Human Genome news

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

ISSN: 1050-6101 2285

Vol. 5, No. 1, May 1993

DOE Holds Contractor-Grantee Workshop To Assess Progress

Over 400 researchers, program managers, and invited guests gathered February 7–10 in Santa Fe, New Mexico, for the Third Contractor-Grantee Workshop, sponsored by the DOE Human Genome Program. The meeting was held to review current research and assess the progress and direction of the genome program.

In his welcoming remarks David Galas, DOE Associate Director for the Office of Health and Environmental Research, recognized mapping contributions made by researchers worldwide and said that major changes now in the offing may alter the program, particularly in the area of sequencing.

U.S. Senator Pete Domenici (R-NM), a major congressional supporter of the project, told attendees that genome researchers "stand poised to rewrite the health care delivery system, and those in policy positions need to be both helpful to and demanding of the Human Genome Project." He stressed the importance of developing policies on ethical, legal, and social issues (ELSI) related to genomics and agreed with Galas that researchers and policymakers need to communicate on these topics. Domenici also urged scientists to establish closer contact with the U.S. marketplace to facilitate the application of new discoveries.

The meeting was organized by Sylvia Spengler [Lawrence Berkeley Laboratory (LBL)], the DOE Human Genome Coordinating Committee, and the DOE Human Genome Program staff. Two comprehensive informatics resource rooms containing a local area network with Internet access were expertly set up and maintained by Claudia Sanders and other Los Alamos National Laboratory (LANL) computing staff.

Plenary sessions included ELSI topics, sequencing, informatics, and chromosome

Workshop participants (l-r) Jane Lamerdin and Anne Olsen (both from LLNL) discuss research results at the 1993 DOE Contractor-Grantee Workshop in Santa Fe, New Mexico.

Major Changes Projected for Sequencing Technology

In This Issue. . .

Genome News

and cDNA mapping and sequencing. All projects were represented in informative poster sessions during the meeting.

The number of funded projects has almost doubled since the last workshop in February 1991, reflecting the rapid growth of the genome program. Almost 200 projects were represented by investigators from DOE human genome centers at LBL, Lawrence Livermore National Laboratory (LLNL), and LANL; other DOE-supported laboratories; and more than 40 universities and research organizations. Some of the presentations are highlighted below.

ELSI Presentations

Michael Yesley (LANL) described the direction of the DOE ELSI program and the issues being explored, including availability and disclosure of genetic information and the possible effects of such knowledge. Joe McInerney (Colorado College), Troy Duster [University of California, Berkeley (UCB)], and Philip Reilly (Shriver Center) addressed education and policy-oriented aspects of DOE research.

McInerney discussed the objectives of the Biological Sciences Curriculum Study module, which was widely distributed to U.S. high school biology teachers. He explained that the program attempts to minimize the resurgence of biological determinism by focusing on interactions between environmental and genetic factors. The program also emphasizes the role played by ethics in the health sciences.

Duster described the assimilation of genetic information by consumers, who often attempt to process it along with their own sets of long-held beliefs. He said that their nongenetic explanations for disorders will not be discarded until more-complete explanations are available.

Reilly discussed genetic information and its potential availability to life and health insurers through databases containing medical records; insurers are not yet requiring routine genetic testing because of cost. He challenged the group to become more involved in ELSI issues and public policy matters and not to permit others to misinterpret discoveries made by individuals in the group.

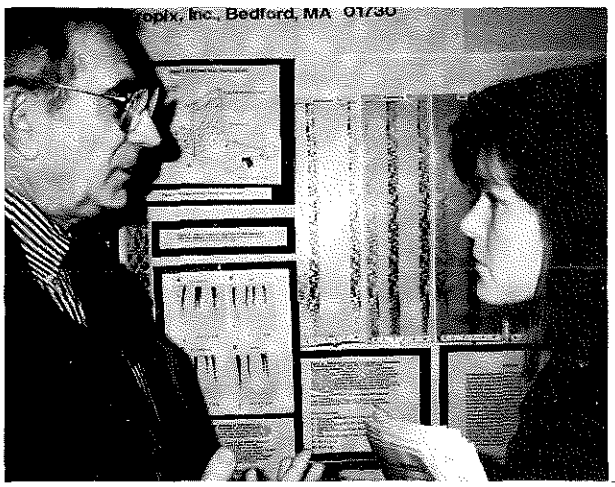
Sequencing Presentations

Lloyd Smith (University of Wisconsin, Madison) opened the session by observing that complete sequencing of the human genome by 2006 will require a 500-fold rate increase (to 500 Mb/yr) and a 5-fold cost decrease over the next 7 years to meet large-scale sequencing goals. This observation led to a lively discussion of sequencing issues.

Problems of scaling up chromosome sequencing rates were discussed by several speakers. William Studier (Brookhaven National Laboratory) proposed the use of primer walking to assemble elementary sequence reads into extended sequence; the approach would eliminate subcloning and assembly steps required by other strategies [*HGN* **4**(5), 1–4 (January 1993)]. Challenges in developing this approach include analyzing the priming capacity of hexamer strings and integrating an appropriate detection system.

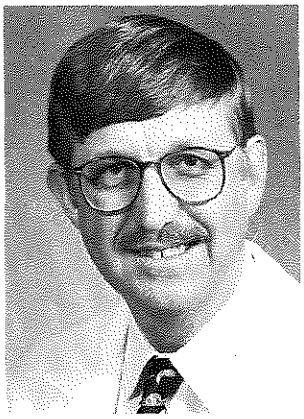
Michael Palazzolo (LBL) and Robert Weiss and Raymond Gesteland (University of Utah) are generating preliminary dense maps to guide sequence assembly. Palazzolo described the production of "DOG-tag" physical maps, in which the distance (D) and orientation (O) of each gene-sized (G) sequencing template (about 3 kb from P1 clones) are known; each is tagged by a sequence tagged site (STS) from each end. No sequence assembly is required. The Utah group is developing automated processes to map transposon-based multiplex priming sites on large-insert (20 kb) cloned DNA.

Promising instrumentation includes automated colony pickers and semiautomated systems that produce data on fragment length for map construction [Joseph Jaklevic (LBL)] and capillary and ultrathin gel electrophoresis systems [Smith, Richard Mathies (UCB), Norman Dovichi (University of Alberta), Edward Yeung (Iowa State University), Barry Karger (Northeastern University)]. Mathies described his system in which fragments are separated on capillary arrays and distinguished by a binary coding scheme using only two different fluorescently labeled dye primers to identify four sets of fragments. He predicted a



Paul Silverman (Beckman Instruments) and Corinne Olesen (Tropix, Inc.) with her poster presentation at the DOE Contractor-Grantee Workshop in Santa Fe, New Mexico.

NIH Names Francis Collins Director of NCHGR



Francis S. Collins

In April NIH Director Bernadine Healy appointed Francis S. Collins to succeed James Watson (President, Cold Spring Harbor Laboratory) as Director of the National Center for Human Genome Research (NCHGR). Michael Gottesman (National Cancer Institute) served as Acting Director in the interim since April 1992.

Healy said, "Dr. Collins will bring to NIH world-class talent and experience in human genetics research. He will provide outstanding leadership to the Human Genome Project as well as critical resources to the NIH community of researchers who are pursuing the genetic basis of human illness."

Collins comes to NCHGR from the University of Michigan Medical Center in Ann Arbor, where he was a Howard Hughes Medical Institute investigator and Director of the NCHGR-supported human genome center. Collins' work at Michigan focused on developing large-scale technologies to identify genes responsible for human illnesses. He is noted for pioneering positional cloning to pinpoint gene location through the study of disease-inheritance patterns. Collins is a codiscoverer of the cystic fibrosis, neurofibromatosis type 1, and Huntington's disease genes. He and his team are currently pursuing genes for early-onset breast cancer and a common form of adult leukemia.

As NCHGR Director, Collins will oversee the center's two divisions. The Division of Extramural Research will continue to administer the NIH component of the Human Genome Project, funding research throughout the country in chromosome mapping; DNA sequencing; database and technology development; and studies of the ethical, legal, and social implications of the availability of genetic data. Collins will also head the Division of Intramural Research, which will focus on finding disease genes, developing DNA diagnostics, and exploring gene therapies. The intramural division will also serve as a hub for NIH-wide human genetics research and enhance the work of investigators in other institutes who are searching for specific genes and studying gene function in health and disease.

