BIOLOGY PHYSICS CHEMISTRY ENGINEERING MATHEMATICS

Human Genome news

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

ISSN: 1050-6101

Vol. 6, No. 5, January–February 1995

DOE Holds Contractor-Grantee Workshop

LANL and LLNL Present High-Resolution Maps of Chromosomes 16 and 19

The fourth DOE Human Genome Program Contractor-Grantee Workshop was highlighted by the announcement that Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL) had completed the world's highest-resolution physical maps, which are for human chromosomes 16 and 19. Summaries of other significant progress and assessments of the program's future direction were also presented November 13–17, 1994, to over 400 researchers, program managers, and invited guests from other agencies and NIH GESTECs in Santa Fe, New Mexico.

More than 50 presentations covered topics on mapping; sequencing; informatics; and ethical, legal, and social issues (ELSI). Poster sessions provided details of about 200 projects from DOE human genome centers at Lawrence Berkeley Laboratory (LBL), LLNL, and LANL; other DOE-supported laboratories; and more than 40 universities and research organizations. Following a successful debut at the last workshop, a large informatics resource room was set up and maintained by LANL staff to demonstrate new resources and software capabilities. Some highlights of the plenary sessions follow.

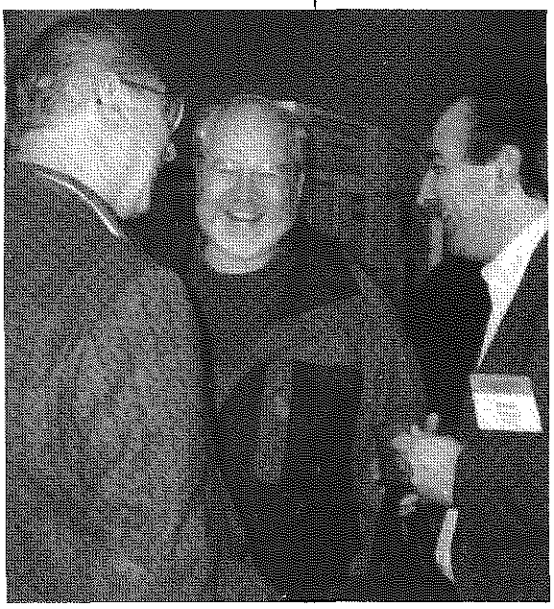
Program Overview

David Smith, Director of the DOE Health Effects and Life Sciences Research Division, observed that the Human Genome Project is ahead of schedule in many areas and increasingly international in scope. He cited achievements in mapping, resource development, instrumentation, and robotics, particularly noting the new maps and the completion of chromosome-specific libraries by the National Laboratory Gene Library Project (NLGLP) at LANL and LLNL.

Smith praised establishment in Santa Fe of the National Center for Genome Resources (NCGR), which now houses the Genome Sequence Data Base (GSDB), initially developed at LANL. The new center is poised to become a major distribution source of genome and related information for wider scientific and general use. An important goal is to maintain and improve transparent linkages between sequence data and the mapping data in the Genome Data Base (GDB).

Accomplishments of researchers and groups described in this article have been greatly facilitated by resources and data from investigators at numerous institutions. This cooperation demonstrates a remarkable sense of community among researchers worldwide.

Anthony Carrano (Lawrence Livermore National Laboratory), David Smith (DOE Office of Health and Environmental Research), and Radomir Crkvenjakov (Hyseq, Inc.) discuss research progress at a poster session of the DOE Contractor-Grantee Workshop.



In This Issue

Genome News

Online Text Available

This text and full abstracts from the Santa Fe meeting are available on the DOE Human Genome Program WWW site:

http://www.er.doe.gov/production/oher/hug_top.html

The material can also be accessed through the HGMIS Home Page:

http://www.ornl.gov/TechResources/Human_Genome/home.html

Strategies for placing genes on chromosomal maps are being discussed intensively within the worldwide research community. Such annotation will be important for understanding human biology as well as for eventual commercial applications. While noting that DOE will continue to play an active role, Smith also mentioned the possibility of substantial funding from the private sector.

Technologies arising from the Human Genome Project are enabling others outside the project to make progress in research and development. For example, DOE-supported basic research and technologies are being modified and developed by several companies to make more-rapid diagnostics that may be useful for clinical, agricultural, and environmental applications. Also, investigators have begun sequencing DNA from microbes of industrial, environmental, and phylogenetic interest in the DOE Microbial Genome initiative [see *HGN* 6(3), 7 (September 1994).]

Smith reminded attendees that, although the project is on time and under budget, the greatest challenge—sequencing the human genome—still lies ahead. DOE and NIH genome program funding is leveling out at $70 million and $110 million, respectively, and the project faces critical challenges in large-scale sequencing. In the next 2 to 3 years, he expects sequencing technology to increase in speed and cost-effectiveness and believes that DOE must shift major resources into high-throughput DNA sequencing. An important milestone in automated large-scale sequencing was reached in 1994 when LBL produced 2 Mb of sequence using a directed strategy.

Smith concluded that the time is right for blending genome technologies into other DOE activities, including the health effects program. This program provides basic understanding of risks associated with exposure to radiation or other environmental agents—part of the original mission that fostered the agency's interest in human genome analysis. DOE will hold several workshops this year to consider how other biology programs and the genome project can share and exploit tools, technologies, and resources developed at universities and the national laboratories.

Mapping Sessions

Physical Mapping

Chromosome 16. Norman Doggett (LANL) summarized the 5-year chromosome 16 mapping effort that resulted in an integrated physical-genetic-cytogenetic map. A high-resolution (1-Mb) cytogenetic breakpoint map provided the framework for constructing all levels of the integrated map. The physical map consists of both a low-resolution (0.25-Mb) YAC contig map and a high-resolution (<0.01-Mb), "sequence ready" cosmid contig map. STSs anchor the cosmid map to the YAC and cytogenetic maps, and highly informative microsatellite-based genetic markers are tightly integrated. All available markers and cloned genes have been positioned on the map. (More map details can be found in the sidebar below.) The integrated map facilitates disease-gene and fragile-site cloning on the chromosome; genes associated with Batten's Disease and breast and prostate cancers are among those on chromosome 16.

High-Resolution Physical Maps of Chromosomes 16 and 19 Completed

Chromosome 16.

The LANL integrated chromosome 16 map announced at the Santa Fe workshop is composed of a cytogenetic breakpoint map, a low-resolution physical map in mega-YACs, and a high-resolution physical map in cosmids and mini-YACs. Over 600 genes, DNA markers, and microsatellite repeats generated by investigators worldwide have been integrated into the new map. The magnitude of this accomplishment is underscored by comparisons with physical maps already available: chromosome 16 is twice the size of chromosome 21 and 3 times the size of the Y euchromatic region, and resolution is over 100 times greater than previously achieved for those chromosomes.

Cytogenetic map. The backbone of the new map is a cytogenetic breakpoint map developed with DOE funding by David Callen's group at Adelaide Women and Children's Hospital (AWCH), Australia. The breakpoint map is derived from 78 mouse-human somatic cell hybrids representing 90 independently ascertained chromosome breakpoints. With the 4 fragile sites, these divide the 85 Mb of euchromatin into 91 intervals (average resolution about 1 Mb).

YAC map. At the second level is a low-resolution YAC contig map composed of about 600 CEPH mega-YACs and 220 flow-sorted chromosome 16–specific YACs that are smaller but of higher quality, with a very low chimeric rate and deletion frequency; these YACs are localized to and ordered within the breakpoint intervals with 350 STSs; average interval between markers is about 260 kb. The chromosome 16 mega-YAC map provides clonal coverage of 98% of the 90-Mb euchromatin regions, with 2 gaps on the p arm and 4 on the q arm.

Cosmid map. A high-resolution "sequence-ready" cosmid contig map comprises the next resolution level (average resolution, 5 kb) and consists of 4000 fingerprinted cosmids assembled into contigs covering 60% of the chromosome. The cosmids are anchored to the YAC and cytogenetic breakpoint maps via STSs developed from cosmid contigs and by hybridizations between YACs and cosmids. About 2000 cosmids in 320 contigs have been localized to the cytogenetic map. The group has 90% coverage for the chromosome in fingerprinted cosmids, but not all have been placed on the integrated map.

Genetic map integration. Correlated with the physical map is a highly informative microsatellite-based genetic map developed at AWCH; it consists of 78 PCR-typable markers at 2.6-cM median intervals (3.2 average). Almost 200 genetic markers provide a resolution of <1 cM on the genetic map.

In 1988 the LANL group chose to map chromosome 16 for several reasons. Among them is the chromosome's 98-Mb size, which is compatible with a cosmid-fingerprinting mapping approach such as that used for the 100-Mb *Caenorhabditis elegans* genome. Resources critical to the project include a monochromosomal hybrid containing 16 as a single human chromosome; flow-cytometry instrumentation for constructing specific cosmid and YAC libraries; and robotics for making high-density gridded arrays from the libraries, pooling YAC libraries, screening by PCR, and setting up reactions in other mapping steps. Other essential resources are a mapping panel containing 78 hybrid cell lines, 90 independently ascertained breakpoints on the chromosome, and the Généthon YAC map. LANL collaborated closely with AWCH investigators in integrating the physical, genetic, and cytogenetic maps.

Genome News

The LANL group is now identifying DNA expressed sequences and integrating them into the map. Mike Altherr explained their choice of an exon-amplification strategy to identify the sequences, which are then mapped to specific chromosomal locations using a panel of somatic cell hybrids and clones. About 1800 exon clones have been generated, and over 800 clones have already been sequenced. Almost 700 sequences have been subjected to database analysis with cDNA INFORM to identify previously characterized genes or conserved motifs that may provide insight into their biological function. This gene annotation map is expected to facilitate identification and isolation of disease genes and functional analysis of genes for which no biological function is known.

Chromosome 5. Deborah Grady (LANL) described progress in constructing a low-resolution physical map of chromosome 5. A YAC contig centers on regions associated with Cri-du-Chat syndrome; with a frequency of 1 in 45,000 births, this is the most common human terminal-deletion disorder. Common clinical features include mental retardation and a characteristic high-pitched, cat-like cry. A complete, nonchimeric YAC contig of 5p15.2 has been identified and characterized; deletion of this region (about 2 Mb) correlates with all clinical features except the cry, which maps to 5p15.3. A YAC contig of this latter region has been constructed and is being characterized. [*Genomics* **24**(1), 63–68 (November 1994).]

Chromosome 19. Linda Ashworth (LLNL) described the 4-year strategy for constructing the high-resolution metric map, which currently consists of 63 islands of known order and distance that span 90% (45 Mb) of the noncentromeric portion of chromosome 19. The map provides clonal continuity in YACs, BACs, PACs, and cosmids. Brigitte Brandriff discussed the map foundation, which consists of 802 fluorescently fingerprinted cosmid contigs. These contigs were anchored to the map by FISH, allowing determination of order and distance and providing a "to-scale" framework (metric) map. Over 75% (>38 Mb) of the chromosome has already been defined to the level of *Eco*R I restriction sites. The group has also assigned genes, cDNAs, STSs, polymorphic markers, and ESTs to the map. Overall coverage is about 1 marker per 147 kb, with the goal of 1 marker per 100 kb (see sidebar below for more map details).

Future LLNL plans outlined by Harvey Mohrenweiser include (1) in collaboration with Oak Ridge National Laboratory (ORNL), construction of a complete high-resolution transcript map of chromosome 19 and genomic regions of special interest in both human and mouse; (2) continued development of high-throughput sequencing methods and technologies and sequencing of selected human and mouse genomic regions; and (3) application of resources and techniques to relevant issues in disease susceptibility, biological structure and function, and environmental sciences.

Tom Slezak noted challenges in providing informatics tools to support mapping efforts at LLNL, including automation of as much map construction and integration as possible and construction of flexible tools to handle multiple viewpoints and allow scaleup. He emphasized the importance of active support and participation by

See p. 14 for a report on the Chromosome 22 map.

Many other LANL physical-mapping resources provide powerful tools for using mapping data, including a mega-YAC contig arrayed on gridded membranes, with YACs ordered as they occur on the chromosome; two flow-sorted chromosome-specific *Cla* I and *Sac* II YAC libraries arrayed on gridded membranes and pooled for PCR-based screening; 10x-coverage cosmid libraries arrayed on gridded membranes and 90% coverage in *Eco*R I and *Hind* III restriction-mapped clones; and a high-resolution somatic hybrid cell mapping panel that provides a standard for localizing new clones to the map.

All chromosome 16 mapping data are available worldwide on Internet (*http://www-ls.lanl.gov*) or through Norman Doggett (505/665-3024, *doggett@gnome.lanl.gov*).

Chromosome 19

The LLNL high-resolution chromosome 19 cosmid-based physical map spans 90% of the euchromatin in 63 contiguous islands consisting of BACs, PACs, YACs, and cosmids where size, order, and distance are known. Only 50 Mb of the 60-Mb chromosome contain unique sequence, with the remainder consisting of centromeric and telomeric repeats.

