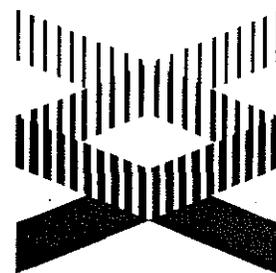


Human Genome news



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

ISSN: 1050-6101

Vol. 7, No. 1, May-June 1995

Education Key to Understanding, Acceptance

The increasing abilities to manipulate and analyze DNA are bringing profound changes to society, particularly in approaches to human health problems, personal identification, and agricultural development. To reap the benefits and avoid pitfalls inherent in DNA technology, the general public must have some understanding of DNA, how it is involved in heredity, and how it works in the cell, as well as the methods used to analyze and manipulate it. With complex genetic concepts and discoveries coming at an ever-increasing pace, what the lay person understands or believes to be true now will help determine how such scientific advances are evaluated and whether they are accepted by the public or not. Clearly, education is the key.

Education in the United States faces a number of challenges in promoting science literacy for the public, students, and teachers. Some public high schools do not even offer a biology course, and most high school and many college science teachers received their degrees before DNA technology was added to the college curriculum. Confident, enthusiastic, and knowledgeable teachers are desperately needed at all levels to convey the latest information on genetics and molecular biology to the first generation that will be influenced by the new genetics and the technologies springing from it.

One of the most efficient ways to foster productive interactions and update educators is to provide them with short courses and workshops in molecular genetics. Several educational programs sponsored by the Human Genome Project have developed effective, field-tested workshops for just this purpose. In addition, many scientists in public and private institutions serve as resources for the general community and help teachers understand molecular genetics and obtain necessary equipment, supplies, and know-how to incorporate genome technology into everyday classroom teaching.

This is a significant beginning, but much more needs to be done by investigators involved in genetics. The types of outreach described in the following article could serve as models for other education activities.

The following article, which is not comprehensive, highlights some outreach efforts to meet the critical need for genetic education throughout the population. Education projects of the NIH National Center for Human Genome Research (NCHGR) and the Ethical, Legal, and Social Issues (ELSI) Programs of NIH and DOE are listed first. These projects are followed by activities at the research centers, colleges, universities, and state departments of education; and programs affiliated with corporations and private industry.

NCHGR Outreach and Education

Office of Communications, Leslie Fink (301/402-0911, Fax: -2218 or -4570, LeslieF@od.nchgr.nih.gov): Prepares reports, publications, and press releases and responds to inquiries about NCHGR research programs and policies; collaborates with professional, community, and other



Exploratorium visitors experiment with the protein production line at the "Diving into the Gene Pool" exhibition, on view in San Francisco through September 4.

In This Issue

Page	
1	Education Key to Understanding
3	Notice to DOE Contractors and Grantees
3	NCHGR Publishes Brochure, Report
5	CSHL Mapping and Sequencing Meeting
5	Two Bacterial Genomes Sequenced
7	Human Gene Map Workshops Held
11	Advisory Council Notes Progress
11	Database Workshops Held
13	GDB Forum: Genetic Maps on Web
<i>Information/Resources</i>	
4	Publications/Videotapes; also 10, 11, 15
8	Web Activities; also 11
9	Biological Resources
12	Informatics Resources
14	Calendars
15	Funding Guidelines/Grants
16	Subscriptions; Selected Acronyms

organizations to develop avenues for dialogue and dissemination of information about the Human Genome Project and other NCHGR programs (see box, p. 3, for recent publications).

Genetics Education Office, Paula Gregory (301/594-0654, Fax: /402-2218, edcore@nchgr.nih.gov): Lectures to teachers, students, physicians, and community groups. Placement of students and interns in Division of Intramural Research laboratories. DNA sequencing partnership involving human genome sequencing by high school students (modeled on University of Washington project, see p. 3). Workshops and short courses for counselors, science writers, teachers of deaf students, and minority faculty. Newsletters, such as *Genome*, for educators and directors of genetics education programs. Training sessions and educational resources for scientists. NCHGR WWW Home Page (<http://gaea.nchgr.nih.gov>) extension to include information on activities of the education office.