Born in Staunton, Virginia, Collins received his bachelor's degree with highest honors from the University of Virginia and M.S. and Ph.D. degrees in physical chemistry from Yale University. He was awarded an M.D. degree from the University of North Carolina School of Medicine and completed his internship and residency in internal medicine at the North Carolina Memorial Hospital. He was a fellow in human genetics and pediatrics at Yale from 1981 to 1984, after which he joined the department of internal medicine and human genetics at Michigan.

Collins is a diplomate of the American Board of Internal Medicine, the American Board of Medical Genetics, and the American College of Medical Genetics. Elected to the Institute of Medicine in 1991 and recently to the National Academy of Sciences, he is a member of the boards of directors of several societies and associate editor for a number of publications. Collins has received numerous awards and honors for his work.◊

throughput of 200 kb per day of raw sequence in about 1 year, running 100 capillaries in parallel.

DNA chain breakage occurring at guanidine residues during DNA vaporization by laser-driven mass desorption was identified as a problem in developing mass spectroscopy as a very fast sequencing tool. These results are stimulating the search for less-harsh vaporization procedures [Klaus Schneider and Brian Chait (Rockefeller University), Winston Chen (Oak Ridge National Laboratory {ORNL}), David Schieltz and Peter Williams (Arizona State University)].

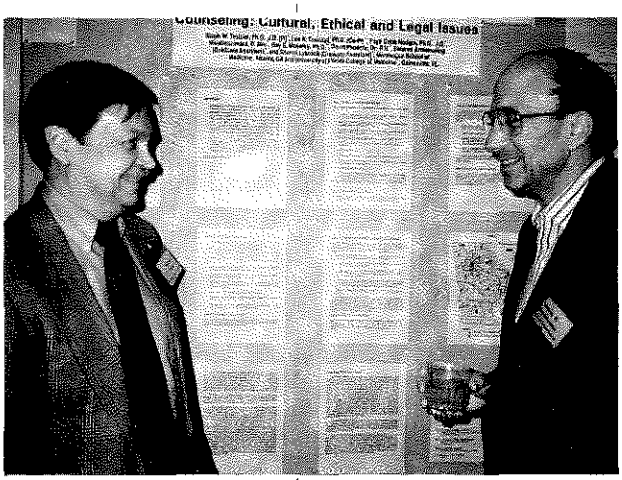
The production of regular anchored arrays of oligonucleotides has allowed the sequencing-by-hybridization technology to begin efficacy testing [Stephen Fodor (Affymax), Robert Foote (ORNL), Michael Pirrung (Duke University)]. Fodor reported using light-directed chemical synthesis to fabricate high-density probe arrays and the construction of partial sequencing chips for probing, sequence checking, and mutation screening [*HGN* **4**(5), 3–4 (January 1993)].

Informatics Presentations

Chair David Kingsbury (Johns Hopkins University) commented on the 1992 review of DOE-funded informatics projects and recommendations by a group of genome informatics experts. The purpose of the review was to evaluate the (1) balance of support provided by informatics programs to various genome centers and the genomics community, and (2) responsiveness of the informatics program to changes

Informatics Oversight Group Identifies Need for Central Direction, Enhanced Communication

Genome News



Above (l-r): Ralph Trottier (Morehouse School of Medicine) and Harvey Mohrenweiser (LLNL) exchange views on ELSI issues during the DOE Santa Fe Workshop poster session.

At right (l-r): Leroy Hood (University of Washington) and Ray Gesteland (University of Utah) with meeting organizer Sylvia Spengler (LBL).

in mapping and sequencing. The oversight group identified the need for central direction and enhanced communication among developers and consumers of informatics tools. Cohousing bench scientists and informatics groups was recommended for better integration of informatics activities with centers. The group also recommended that guidelines for review of informatics proposals be developed.

Kingsbury noted that the reviewers found many impressive off-site projects and applauded DOE for supporting activities having broad, long-term views of genome project informatics needs.

Several researchers described or demonstrated software for mapping and sequencing analysis [Michael Cinkosky (LANL), Ed Thiel (LBL), Ed Uberbacher (ORNL)]. Hands-on demonstrations of popular software were available throughout the meeting. Tom Marr [Cold Spring Harbor Laboratory (CSHL)] described an interactive, unified database that will link Genome Data Base (GDB), the Jackson Laboratory mouse database, FlyBase, ACEDB for *Caenorhabditis elegans*, CSHL Fission Yeast database, and others. The object-oriented computer program uses an organism-independent underlying data model to unify data across eukaryotic organisms. Marr explained that this will help researchers who study gene function by analyzing conserved genes.

A recurrent theme in the informatics session was the need to improve map representation and define map types needed by investigators. Ken Fasman (GDB) proposed a three-tiered organization of map information: (1) "nascent" maps containing low-level map data, (2) individual observed maps constructed from nascent maps, and (3) consensus maps containing data about the whole population. Setting up such maps will require coordinate systems to describe genome locations and provide two resolution levels: (1) global coordinates (reference markers representing order and genetic distance) overlaid by (2) local coordinates pinned to local landmarks (defined relative to markers at the base-pair level) representing genomic variation.

Increasing sequence-analysis speed will become more critical as the rate of data doubling increases. GenBank® now receives 3 Mb/yr, with the rate of data receipt doubling every 15 months. This increase presents challenges for managing data output and improving processing speed. Manfred Zorn (LBL) discussed a toolkit [parallel object-oriented environment and toolkit (POET)] that uses existing software to exploit the power of parallel processing and accomplish sequence analysis more quickly. The output display module shows sequence matches and provides a graphical user interface, while hiding implementation details.

Mapping Presentations

Groups led by Anthony Carrano at LLNL and Robert Moyzis at LANL are approaching closure of chromosomes 16 and 19. Many talks focused on vector development, with bacterial artificial chromosomes (BACs) representing a major new resource. With an average insert size of 200 kb, BACs are some 4 to 5 times larger than cosmids, have no detected instability and chimerism problems, and are much easier to purify than yeast artificial chromosomes (YACs) [Bruce Birren (California Institute of Technology), Pieter de Jong (LLNL)]. In describing the large-insert

cloning systems at LANL, MaryKay McCormick noted the extremely low frequency of chimeras (0 to 10%) found in the YAC libraries for chromosomes 5, 9, 16, and 21 and in the total genomic library.

Glen Evans (Salk Institute) indicated that chromosome 11 will be sequenced directly from cosmid templates with no subcloning. Evans described his group's plans to complete low-resolution YAC contig maps of this chromosome and begin assembling high-resolution clone maps as a prelude to sequencing. The Salk team is collaborating with Harold Garner (General Atomics) to develop high-throughput robotic and informatics tools that address known bottlenecks.

Reagents and technologies developed in the Human Genome Project are now used for human-disease analysis. Daniel Pinkel (University of California, San Francisco) discussed medical diagnostic applications of hybridization technologies, probes, and digital microscopy for clinically important problems such as prenatal and neonatal disorders and cancer. The chromosome painting technology of Pinkel's group was transferred to industry 2 years ago. The team recently developed a method using comparative genomic hybridization to measure chromosomal deletions and duplications in tumor cells and provide a "copy number karyotype."

cDNA Presentations

Sequences from several thousand cDNAs are now available for generating STSs [James Sikela (University of Colorado Health Sciences Center), J. Craig Venter (The Institute for Genomic Research)]. cDNA mapping is proceeding by fluorescence in situ hybridization (FISH) [Julie Korenberg (Cedars-Sinai Medical Center), Joseph Gatewood (LANL)] and polymerase chain reaction analyses on somatic hybrid panels [Mihael Polymeropolous (National Institute of Mental Health), Donna Maglott and William Nierman (American *Type Culture* Collection)]. Sikela is using a two-step cDNA mapping approach, in which he assigns cDNAs to gridded YACs and maps the YACs by FISH. Massively parallel oligomer fingerprinting of cDNA clones [Radomir Crkvenjakov and Radoje Drmanac (Argonne National Laboratory)] is providing a means for rapidly distinguishing cDNA clones and identifying related cDNAs and uncharacterized cDNAs. Mapping of cDNAs, which is preceeding more slowly than cDNA sequencing, continues to be challenging. *[Reported by Denise K. Casey (HGMIS, ORNL) and Marvin Stodolsky (DOE Human Genome Program)]* ◊

Gene for Huntington's Disease Discovered

An international research group composed of 6 teams of researchers affiliated with 11 institutions has identified the gene that causes Huntington's disease (HD). As reported in the March 26 issue of *Cell*, the Huntington's Disease Collaborative Research Group used cloned "trapped exons" to isolate the gene IT15. The collaborative group was formed after James Gusella's team at Massachusetts General Hospital (MGH) mapped the gene in 1983 to the distal portion of the short (p) arm of chromosome 4; it was one of the first uses of DNA marker technology for this purpose.