Metric map construction. The chromosome 19–specific cosmids were clustered into contigs using automated restriction-fragment fingerprinting analysis. At the next mapping level, contigs were ordered relative to each other, and physical distances between certain clones within contigs were measured by FISH to metaphase chromosomes (resolution, 400 to 500 kb), somatic interphases, and sperm pronuclear interphases (resolution, 50 kb). LLNL's goal was to obtain sequentially ordered contigs spanning telomere to telomere, with known distances separating cosmid pairs and the resulting "metric" map serving as a backbone for further map building and other efforts. The average distance separating pairs of cosmids is 250 kb for 80 cosmids from the p arm (range, 50 to 700 kb) and 270 kb for 126 cosmids on the q arm (range, 50 to 840 kb). [*Genomics* **23**(3), 582–91 (October 1994).]

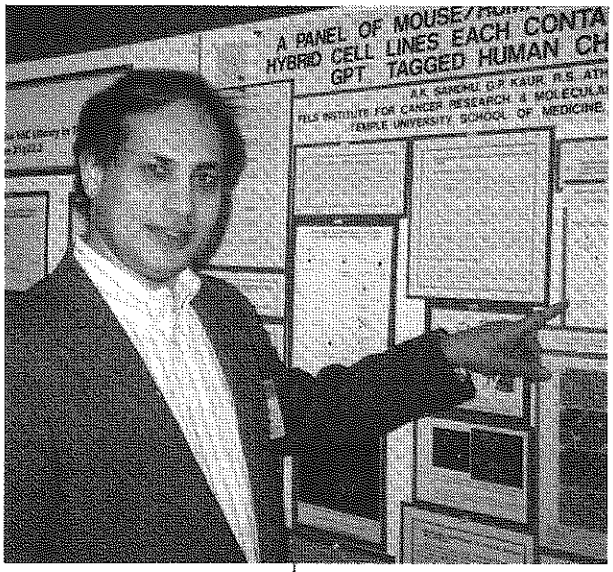
Closing gaps. Gaps between ordered cosmid contigs were spanned using the larger insert libraries with a variety of hybridization and STS technologies; simple cosmid walking achieved continuity at the cosmid level. Over 38 Mb of *Eco*R I restriction maps were used to close gaps in some regions. Remaining gaps will be closed by those with a particular interest in a region.

Integrating genes and genetic markers. Significant strides have been made toward integrating genetic and physical maps. Over 400 genes and genetic markers have been localized on cosmids, and almost 300 have been incorporated into the metric map. More than 340 STSs, ESTs, and clone markers have been mapped and ordered.

Chromosome 19 may be far richer in genes than the previous estimates of about 2000. Researchers are now finding a gene every 25 kb, and at least 20 to 25 gene families with a minimum of 3 to 5 members are represented on the chromosome. These biologically interesting genes will be targeted for study as the high-resolution *Eco*R I map is being completed. Genomic sequencing is now being done in selected regions of interest, especially on three DNA repair genes. Software has been designed to integrate and display cosmid, YAC, FISH, and restriction maps, as well as sequence, hybridization, and screening data.

Detailed information on the chromosome 19 map is available through WWW (*http://www-bio.llnl.gov/bbrp/genome.html*) and from Anthony Carrano (510/422-5698, *carrano@llnl.gov*).◊

Genome News



Raghbir Athwal (Temple University) presents his work on mouse-human cell lines containing purified human chromosomes.

biologists in informatics design. Future challenges include building informatics tools to support sequencing, expanding the system to query and navigate across multiple chromosomes and species, actively participating in the federation of genomic databases, and expanding external collaborations.

Lisa Stubbs and Richard Woychik (ORNL) discussed the mouse as a model system for predicting and identifying genes in humans and for studying gene function. Stubbs reported on collaborative efforts with LLNL that have revealed striking similarities between sections of mouse chromosome 7 and human chromosome 19. Woychik stressed the importance of developing a high-throughput transgenic and targeted mutagenesis strategy in mice to accommodate the future study of health effects caused by genes identified through human genome mapping.

Chromosome X. David L. Nelson (Baylor College of Medicine) discussed use of LLNL flow-sorted cosmids to increase resolution of the X chromosome short-arm YAC map; he believes these cosmids will become common currency for exchanging X chromosome mapping information. His group is using WWW-based browsers for entering, annotating, and correcting information. Nelson stressed the need to develop joint databases that provide researchers worldwide with access to information on each flow-sorted chromosome library (*http://gc.bcm.tmc.edu:8088*).

Chromosome 11. Glen Evans (University of Texas Southwestern Medical Center at Dallas) reported that YAC contigs spanning chromosome 11 are nearly complete. His laboratory has begun constructing higher-resolution maps based on a sampled-sequencing approach to ordered cosmids that generates maps displaying likely locations of gene hits, STSs, polymorphisms, and other sequence features. [Genome center and chromosome 11 information (*http://mcdermott.swmed.edu*).]

Resources and Technology Development

Several speakers discussed ongoing work in constructing genomic reagents and technologies critical for mapping, gene finding, sequencing, expression studies, and molecular cytogenetics.

NLGLP. Larry Deaven (LANL) and Jeff Gingrich (LLNL) announced the successful completion of NLGLP Phase II; libraries are now available for each chromosome, with 40-kb inserts in cosmids and 10-kb inserts in lambda vectors. Lambda clones are distributed through the American *Type Culture* Collection in Rockville, Maryland. Cosmid arrays, which have been available on a limited basis, soon will be released nonexclusively to industry to establish screening and clone-distribution services for the entire scientific community. Keeping up with demand for library copies has been difficult, Deaven noted, and the national laboratories are now seeking to share distribution with private-sector companies. LANL also completed YAC libraries with an average insert size of about 200 kb for some chromosomes, and both laboratories are exploring use of PAC and BAC vectors. [Library information (*http://www-bio.llnl.gov/bbrp/genome.html*).]

BACs, PACs, HAECs. Melvin Simon (California Institute of Technology) and Pieter de Jong (Roswell Park Cancer Institute) discussed the construction of libraries in BAC and PAC vectors, respectively, and their usefulness in providing large-fragment, stable genomic resources for mapping and sequencing. A mouse BAC library is available from Research Genetics that consists of about 145,000 clones with average insert size of 135 kb for 4× coverage and a human BAC library that provides more than 6× coverage.

BACs have been used to form contigs covering a number of regions of interest and to fill holes in YAC maps. Extensive contigs have been formed over multimegabase regions on chromosome 22; eventually, these will provide excellent substrates for sequencing chromosomes. Hundreds of cDNA, *Alu*-PCR, and STS probes have been used to interrogate the BAC library by hybridization and pooled PCR. All of them have successfully generated appropriate BAC contigs that can be used in gene finding and chromosome mapping.

Simon encouraged researchers to send him problem chromosome segments. Copies of a human PAC library with average insert size of 120 kb and 3× coverage have been distributed to more than 30 genome centers in North America and Europe and to commercial screening companies, including Genome Systems. A stable PAC contig has been prepared for the spinal muscular atrophy region.

Jean-Michel Vos (University of North Carolina, Chapel Hill) described a novel system to clone DNA in human cells as human artificial episomal chromosomes (HAECs) with average insert size 150 to 200 kb for mapping and sequencing and functional studies of large mammalian DNA regions. He presented data on a recently published HAEC-based human genomic library.

ATP Program Accelerating Genome Research Payoffs

The Tools for DNA Diagnostics component of the Advanced Technology Program (ATP) is accelerating payoffs from genome research, and several companies funded by ATP have built on DOE-sponsored research. Stanley Abramowitz (National Institute of Standards and Technology) explained that ATP's mission is to stimulate U.S. economic growth and industrial competitiveness by encouraging the development of high-risk but powerful new technologies. David Smith urged researchers with promising technologies to seek industrial partners, perhaps through the ATP program. [ATP Contact: Abramowitz (301/975-2587, Fax: /869-1150, *stan@micf.nist.gov*).] ◊

Probe hybridization. Joe Gray [University of California, San Francisco (UCSF) and LBL] reported on the LBL-UCSF Resource for Molecular Cytogenetics, which was created to develop and make available technology and reagents for molecular cytogenetic characterization of genetic diseases. Technologies include comparative genomic hybridization, FISH, and digital-imaging microscopy. These are being applied to detect and characterize DNA sequence-copy-member aberrations in human malignancies and other genetic conditions. The group is also developing a set of reagents from chromosome-specific YAC, cosmid, and P1 libraries for use as FISH probes. The goal is to develop probes distributed at about 5-Mb intervals over the entire human genome that contain STSs defining genes or genetically mapped polymorphic loci. Information about technologies, probes, and software is available at *http://rmc-www.lbl.gov* and will be accessible through GDB when appropriate.

Promising technologies. David Barker (Molecular Dynamics) spoke about his group's capillary-array electrophoresis instrumentation for high-throughput DNA analysis using confocal fluorescence detection. Fragments are labeled with thiazole orange and run in eight groups of six capillaries with a replaceable hydroxyethylcellulose matrix. Total run time from sample injection to fragment resolution (up to 2 kb) is 20 to 30 min, with single-bp resolution of double-stranded restriction fragments, PCR products over 100 to 400 bp, and a total size range of 50 bp to 12 kb. Applications include mapping, sequencing, and disease diagnostics.

Another promising technique for sizing DNA fragments is single-molecule detection of a fluorescently labeled sample in a flowing stream intersected by a laser beam. Jeffrey Petty (LANL) described early results with this method for sizing larger (2- to 150-kb) DNA fragments more rapidly and accurately than is now possible using gel electrophoresis. (Fragment sizes are inferred from the fluorescence intensity of samples stained with thiazole orange.) Petty also reported progress in detecting signals from individual nucleotides; the group is working to extend the technique for rapidly sequencing 40-kb or larger DNA fragments at a rate of 100 to 1000 bases/s. This technology is being developed under the first (1991) Cooperative and Research Development Agreement (CRADA) between a DOE human genome center and a private company. John Harding leads the research group at GIBCO BRL.

Robert Lipshutz (Affymetrix) discussed applications of high-density oligonucleotide arrays (DNA probe arrays) to screen and physically map cosmid clones into contigs and to detect mutations for clinical diagnosis. For mapping applications, a fluorescently labeled target from a clone (a four-base sequence adjacent to a restriction site) is hybridized to a chip containing spatially arranged probes; this generates a pattern corresponding to a unique marker subset in the clone. Overlapping clones are detected by correlating hybridization patterns. Scanner instrumentation for commercial use is being developed by Molecular Dynamics and Hewlett Packard.

Maynard Olson [University of Washington, Seattle (UWS)] noted the importance of automating production of high-accuracy and -resolution restriction maps as a front end for genomic sequencing. He emphasized the need for substantially eliminating stepwise human decision making to achieve major throughput increases and meet Human Genome Project goals.

Mapping Informatics

Ken Fasman (GDB) discussed the federated approach being used to guide improvements in the next release of GDB (V6.0). A new focus is an improved representation of physical and genetic maps in the Genome Data Base; the database core will tightly couple maps, map objects, and mapping reagents. The new version allows direct user submissions, updates, and third-party data annotations and will be organized on a revised editorial model that differentiates between original and consensus data. GDB V6.0 is implementing an object-oriented data model on top of a Sybase relational database using the Object Protocol Model tools developed by Victor Markowitz and his group at LBL (*http://gizmo.lbl.gov/opm.html* or *ftp://gizmo.lbl. gov/pub/DM_TOOLS/DMTools.html*).

V6.0 will feature increased modularity of both applications and databases, replacing the current monolithic structure with a "minifederation" of separate databases. The key component is object-broker technology, which enables interconnections to databases such as OMIM and eventually will extend across GDB, GSDB, and other public and genome center databases.

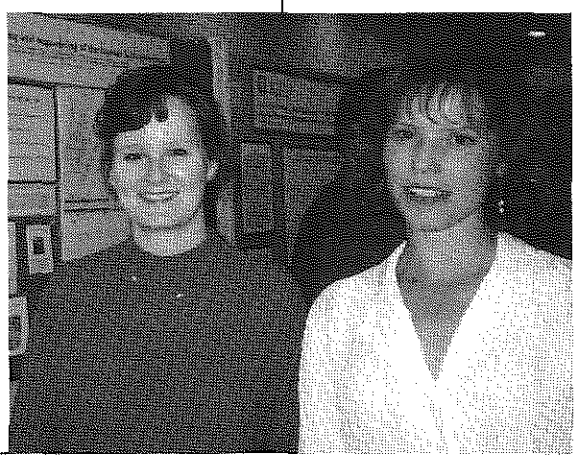
Fasman noted the increasing importance of WWW for federated modular databases, observing that GDB has developed an extensive WWW interface. He strongly encouraged researchers to explore the Web and obtain Mosaic or other clients for accessing GDB data.

A. Jamie Cuticchia (GDB) explained GDB data acquisition and curation. GDB acquires data by paper and electronic submissions and

Need Information?

- Have questions about an article?
- Need contact information?
- Want to subscribe or view newsletter on the Web?

Contact HGMIS
615/576-6669, Fax: -9888
http://www.ornl.gov/TechResources/Human_Genome/home.html



Julia Parrish (Baylor College of Medicine, 1992 DOE Human Genome Program Distinguished Postdoctoral Fellow) and Lisa Stubbs (Oak Ridge National Laboratory) exhibit their respective posters during a poster session.

