -8521, or Internet: *#genome%er@mailgw.er.doe.gov* or *genome@oerv01.er.doe.gov*

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).
- Bettie Graham; Bldg. 38A, Rm. 610; NIH; 9000 Rockville Pike; Bethesda, MD 20892 (301/496-7531, Fax: /480-2770).

National SBIR conference: Houston, TX (April 26–28, 1994). Conference Hotline: 407/791-0720.◊

July

18–29. 35th Ann. Short Course in Medical & Experimental Mammalian Genetics; Bar Harbor, ME [Jackson Lab., 207/288-3371, ext. 1253]

August

8–12. Tissue Culture/Baculovirus Expression Systems; Exon-Intron, Columbia, MD [see contact: Feb. 28–Mar. 4]◊

Human Genome Management Information System Subscription/Document Request* (Vol. 5, 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. ___ *Understanding Our Genetic Inheritance, The U.S. Human Genome Project: The First Five Years, FY 1991–1995* (Joint DOE-NIH 5-Year Plan)
3. ___ DOE *Human Genome 1991–92 Program Report* (includes "Primer on Molecular Genetics") ___*Primer* as a separate document
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.

☆U.S. GOVERNMENT PRINTING OFFICE:1993-550-220/80002

CALENDAR ACRONYMS

AACR Am. Assoc. for Cancer Res.
ATCC Am. *Type Culture* Coll.
BTP Biotechnology Training Programs
CATCMB/CUA Ctr. for Advanced Training in Cell and Mol. Biology/Catholic Univ. of Am.
CHI Cambridge Healthtech Institute
CIMB Ctr. for Intl. Meetings on Biol.
ELSI Ethical, Legal, and Social Implications
GDB/OMIM Genome Data Base/Online *Mendelian Inheritance in Man*
GRC Gordon Res. Conferences
HGP Human Genome Project
IEE Institute of Electrical Engineers
LTI Life Technologies, Inc.
NCHGR Natl. Ctr. for Human Genome Res.
NIH Natl. Inst. of Health
SBIR Small Business Innovation Research
UCLA Univ. of Calif., Los Angeles

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

HGMIS MAILING ADDRESS
Betty K. Mansfield
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6050, 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