"The cooperative effort was important," said Gusella, whose laboratory also pinpointed the HD gene. "This has been a very, very hard mystery to solve. If the separate research teams had been competing rather than cooperating, the search would have gone on a lot longer." About a year before the gene was discovered, the MGH group took a considered risk in concentrating on one promising 500-kb segment while the rest of the consortium searched elsewhere in the region.

As in the case of fragile X syndrome and myotonic dystrophy, the HD mutation involves a polymorphic trinucleotide repeat, an unstable DNA segment in which the sequence is copied many more times than normal. This defect was found in all 75 HD families studied. In people without the disease the number of repeats ranges from 11 to 34; those affected by HD showed a minimum of 42, with an estimated 100 copies in a severely affected patient. Preliminary evidence indicates that the number of repeats may relate to the severity of the disease and the age at which it becomes apparent: the shorter the repeat, the older the individual when symptoms first appear; the longer the repeat, the earlier the onset of symptoms.

HD, a fatal progressive disease that attacks both mind and body by killing brain cells, is usually manifested by the age of 35 with small involuntary movements (chorea) that gradually overwhelm all parts of the body. HD is also characterized by cognitive decline leading to dementia and psychiatric manifestations. Both sexes are equally affected, and each offspring of an affected person has a 50% chance of inheriting the disease, for which no effective treatment is known. About 1 in 25,000 Americans carries the gene, and another 150,000 are at risk. Rare spontaneous mutations can also occur in individuals whose parents are not affected.

In addition to the MGH team, the Huntington's collaboration included groups led by Hans Lehrach of the Imperial Cancer Research Fund (U.K.); David Housman at Massachusetts Institute of Technology; John Wasmuth of the University of California, Irvine; Francis Collins at the University of Michigan at Ann Arbor; Peter Harper of the University of Wales College of Medicine (U.K.); and others. The work was supported by grants from NIH and a number of other foundations and institutions.

"The search for the Huntington's disease gene has been the most difficult gene hunt yet," said Collins. "Its success now means that accurate diagnosis of Huntington's disease will be available almost immediately, the basic biology of the disease can at last be understood, and the potential for new treatments can be vigorously pursued." ◊

Genome News

YAC Germ Line Transmission Achieved

Investigators at the Whitehead Institute for Biomedical Research have successfully introduced a yeast artificial chromosome (YAC) into the mouse genome and demonstrated that the chromosome and gene expression were transmitted in two "transgenic" mouse strains. This achievement was reported in the March 26 issue of *Science* (**259**, 1904–08).

A team headed by Rudolf Jaenisch, of the Whitehead genome center, inserted YAC DNA carrying a normal mouse collagen gene into embryonic stem (ES) cells derived from very early mouse embryos lacking a functional collagen gene. ES cells have the potential to develop into any type of cell in the body. Injecting the altered ES cells into intact host embryos produced chimeras (mixed animals), some of whose offspring carried functional YAC DNA in every cell.

The ability to transfer physically intact YACs into the germ line will permit the identification of regulatory sequences that control gene function. Many early gene-therapy experiments failed in mice because investigators could not control the level or timing of gene expression. With YAC technology, scientists can design a series of gene-transfer experiments to identify the combination of genetic elements most likely to produce acceptable levels of gene expression. Once the regulatory elements have been identified, human counterparts can be incorporated into vectors for human gene therapy.

YAC vectors have the advantage of being able to carry pieces of foreign DNA more than 20 times larger than those incorporated by conventional cloning vectors, which are limited to genes of less than 50 kb. This limitation impeded the development of accurate disease models because some human disease genes, such as those for neurofibromatosis, muscular dystrophy, and hemophilia, approach 100 kb.

The new YAC technology can also be used to pinpoint new disease genes whose approximate locations have been determined by gene mapping. Gene function and precise location can be studied in mice by inserting all or part of the gene-carrying DNA fragment and observing the results.

Jaenisch said that the ability to introduce YAC clones into the mouse germ line will allow quick and efficient application of new information from the Human Genome Project to the study of human disease. ◊

Other Groups Achieve Similar Results

Insertion of YACs has also been reported by Cell Genesys Inc. (Foster City, California) in the March 18 issue of *Nature* (362, 255–58) and by a group led by T. Choi (in press at *Nature Genetics*).

EMBL Introduces BLITZ E-Mail Server

The European Molecular Biology Laboratory (EMBL) and the Biocomputing Research Unit at the University of Edinburgh have introduced BLITZ, a new e-mail server for comparing newly determined protein sequences with the SWISS-PROT protein sequence database. BLITZ is based on the MPsrch program developed by Shane Sturrock and John Collins (University of Edinburgh). MPsrch is the fastest implementation of the exhaustive local similarity algorithm for sequence database searching. It runs on the 4096-processor MasPar MP-1 machine at EMBL, using the Smith & Waterman algorithm and performing up to 160 million cell updates per second.

For detailed help, send the mail message *help* to the Internet address *blitz@embl-heidelberg.de*. To run a search using default-weight matrix and INDEL cost settings, send an e-mail message with the word *seq* on the first line and the sequence in free format on succeeding lines. The end of the sequence is denoted by ending the mail message or by inserting a new line beginning with the word *end*. A similar service for nucleotide database searching is forthcoming.

For further information, contact EMBL Data Library; Postfach 10.2209; W-6900 Heidelberg, Germany (Int. 49/62-21-387258, Fax: -387519; Internet: *datalib@embl-heidelberg.de*). For information on MPsrch, contact John Collins; Biocomputing Research Unit; University of Edinburgh; U.K. (Int. 44/31-650-5365, Fax: /668-3870; Internet: *jfc@biocomputing.edinburgh.ac.uk* or *mpsrch_help@biocomp.ed.ac.uk*). ◊

NIH Advisory Groups Seek To Merge

Since the initiation of the genome program, the NIH National Center for Human Genome Research has operated with two advisory bodies. They were (1) the Program Advisory Committee on the Human Genome (PAC or PACHG), which focused on strategic planning, and (2) the National Advisory Council for Human Genome Research (NACHGR), which was charged with reviewing grants and making funding recommendations. To achieve closer coordination of these increasingly interrelated functions, the two committees will merge into NACHGR. NACHGR, which meets in January, May, and September, will continue to have working groups and subcommittees as needed.

To facilitate interaction between NIH and DOE, a proposed charter amending NACHGR calls for a joint coordinating committee composed of members of NACHGR and the DOE Health and Environmental Research Advisory Committee.◊

GDB Adds Nodes in Sweden and the Netherlands

To increase Genome Data Base (GDB) accessibility outside the United States, additional nodes have been added in Sweden and the Netherlands. These nodes offer database and user support services equivalent to those available from GDB in Baltimore.

SWEDEN: The Swedish node is run by the Swedish Medical Research Council Genome Initiative, located at the Biomedical Center of Uppsala University.

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

NETHERLANDS: GDB service is made possible with the financial support of the Netherlands Organization for Scientific Research. The node is run by the CAOS/CAMM Center (Dutch National Expertise Center for Computer-Assisted Chemistry and Bioinformatics). Services are intended primarily for users in the Netherlands, but international accessibility is provided via Internet.

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

2290

GDB Changes Submission Forms, Provides Electronic Templates

GDB recently made important changes in paper data-submission forms:

- replacement of the D-Segment Submission Form with the Cloned Reagent Submission Sheet to simplify submission of data associated with mapping reagents and D-segment assignment, where applicable;
- introduction of the Contig Submission Form to assist physical mapping laboratories in transmitting data to GDB; and
- slight changes in the PCR Submission Form to reflect modifications in GDB 5.0.

Many investigators recognize the advantages of electronic submissions (via the Internet or diskettes) over paper submission forms: (1) data enters the database much more quickly and (2) laboratories that have repetitive information for submitted elements can enter common conditions in the template, which can be copied for each element. GDB provides a suite of templates that can be imported into word processors, spread sheets, and text editors.