Genome News

literature scanning; scanning is being phased out as GDB moves toward more community submission and editing. Cuticchia emphasized the responsibility of researchers to submit their data promptly and of databases to facilitate submission preparation. Data is curated by Human Genome Organisation chromosome-mapping editors, who approve and produce consensus marker localizations and maps for each chromosome (GDB V6.0 will display both raw and consensus data). [GDB (*http://gdbwww.gdb.org*).]

Manfred Zorn (LBL) described SubmitData, a tool that allows researchers to create valid database submissions by merging their own data with a defined template. SubmitData is now available for GSDB submissions, and a GDB version will be ready soon. Future improvements will add a more-complex set of variables, a WWW server to make Mosaic a user interface, and integration with data-analysis tools such as GRAIL.

ELSI

DNA and the courts. Genetic advances pose enormous challenges to state and federal courts, where judges are struggling to understand and assess new information. Franklin Zweig (Einstein Institute for Science, Health, and the Courts) discussed some of these issues, noting that while the impact on the judicial process has been dramatized recently by DNA evidence in the O.J. Simpson case, future impact will be far more dramatic as courts confront noncriminal topics. These issues may be as large as proposals to enact eugenic statutes and as small as a laboratory's failure to diagnose a catastrophic disease in a prenatal test. Zweig's group is developing a reference text with companion videotape and CD ROM to help judges understand genetic evidence [see *HGN* 5(6), 1–3 (March 1994)]. He observed, however, that more-complex cases will require the help of neutral, court-appointed expert witnesses. Zweig is compiling a roster of scientists to serve as expert witnesses and encouraged workshop attendees to add their names.

Hispanic educational program. Margaret Jefferson (California State University, Los Angeles) and Mary Ann Sesma (Los Angeles Unified School District) are developing a culturally and linguistically appropriate curriculum on genetics for Hispanic students and their families. Materials include a Spanish translation of the Biological Sciences Curriculum Study module and supplemental references on ELSI and the genetics of New World Hispanics.

Privacy studies. Alan Westin (Center for Social and Legal Research) is analyzing the effects of new genetic technologies on individuals and institutions. He observed that society is now past the issue-identification stage and

LBL Completes 2 Mb of DNA Sequence with Directed Strategy

By early December 1994, the LBL Human Genome Center had completed a total of 2 Mb of genomic sequence from human and *Drosophila* DNA using a directed strategy. Acting Center Director Mohandas Narla observed that this accomplishment, an important milestone toward cost-effective and accurate large-scale DNA sequencing, validates the DOE strategy of establishing genome facilities in national laboratories where interdisciplinary scientific teams successfully attack large problems.

LBL's directed DNA-sequencing strategy, developed by Michael Palazzolo and Christopher Martin, produces high-resolution physical maps of P1 clones with a small set of standard primer-binding sites every 300 bp. A mapped (ordered) minimum template set is then sequenced. Compared with random strategies, directed sequencing reduces template preparation by tenfold and sequencing by fivefold. Most important, sequence assembly is straightforward because all templates are mapped relative to each other before sequencing. The approach allows investigators to organize the work into six separate modules: physical mapping, high-resolution physical mapping, transposon insertion, sequencing, data assembly, and editing.

Automation. Joseph Jaklevic leads the LBL instrumentation group in automating the new sequencing strategy as an assembly line of individual steps. These steps include clone picking and arraying, robotic manipulation of libraries, preparation of PCR and sequencing reactions, custom oligonucleotide synthesis of PCR primers, thermal cycling for PCR-amplification reactions, large-scale gel loading and running, image processing, and plasmid preparation.

The instrumentation group has developed modules that are already everyday tools of the sequencing-production group. An automated system captures and analyzes information from agarose gels and automates data acquisition, interpretation, and decision making (e.g., picking valid clones to build a minimum subset), thus reducing fragment-sizing time from 8 h to between 4 and 8 min. A robotics-compatible, multistation thermal cycler is now performing 600 PCR amplifications/h in a 3-plate, 96-well format; a larger version is under development. An automated 12-channel oligosynthesizer (expandable to a 96-well format) produces 12 custom oligos in 2.5 h with minimal reagent use at a significantly lower cost than commercial systems.

The group also has a semiautomated system for running gels and is developing instrumentation to simplify plasmid preparation. Other instrumentation includes a colony picker reconfigured to pick smaller libraries and a robotic library-management system. Current modules are being integrated into second-generation systems.

Informatics. The directed strategy also has the advantage of facilitating automated data handling. The LBL informatics group, headed by Frank Eeckman, aims to collect and assemble all information from different work modules without human intervention. To accomplish this goal, the group has developed a series of computer programs that will speed construction of high-resolution physical maps, move files among different systems, and provide a visual interface for template selection. Software is being developed to support manipulation and analysis of long sequence tracts. The database Syndb is replacing 21Bdb as the primary mechanism for sharing physical-mapping and sequencing data, and annotated sequences are being submitted to GSDB.

Template-preparation and function studies. Edward Rubin and Jan-Fang Cheng head the human biology group, which develops DNA-sequencing templates from human regions selected for their probability of containing medically important genes. A P1 contig covering 1.2 Mb of human 5q31 and containing the entire interleukin gene cluster is now the target of directed sequencing, and genetically engineered mice are being used to study the function of sequenced regions. Coupling sequence analysis with biological studies can dramatically increase knowledge of the biological and medical significance of chromosomal regions.

For more information, see *http://www-hgc.lbl.gov/GenomeHome.html* or contact Narla (510/486-7029, Fax: -6746).◊

noted that ELSI studies give issues a high visibility that has helped stave off for the past 3 to 4 years the tendency to pass laws before issues are understood.

Westin suggested empirical studies to analyze early uses of genetic discoveries in sectors where social issues are sharply defined. Other study topics include institutional policies and standards, duties and responsibilities of professionals regarding confidentiality and disclosure, and value questions involving privacy.

Sequencing Sessions

Sequence Patenting and Technology Transfer. Rebecca Eisenberg (University of Michigan Law School) focused on technology transfer, proprietary rights to DNA sequences, and access to sequence databases. She pointed out that the U.S. government is uniquely situated to enrich the public domain and that, by restricting access to valuable discoveries, the current federal pro-patent policy may sometimes undermine rather than support incentives to develop new products. Eisenberg also observed that companies may not need to protect a potentially lucrative product by obtaining patents for every step of development.

Large-Scale Sequencing

Human T-cell receptor (TCR). Leroy Hood (UWS) reported analyzing more than 685 kb of contiguous sequence from the human TCR beta locus; this amount represents the longest stretch of human sequence yet obtained. The power of large-scale genomic sequencing and evolutionary comparisons, he noted, lies in the ability to reveal crucial genomic features undetectable by cDNA analysis alone; this approach will also generate reagents needed for access to all genomic regions. The complexity of mapping the human TCR beta locus to obtain sequencing substrates emphasizes the importance of generating rigorous sequencing technologies that can handle all problematic genomic regions. [See *HGN* 6(2), 1–2 (July 1994).]

Hood described his group's efforts to train high school teachers to lead their students in the shotgun sequencing of small DNA fragments encoding genes. He asked investigators to send appropriate DNA fragments if they would like to become involved with the project.

Caenorhabditis elegans. Richard Wilson [Washington University (WU) School of Medicine Genome Sequencing Center, St. Louis], who is funded by NIH, was invited to speak on his laboratory's collaboration with the Sanger Centre (U.K.) to sequence the 100-Mb *C. elegans* genome. Over 9 Mb of sequence from the chromosome III central region are finished, and sequencing of chromosomes II and X is in progress. Project completion is expected in 1998. The group's sequencing strategy includes shotgunning M13 or phagemid subclones from cosmids, followed by a directed-walking finishing phase. Sequence analysis is revealing a gene about every 5.1 kb, and 42% of predicted genes have homologies to previously identified genes in databases.

Production sequencing of human and *Drosophila.* The successful application of a directed-sequencing strategy was described by Michael Palazzolo (LBL). The LBL production group completed 2 Mb of sequence, representing a 23-fold throughput increase in just a year and a half. The group has achieved a sequencing rate of about 120 kb/month/person and is scaling up to 170 kb/month; this would translate to over 1 Mb/month for a six-person team (see sidebar, p. 6).

At the National Center for Genome Resources open house, Gifford Keen shows Margaret Jefferson (California State University, Los Angeles) the mainframe computer that manipulates all data queries coming into Genome Sequence Data Base.

Palazzolo praised collaborations among LBL, the *Drosophila* genome center, and University of California, Berkeley (UCB), as well as the cooperative support of DOE and NIH. DOE supports organism-independent technology development for directed and human genomic sequencing, while NIH supports *Drosophila* sequencing. The 120-Mb euchromatic *Drosophila* genome is the first to be physically mapped with single-copy-vector large clones (PACs), which have proven stable and nonchimeric. The project has generated about 1600 STSs for an average of 1 per 55 kb and has assigned to contigs over 61% of the 6000 P1 clones for an estimated 85% coverage of the *Drosophila* genome.

Joseph Jaklevic reported recent progress in automating scaleup of the LBL directed-sequencing strategy, which uses gel-based PCR assays extensively for robust contig mapping. Jaklevic described work on the group's 12-channel oligosynthesizer and 3- and 4-channel thermal cyclers. The thermal cyclers are being combined with Biomek and ORCA robots to reduce material handling and tracking.

Sequencing Technologies

One of the genome project's major challenges is the need for increased automation in DNA-sequencing technologies to increase speed and reduce costs. Standard methods based on gel-electrophoresis separation of nested fragment sets are considered too slow and expensive. Progress in automation ranges from enhancements of conventional gel-based technologies to novel, gel-less, automatable approaches. Some topics represented at the meeting are outlined below.

Clone processing. Richard Guilfoyle [University of Wisconsin, Madison (UWM)] discussed improvements in front-end strategies and progress toward automation. Triple-helix-affinity

Genome News

capture techniques facilitate (1) purification (95%) of cosmid inserts (reducing 20-fold the amount of sequencing-vector DNA) and (2) *CviJ* 1 digestion for random fragmentation and cloning into M13-100, a direct-selection vector. The group is exploring a promising flow-cytometric sorting approach for isolating M13 clones that could deliver 7000 clones/h and is optimizing procedures for ordering M13 clones and selecting minimally overlapping inserts. This will drastically reduce redundancy in the shotgun approach and the postsequencing fragment-assembly process. The UWM group is also quantitating M13 sequencing-template concentrations to identify unsequenceable clones and normalize base calling.

Sequencing chemistries. Current widely used procedures for automated DNA sequencing all involve the use of substrates synthesized by DNA polymerases and terminated by a dideoxynucleotide analog. Stanley Tabor (Harvard Medical School) identified a site in many DNA polymerases that can be modified to incorporate these analogs more efficiently. This is critical for obtaining bands of uniform intensity in the gel electrophoresis step, increasing sequencing sensitivity and accuracy. The presence or absence of a single hydroxyl group (tyrosine vs phenylalanine) at a highly conserved position on *E. coli*, T7, and *Taq* polymerases makes more than a 1000-fold difference in their ability to discriminate against these analogs. Another advantage of these modified DNA polymerases is their requirement for much lower amounts of the fluorescent dideoxynucleotide analogs used in automated DNA sequencing; this reduces the cost of reagents and also the fluorescent background caused by unincorporated analogs.

Barbara Ramsay Shaw (Duke University) described a method that allows direct sequencing of PCR products, bypassing the need for cycle sequencing. The method is based on efficient and stable incorporation into DNA of a new class of boronated triphosphates that permit exponential amplification, unlike chain terminators. Sequence is revealed by simple exonuclease digestion. The method should be completely automatable and requires much less DNA template.

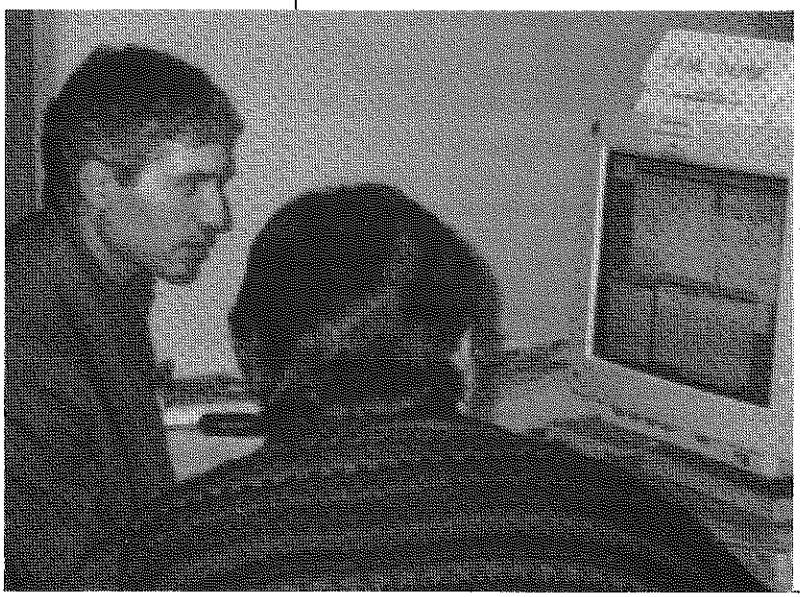
Capillary electrophoresis (CE), ultrathin gels, new gel matrices, enhanced detection. Application of higher electric fields in electrophoretic separation of DNA fragments can increase sequencing speed and efficiency. Conventional slab gels cannot dissipate adequately the additional heat produced, but capillary arrays or ultrathin (50- to 100 μm) slab gels can be used for more efficient heat transfer. Progress on these systems is reported below.