NCHGR ELSI Program

Elizabeth Thomson, Acting Chief, NCHGR ELSI Branch (301/402-4997, Fax: -1950, exx@cu.nih.gov)

Foundation for Blood Research, Paula Haddow: Existing high school biology curriculum unit revised to include societal implications of genetic advances. Field test of new materials and pilot test of an experimental theater component.

Georgetown University (GU), Virginia Lapham: Human Genome Program Education Model Project to develop programs for training individuals, family members of 90 voluntary support groups, and related multidisciplinary health professionals to educate others.

High school biology students sequence human cosmid DNAs in a University of Washington, Seattle, project.



GU, LeRoy Walters: Further development of a database of ELSI literature in the Bioethicsline online bibliography (medethx@guvm.ccf.georgetown.edu or 800/633-ethx). Portions of *ScopeNotes* and *Bibliography of Bioethics* from the GU Kennedy Institute of Ethics available through National Center for Genome Resources Home Page (<http://www.ncgr.org>).

University of Virginia, John Fletcher: Integrated textbook, case book, teaching manual, and workshops on genetic evidence for appellate judges and journalists.

Massachusetts Corporation for Educational TV, Cardie Texter: Semester-long course through a public broadcasting network using satellite, computer, audio, and print materials.

University of Washington, Seattle (UWS), Albert Jonsen: Five-day, advanced-level course on ethical and social implications for doctoral and post-doctoral students and professionals trained in bioethics and genetics.

Other education efforts by the NCHGR ELSI component include the following: **1993 meeting of directors** of counselor and nurse-specialist training programs to develop new recommendations for training genetic counselors. **Programs to recruit talented nonbiologists** into the Human Genome Project and disseminate information to scientists about the uses of new genomic tools. **Short courses about genomic sciences** for the general scientific and ELSI communities.

DOE Human Genome Program Outreach and Education

DOE ELSI Program: Daniel Drell (301/903-6488, Fax: -8521, Daniel.Drell@mailgw.er.doe.gov)

American Association for the Advancement of Science, Maria Sosa: Booklets and a database of health and science resources on genetics for adults in basic literacy classes.

American Society for Human Genetics, Stephen Goodman: Five-year program giving one fellowship each year to a mid-career geneticist for working on Capitol Hill.

Biological Sciences Curriculum Study (BSCS), Joe McInerney: Curriculum units for 55,000 high school biology teachers and students, focusing on human genome mapping and sequencing; information management, access, and regulation; and nontraditional forms of inheritance.

California State University, Margaret Jefferson, and Los Angeles Unified School District, Mary Ann Sesma: BSCS module translated into Spanish and adapted for introduction into public schools and the Hispanic community.

Cold Spring Harbor Laboratory, Jan Witkowski: Workshops on human genetics for policymakers and opinion leaders.

Einstein Institute, Franklin Zweig: Deskbook for federal and state courts to assist judges in comprehending and applying genetic evidence.

Exploratorium, San Francisco, Charles Carlson (charliec@exploratorium.edu): Museum exhibits, including "Diving into the Gene Pool," an extensive hands-on, 25-exhibit display examining DNA structure and function and the Human Genome Project from a variety of perspectives (open through September 4).

Science and Technology Radio Project, Barinetta Scott: Two-year project to produce over 17 hours of radio programs and printed and electronic materials on science and ethical issues of the genome project.

University of Chicago, Mary Mahowald: Training for physicians and nurses who will educate practitioners introducing new genetic services.

University of Kansas, Debra Collins: Series of educational workshops held each summer for about 50 high school science teachers (<http://www.kumc.edu/instruction/medicine/genetics/homepage.html>).

UWS, Maynard Olson, Leroy Hood, and Maureen Munn: Pilot program allowing high school biology students to sequence human cosmid DNAs and place them in sequence databases.

WGBH Educational Foundation, Paula Apsell and Graham Chedd (Chedd-Angier Productions): "The Secret of

Life"; WNET/Thirteen, George Page and Stefan Moore: "Medicine at the Crossroads," partially supported by NCHGR. TV documentaries broadcast several times, some available on videotape. ["Secret of Life": Films for the Humanities & Sciences; Princeton, N.J. (800/257-5126, Fax: 609/275-3767)]

Oregon Public Broadcasting and Graham Chedd: TV documentary on ELSI issues, "Genetics in Society," planned for release in 1997.