Electronic templates and postscript files for paper submission forms are available through the GDB anonymous ftp server in the gdb/submit-data directory [*mendel.welch.jhu.edu* (128.220.59.42)]. After logging in, requestors should type *anonymous* at the login prompt and their e-mail address at the password prompt. Mac or PC diskettes and copies of paper forms can be requested. [Contact: GDB Data Maintenance and Acquisition Core; Johns Hopkins University School of Medicine; 2024 E. Monument Street; Baltimore, MD 21205-2100 (410/955-9656, Fax: /614-0434, Internet: *data@library.welch.jhu.edu*).] ◊

GDB USER SUPPORT, REGISTRATION

To become a registered user of GDB and OMIM, contact one of the User Support offices listed at right (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.

GDB and OMIM Training Schedule

Comprehensive hands-on training courses on the use of GDB and OMIM will have at least one computer workstation for two participants. Registrants will receive at least 3 weeks notice if insufficient registration causes class cancellation.

- The general course for scientific users provides a basic understanding of the databases and relationships among different types of data.
- The course for users with editing privileges includes instructions on adding, modifying, and deleting GDB data.

Class frequency and location will be determined by demand (schedule below). Courses are free, but attendees must pay their own travel and lodging expenses. Hotel information and directions will be mailed with registration materials.

As interest in GDB continues to grow, organizations around the world will offer training that requires access to GDB in Baltimore. Notifying GDB User Support about planned training activities will enable the staff to ensure database availability by scheduling maintenance and repairs at other times.

COURSE REGISTRATION INFORMATION

Contact U.S. GDB User Support Office (at top right).

COURSE SCHEDULE

General User: June 21–22, Baltimore.

USER SUPPORT OFFICES

United States

GDB User Support
Applied Research Laboratory
William H. Welch Med. Library
Johns Hopkins University
2024 E. Monument Street
Baltimore, MD 21205-2100
410/955-7058, Fax: 410/614-0434
Internet: *help@welch.jhu.edu*

The Help Line is staffed from 9 a.m. to 5 p.m. EST for information on accounts and training courses, technical support, and data questions. Calls received after hours will be forwarded to the appropriate voice mail and returned as soon as possible. To obtain a user's local SprintNet (Telenet) number for locations within the United States: 800/736-1130.

United Kingdom

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

Germany

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

Australia

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

Meeting Reports

- Chromosome 13
- Chromosome 22
- Computational Molecular Biology
- Genome Sequencing and Analysis
- Human Genome '92
- Chromosome 18

Because of limited space, *HGN* is unable to print full reports of six meetings held since July 1992. Citations for the proceedings of two meetings and excerpts from four reports are shown below . For more information, contact HGMIS or the report's author.

✦ ✦ ✦

✦"Report of the First International Workshop on Human Chromosome 13 Mapping"

September 21–22, 1992, Dallas, Texas
Cytogenetics and Cell Genetics **62**(2–3), 89–107 (1993).

✦"Proceedings of the Third International Workshop on the Mapping of Chromosome 22"

September 17–20, 1992, Philadelphia
Genome Priority Reports **1**, 574–87 (1993) (Karger Publishing, Farmington, Connecticut)

✦*Second International Workshop on Open Problems in Computational Molecular Biology*

July 19–August 2, 1992, Telluride, Colorado

ORGANIZERS: Andrzej Konopka [National Cancer Institute (NCI)] and Peter Salamon [San Diego State University (SDSU)].

COORDINATOR: Danielle Konings (University of Colorado at Boulder).

SPONSOR: DOE Human Genome Program.

COSPONSORS: CONVEX, SUN, GCG, IntelliGenetics, MasPar, and SiliconGraphics.

FOCUS: Computational experiments in molecular biology and computational methods to classify large amounts of sequence data.

SELECTED PRESENTATIONS AND DISCUSSIONS

New Research Results: Computational analysis of the evolution of G-proteins; possible roles of tandem repeats polymorphism in recombination; neighbor-dependent mutation rates and their relationship to the bias toward low-complexity sequences in eukaryotic genomes; and maximum entropy relationships in nucleotide and protein sequences.

Research Approaches to Molecular Biology Software Development: Methods for automated sequence annotation and database organization; segmenting nucleotide and protein sequences; determining biological (as opposed to statistical) significance of oligopeptide and oligonucleotide matches from comparisons of "all-against-all" database entries; computational techniques to determine genome organization; protein motif representation that takes into account conservation of structure and/or function; formalization of heuristic protocols for genome map assembly; and progress in predicting protein structure from sequence data.

Methodology of Computational Molecular Biology: Machine metaphors and the methodological consequences for sequence research; logical aspects of inductive inference from incomplete, inexact, and inaccurate data; language metaphors that take into account three-dimensional conformations of meaningful texts; paradoxes resulting from simultaneous use of language and mechanism metaphors for DNA sequences; specification of database properties leading to maximum data integrity; and factors contributing to the paradigm of computational molecular biology. *[Reported by Andrzej K. Konopka (NCI) and Peter Salamon (SDSU)]* ◊

The Third International Workshop on Open Problems in Computational Molecular Biology will be held July 12–26 [contact: A. Konopka, 301/846-5396].

Peer-reviewed papers from the 1992 workshop will appear in a special issue of *Computers and Chemistry* (Vol. 17, 1993); papers from the first (1991) workshop were published in the journal's April 1992 [16(2)] issue.

✦*Genome Sequencing and Analysis Conference IV*

September 26–30, 1992, Hilton Head, South Carolina

ORGANIZERS: J. Craig Venter [The Institute for Genomic Research (TIGR)] and C. Thomas Caskey (Baylor College of Medicine).

SPONSORS: NIH National Center for Human Genome Research, DOE Office of Health and Environmental Research, and TIGR.

FOCUS: Sequencing results and technology development.

SELECTED MEETING TOPICS

Sequencing and Mapping Results: Data obtained by multiplex sequencing of *Escherichia coli, Mycoplasma leprae*, and human sources; new genomic sequences from *E. coli, Caenorhabditis elegans*, and *Drosophila melanogaster;* sequencing of a 90-kb P1 insert from the bithorax region of the *Drosophila* genome; update on whole-genome physical mapping using mega–yeast artificial chromosomes; identification of many of the 13,000 available human expressed sequence tags (ESTs) by sequence similarity to genes identified in bacteria, yeast and *C. elegans*; demonstrations that sequencing and analysis of large regions of anonymous human DNA are now reasonably straightforward.

Technology Development: New or improved mapping, sequencing, and functional characterization methods; use of single-band-resolution fluorescence in situ hybridization to map cDNA fragments less than 1 kb long; use of tandem hexamers as primers for chromosome walking without ligation; sequencing by hybridization as a fast and inexpensive method for fingerprinting cDNAs and genomic fragments; collection of transposon-insertion mutants that facilitate positional cloning strategies for many genes in *C. elegans*; new sequencing technologies, including multiplexing and hybridization using high-density filters; database designs for tracking sequencing and mapping information from genomic and EST projects.

Informatics: The Data Fair, a new conference feature, allowed comparison and on-site analysis of unpublished sequence data, using computers ranging from PCs to supercomputers. Participants analyzed about 20,000 unpublished EST sequences (6,200,000 nucleotides) from human brain, lymphocyte, and liver cells; mouse testes; *C. elegans*; and *Plasmodium falciparum*; as well as 438 kb of new genomic sequences from *D. melanogaster*, *C. elegans*, *M. leprae*, and *E. coli*. Several new genes were identified during comparison of ESTs and genomic sequences, especially in *E. coli*, providing further evidence of the complementary nature of genome projects in all organisms. Several groups demonstrated software packages for sequence analysis and project support. *[Reported by Alison Hay Tinsley (TIGR)]* ◊

✦*Human Genome Project International Conference: Human Genome '92*

October 14–17, 1992, Nice, France

ORGANIZERS AND COSPONSORS: Human Genome Organization and *Science*.

FOCUS: Highlights of the latest technologies, progress, and social issues of the genome project for the research community and the general public.

SELECTED MEETING TOPICS

Whole-Genome Mapping: The microsatellite mapping approach at Genethon is expected to produce a 2-cM-resolution map by 1994. Aimed toward assembling the genome as yeast artificial chromosome (YAC) contigs, roughly half the genome has already been covered in contigs with a mean length of 1.5 Mb. The potential of large-scale cDNA sequencing for generating maps based on gene function was explored.