A bottleneck in gel-based systems is the manual gel-preparation step, which is both time consuming and a potential source of variability in DNA sequencing. Barry Karger (Northeastern University) discussed a fully automated, closed-end, high-throughput CE system to minimize human intervention, with a noncross-linked polymer matrix that is replaced after each run. The group is also attempting to incorporate a library-based primer-walking system. Karger observed that mutation detection will be another important use for capillary technology.

Norman Dovichi (University of Alberta) noted progress in constructing a CE sequencing system that requires a single laser for simultaneously exciting CE fluorescence signals from many capillaries. As they exit capillaries, DNA fragments with fluor labels are smoothly entrained in an optically clean fluid sheath flow in which as few as 120 fluorescein-labeled molecules can be detected. The capillary ends are physically staggered to distinguish fluorescence signals of the exiting DNA fragments. A 32-capillary system has been demonstrated, and Dovichi projected expansion to 864 within a sheath flow cuvette.

Richard Mathies (UCB) spoke of recent efforts to develop and combine improved high-sensitivity fluorescent reagents and new instrumentation for capillary array electrophoresis (CAE) coupled with confocal fluorescence detection. His group has developed sequencing primers labeled with pairs of dyes coupled by fluorescence energy transfer and has obtained dye signals 2 to 6 times more intense than for single-dye-labeled primers, thus decreasing the amount of template DNA needed. Another project involves fabrication of a miniaturized capillary system using photolithography to etch very tight injection zones onto glass microscope slides; up to 100 channels can be placed on a single slide. Use of these CAE chips has enabled the group to achieve high-resolution separations of double-stranded DNA from 70 to 1000 bp in only 120 sec.

Michael Westphall (UWM) reported on a new fluorescence-based detection system for use with multiple fluorophore sequencing in horizontal ultrathin slab gel electrophoresis (HUGE).



Norman Doggett (Los Alamos National Laboratory) and David Callen (Adelaide Women and Children's Hospital) interpret chromosome 16 map data.

Genome News

The system uses laser through-the-side excitation and a cooled CCD detector, which allows for parallel detection of up to 24 sets of 4 fluorescently labeled DNA-sequencing reactions during their separation in HUGE gels. The automated sequencing system is capable of producing 500 bp of raw data from 24 samples in less than 70 min. The group is exploring dense loading and is building an automated gel loader.

As an alternative to slab and capillary systems, Joe Balch (LLNL) and colleagues at Perkin Elmer Corporation are investigating a hybrid technique based on a high-density array of electrophoresis channels micromachined on a single, large substrate at fixed locations. DNA sequencing in both mechanically polished and chemically etched microchannels allows base calling to about 500 bp/channel (comparable to slab gels) for a 25-cm load, and current efforts are focused on developing larger channel arrays for high-throughput sequencing. A 1993 CRADA established this collaboration between LLNL and Perkin-Elmer to develop analytical instrumentation for faster DNA sequencing via electrophoresis.

Robert Weiss (University of Utah, NIH GESTEC and DOE funded) described instrumentation for automated hybridization and detection of DNA hybrids on nylon membranes to be used in multiplex mapping and sequencing of transposon inserts in large plasmid templates. The prototype system features a pair of nested Plexiglas cylinders with a heated inner drum having nylon membranes fixed to its outer surface; the drum rotates through a fluid puddle formed by an outer-drum enclosure. Fluorescent-light output on nylon membranes is amplified with a conjugated probe system of alkaline phosphatase combined with a fluorogenic alkaline-phosphatase substrate. The amplified signal allows detection of DNA hybrids in the subfemtomole band range. The group also began large-scale genomic DNA sequencing of the *Pyrococcus furiosus* genome for the DOE Microbial Genome Initiative.

Primer walking. Directed strategies using presynthesized libraries of short (5 to 7) oligonucleotides may eliminate subcloning as well as the redundancy, sequence-assembly, and gap-closure problems associated with random-sequencing methods. Primer availability reduces the time and cost of synthesizing individual primers and will enable complete automation of this approach.

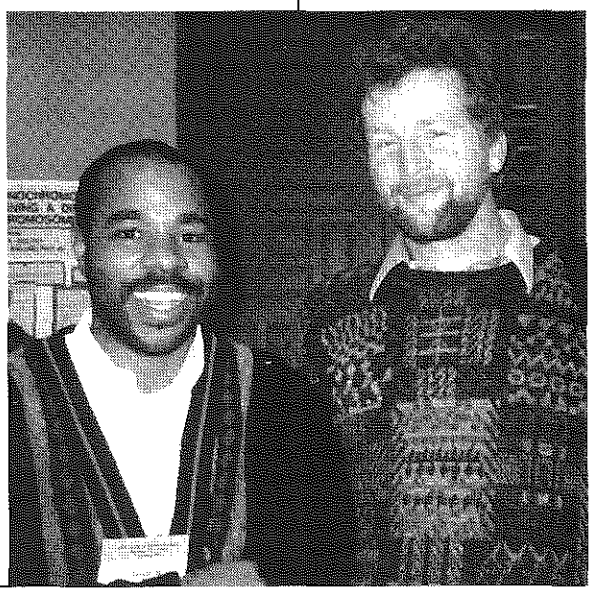
Levy Ulanovsky (Weizmann Institute of Science) and William Studier (Brookhaven National Laboratory) reported progress in designing primers from oligo libraries and developing closed-end, automatable systems. Ulanovsky assembles primers without ligation using a 5+7+7 structure with purine base stacking between the 5-mer and the adjacent 7-mer. The modular primers, which can be used with dye terminators on the ABI 373A automated sequencer and on the replaceable matrix CE system of Barry Karger, are extended by polymerase selectivity. The 1000-primer library contains 500 each of pentamers and degenerate hexamers.

The workshop was organized by Sylvia Spengler and staffed by Kati Markowitz, Catherine Pinkas (all at LBL), and Leilani Correll (LLNL). Computing support was coordinated by systems administrator Beth Dermer (LANL), and Steve Howard (LANL) was responsible for Internet lines. Computers and other electronic equipment were supplied by Digital Equipment Corporation, Sun Microsystems, and LANL. Human Genome Management Information System (HGMIS) designed and compiled the program abstract booklets, which are available from HGMIS (see address on p. 12). Meeting abstracts and this text are accessible at the DOE Human Genome Program WWW site (*http://www.er.doe.gov/production/oher/hug_top.html*).

Studier summarized efforts to construct a closed system that can be scaled up for production sequencing. His group assembled an efficient priming complex containing 3 hexamers from a library of 4096 hexamers. With dye terminators and an ABI 373A automated sequencer, primer walking with hexamer strings rapidly determined the entire sequence of two different M13 templates (6.4 and 7.2 kb), at a success rate comparable to that obtained with conventional primers. To extend primer walking to cosmid-sized (40-kb) DNA, new fesmid cloning vectors were developed (based on Simon's fosmids), whose inserts are flanked by replication and packaging signals recognized by bacteriophage T7. When infected by T7, the cloned fragment is amplified and packaged into phage particles, leaving most of the vector sequence behind. Fesmids should provide adequate amounts of DNA for performing several primer-walking steps directly on each 40-kb template. The group (with Karger's laboratory) is integrating this walking approach into a CE system containing multiple capillaries with a replaceable polyacrylamide matrix.

Mass spectrometry (MS). An innovative approach for high-throughput DNA sequencing uses MS instead of gel electrophoresis for very fast separation and detection of nested sequencing fragments. MS has the potential to reduce fragment separations from about 1 h using gel-based methods to mere seconds.

Winston Chen (ORNL) reported detection of single- and double-stranded DNA up to 500 bp and enhanced detection sensitivity to the femtomole region using matrix-assisted laser desorption ionization (MALDI) MS. Current mass resolution corresponds to about 10 bases, with single-base resolution the immediate goal.



Jeffrey Garnes (Lawrence Livermore National Laboratory) and Pieter de Jong (Roswell Park Cancer Institute) tour the more than 200 posters at the workshop.

Genome News

New WWW Tools for Sequence Data Submission

WebSub

GSDB recently announced the beta version of WebSub for submitting sequence data via WWW:
http://www.ncgr.org/gsdb/WebSub.html

WebSub works with Mosaic and Netscape on Sun platforms and with Netscape for Mac and IBM users. Those without access to a forms-supporting Web browser can download applications via ftp.

BankIt

BankIt, a new GenBank® forms approach, is available for use with Netscape for UNIX, Macintosh, and PCs; Mosaic for UNIX; and MacWeb for Macintosh. BankIt access: *http://www.ncbi.nlm.nih.gov*, questions: *info@ncbi.nlm.nih.gov* or 301/496-2475.

GenBank release 87.0 is available via ftp to *ncbi.nlm.nih.gov* in the *genbank* directory. The ASN.1 version is in the *ncbi-asn1* directory.

Data Sharing

EBI and DDBJ also add data to the international collection of DNA sequences, with all four sites sharing data. An article on sequence databases appeared in *HGN* 6(3), 1–6 (September 1994) and is also accessible via the HGMIS Home Page:
http://www.ornl.gov/TechResources/Human_Genome/home.html

Peter Williams' team (Arizona State University) launches DNA into time-of-flight MS from an ice matrix. The ice is transparent to the laser beam, which impacts the underlying copper support and shock-vaporizes the ice-DNA mixture. DNA mixtures with lengths up to 90 nucleotides have been analyzed with near-single-base resolution. The immediate challenge is to achieve high reproducibility of still-rare successes.

Christine Nelson (UWM) discussed studies on DNA fragmentation that may be responsible for MALDI's limitations in size and base composition. The group found that base protonation seems to initiate fragmentation.

In another MS approach, Richard Smith (Pacific Northwest Laboratory) described analysis of single, large DNA fragments by using electrospray ionization (ESI) combined with Fourier transform ion cyclotron resonance (FTICR) MS. In this approach, large (up to 25-kb) DNA segments are transferred to the gas phase using ESI; multiply charged ions are trapped in an FTICR mass spectrometer cell, where single ions can be selected for precise and rapid mass measurements. Thus, large DNA segments can be ionized intact and detected with high sensitivity by a nondestructive process, due to the high charge that results from ESI, allowing reactions to be followed for effectively unlimited time (>hours). Current research is aimed at developing the gas phase chemistry necessary to obtain sequence by inducing a step-wise dissociation of the trapped ions.

Sequencing by Hybridization (SBH). SBH provides sequence information by specifically hybridizing small probes to a target sequence fixed to a solid support. Although initially conceived as a direct DNA-sequencing procedure, the SBH format is also practical for mutation and polymorphism detection, HLA typing, and other uses. Andrei Mirzabekov (Argonne National Laboratory) spoke about developments in the DNA-hybridization microchip and its use as a versatile tool for DNA analysis. His group successfully identified mutations in blood samples of beta-thalassemia patients.

In collaboration with Nanogen, Inc., Glen Evans' team has developed Active Programmable Electronic Matrix (APEX) to provide exquisite electronic control of individual oligomer sequences in oligomer arrays. An electrode under each patch permits selective attraction and repulsion of interrogating sequences as well as transfer of test DNA strands. The matrix is an electronic device like a computer chip capable of interacting with DNA. This novel technology, which has potential as a minilab on a chip, allows very fast hybridizations under attraction conditions and more sensitive mismatch discrimination in oligomer and test strands under repulsion conditions.

Charles Cantor (Boston University) described the use of DNA sequences in capture and detection methods to facilitate genome analysis. In his group's enzyme-enhanced SBH approach, a partially duplex DNA probe captures the five complementary bases at the end of a single-stranded DNA target. Additional bases can be read and accuracy of the original detection ensured by DNA-polymerase extension of the probe serving as the primer along a target serving as template. Cantor discussed using the SBH format to prepare DNA samples rapidly for such fast serial-sequencing methods as CE or MS; an array of 1024 probes could capture and generate sequence ladders from any arbitrary DNA sequence. He also described an efficient single-sided *Alu*-PCR procedure that provides larger samples for mapping and other purposes.

Sequencing Informatics

GSDB. Sequence-database challenges and features of the new GSDB schema were among topics addressed by Michael Cinkosky (NCGR), who observed that the GSDB staff's responsibility is to help the community keep the database complete, accurate, and up to date. (See information in left margin.) The new system will support direct client-server inserts and updates; entry versioning, which retains all versions of public entries; and third-party annotation, in which the core entry belongs to the original author. Entries are recast so that almost all data can point to links in other databases. Efforts are directed toward achieving "anonymous interoperability," with GSDB as one component in a biology-wide database federation.

GSDB staff is designing a sequence editor as an interactive tool for sequencing laboratories to view and edit large regions having complex annotation; design goals include online intuitive, graphical editing. The editor will be freely distributable and integratable. Anyone interested in participating in the design process can obtain the prototypes that serve as a basis for discussion (*http://www.ncgr.org*).

Chris Fields (NCGR) outlined the emerging informatics challenges guiding NCGR's strategic planning and offered some ideas on a productive new direction for genome informatics.