Technologies for Visualizing DNA: Combining fluorescence in situ hybridization with fluorescence-activated cell sorting or atomic force microscopy with light microscopy may make possible the improved resolution of YAC or cosmid separations to the 10- to 200-kb range.

Human Diversity and Variation: Although populations show large variations, the homogeneity of Europeans when compared to the global population is surprising.

Mapping Model Organisms: The impact of the *Drosophila* genome project on genome analysis of the malaria-causing parasite was discussed. Mouse maps are used to dissect multifactorial traits whose genetics would be inaccessible in humans.

Patenting cDNA Fragments: A spirited debate took place on the ethics and practicality of patenting cDNA fragments. Attendees felt that the real challenge is to encourage industry to translate the products of human genome research most effectively into powerful tools that will aid in improving the quality of human health throughout the world. *[Reported by Barbara R. Jasny (*Science*) and Robin Yeaton Woo (American Association for the Advancement of Science)]* ◊

✦*First International Workshop on Chromosome 18*

July 21–22, 1992, Chicago

SPONSORS: NIH, DOE, the European Community through the Human Genome Organization, and the Dutch Research Organization.

FOCUS: Preparation of consensus genetic linkage and physical maps of chromosome 18, assessment of available mapping resources, and facilitation of entry into the Genome Data Base.

SELECTED MEETING TOPICS

Mapping Information Presented

New Genes: GOLF (G olfactory protein, 18p11.3-11.2); PACAP (pituitary adenylate cyclase activating protein, 18p11.32); and two desmogleins (DSGII, DSGIII, 18). The pMCT108.2 (D18S24) probe appears to contain DNA sequences derived from both chromosomes 9 and 18.

Genetic Mapping Data: Over 50 polymorphic markers; a framework map including 21 markers [4 restriction fragment length

A complete chromosome 18 meeting report is in *Cytogenetics and Cell Genetics* [63, 77–96 (1993)].

Meeting Reports

The second chromosome 18 workshop is scheduled for July 19–20 in Nijmegen, Netherlands [Organizers: Ad Geurts van Kessel (Int. 31-80/61-41-07, Fax: /54-21-51) and Joan Overhauser (215/955-5188, Fax: -5393).]

polymorphisms (RFLPs) and 13/17 microsatellite markers having PIC values >70%; the markers were tested on 19 families, and 13 were also mapped physically using a panel of somatic cell hybrids]. A Centre d'Etude du Polymorphisme Humain (CEPH) framework map of 33 markers (1000:1 odds; 7/33 with PIC values >70%). Three markers are common to both framework maps.

Physical Mapping: A somatic cell hybrid panel was developed using human cells with various structural rearrangements of chromosome 18 [includes 1 hybrid retaining only chromosome 18 (HHW324) and 26 additional hybrids that distinguish 24 intervals of chromosome 18]. A lambda phage library was prepared from HHW324 (205 clones selected and sublocalized to chromosome 18 by Southern blot analysis); sequence tagged sites (STSs) were prepared for a number of genes and D segments and localized to this panel of hybrids. A 7-hybrid panel was developed dividing chromosome 18 into 8 regions (used to map 22 genes and D segments); a number of probes refined by using DNA with chromosome 18 deletions were localized.

New RFLP Probes, YAC Library: A hybrid panel is being used to determine the allele frequencies and cytogenetic location of RFLP probes. A yeast artificial chromosome (YAC) library prepared from a hybrid cell line retaining 2 copies of chromosome 18 as the only human chromosome material has yielded 84 human-specific YACs. These were characterized, fingerprinted, and screened for STS content; initial contigs were established. Fluorescence in situ hybridization was used to localize 60 of these YACs to chromosome 18.

Radiation Hybrids: Radiation-reduced hybrids were generated from a hybrid cell line retaining a single copy of chromosome 18 as the only human material; a set of 92 radiation hybrids were tested for 50 markers; a radiation hybrid map is under construction.

Disease Studies: Trisomy 18 parental origin studies showed the additional chromosome to be maternal in most cases (85 to 95%); two panels of somatic cell hybrids prepared from cells containing deletions of 18q or partial duplications of chromosome 18 enabled construction of a preliminary phenotypic map (molecular analyses performed with probes from the lambda phage library described below); probes were ordered cytogenetically by loss-of-heterozygosity analysis of colon carcinomas; YAC technology was used to clone the synovial sarcoma–associated translocation breakpoint of the t(X;18) (p11.2;q11.2) regions.

Resources Available

Somatic Cell Hybrids: Panel of six somatic cell hybrids defining seven intervals on chromosome 18; hybrids selected to form a reference panel for mapping chromosome 18 (both deposited in the Human Genetic Mutant Cell Repository at the Coriell Institute in Camden, New Jersey).

STSs: 156 STSs available in a compendium from Gary Silverman (Harvard University), including 75 from 20 cloned genes and 81 from 76 anonymous DNA markers (44 of these are CA repeats or other highly polymorphic loci).

YACs and Cosmids: A YAC library prepared from the HHW324 hybrid (average insert size of about 300 kb) and other YAC clones; CEPH YAC library for screening chromosome 18–specific sequences [Eric Lander (Whitehead Institute)]; a flow-sorted chromosome 18 cosmid library prepared by Dean Nizetic and Hans Lehrach (Imperial Cancer Research Fund), of which about 20,000 clones have been gridded (this fraction of about 5 to 6 chromosome equivalents will be stamped onto high-density filters for hybridization screening with low-copy probes).

Phage Clones and Libraries: A lambda phage library prepared from the HHW324 hybrid is available from Joan Overhauser (Thomas Jefferson University); a 24X genomic-equivalent lambda library (average insert of about 15 kb) was deposited with the American *Type Culture* Collection by Pieter de Jong (Lawrence Livermore National Laboratory).

Needed Resources: New markers, particularly new microsatellite markers; telomeric YAC clones; a gridded and arrayed chromosome 18 cosmid library; and reference markers converted to those based on the polymerase chain reaction.

[Reported by Michelle M. Le Beau (University of Chicago), Ad Geurts van Kessel (University of Nijmegen), and Joan Overhauser (Thomas Jefferson University)] ◊

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 and Safety 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.◊

Sixth International Mouse Genome Conference

The Sixth International Mouse Genome Conference was held in October 1992 in Buffalo, New York. Substantial progress was reported toward (1) the production of a very high density genetic map and (2) physical mapping in chromosome regions having a large number of genetic markers. Some meeting highlights follow.

Genetic Mapping

Eric Lander reported the mapping of more than 1000 microsatellites toward a 2- to 3-year target of 6000 at the Whitehead–Massachusetts Institute of Technology (MIT) genome center. Neal Copeland [National Cancer Institute (NCI)] noted that 1000 genes had been mapped, integrating a number of microsatellites from Lander. Michael Seldin (Duke University Medical Center) reported the mapping of 500 markers across the mouse genome. Restriction landmark genome scanning reported by Yoshihide Hayashizaki (RIKEN Tsukuba Life Science Center, Japan) has also added a large number of new loci to the genetic map.

The target of producing a marker every 1 cM or less is close to being achieved. A 1000-animal C57BL/6-*spretus* backcross, the European Collaborative Interspecific Backcross (EUCIB), is being conducted at the U.K. Human Genome Mapping Project Resource Center and the Pasteur Institute of Paris. This backcross provides genetic resolution of 0.3 cM at the 95% probability level and will contribute to construction of yeast artificial chromosome (YAC) contig maps based on sequence tagged sites (STSs).

Uses of the Mouse Genetic Map

Increases in mapped genetic markers across the mouse genome have stimulated investigations in analyzing multifactorial traits and identifying candidate genes for mouse or human mutations. Reports by Phil Avner (Pasteur Institute), Mazakasu Hattori (Harvard Medical School), and Jan Prins (Oxford University) emphasized continued interest in loci that predispose to Type I diabetes; different loci are revealed in mouse crosses of varying genetic backgrounds. The advantages of a dense microsatellite map of the mouse were further illustrated by the genetic analysis of modifying loci for Familial Adenomatous Polyposis (Bill Dietrich, MIT). Nancy Jenkins (NCI, Frederick) reported mapping a maternally imprinted gene for a small nuclear ribonucleoprotein particle (snRNP). This gene maps to a mouse chromosome 7 region homologous to a region of human chromosome 15q11-13 implicated in Prader-Willi syndrome.

Physical Mapping

With the increased density of mapped markers across the mouse genome, efforts are now under way to create STS-based YAC contig maps of some chromosome regions. Kent Hunter (MIT) reported progress on distal chromosome 1, as did Gail Herman (Baylor College of Medicine) on the X chromosome. She has constructed a 2.5-Mb contig spanning the bare patches (*Bpa*) locus—a putative homologue to *Chondrodysplasia punctata* on the human X chromosome. Efforts to assemble chromosome 17 YAC contigs in the regions of the t^{lub2} deletion (Roni Bollag, Princeton University) and *qk* mutation [Roger Cox, Imperial Cancer Research Fund (ICRF)] were also discussed.

Aiding these efforts has been the availability of mouse YAC libraries from Princeton University, ICRF, and St. Mary's Hospital Medical School. Completion of the high-density mouse genetic map and improvement of YAC library resources will enable researchers to begin a concerted effort toward a complete STS-based YAC contig map in about 2 years.

Positional Cloning of Mouse Mutations

Mapping and characterizing mouse mutations are important to understanding human development and disease processes. Jeff Friedman (Rockefeller University) reported substantial progress in defining YAC clones carrying the mouse obese (*ob*) gene. Physical mapping is proceeding toward identifying genes involved in pivotal stages of development, including gastrulation (Bernadette Holdener-Kenny, Case Western Reserve University) and skeletal development—the fused locus (Janice Rossi, Princeton University).

Progress toward a fully integrated mouse genome database was reported by the Jackson Laboratory. A number of presentations reviewed improved software tools for constructing genetic maps.◊

Reported by Steve Brown
St. Mary's Hospital Medical School, London

Substantial Progress Reported:

- **Very High Density Genetic Maps**
- **Physical Maps of Marker-Rich Regions**

The Seventh International Mouse Genome Conference will be held November 7–11 in Hamamatsu, Japan. Contact: Kazuo Moriwaki; Department of Cell Genetics; National Institute of Genetics; Yata 1111 Mishima; Shizuoka-ken 411, Japan (Int. 81-559/75-0771, Fax: -6240).

Chromosome committee reports detailing loci and maps for each mouse chromosome are available as a special volume of *Mammalian Genome* [3(8), S1–294 (July 1992)].

The addition of 237 polymorphic mouse markers from the 1992 mouse conference brings to 988 the number of markers available under the tradename MapPairs™. Each marker consists of two PCR primers mapped to a specific location in the mouse chromosome and capable of being run under identical PCR conditions. [Research Genetics; 2130 Memorial Parkway SW; Huntsville, AL 35801 (800/533-4363, Fax: 205/536-9016).]

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

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

Sponsor Contacts:

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

Leslie Fink
NIH National Center for Human Genome Research
Bethesda, MD 20892
301/402-0911
Fax: 301/480-1950
Internet: *lsf@cu.nih.gov*

Second Transcribed Sequences Workshop

The DOE-sponsored Second International Workshop on the Identification of Transcribed Sequences was held in San Francisco on November 7–8, 1992. The purpose of the workshop, which was attended by 44 scientists from 7 countries, was to discuss and evaluate techniques for developing a complete transcriptional map of the human genome.

Such a map requires the positions, sequences, and expression patterns of all genes. This goal is being approached from two different directions, each with strengths and weaknesses. One method is to identify the transcribed sequences from genomic DNA of a given region; the other is to systematically sequence and map cDNAs. The cDNA approach yields sequence information rapidly, but mapping each cDNA is a technical challenge. In the first approach, the map locations of genomic sequences are known at the outset, and the challenge is to identify exons. The efficient construction of a transcriptional map will require a diverse array of techniques.

cDNA Sequencing and Libraries

Charles Auffray (Genethon, France), James Sikela (University of Colorado), and Mihael Polymeropoulos [NIH National Institute of Mental Health (NIMH)] presented large-scale, partial cDNA sequencing results. A number of methods are being used to integrate these sequences with physical or genetic maps. Minoru Ko (Wayne State University) used polymorphic sequences in 3' untranslated regions of mouse cDNAs as genetic markers in interspecific mouse crosses. Polymeropoulos employed polymerase chain reaction (PCR) primers with a panel of hamster-human cell hybrids to assign human cDNAs to chromosomes. Sikela reported use of fluorescence in situ hybridization to map some cDNA sequences.

The unequal representation of clones in cDNA libraries poses another problem for the cDNA sequencing approach. With 10,000 to 20,000 genes expressed in a given tissue, their unequal representation would require the sequencing of 1 million clones to approach complete representation for a tissue. An important goal is to achieve equal representation of both abundant and rare species in cDNA libraries (i.e., a "normalized" library). Bento Soares (Columbia University) has produced normalized cDNA libraries by reannealing cDNA at a high C_0t value and selectively cloning the remaining single-stranded fraction. Cheng Chi Lee [Baylor College of Medicine (BCM)] proposed normalizing arrayed cDNA libraries by rescreening with pools of clones already sequenced to eliminate redundancies.

Another approach to defining a minimal set of cDNA clones is the identification of each clone according to its hybridization "fingerprint" against a panel of oligonucleotides. Radoje Drmanac (Argonne National Laboratory) and Sebastian Meirer-Ewert (Imperial Cancer Research Fund) reported progress in rapidly screening cDNA libraries with short oligonucleotide probes.

cDNA libraries from rare cell types are necessary for identifying the genes expressed in these cells. Barbara Knowles (Wistar Institute) has produced cDNA libraries from specific stages of early mouse embryos. James Eberwine (University of Pennsylvania) has developed techniques for producing cDNA libraries from single cells. His method relies on amplification by prokaryotic RNA polymerases rather than on PCR amplification, which can severely skew representation in heterogeneous mixtures.

Genomic Reference Libraries and Exon Identification

While mapping and normalization are challenges in the cDNA sequencing approach, the genomic approach requires identification of transcribed sequences among the surrounding untranscribed sequences. Genomic DNA for this strategy is made available by advances in physical mapping, especially large arrayed cosmid and lambda libraries from specific chromosomal regions and yeast artificial chromosome (YAC) clones. Ute Hochgeschwender (NIMH), Anne Marie Poustka (German Cancer Research Center), and Geoffrey Falk (Scripps Research Institute) reported using labeled cDNA for probing large arrayed genomic libraries to identify transcribed sequences. This technique is suited for very large regions covered by arrayed libraries; about one-half the genes in a given tissue are expressed at levels high enough to be detected with these probes. Furthermore, expression patterns can be determined by repeated cDNA probe hybridizations of different tissues at different developmental stages.

Exon Identification by Computer

Richard Mural [Oak Ridge National Laboratory (ORNL)] and Gordon Hutchinson (Canadian Genetic Disease Network) have developed software that can identify coding exons from genomic sequences with over 80% success and a low false-positive rate. An additional feature of the ORNL GRAIL program is a gene-assembly program (GAP) that produces complete gene sequences from the GRAIL output, resulting in an even lower false-positive rate. Susan Berget (BCM) discussed her work on splice-site selection, which is based on defining the nucleic acid "signals" associated with splicing. David Searls (University of Pennsylvania) presented a linguistic approach to sequence analysis based on conceptual similarities between meaningful language and expressed sequences.

Direct Exon Identification

Exons can also be identified by procedures in which a genomic fragment is cloned into an intron of a mammalian expression vector. A fragment containing an exon will be included in the mRNA expressed from the vector. Alan Buckler [Massachusetts General Hospital (MGH)] reported improvements to the splicing vector pSPL1. Paul Nisson (Gibco BRL/Life Technologies, Inc.) discussed a potential problem in transforming pools of genomic fragments cloned in splicing vectors: unequal exon-amplification rates in RT/PCR reactions can lead to the loss of all but the most efficiently amplified exons. This problem is also encountered in amplifying cDNA libraries. Geoffrey Duyk (Harvard Medical School) and Nicole Datson (Leiden University, Netherlands) reported on modifications of exon-trapping systems. While most exon-identification protocols are designed to detect internal exons, David Krizman (BCM) described a new system for detecting 3′ terminal exons in which polymorphic sequences in the 3′ untranslated regions may facilitate exon mapping.

David Kurnit (University of Michigan) described his system for identifying exons cloned in a plasmid vector by their ability to recombine homologously with cDNA clones in lambda. The system detects the exon, isolates the cDNA, and (when tested with multiple cDNA libraries) supplies information on expression patterns.