☛ Editing Program for Protein, DNA, RNA Alignments

DCSE (Dedicated Comparative Sequence Editor) is a program for editing protein, DNA, or RNA alignments. The new version 3.3, with complementary programs Convers and tkDCSE for X-Windows, can be obtained by anonymous ftp to *www-rma.uia.ac.be* under the directory *pub/dcse* and through WWW from the DCSE Home Page *(http://www-rma.uia.ac.be/~peter/dcse)*. A paper about DCSE by Peter De Rijk and Rupert De Wachter (University of Antwerp) appeared in *Comput. Applic. Biosci.* 9, 735–40 (1993).◊

A key technical challenge is the diversity of data applications, which will require the connection of genome data with information generated from various disciplines and maintained in different databases. Fields observed that the community will be forced to reduce data-maintenance costs by moving from centralized data banks and databases to interoperable data resources joined by the Internet.

An even greater challenge, Fields continued, will be the need for precise description of the same biological data at successive levels of complexity. He said computer science has developed conceptual tools, including the Virtual Machine, for describing the precise context in which data occurs. The bioinformatics community should consider using these tools to interrelate such data.

Sequence annotation. Automating sequence-data annotation is becoming increasingly important as volume overwhelms manual annotation. David States (WU) is developing improved methods for analyzing nucleic acid sequences based on sequence similarity and very large scale classification techniques. The WU group has developed an improved user interface that uses WWW, "perl," and the html protocol to link classification data with other databases (*http://ibc.wustl.edu*). Goals include identifying unannotated reading frames in nucleic acid databases.

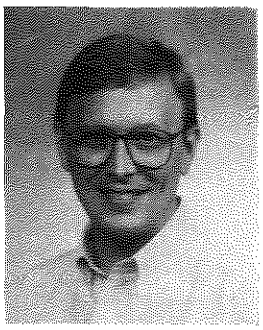
Fragment assembly of shotgun-sequencing data involves deciding which overlaps or melds to use in reconstructing the original strand. Eugene Myers (University of Arizona) described an algorithm that greatly simplifies the problem by first identifying all overlaps and melds that *must* occur in an optimal solution. For highly repetitive target DNA sequences, he further proposed use of a maximum-likelihood estimator based on the two-sided Kolmogorov-Smirnov statistic.

Sequence analysis. Edward Uberbacher (ORNL) described the latest releases of GRAIL (version 1.2) and genQuest (version 1.1). These analytical systems take sequences through a series of recognizers that pick out such features as exons, promotors, and genes; search seven databases for similarities; and incorporate the information in an annotated report. Exon-recognition rate is 94% for sequences that are correct in the database. About 20 million bases/month are processed from 4700 transactions, including e-mail GRAIL and genQuest and XGRAIL and XgenQuest. In collaboration with Fasman and NCGR, Uberbacher's group is working on faster sequence searches and parallel queries of some relational database systems.

Uberbacher presented the new GRAIL 1A, which is designed to process large files of cDNAs, ESTs, or genomic fragments to predict coding regions, search databases, and produce a summary report. GRAIL-ET, a new technology that will be useful for analyzing very low pass sequence, detects errors in coding sequences using a coding-recognition and dynamic-programming method. With a 1% indel error rate, the system found 94% of the exons (89% of the gene message after the model was made). GRAIL and genQuest can be accessed by e-mail server (*GRAIL@ornl.gov* and *Q@ornl.gov*, respectively), graphical client tools obtained by ftp at *arthur.epm.ornl.gov* or via Mosaic (*http://avalon.epm.ornl.gov*). GRAIL has been licensed to ApoCom, Inc., for use by researchers in proprietary pharmaceutical and biotechnology companies who cannot use Internet because of data-security concerns. Uberbacher's group is also supporting mouse-human mapping research at ORNL and has constructed ACEDB implementation containing mapping and phenotype data.

Participants look forward to the next DOE Contractor-Grantee Workshop, scheduled for January 28–February 1, 1996. *[Denise Casey, HGMIS]* ◊

Nelson Heads Baylor Genome Center



On January 1, David L. Nelson assumed leadership of the Baylor College of Medicine (BCM) Genome Center, succeeding C. Thomas Caskey, who has joined Merck Research Laboratories in Westpoint, Pennsylvania. Nelson is Associate Professor in the Department of Molecular and Human Genetics at BCM and has served as Associate Director of the genome center since 1992. He joined the BCM staff in January 1986 after receiving his B.A. degree from the University of Virginia and Ph.D. from the Massachusetts Institute of Technology. ◊

☛ *Saccharomyces* Database Available

The first public release from the *Saccharomyces* Genome Database (SacchDB) at Stanford University Genome Center is now available using ACEDB software for Macintosh and a variety of UNIX systems. Version 2.3 of SacchDB includes all *Saccharomyces* genes contained in the "Registry of Gene Names" on January 21; results of completed sequencing projects for chromosomes I, II, III, V, VIII, IX, and XI; physical maps; remapped Olson prime clones for the completed chromosomes; all *Saccharomyces* DNA sequences in GenBank® release 86; literature references, most including abstracts; gene-protein product information from the yeast protein database; and genetic maps with underlying two-point tetrad data.

SacchDB is accessible via Internet (*http://genome-www.stanford.edu*) or Gopher to *genome-gopher.stanford.edu* (port 70). The complete ACEDB database for Macintosh or UNIX can be obtained via anonymous ftp from *genome-ftp.stanford.edu* (*pub/yeast/SacchDB*) or *ncbi.nlm.nih.gov* (*repository/SacchDB*). The file <README.installation> should be retrieved first to determine files needed for a particular system. Suggestions or corrections (*yeast-curator@genome.stanford.edu*). [Contact: Mike Cherry (415/725-8956, Fax: /723-7016, *cherry@genome.Stanford.edu*).] ◊

Genome News

Human Genome news

National Center for Human Genome Research

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. Mavournin

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*
WWW: *http://www.ornl.gov/TechResources/Human_Genome/home.html*

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*

Fourth Transcribed Sequences Workshop Held

The Fourth International Workshop on the Identification of Transcribed Sequences, sponsored by the Canadian Genome Technology and Analysis Program, DOE, and Amgen, Inc., was held October 16–18, 1994, in Montreal, Canada. Some 80 participants from 12 countries, including more than 50 speakers, gathered to discuss methods for gene isolation and identification.

For the first time in this workshop series, significant progress in actual transcriptional map construction was presented for several chromosomal regions, and transcriptional mapping in the mouse system made a strong showing. Although protocols are becoming standardized for gene-identification techniques such as cDNA selection and exon trapping, an alternative approach—large-scale genomic sequencing coupled with "software trapping"—is now considered an attractive option. Functional analysis poses a growing challenge as large numbers of transcribed sequences become available. Some highlights of workshop presentations and discussion groups follow.

cDNA Hybridization Selection

This approach has proved highly efficient and is clearly the most popular gene-identification technique. Although the technique itself is well standardized, new improvements and variations were presented. S. Patanjali (Yale University) has demonstrated that whole-yeast DNA from YAC-containing strains can be used effectively without prior YAC purification; enrichments of $>10^4$ can still be obtained without increases in ribosomal contaminants. As alternative genomic material, Michel Fontes [Institut National de la Santé et de la Recherche Medicale (INSERM)] used IRS-PCR material amplified with degenerate *Alu* primers from whole-yeast DNA.

An ongoing problem with cDNA selection and exon trapping is the isolation of full-length cDNAs corresponding to the short cDNA fragments typically obtained. Sherman Weissman (Yale University) showed that selected cDNAs can be used without prior cloning to screen full-length cDNA libraries directly. Bernard Korn (Germany Cancer Research Center), Cynthia Jackson (Rhode Island Hospital), and Greg Lennon [Lawrence Livermore National Laboratory (LLNL)] carried out cDNA selection from more-complex genomic sources—flow-sorted X chromosome, a chromosome 9 somatic cell hybrid, and flow-sorted chromosome 19, respectively. Selection specificity was decreased, but processing simplicity and rapidity may compensate for this.

Exon Trapping

To increase the size of trapped exons, Johan den Dunnen (Leiden University) proposed a cosmid-based exon-trapping vector. Providing large genomic clones should allow for processing of full-length or nearly full-length cDNAs rather than the one or two exons typically trapped. Greg Landes (Integrated Genetics) used the pSPL3-CAM vector to trap 17 P1 clones from the PKD1 region and obtained 4 to 20 exons per clone. Yun-Fai Chris Lau (University of California, San Francisco) used 3′ exon trapping with pTAG4 and a pool of 4600 Y chromosome cosmids. Products were 40% artifacts with 60% of the unique clones mapping back to the Y chromosome.

Genomic Sequencing

Marcia Budarf [Children's Hospital of Philadelphia (CHOP)] sequenced >200 kb from the DGCR of 22q11.2 and used GenBank® searches, GRAIL predictions, and RT-PCR to identify nine new genes. B. Rajendra Krishnan (Washington University School of Medicine) applied a sample sequencing technique to segments of the HLA-C region and also identified new genes using both computer and laboratory techniques.

Regional Maps

Using predominantly cDNA-selection techniques, researchers demonstrated significant progress in constructing high-density transcriptional maps of several chromosomal regions. These included the MHC class I region of 6p (Wufang Fan, LLNL); A-T region of 11q22-q23 (Anat Bar-Shira, Tel-Aviv University); 7q21-q22 (Johanna Rommens, The Hospital for Sick Children); regions of 21q (Hongxia Xu, Yale University, and Katheleen Gardiner, Eleanor Roosevelt Institute); 22q11 (Weilong Gong, CHOP, and Howard Sirotkin, Albert Einstein College of Medicine); and Xq28 and the entire X chromosome (Korn). The percentage of putative new cDNAs that map back to the correct chromosomal region varies considerably, but all are >50% and the best is >90%. After sequencing, further analyses include fine mapping to YAC and cosmid contigs, determination of transcription orientation, Northern analysis, and cDNA library screening.

cDNA Library Analysis

Projects to sequence, map, and further analyze random cDNAs are ongoing. In sequencing >600 clones from a human testes cDNA library, Michael Jones (Cambridge University) found that 3% were identical to known genes and 70% novel. Of 200 cDNAs examined, 15% mapped to multiple chromosomes, and 70% showed significant cross hybridization to rodent DNA. To characterize further the apparently widely expressed novel ESTs, Donna Maglott (American *Type Culture* Collection) used Northern analysis to show that 20% were brain specific. Donald Moir (Collaborative Research, Inc.) mapped 138 infant brain ESTs to a grid of 2800 mega-YAC clones and found that >50% identified one or more YACs.

Gridded arrays of cDNA libraries are becoming more widely used. Catherine Nguyen (INSERM and Centre National de la Recherche Scientifique) hybridized a grid of 80,000 mouse thymus cDNAs with labeled cDNA from various tissues. Signals can be related quantitatively to the tissue-specific levels of expression. Daniela Toniolo (Consiglio Nazionale delle Richerche) used a similar approach with a mouse 10-day-embryo central nervous system library.

Greg Lennon (LLNL) discussed the multiuser analysis of the Bento Soares (Columbia University) gridded infant brain, liver, and spleen libraries under the auspices of the Integrated Molecular Analysis of Gene Expression (IMAGE) consortium. Filters of libraries are available from Lennon (*greg@mendel.llnl.gov*).

Because of the potential relevance to neurodegenerative disorders, Christian Neri (CEPH) screened a human fetal brain library with CAG and CCG repeats and is cataloguing those with >5 repeat units. Of the 114 analyzed, 17.5% have homologies in the EST databases.

Susan Ackerman (Jackson Laboratory) constructed a cDNA library from differentiated and nondifferentiated NT2 cells and is using a subtractive library and differential display to identify cDNAs expressed specifically in NT2 neurons. So far, most differences appear related to expression level and not absolute specificity.

Mouse and Functional Analysis

Toniolo had previously identified a number of genes and CpG islands in two regions of human Xqter. She presented data showing that the order and orientation of the same genes is preserved in the mouse. Transcriptional patterns in mouse development are being investigated.

Kevin Brady (Harvard Medical School) discussed the ease and rapidity of mapping expressed sequences by single-strand confirmational polymorphism in recombinant inbred strains. Filters are available for such analyses at an estimated cost of $100.00 per locus.

Catherine Lambert (FUNDP School of Medicine) used cDNAs from 14-day embryos and embryonic brains to select candidate reeler genes among YAC clones from the mouse chromosome 5 centromeric region. Selected cDNAs and a random sequencing approach have thus far yielded no homologs in GenBank searches and no exons by GRAIL analysis.

Miriam Meisler (University of Michigan) is isolating genes in the distal mouse chromosome 15 region of A4, a transgenic line associated with a neuromuscular disorder. Monica Justice (Kansas State University) cloned a retroviral insertion site, Evi3, associated with mouse B-cell lymphomas. Reverse-transcription PCR and genomic sequencing were needed to analyze the complex gene organization in the region. In both cases, understanding the mouse genes involved is expected to facilitate the cloning and characterization of the homologous human genes, mapping to human chromosomes 12 and 18, respectively. This can be particularly important in such cases where tissue and timing of expression make human studies difficult.

In the only presentation to address functional analysis of new genes directly, Russ Finley (Massachusetts General Hospital) described use of a yeast-interaction mating system to identify and characterize interactions among cell-cycle regulatory proteins. Such systems will be increasingly important for defining gene functions.