Another general approach to identifying transcribed sequences from genomic clones involves hybridizing cDNA to genomic DNA and eluting the specifically bound "selected" cDNA. Michael Lovett (University of Texas), Poustka, Danilo Tagle (University of Michigan), and Sherman Weissman (Yale University) presented results using variations of this procedure with pools of cosmid clones and with whole YAC DNA. To map these genes within the YAC, Ruchira DasGupta (Albert Einstein College of Medicine) is using cDNAs isolated by this method and cloned in a yeast vector to truncate the YAC at the region of homology.

Perspectives

The usefulness of the different approaches can best be assessed by applying them to large regions. Gail Bruns (Children's Hospital, Boston) reported searching through the 15-Mb WAGR region for conserved sequences associated with HTF (*Hpa* II twin fragment) islands. Bernhard Weber (University of British Columbia) used a cDNA selection strategy and single-strand–conformation polymorphism technique to isolate new genes from the Huntington's disease region and to search these for mutations.

Giorgio Bernardi (Institut Jacques Monod, France) and Katheleen Gardiner (Eleanor Roosevelt Institute) discussed long-range differences in genome structure. While gene density is highest in regions richest in GC (in particular, most telomeres), one-third of genes are in AT-rich regions. Gene content as measured by CpG-island density does not appear to vary with GC content in a predictable manner. The applicability of techniques to the gene-poor regions was raised, and Mural and Hutchinson requested sequences of verified AT-rich exons for teaching neural networks.

The functional analysis of products encoded by novel genes was discussed. Roger Brent (MGH) described a yeast interaction trap system for identifying protein regions that interact in vivo. Miles Brennan (NIMH) reported on the specific integration of a mammalian cDNA at a homologous gene in yeast, a system that may allow direct selection and functional analyses of such cDNAs. The need for more functional assays was noted by Weissman.

Reported by Miles B. Brennan
National Institute of Mental Health
and
Katheleen Gardiner
Eleanor Roosevelt Institute

The Third International Transcribed Sequences Workshop is planned for October 2–4 in New Orleans. Contact: Ute Hochgeschwender; Unit on Genomics, NIMH; Bldg 10, Room 4N 320; Bethesda, MD 20892 (301/402-1769, Fax: -2140).

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

June

17. †Jasper Rine: Dog Genome Initiative—Towards the Genetics of Morphology and Breed Traits; Bethesda, MD [NCHGR Lecture Series, C. Dahl, 301/402-0838]

21–23. 2nd Int. Workshop on Chromosome 6; Berlin [A. Ziegler, (Int.) 49/30-3035-2617, Fax: -3778]

26–29. Human Gene Therapy; Washington, DC [NYAS, 212/838-0230, Fax: -5640]

July

7–9. 1st Int. Conf. on Intelligent Systems for Molecular Biology; Bethesda, MD [J. Shavlik, Fax: 608/262-9777, Internet: *ismb@nlm.nih.gov*]

10–11. Chromosome 4 Workshop; Stanford, CA [R. Myers, 415/476-8138, Fax: -8217]

12–26. *Open Problems in Computational Molecular Biology: 3rd Int. Workshop; Telluride, CO [A. Konopka, 301/846-5396; *konopka@fcrfv1.ncifcrf.gov*]

19–20. 2nd Int. Workshop on Chromosome 18; Nijmegen, Netherlands [A. Geurts van Kessel, (Int.) 31/80-614105, Fax: -542151]

23–24. ‡Maximizing the Return from Human Genome Research; Concord, NH [C. Ruh, 603/228-1541, Fax: /224-3342]

31–Aug. 4. DNA Damage: Effects on DNA Structure and Protein Recognition; NYAS, Burlington, VT (poster abstracts by May 1) [see contact: June 26–29]

August

2–7. *6th Annual Summer Seminar in Health Care Ethics; Seattle [Univ. of Wash., 800/869-2633, Fax: 206/543-3195]

3–6. 6th Int. Workshop on the Fragile X and X-Linked Mental Retardation; Cairns, Australia [G. Sutherland, (Int.) 61/8-204-7284, Fax: -7342]

6–10. Sci. Innovation 1993: New Techniques in Biomolecular Research; Boston [AAAS, 202/326-6450, Fax: /289-4021]

9–13. STM '93: Int. Conf. on STM; Beijing, China [C. Bai, (Int.) 86/1-256-8158, Fax: -9564]

15–21. 17th Int. Congress of Genetics; Birmingham, UK [D. Smith, (Int.) 44/21-414-5888, Fax: -3850]

17–22. MacroMolecules, Genes, and Computers: Chap. Three; Waterville Valley, NH [D. Madden, 415/570-6667, ext. 8803, Fax: /572-2743, Internet: *dawn@apldbio.com*]

24–29. Molecular Genetics of Bacteria & Phages; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845]

29–Sept. 3. *Artificial Intelligence and the Genome at IJCAI '93; Chambery, France [J.-G. Ganascia, (Int.) 33/1-44-27-4723, Fax: -7000, Internet: *ganascia@laforia.ibp.fr*]

September

8–12. Eukaryotic DNA Replication; CSHL [see contact: Aug. 24–29]

9–13. *E. Coli* Genome Meeting; Madison, WI [M. Ellingson, 608/262-2755, Fax: -3453]

20–21. *National Advisory Council for Human Genome Research; Bethesda, MD [J. Ades, 301/402-2205, Fax: -2218]

20–24. 11th Australian Biotechnology Assoc. Conf.; Perth, Australia [J. Sargeant, Fax: (Int.) 61/9-310-3505, Internet: *mgkjones@murdoch.edu.au*]

27–29. Fetal Cells in Maternal Blood: Prospects for Noninvasive Prenatal Diagnosis; NYAS, Arlington, VA (poster deadline: June 1) [see contact: June 26–29]

October

2–5. NSGC 12th Annual Education Conf.; Atlanta (application deadline: April 15) [B. Leopold, 215/872-7608, Fax: -1192]

5–9. ASHG 43rd Annual Meeting; New Orleans (abstract deadline: June 1) [M. Ryan, 301/571-1825, Fax: /530-7079]

13–16. DNA: The Double Helix. Forty Years: Perspective and Prospective; NYAS, Chicago (poster deadline: July 2) [see contact: June 26–29]

21–22. Law and Science at The Crossroads: Biomedical Technology, Ethics, Public Policy, and the Law; Boston [SULS, 617/573-8627, Fax: /248-0648]

23–27. Genome Sequencing and Analysis V; Hilton Head, SC [S. Wallace, 301/216-9567, Fax: /977-7233]

November

5–6. First Genetic Marker—Blood Group Research, "Race," and Disease: 1900–1950; Indianapolis, IN [W. Schneider, 317/274-3811, Fax: -2347]

7–10. Electrophoresis '93; The Electrophoresis Society, Charleston, SC (paper deadline: June 1) [J. Cunningham, 301/898-3772, Fax: -5596]

7–11. MGC 93; Hamamatsu, Japan [K. Moriwaki, (Int.) 81/559-75-0771, Fax: -6240]

10–12. *CCM 93; Tsukuba, Japan [N. Shimizu, (Int.) 81/333-53-1211, ext. 2721, Fax:-51-2370]

14–17. HGM 93; Kobe, Japan [HGM Secretariat, (Int.) 81/6-454-4811, Fax: -4711]

15–19. Nanometer Scale Biotechnology: DNA Reconstruction, at the 1993 AVS Natl. Symposium; Washington, DC [D. Manos, 804/221-3525, Fax: -3540]

20–21. Healthcare Financing in a Changing World; Washington, DC [J. Weiss, 800/336-GENE]

December

8–9. Mathematical and Statistical Aspects of DNA and Protein Sequence Analysis; London [Scientific Meetings Secretary, (Int.) 44/71-839-5561, ext. 278, Fax: -930-2170] ◊

Training Calendar**

June

7–8. Basic IVD Approval Process; Washington, DC [Bioconferences Intnl., 301/652-3072, Fax: -4951]

7–11. Expression of Recombinant DNA in Mammalian Cells; Washington, DC [CATCMB/CUA, 202/319-6161, Fax: -4467]

7–11. Recombinant DNA: Techniques & Applications; Rockville, MD [ATCC, 301/231-5566, Fax: /770-1805]

9–11. Recombinant DNA for Chemists; Washington, DC [ACS, 800/227-5558, Fax: 202/872-6336]

13–26. Human Gene Mapping and Sequencing; Salt Lake City, UT [Genome Tech. Workshop, 801/585-5606, Fax: /581-7796]

14–18. Recombinant DNA Techniques I; LTI, Germantown, MD [L. Kerwin, 800/952-9166, Fax: 301/258-8212]

14–25. Workshops for Secondary School Biology Teachers: Project Genethics; Winter Park, FL [J. Hendrix, 800/537-9604]

15–18. PCR/Cycle DNA Sequencing; ATCC, Rockville, MD [see contact: June 7–11]

21. Intro. to PCR; Chicago [BTP, 800/821-4861, Fax: 515/232-8306]

21–22. GDB/OMIM Training Course, [see schedule, p. 7]

21–26. Recombinant DNA Techniques II; LTI, Germantown, MD [see contact: June 14–18]

21–26. Natl. HGP Workshop for Secondary Science Teachers; Kansas City [D. Collins, Fax: 913/588-3995]

22–23. Quantitative RNA-PCR; BTP, Chicago [see contact: June 21]

23–25. Advanced Automated/Diagnostic DNA Sequencing Workshop; Houston [L. Lawler-Lopez, 713/798-5393, Fax: -5741]

24–25. Clinical Applications to PCR; BTP, Seattle [see contact: June 21]

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

For Your Information

Foreign Subscribers Asked To Confirm Addresses

To update the *HGN* mailing list and conserve resources, HGMIS requests that foreign subscribers confirm their addresses by faxing or mailing a copy of their newsletter label to the HGMIS address on p. 12. Corrections should be made as needed. Names not confirmed by September 1 will be removed from the mailing list.◊

24–25. DNA Sequencing without Radioactivity; BTP, Chicago [see contact: June 21]

27–July 2. Chromatin and Transcription; Copper Mountain, CO (application deadline: April 16) [FASEB, 301/530-7093, Fax: -7014]

28–July 2. In Situ Hybridization; CATCMB/CUA, Washington, DC [see contact: June 7–11]

30–July 1. Basic Cloning Techniques; BTP, Chicago [see contact: June 21]

July

3–8. Restriction Endonucleases and Modification Methyltransferases: Structures and Mechanisms; FASEB, Saxtons River, VT [see contact: June 27–July 2]

5–25. *Arabidopsis* Molecular Genetics; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845]

6–15. Experimental Mammalian Genetics: Gene Targeting and Trapping; UMDS, London [P. Faik, (Int.) 44/71-403-6998, Fax: -407-5281]

8–9. *Intro. to Molecular Cytogenetics; Oncor, Inc., Gaithersburg, MD [M. Williams, 800/776-6267, Fax: 301/926-6129]

12–16. *Genome Technology and its Implications: A Hands-On Workshop for Educators; Ann Arbor, MI (also offered Aug. 16–20) [P. Gregory, 313/747-2738, Fax: /763-4692]

12–23. Summer Institute of Supercomputing 1993; PSC, Pittsburgh (application deadline: May 10) [A. Conniff, 412/268-6800, Internet: *conniff@psc.edu*]

18–31. Transcription; Chapel Hill, NC [W. Litaker, 919/966-1730, Fax: -6821]

19–20. Basic Cloning & Hybridization Techniques; BTP, San Francisco [see contact: June 21]

19–30. Medical & Experimental Mammalian Genetics; Bar Harbor, ME [Jackson Laboratory, 207/288-3371, ext. 1253, 7:30 am–3:30 pm EST]

19–30. Recombinant DNA Methodology and Applications; UMBC, Baltimore [C. Harriger, 410/455-2336, Fax: -1074]

25–30. Genetic Recombination and Genome Rearrangements; FASEB, Copper Mountain, CO [see contact: June 27–July 2]

28–30. 12th Summer Symposium in Molecular Biology: Structure/Function Relationships in Proteins and Enzymes; University Park, PA [P. Phillips, 800/833-5533]

29–Aug. 6. Human Genome Analysis: From YAC to Gene; UMDS, London [see contact: July 6–15] ◊

U.S. Genome Research Funding Guidelines

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

NIH National Center for Human Genome Research (NCHGR)

Application receipt dates:

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

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

Program announcements are listed in the weekly *NIH Guide for Grants and Contracts*,* which is available through

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

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

DOE Human Genome Program – Proposals Due July 15

Solicitations for proposals were announced in the *Federal Register* **58**(30), 8746–48 and in *Science* and other publications. Preproposals were due in April; proposals are due July 15.

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

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

SBIR Grants

DOE also invites small business firms to submit grant applications addressing the human genome topic of SBIR programs, which are designed to strengthen innovative firms in research and development and to contribute to the growth and strength of the nation's economy. Applications are invited in the following three subtopics only: (1) Development of Improved DNA Sequencing Technologies; (2) Improvements in Genetic Data Storage, Processing, and Analysis; and (3) Development of Innovative Materials or Dissemination Techniques to inform students and the lay public about benefits, opportunities, and challenges arising from the Human Genome Project. For more information, contact

- Samuel Barish; SBIR Program Manager, ER-16; DOE; Washington, DC 20585 (301/903-5707).

Human Genome Distinguished Postdoctoral Fellowships

Most recent deadline: February 1. For further information, contact

- Linda Holmes, Oak Ridge Institute for Science and Education (615/576-4805).◊

Baylor Automated Sequencing Workshop

Baylor College of Medicine will hold a workshop June 23–25 on state-of-the-art automated DNA sequencing techniques. Participants are limited to ten experienced scientists, who will use fluorescent DNA sequencing instruments for large-scale and diagnostic DNA sequencing at a core facility. Topics will include sequencing by polymerase chain reaction for mutation detection and human immunodeficiency virus characterization. Semiautomated template preparation, robotic DNA sequencing (Beckman Biomeck and ABI Catalyst systems), data management and analysis, and overall core lab organization will also be covered. Tuition of $750 includes accommodations. Application deadline: June 1. [Contact: Lori Lawler-Lopez; Institute of Molecular Genetics; Baylor College of Medicine; Houston, TX 77030 (713/798-5393, Fax: -5741, Internet: *llawler@bcm.tmc.edu*).] ◊

Human Genome Management Information System

Subscription/Document Request (Vol. 5, No. 1)

Please type or print legibly. Enclose a business card, if possible. To change your name/address/affiliation or to be dropped from the mailing list, you must enclose your current newsletter address label to ensure that the proper entry is modified.

Name _____
(First) (MI) (Last)

Affiliation _____

Department/Division _____

Street and/or P.O. Box _____

City/State/Zip Code/Country _____

Phone _____ **Fax** _____

E-Mail Address _____ **Area of Interest** _____

1. ___ *Human Genome News*
   ___New Subscriber ___Change of Name/Affiliation/Address (circle all that apply) ___Drop Name from Mailing List
2. ___ *Understanding Our Genetic Inheritance, The U.S. Human Genome Project: The First Five Years, FY 1991–1995* (Joint DOE-NIH 5-Year Plan)
3. ___ DOE *Human Genome 1991–92 Program Report.*
4. ___ *Primer on Molecular Genetics.* (Included in program report above. Extracted as separate document for educational use.)

CALENDAR ACRONYMS

AAAS Am. Assoc. for the Advancement of Sci.
ACS Am. Chem. Soc.
ASHG Am. Soc. of Human Genetics
ATCC Am. *Type Culture* Coll.
AVS Am. Vaccum Soc.
BTP Biotechnology Training Programs
CATCMB/CUA Ctr. for Advanced Training in Cell and Mol. Biology/Catholic Univ. of Am.
CCM Chromosome Coordinating Meeting
CSHL Cold Spring Harbor Lab.
FASEB Fed. of Am. Societies for Experimental Biol.
GDB/OMIM Genome Data Base/Online *Mendelian Inheritance in Man*
HGM Human Genome Mapping
HGP Human Genome Project
IJCAI Int. Joint Conf. on Art. Intell.
IVD In Vitro Diagnostic
LTI Life Technologies, Inc.
MGC Mouse Genome Conference
NCHGR Natl. Ctr. for Human Genome Res.
NSGC Natl. Soc. of Genetics Counselors
NYAS NY Acad. of Sci.
PSC Pittsburgh Supercomputing Ctr.
STM scanning tunneling microscopy
SULS Suffolk Univ. Law School
UMBC Univ. of MD, Baltimore County
UMDS United Medical & Dental Schools

Reader Comments:

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