Informatics

As more transcribed sequences are isolated, the ability to interpret sequence information becomes increasingly important. Jean-Michel Claverie (NIH) discussed the expanding problem of analyzing novel genes with no homologies in the databases. Currently, 69% of ESTs are "unknown" cDNAs. A cross-species comparison strategy of these sequences identified 180 previously uncharacterized protein domains, many of which are likely to be involved in as-yet-uncharacterized basic cellular functions.

Richard Mural (Oak Ridge National Laboratory) discussed new enhancements to GRAIL that will aid in gene discovery through large-scale genomic sequencing coupled to "software trapping" and new tools for sequence annotation in the GRAIL system.

Progress and Future Perspectives

Martin Ringwald (Jackson Laboratory) discussed a project to develop a gene-expression information resource for mouse embryonic development. This would include both textual descriptions and 3-D images of gene-expression patterns during mouse development.

The gene-finding techniques of cDNA selection, exon trapping, and genomic sequence analysis are being widely and successfully applied with increasingly standardized protocols. Many groups find that focusing on one of these approaches quickly provides a vast resource of new genes. Comprehensive transcriptional mapping will probably employ combinations of all three strategies. Old problems of how to isolate complete cDNAs from exons and small cDNA fragments and identify pseudogenes remain to be solved efficiently. Clearly, future workshops will focus increasingly on two additional problems—interpretation of sequencing information of both genomic and cDNA clones and functional analysis of novel genes. *[Richard Mural, Oak Ridge National Laboratory, and Katheleen Gardiner, Eleanor Roosevelt Institute]* ◊

Introductory Linkage Course

A course for researchers who have a basic understanding of linkage analysis but little or no experience in using linkage programs will be held June 12–16 at Columbia Presbyterian Medical Center in New York City. Topics will include an introduction to linkage analysis, practical aspects of data collection, strategies and methods of linkage analysis, incomplete penetrance (narrow and wide definition), inbreeding loops, simple risk calculations, and introduction to computer simulation. A major part of the course will consist of exercises using LINKAGE software programs. Attendance is limited to 30. [Contact: Katherine Montague (212/960-2507, Fax: /568-2750, *jurg.ott@columbia.edu*).] ◊

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.◊

Genome News

Groups Publish Detailed Chromosome 22 Map

Research groups led by Beverly Emanuel at the Children's Hospital of Philadelphia (CHOP), the University of Pennsylvania, and the Fox Chase Cancer Center and by Tom Hudson and Eric Lander at Whitehead Institute–Massachusetts Institute of Technology published a detailed physical map of human chromosome 22 in the January 11 issue of *Human Molecular Genetics*. The map, which provides about 70% coverage of the chromosome in 15 YAC contigs, is the product of an ongoing collaboration between two genome science and technology centers (GESTECs) supported by the NIH National Center for Human Genome Research (NCHGR). Similar low-resolution comprehensive maps have been published for chromosomes Y and 21 (see *HGN* 4(4), 1–4 (November 1992).

Emanuel said, "Our map provides information of immediate use to anyone who is looking for genes and disease-causing rearrangements on this chromosome." The new map also provides a framework for constructing higher-resolution physical maps for eventual sequencing of the chromosome.

The 587 YAC clones were assembled into contigs by STS-content mapping and ordered along the chromosome using cytogenetic breakpoint, meiotic, and pulsed-field gel maps. Most YACs were identified from the CEPH-Généthon original and mega-YAC libraries (average insert sizes, 470 and 900 kb, respectively); additional YACs were isolated from the Washington University YAC library and from a chromosome 22–specific YAC library generated at the Philadelphia GESTEC from GM10888, a chromosome 22–only hybrid cell line.

Callum Bell (CHOP), the paper's lead author, pointed out that contig construction was complicated by deletions and permutations caused by chromosomal regions that are unstable when cloned into YACs. To overcome this problem, researchers assembled contigs using a computer algorithm developed by David Searls (University of Pennsylvania Medical School) to show a series of possible marker orders. This approach helped to minimize gaps in the map and reveal possible ambiguities in assigning marker order.

YAC contigs provide extensive coverage of the chromosome long arm, with the highest marker density in the 22q11.2–22q13.1 region. The distal end of the long arm appears to be resistant to YAC cloning; for this region, the group is exploring use of other cloning vehicles such as BACs, P1 phage clones, PACs, and cosmids. Other efforts to achieve complete coverage of the long arm include generating additional STSs for further screening of large genomic clones and the use of *Alu*-PCR hybridization methods.

The third-smallest human chromosome at about 50 million bp, chromosome 22 is thought to contain as many as 2000 genes and pseudogenes. More than 100 have been mapped, including genes for cancers such as Ewing's sarcoma, Burkitt's lymphoma, meningiomas, acoustic neuromas, and acute lymphoblastic leukemia. Recent reports suggest that the chromosome may contain a tumor-suppressor gene involved in breast cancer and schizophrenia. Also linked to chromosome 22 are Cat Eye syndrome, DiGeorge syndrome, isolated congenital heart defects, velocardiofacial syndrome, and neurofibromatosis 2. *[Article adapted from press release by Bob Kuska, NCHGR]* ◊

Extensive details of the STSs, probes, and YACs they detected may be retrieved by anonymous ftp from

- *cbil.humgen.upenn.edu/pub/22* and
- *genome.wi.mit.edu/distribution/human_STS_releases.*

Data are also accessible via WWW at

- *http://www.cis.upenn.edu/~cbil/chr22db/chr22dbhome.html* and
- *http://www-genome.wi.mit.edu.*

Association Considers Genetics, Justice, Minorities

With a program entitled "DNA: Genetics, Criminal Justice, and the Minority Community," a large professional association considered for the first time the implications of genetics for law enforcement and civil rights. On September 23–24, 1994, the Justice George Lewis Ruffin Society of Massachusetts marked its tenth anniversary by bringing Boston law-enforcement representatives together with minority professionals to examine issues of genetics; race; criminal behavior; and civil discrimination in employment, insurance, and health care.

Ruffin Society President Judge Julian R. Houston (Massachusetts Superior Court) heralded the 150-person convocation as "an opportunity for us to confront and consider the future as it rolls out of the laboratory and into the courtroom and the streets." Aided by a grant from the Ethical, Legal, and Social Issues (ELSI) component of the DOE Human Genome Program, the conference received high marks from participants and local news media. "We planned carefully," Houston said, "and it paid off with a program and a process in which I think we justifiably can take pride." The project is managed by Robert D. Croatti, Associate Dean of the College of Criminal Justice, Northeastern University.

The conference planning committee included a representative from each of the collaborating organizations: College of Criminal Justice of Northeastern University; Massachusetts State Police; Boston Police Department; Massachusetts Division of Probation and Parole; Massachusetts Superior Court; Harvard University Law School; Franklin Flaschner Judicial Institute; Whitehead Institute; and the Washington-based Einstein Institute for Science, Health, and the Courts.

"You can't depend on just one sector when we see the spillover of genetic discovery into criminal proceedings, discrimination, and health care," Houston observed. "It takes a whole team, and we were very lucky to have had one."

Convocation plenary and discussion sessions surveyed every high-profile issue confronting forensic and medical molecular biology, from criminal identification and paternity establishment to the possible role of genetics in criminal predisposition. The educational program featured plenary sessions in DNA analysis, behavioral predisposition limitations, and technologies generated and promised by the Human Genome Project.

The unique program featured five small-group discussions with trained leaders and recorders. These groups focused on hypothetical cases involving the day-to-day implications of scientific

(see Minority, p. 15)

Genome News

NCHGR Plans New Approach to SCWs in 1996

A National Center for Human Genome Research (NCHGR) staff report has recommended that NIH and DOE terminate U.S. funding at the end of 1996 for the international single-chromosome workshop (SCW) program but continue to encourage applications for individual SCWs as needed. The overall SCW program is coordinated by the Human Genome Organisation (HUGO).

The report was presented January 30 to the NIH National Advisory Council by NCHGR Program Director Elise Feingold. She also recommended that the two agencies expend a maximum of $20,000 per workshop and withhold 50% of travel reimbursement until a meeting report is submitted for publication. Data submission to GDB would be required when abstracts are due to organizers.

The SCW program was initiated early in the Human Genome Project to further mapping goals by bringing together investigators to pool and share up-to-date research data. To improve workshop consistency and productivity, funding agencies and HUGO later developed and refined a set of guidelines that generally have been well accepted in the scientific community.

Feingold noted to the council that the SCW program needed reassessment in light of such mapping achievements as the 2.5-cM human genetic linkage map and rapid progress in human chromosome physical maps.

Chromosome Workshops Planned*

Chromosome	Date	Place	Contact
1	Sept. 22–24	Vienna	A. Weith, +43-1/797-30-625
3	Oct. 22–23	Minneapolis	S. Naylor, 210/567-3842
9	April 22–25	Williamsburg, Va.	M. Pericak-Vance, 919/684-6274
10	Sept. 29–Oct. 1	Crete, Greece	J. Mao, 617/893-5007 ext. 242
12	Nov. 17–19	Leuven, Belgium	P. Marynen, +32-16/34-5891
13	Oct. 29–31	Ossining, N.Y.	D. Warburton, 212/305-7143
16	Nov. 13–14	Leiden, Netherlands	M. Breuning, +31-71/276-293
18	May 8–10	Philadelphia	J. Overhauser, 215/955-5188
Y	Sept. 17–20	Pacific Grove, Calif.	C. Lau, 415/476-8839

*Chromosome workshops are listed in the Calendar of Genome Events on page 14. Workshops to be held this year between April and November are also listed above. Readers should inform HGMIS of other meetings to include in the calendar.

Daniel Drell, speaking for the DOE Human Genome Program, totally concurred with the report's recommendations.

The two agencies contribute equally to the SCW program, which supports U.S. participants' travel to workshops as well as local arrangements for meetings held in the United States. NIH and DOE spent about $346,000 in 1993 and $443,000 in 1994 for SCWs, with the average cost per meeting around $31,500 and $26,000, respectively. Some 44 investigators participated in each workshop, with an average of 19 from the United States.

SCW Contacts

Feingold
NIH NCHGR
Bldg. 38A, Room 610
Bethesda, MD 20892
301/486-7531
Fax: /480-2770
Elise_Feingold@nih.gov

Drell
DOE, ER-72
19901 Germantown Rd.
Germantown, MD 20874
301/903-4742
Fax: -8521
Daniel.Drell@mailgw.er.doe.gov ◊

Citations for 1994 Single-Chromosome Workshops

Chromosome	Meeting Date	Citation
Cytogenetics and Cell Genetics		
1	March 25–27	**67**, 143–65
2	June 24–26	**67**, 215–44
3	May 8–9	**68**, 125–46
8	Sept. 16–18	**68**, 147–64
11	Sept. 25–28	**69**, 127–58
12	June 20–22	**67**, 245–76
13	May 20–22	**70**, 1–22
14	Sept. 1–3	**69**, 159–74
15	Feb. 18–20	**67**, 1–22
16	May 7–9	**68**, 165–84
22	July 2–5	**67**, 277–94
X	April 24–27	**67**, 295–358
Y	April 2–5	**67**, 359–402
Annals of Human Genetics		
9	April 9–11	**58**, 177–250

Minority *(from p. 14)*

progress. Postdoctoral fellows from the Whitehead Institute served as science advisors, who were trained by Franklin Zweig (Einstein Institute) and Judge Rosalyn Bell (Maryland Court of Special Appeals).

The Ruffin Society, an association of minority criminal justice professionals employed in Massachusetts law enforcement, is named for the first African American appointed to the Massachusetts courts (in 1883). A 57-page report drafted by Einstein Institute personnel is being converted into a booklet highlighting the Ruffin Convocation (book requests, Fax: 301/913-5739).

In the report Zweig and Bell commented, "The ELSI program has shown in this convocation that minority engagement with issues raised by genetics can produce a real and durable agenda that in turn can undergird future policy considerations. The Ruffin Society has shown that these issues can be anticipated, discussed, and interpreted within a scientific, legal, and social context." *[Franklin M. Zweig, Einstein Institute]* ◊

Genome News

NIH Advisory Council Meets

The NIH National Advisory Council for Human Genome Research was convened for its twelfth meeting on September 22, 1994, in Washington, D.C. Francis Collins, Director of the NIH National Center for Human Genome Research (NCHGR), presided. Selected highlights of the meeting follow.

Jeffrey Trent, Scientific Director of the NCHGR Division of Intramural Research, updated the council on the division's activities. The intramural division was established in 1993 to study genes that cause diseases, including cancer, and to focus on medical genetics, clinical gene-therapy research, and development of clinical diagnostic tests. Trent also reported overwhelmingly positive response to the new visiting investigator program, which allows university scientists to use NCHGR resources for 3 to 12 months [*HGN* **6**(2), 7 (July 1994)].

The council discussed a draft mission statement calling for the DOE-NIH Working Group on Ethical, Legal, and Social Issues (ELSI) to report to both the NCHGR advisory council and the DOE Health and Environmental Research Advisory Committee. Collins noted that the White House Office of Science and Technology had proposed establishing a National Bioethics Advisory Commission within the Executive Branch [*Federal Register* **59**(155), 41584–86 (August 12, 1994)]. This advisory commission would consider bioethics issues arising from research on human biology and behavior and applications of that research.

Jane Peterson (Chief, NCHGR Mammalian Genomics Branch) presented a set of guidelines for assessing physical-mapping progress toward Human Genome Project 5-year goals. David Cox (Stanford Human Genome Center) stated that a meeting of center directors had developed reporting standards, and the goal now is to inform the scientific community about reporting mapping data in a common context. The guidelines were published in *Science* (September 30, 1994).

Peterson led the discussion on limiting ELSI components of NCHGR genome science and technology centers (GESTECs) to 5% of the total budget. She said some GESTEC proposals contain expanded ELSI components, which are reviewed differently from individual ELSI proposals. GESTEC reviewers do not see the full range of ELSI proposals, nor do ELSI review groups see the scientific context of education applications. Rather than establish a policy, the council recommended approaching each GESTEC review on a case-by-case basis. They requested a report on the ELSI portfolio at the January meeting, focusing particularly on education components in ELSI grants and GESTECs.

David Benton (Director, NCHGR Genome Informatics Program) described the informatics meeting of GESTEC directors at which 10 centers were represented [meeting report (*http://www.gdb.org/Dan/nchgr/intro.html*)]. Of issues identified at the meeting, Benton reported significant action to disseminate information on software, database integration, and priority software needs. He also announced that GESTEC informatics staff would hold two meetings during the next 18 months. David Botstein (Stanford University) noted enormous progress in identifying and addressing informatics problems.

Linda Engel (NIH Office of Scientific Review) described NIH and NCHGR reinvention initiatives, including the use of triage for quick response to funding applications and summaries of reviewers' unedited comments and discussions. A proposal has been made that the scientific review office stop calculating recommended budgets, except for GESTECs with priority scores of 200 or higher. Reviewers would give overall recommendations, and applicants would provide detailed budgets at the time of award. This procedure would save time and work for reviewers, applicants, and staff.

The council reviewed 123 applications requesting $31,650,501 and recommended approval for 81 applications totaling $19,437,332.◊

Hilton Head Sequencing Meeting

Lloyd Smith (University of Wisconsin, Madison) reported to the NIH Advisory Council that encouraging progress was demonstrated at the September 17–21, 1994, Hilton Head sequencing meeting. As part of that meeting, Robert Strausberg and Carol Dahl (Chief and Program Director, respectively, NCHGR Sequencing Technology Branch) conducted a useful workshop on a range of sequencing issues, particularly those related to closure. Smith noted that community members were cautiously optimistic about meeting sequencing goals, but they felt progress continues to be hampered by the shortfall in grant support. He also pointed out that industry is becoming more involved in sequencing efforts.◊

Whitehead-MIT Announces Release Nine of Mouse Genetic Map

Release Nine of the Whitehead Institute–Massachusetts Institute of Technology Genome Center Genetic Map of the Mouse is now available. The map consists of randomly chosen simple sequence length polymorphisms (microsatellites) that can be analyzed using PCR as described in W. Dietrich et al., *Genetics* **131**, 423–47 (1992) and *Nature Genetics* **7**, 220–45 (1994).

Release Nine contains 5752 markers in 20 linkage groups spanning about 1400 cM (average spacing, about 0.25 cM). The map can be accessed in the following ways:

- WWW to *http://www-genome.wi.mit.edu*. This method is preferred.
- Internet e-mail for the most-current e-mail query form. Send a message to *genome_database@genome.wi.mit.edu* with *help* in either the subject line or body text.
- Anonymous ftp to *genome.wi.mit.edu* in directory */distribution/mouse_sslp_release/jan95/*.

This project is ongoing, with new markers released at the beginning of each quarter. [Contact: Ert Dredge; Whitehead Institute Center for Genome Research; One Kendall Square; Bldg. 300, 5th Floor; Cambridge, MA 02139 (617/252-1922, Fax: -1902, *ert@genome.wi.mit.edu*).] ◊

HUGO Chromosome Committee Report Planned for Spring

In the two decades since the first Human Gene Mapping (HGM) conference in 1973, the number of known human genes, as reflected in the Genome Data Base (GDB), has grown to more than 5000. An annual compendium of these data is usually coupled with a conference such as HGM or the Chromosome Coordinating meeting. Following the trend of the past 2 years, *Human Gene Mapping 1994* will be published this spring.

The report will reflect GDB data through mid-December 1994 from each Human Genome Organisation committee: chromosome, nomenclature, comparative mapping, human mitochondrial DNA, and neoplasia. As with the previous two reports, A. Jamie Cuticchia (GDB and Johns Hopkins University School of Medicine) will serve as primary editor.◊

GDB 5.5 Provides Links for Probe-Probe, Probe-Library Data

GDB 5.5, released in December 1994, includes relational information under Probe and Library Managers. For probe-probe links, the type of interaction between the two probes is described. For probe-library links, the library location is shown and includes Plate, Row, and Column. A detailed description of all new features is available online in "Release Notes" under "News."◊

New WWW Queries for GDB ID, Submissions, SQL

Four new applications are available from the WWW GDB Browser through the GDB Home Page (*http://gdbwww.gdb.org*).The first three applications are located near the top of the GDB Browser page, and the last is near the bottom.

- GDB ID Query allows searching by single number, list or range of numbers, and GDB data type. Retrieved GDB ID numbers are linked to detailed object data.
- Submission ID Query allows direct searching for a GDB Submission ID as given in a publication or linked to a GDB contact. Retrieved data includes information about the submission itself and the GDB ID numbers for all objects that are part of the submission set.
- Ad Hoc SQL Query provides access to publicly available GDB databases.
- GDB Schema Browser, which includes a link to GDB Schema Diagrams, is a tool for looking at descriptions and definitions of objects in databases associated with GDB.◊

GDB USER SUPPORT, REGISTRATION

GDB and OMIM are available via WWW from the GDB Home Page (*http://gdbwww.gdb.org*). 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" class will be held in Baltimore on June 5–6. 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
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
reisner@angis.su.oz.au

FRANCE
Philippe Dessen
Service de Bioinformatique
CNRS-INSERM
7 rue Guy Moquet - BP8
94801 Villejulf Cedex
France
+33/14559-5241
Fax: -5250
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
dok261@cvx12.dkfz-heidelberg.de

ISRAEL
Jaime Prilusky
Bioinformatics Unit
Weizmann Institute of Science
76100 Rehovot, Israel
+972/8-343456
Fax: -344113
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
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
post@caos.caos.kun.nl

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

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

Calendar of Genome Events*

April 1995

6–9. DNA Vaccines: A New Era in Vaccinology; Arlington, VA [NYAS, 212/838-0230 ext. 324, Fax: -5640]

9–13. Exp. Biol. '95; Atlanta [FASEB, 301/530-7010, Fax: -7014]

10–11. Dev. of Small Mol. Mimetic Drugs; San Francisco (poster deadline: Mar. 10) [CHI, 617/487-7989, Fax: -7937, *chi@world.std.com, http://id.wing.net/~chi/homepg.html*]

10–14. 3rd Intl. WWW Conf.: Technol., Tools, and Appl.; Darmstadt, Germany (abs. deadline: Mar. 1) [R. Doelz, +41-61/267-2247, Fax: -2078, *doelz@ubaclu.unibas.ch, http://www.igd.fhg.de/www/www95/www95.html*]

17. *DIMACS Spec. Yr.: Distinguished Lect. Ser., Charles Cantor (also offered Apr. 18); contact for specific location [M. Farach, 908/445-4580, Fax: -5932, *special@dimacs.rutgers.edu, http://dimacs.rutgers.edu*]

19–20. Adv. in Genet. Diagn. for Infectious Dis.; CHI, Washington, DC (poster deadline: Mar. 24) [see contact: Apr. 10–11]

20. *Jeffrey Murray: Construction and Appl. of High-Res. Hum. Meiotic Maps; Bethesda, MD [NCHGR Lect. Ser., E. Feingold, 301/496-7531, Fax: /480-2770, *Elise_Feingold@nih.gov*]

20. *Michael P. Czech: DNA, Genet., and Biotechnol.; Gaithersburg, MD [TIGR/NIST, D. Hawkins, 301/869-9056, Fax: -9423]

22–25. **4th Intl. Workshop on Chromosome 9; Williamsburg, VA [M. Pericak-Vance, 919/684-6274, Fax: -6514, *mpv@dnadoc.mc.duke.edu*]

26–28. Natl. SBIR Conf.; Chicago [box, p.19]

27–28. Symp. on Bus. and Potentials of Biotechnol.; Hong Kong, CH [W. Ho, +852/609-6345, Fax: /603-5123, *B080707@mailserv.cuhk.hk*]

May 1995

1–3. Adv. in Gene Therapy: New Tech. and Appl.; CHI, Arlington, VA (poster deadline: Apr. 3) [see contact: Apr. 10–11]

1-3. IBC's 5th Annu. Gene Therapy; Bethesda, MD [IBC, 508/481-6400, Fax: -4473]

3–5. **DIMACS Spec. Yr.: HIV Sequence Data Workshop; New Brunswick, NJ [see contact: Apr. 17]

4–5. Pharmacogenet.: Optimizing Drug Discovery and Dev.; IBC, Bethesda, MD [see contact: May 1–3]

8–9. PCR and Alternative Gene Amplification Technol.; IBC, Bethesda, MD [see contact: May 1–3]

8–10. 3rd Intl. Workshop on Chromosome 18; Philadelphia [J. Overhauser, 215/955-5188, Fax: /923-9162, *overha@calvin.jci.tju.edu*]

10–14. Genome Mapp. and Sequencing; Cold Spring Harbor, NY (abs. deadline: Feb. 22) [CSHL, 516/367-8346, Fax: -8845, *meetings@cshl.org, http://www.cshl.org*]

12. **DIMACS Spec. Yr.: Database Aspects of Biol. Data Mini-Workshop; contact for specific location [see contact: Apr. 17]

13–17. BIO 9th Intl. Biotechnol. Meet. & Exhibition; San Francisco [BIO, 202/857-0244, Fax: /331-8132]

14–18. Receptor-Regul. Calcium Influx; Pacific Grove, CA [C. Felder, 301/496–9925, Fax: /402-1748, *feldercc@helix.nih.gov*]

14–27. 4th Intl. ACEDB Conf. and Workshop; Geyserville, CA [J. McCarthy, 510/486-5307, Fax: -4004, *ace95@genome.lbl.gov*]

17–19. High-Throughput Screening; IBC, San Francisco [see contact: May 1–3]

17–21. RNA Proc.; CSHL (abs. deadline: Feb. 22) [see contact: May 10–14]

18. *George Church: Multiplex Technol. for Genome Sequencing and Anal.; NCHGR, Bethesda, MD [see contact: Apr. 20]

18. *Mary Claire King: DNA, Genet., and Biotechnol.; TIGR/NIST, Gaithersburg, MD [see contact: Apr. 20]

18-20. 5th Symp. on Delivery and Targeting of Peptides, Proteins, and Genes; Leiden, Netherlands [F.J. Velthorst, +31-71/274-341, Fax: -277, *velthors@chem.leiden.univ.nl*]

18–21. 3rd Intl. Workshop on Mutation Detection; Visby, Sweden [HUGO, L. Evans, +44-171/935-8085, Fax: -8341, *e_evans@icrf.ac.uk*]

19–30. 2nd Adv. Sch. on Biol. Magnetic Resonance: Dynamics and Prob. of Recognition in Biol. Macromolecules; Erice, Sicily (abs. deadline: Mar. 20) [O. Jardetzky, 415/723-6270, Fax: -2253, *jardetzky@camis.stanford.edu* or J.-F. Lefevre, +33-88/65-52-69, Fax: /65-53-43, *lefevre@bali.u-strasbg.fr*]

20–21. Commun. in Genet.; Arlington, VA [AGSG, J. Weiss, 800/336-GENE, Fax: -0171]

22–23. Natl. Adv. Council for Hum. Genome Res.; Washington, DC [J. Ades, 301/402-2205, Fax: -2218, *ja51b@nih.gov*]

21–25. ASBMB Meet.; FASEB, San Francisco [see contact: Apr. 9–13]

21–25. Genome Imperative: Ethical and Policy Implications in Clin. Med.; Pasadena, CA [M. Cloutier, 510/486-0626, Fax: /540-7643, *BCG@ix.netcom.com*]

29–31. Intelligence in Neural & Biol. Syst.; Washington, DC [IEEE, N. Bourbakis, 607/777-2165, Fax: -4822, *bourbaki@bingsuns.cc.binghamton.edu*]

31–June 2. Nucleic Acid-Based Technol.: Current Challenges, Future Strategies, and End-User Perspectives; CHI, San Francisco (poster deadline: May 6) [see contact: Apr. 10–11]

June 1995

5–7. Bioinformatics; CHI, San Francisco [see contact: Apr. 10–11]

14–16. 2nd Annu. High-Throughput Screening; CHI, San Diego [see contact: Apr. 10–11]

19–20. Nucleic Acid Therapeutics; CHI, San Diego [see contact: Apr. 10–11]

20–22. IGES 4th Annu. Meet.; Snowbird, UT (abs. deadline: Feb. 10) [M. Austin, 206/685-9384, Fax: -3407, *maustin@u.washington.edu*]

July 1995

5–8. 10th Intl. Conf. on Math. and Comput. Model. and Sci. Comput.; Boston (abs. deadline: Jan. 15) [X. Avula, 314/341-4585, Fax: /364-3351, *avula@umr.edu*]

17–18. **DIMACS Spec. Yr.: Geom. Methods for Conformational Model. Mini-Workshop; New Brunswick, NJ [see contact: Apr. 17]

17–28. Med. and Exp. Mamm. Genet.; Bar Harbor, ME [Jackson Lab., 207/288-3371 ext. 1253, from 7:30 a.m. to 3:30 p.m. (EST)]

18–Aug. 1. 11th World Cong. on Med. Law; Pilansberg, South Africa [D. Friedman, +27-140/84-2470/1, Fax: -24894]

20–22. MIMBD '95; Cambridge, UK (abs. deadline: Mar. 10 to *victor@csr.lbl.gov*) [P. Karp, Fax: 415/859-3735, *pkarp@ai.sri.com, http://www.ai.sri.com/people/pkarp/mimbd.html*]

23–28. 8th Intl. Conf. on STM; Snowmass Village, CO [R. Colton, 202/767-0801, Fax: -3321, *rcolton@stm2.nrl.navy.mil*]

August 1995

2–5. 7th Intl. Workshop on Fragile X and X-linked Mental Retardation; Tromso, Norway [L. Tranebjaerg, +47-77/645-410, Fax: -430]

13–18. 23rd Meet. of Fed. of European Biochem. Soc.; Basel, Switzerland [Convention Ctr. Basel, +41-61/686-2828, Fax: -2185]

15–20. Yeast Cell Biol.; Cold Spring Harbor, NY (abs. deadline: May 31) [see contact: May 10–14]

September 1995

11–12. Natl. Adv. Council for Hum. Genome Res.; Washington, DC [J. Ades, 301/402-2205, Fax: -2218, *ja51b@nih.gov*]

16–20. 7th Intl. Genome Sequencing and Anal. Conf.; Hilton Head, SC (abs. deadline: May 1) [Genome Sequencing Conf. Office, 301/869-9056, Fax: -9423, *seqconf@tigr.org*]

17–20. 2nd Intl. Workshop on Hum. Y Chromosome; Pacific Grove, CA [C. Lau, 415/476-8839, Fax: /502-1613, *clau@itsa.ucsf.edu* or N. Affara, +223/333–700, Fax: -346, *na@mbuc.bio.cam.ac.uk*]

19–22. Data Banks and Comput. Support of Hum. Genome Proj.; Moscow [V. Tsitovich, 7-095/135-2311, Fax: -1405, *imb@imb.msk.su*]

22–24. 2nd Single Chromosome Workshop on Hum. Chromosome 1; Vienna [A. Weith, +43-1/797-30-625, Fax: /798-7153, *weith@aimp.una.ac.at*]

29–Oct. 1. 1st Intl. Chromosome 10 Workshop; Crete, Greece (abs. deadline: July 31) [J. Mao, 617/893-5007 ext. 242, Fax: /642-0310, *mao@cric.com*]

October 1995

13–14. Workshop on Gene-Finding and Gene Struct. Prediction; Philadelphia (abs. deadline: May 1) [D. Searls, 215/573-3107, Fax: -3111, *dsearls@cbil.humgen.upenn.edu*]

22–23. Chromosome 3 Workshop; Minneapolis [S. Naylor, 210/567-3842, Fax: -6781, *naylor@thorin.uthscsa.edu*]

24–28. ASHG 95; FASEB, Minneapolis [see contact: Apr. 9–13]

29-Nov. 1. 1995 Natl. Soc. of Genet. Counselors 14th Annu. Educ. Conf.; Minneapolis [B. Leopold, 610/872-7608 voice mailbox #6, Fax: -1192] ◊

*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*

1995 Spring and Summer Gordon Res. Conf.—Frontiers of Sci. [Specific locations and information: contact GRC, 401/783-4011, Fax: -7644, *grc@grcmail.grc.uri.edu*, Gopher to *hackberry.chem.niu.edu* (port 70), WWW: *http://hackberry.chem.niu.edu:70/0/webpage.html*, or anonymous ftp: *hackberry.chem.niu.edu* (*pub/Conferences/Gordon Conferences* directory)]

April 1995

18–21. Basic Tech. in Mol. Mycobacteriology; Rockville, MD [ATCC, 301/231-5566, Fax: /770-1805]

24–28. Recombinant DNA Tech. I; Germantown, MD (also offered June 12–16) [LTI, L. Kerwin, 800/952-9166, Fax: 301/258-8212]

May 1995

1–3. Nonlinear Optics of Bulk Media: Prin. and Appl.; Los Angeles [UCLA, 310/825-1047, Fax: /206-2815, *mhenness@unex.ucla.edu*]

1–5. Anal. of Gene Expression; LTI, Germantown, MD (also offered Aug. 7–11) [see contact: Apr. 24–28]

8–12. Recombinant DNA Methodol.; Washington, DC (also offered July 10–14) [CATCMB/CUA, M. Miller, 202/319-6161, Fax: -4467, *millerm@cua.edu*]

13–15. PCR Tech.; CATCMB/CUA, Washington, DC [see contact: May 8–12]

15–19. Biotechnol. for Bus.; Durham, NC (reg. deadline: Mar. 15) [B. Blackburn, 919/660-1579, Fax: -1591]

15–19. DNA Protein Interactions; LTI, Germantown, MD [see contact: Apr. 24–28]

16–19. DNA Sequencing; CATCMB/CUA, Washington, DC [see contact: May 8–12]

22–26. PCR Tech.; LTI, Germantown, MD (also offered July 10–14) [see contact: Apr. 24–28]

29–June 16. Patterns in Biol. Sequences; Aspen, CO (appl. deadline: Feb. 15) [C. Burks (505/667-6683, Fax: /665-3493, *cb@t10.lanl.gov*]

31–June 3. Site-Directed Mutagenesis; CATCMB/CUA, Washington, DC [see contact: May 8–12]

31–June 7. **Med. Informatics; Woods Hole, MA (appl. deadline: Mar. 14) [MBL, 508/548-3705 ext. 401, Fax: /457-1924, *admissions@mbl.edu*, *http://www.mbl.edu*]

June 1995

4–9. **Nucleic Acid and Protein Sequence Anal. Workshop; Pittsburgh (appl. deadline: Apr. 17) [PSC, N. Blankenstein, 412/268-4960, *biomed@psc.edu*]

4–9. Recombinant DNA Technol.; Bloomington, IN (reg. deadline: May 19) [J. Clay, 812/855-6329, Fax: -8997, *jclay@indiana.edu*]

5–9. In Situ Hybridization Tech.; LTI, Germantown, MD (also offered Aug. 21–25) [see contact: Apr. 24–28]

9–29. Adv. Bacterial Genet. ; Cold Spring Harbor, NY (appl. deadline: Mar. 15) [CSHL, 516/367-8346, Fax: -8845, *meetings@cshl.org*, *http://www.cshl.org*]

11–14. Genomic Info.: Ethical Implications; Seattle (appl. deadline: April 11) [B. Brownfield, 206/616-1864, Fax: /685-7515, *brbrownf@u.washington.edu*]

11–16. Appl. of Recombinant DNA Technol.: RFLP and Fingerprinting Anal., RAPD Anal., and DNA Sequencing; Bloomington, IN (reg. deadline: May 19) [see contact: June 4–9]

11–16. **GRC—Frontiers of Science: Nucleic Acids; New Hampton, NH [GRC, 401/783–4011, Fax: -7644]

11–July 22. **Physiology: Cell. and Mol. Biol.; MBL, Woods Hole, MA (appl. deadline: Mar. 15) [see contact: May 31–June 7]

12–16. **Intro. Linkage Course; New York [box, p. 13]

14–16. Intro. to Mol. Cytogenet.; Gaithersburg, MD (also offered Aug. 2–4) [Oncor, Inc., 800/556-6267, Fax: 301/926-6129]

19–23. Expression of Recombinant DNA in Mamm. Cells; CATCMB/CUA, Washington, DC [see contact: May 8–12]

19–24. Hum. Genome Networking Proj. for Mid. and Sec. Sci. Teachers; Kansas City, KS [D. Collins, 913/588-6043, Fax: -3995,*collins@ukanvm.cc.ukans.edu*, *http://kumchttp.mc.ukans.edu:80/instruction/medicine/genetics/homepage.html*]

25–28. **Structure Determination from NMR; PSC, Pittsburgh (appl. deadline: Apr. 28) [see contact: June 4–9]

26–30. In Situ Hybridization; CATCMB/CUA, Washington, DC [see contact: May 8–12]

*Extended calendars are available at *http://www.ornl.gov/TechResources/Human_Genome/home.html* or from HGMIS (see p. 12 for contact information).

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 due dates: [1] February 1, June 1, and October 1; [2] April 5, August 5, and December 5; [3] January 10, May 10, and December 10: [4] on continuous basis.

Program Categories

Program announcements listed in *NIH Guide for Grants and Contracts* (*gopher.nih.gov* and *http://www.nih.gov* or 301/496-0844).

Research

- Ethical, legal, and social implications (ELSI) of human genome research, Fellowships (PA 92-21) [1].
- Genome science and technology centers (PAR 94-044) [1].
- Informatics (PA 92-59) [1].
- New and improved technologies for genomic research and analysis (PA 94-045) [1].
- Pilot projects or feasibility studies for genomic analysis (PAR 94-046) [1].

Training

- Courses related to genomic analysis (PA 91-88) [1].
- Individual postdoctoral and senior fellowships in genomic analysis and technology (PA 92-21) [2].
- National research service awards: Institutional training grants in genomic science (PA 94-085) [3].
- Special emphasis research career awards in genomic research (PA 91-89) [1].

Special Programs

- International genome research program for Central and Eastern Europe (PA 92-67) [1], Fellowships (PA 92-68) [2].
- Minority institution travel awards (PA 91-17) [4]
- Research supplements for underrepresented minorities and disabled [4].

NCHGR Staff: 301/496-7531, Fax: /480-2770.

- ELSI: *Elizabeth_Thomson@nih.gov* or 301/402-4997.
- Genetic linkage mapping, annotation, and single-chromosome workshops: *Elise_Feingold@.nih.gov*
- Informatics: *David_Benton@nih.gov*
- Large-scale mapping, sequencing of human and mouse genomes: *Jeff_Schloss@nih.gov*
- Physical mapping technology, training, and special programs: *Bettie_Graham@nih.gov*
- Sequencing technology development, technology transfer, nonmammalian model organisms: *Carol_Dahl@nih.gov* or *Robert_Strausberg@nih.gov*

DOE Human Genome Program

- Contact for funding information or general inquiries: *genome@er.doe.gov* or 301/903-6488.
- Relevant documents (ftp: *oerhp01.er.doe.gov* in directory */genome* or *http://www.er.doe.gov/production/oher/hug_top.html*).

DOE Human Genome Distinguished Postdoctoral Fellowships

Next deadline: February 1, 1996.

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

Small Business Innovation Research (SBIR) Grants

DOE and NIH invite small business firms (less than 500 employees) to submit grant applications addressing the human genome topic of SBIR programs. The two agencies also support the Small Business Technology Transfer (STTR) program to foster transfers between research institutions and small businesses. Contacts:

- Kay Etzler; c/o SBIR Program Manager, ER-16; DOE; Washington, DC 20585 (301/903-5867, Fax: -5488). DOE SBIR due March 1, 1996; STTR, early 1996.
- Bettie Graham (see contact, NCHGR). NIH SBIR due April 15, August 15, and December 15. STTR, December 1.

National SBIR/STTR conferences: Chicago, IL (April 26–28); Washington, DC (October 16–18); Salt Lake City, UT (October 30–November 1); Dallas, TX (April 29–May 1, 1996). Conference hotline: 407/791-0720; electronic registration: 203/379-9427.◊

Human Genome Management Information System Subscription/Document Request* (Vol. 6, No. 5)

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.

SELECTED ACRONYMS

ACEDB A *C. Elegans* database

ASHG Am. Soc. of Hum. Genet.

BAC bacterial artificial chromosome

CATCMB/CUA Ctr. for Adv. Train. in Cell and Mol. Biol./Cathol. Univ. Am.

CCD charge coupled device

CEPH Centre d'Etude du Polymorphisme Humain

CHI Cambridge Healthtech Inst.

CSHL Cold Spring Harbor Lab.

DDBJ DNA Data Bank of Japan

DIMACS Discrete Math. & Comp. Sci.

EBI European Bioinformatics Institute

EST expressed sequence tag

FASEB Fed. of Am. Soc. for Exp. Biol.

FISH fluorescence in situ hybridization

IBC Intl. Bus. Comm.

IGES Intl. Genetic Epidemiology Soc.

LTI Life Technologies, Inc.

MBL Maine Biological Lab.

MIMBD Meet. on the Interconnection of Mol. Biol. Databases

NYAS NY Acad. of Sci.

PAC P1 artificial chromosome

PCR polymerase chain reaction

PSC Pittsburgh Super Computing Ctr.

SBIR Small Bus. Innovation Res.

TIGR/NIST The Inst. for Genomic Res./Natl. Inst. of Standards and Technol.

STS sequence tagged site

WWW World Wide Web

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