Human Genome Management Information System (HGMIS), Betty Mansfield (Fax: 615/574-9888, bkq@ornl.gov): *Human Genome News*, cosponsored by NCHGR and DOE, and DOE Human Genome Program *Primer on Molecular Genetics* distributed to more than 30,000 high school, college, and medical educators; accessible through WWW (http://www.ornl.gov/TechResources/Human_Genome/home.html). HGMIS staff answer questions from scientists and teachers and put them in touch with others and with educational organizations and consumer groups.

NCHGR and DOE Centers

Most human genome centers have outreach and education programs. Some examples are given below with contact information.

Baylor College of Medicine Human Genome Center: Presentations for junior and senior high school teachers and students, including video, slide presentation, lecture, discussions, brochures, and information on teaching materials and continuing-education opportunities. Quarterly newsletter. [Contact: Belinda Rossiter (Fax: 713/798-5386, rossiter@bcm.tmc.edu)]

Cooperative Human Linkage Center, University of Iowa: New high school curriculum module on nontraditional inheritance in collaboration with BSCS staff. Teacher workshops on technology and ELSI issues. Science teachers hosted in laboratory for 2 weeks to 3 months. Summer training for high school students. Three-month visiting fellowship program for nonscientist professionals demonstrating strong interest in ELSI issues. Gene-mapping project developed for use in secondary science classrooms. Newsletter. Collaborations in Science and Technology Radio Project (see p. 2). [Contacts: Nancy Newkirk (nancy-newkirk@uiowa.edu) or Jeffrey Murray (jeff-murray@umaxc.weeg.uiowa.edu), Fax: 319/335-6970]

Lawrence Berkeley Laboratory (LBL) Human Genome Program Coordination: ELSI Issues in Science seminars,

☆☆Notice to DOE Contractors and Grantees☆☆

The fifth DOE Human Genome Program Contractor and Grantee workshop will be held January 28-February 1, 1996, in Santa Fe, New Mexico. At least one investigator from each funded project is expected to attend the entire meeting and represent the project at poster sessions. Some projects will also be represented in platform presentations. More information will be forthcoming from Sylvia Spengler; Human Genome Program Coordination; 459 Donner Laboratory; Lawrence Berkeley Laboratory; Berkeley, CA 94720 (510/486-4879, Fax: -5717, sylviaj@ux5.lbl.gov).

NCHGR Publishes Brochure, Progress Report

The Human Genome Project: From Maps to Medicine is a new color brochure from the NIH National Center for Human Genome Research (NCHGR) about the genome project and its expected impact on medical genetics. A four-page foldout illustration depicts the flow of discovery from gene isolation to treatment for inherited illnesses. Suitable for a general audience.

Progress Report: Fiscal Years 1993-1994 reviews progress of the NCHGR Division of Extramural Research toward the goals of the U.S. Human Genome Project. The report, which is intended for a technical audience, lists accomplishments and resources available to the scientific community.

Contact for free copies: NCHGR Office of Communications (301/402-0911, Fax: -2218, SharonD@od.nchgr.nih.gov).

lectures for LBL staff, ELSI curriculum for WWW. Student and teacher training programs and research opportunities for community college students. Lectures at high schools, colleges, universities, and community events. Postdoctoral fellowships. Summer and semester-long research programs. Summer internships for minority high school students and teachers. ELSI in Science Home Page (<http://www.lbl.gov/Education/ELSI/ELSI.html>) [Contact: Catherine Pinkas (Fax: 510/486-5717, pinkas@mh1.lbl.gov)] **LBL Human Genome Center:** Tours for students, teachers, and lay people. [Contact: Jennifer Knox (Fax: -6746, jlknox@lbl.gov)]

Lawrence Livermore National Laboratory Human Genome Center: Center tours for high school and college classes, medical students, and visiting congressional aides; summer training for teachers and internships for high school and college students; 2-week traveling course in molecular biology for local schools. Lectures in high schools and colleges and talks to civic groups. [Contact: Linda Ashworth (Fax: 510/422-2282, Ashworth1@llnl.gov)]

Los Alamos National Laboratory Center for Human Genome Studies: Twelve invited lectures to various New Mexico organizations each year. *Los Alamos Science: The Human Genome Project* distributed to individuals, universities, high schools, medical schools, and private companies. Comprehensive database of ELSI publications developed by

Michael Yesley and staff; three editions of printed bibliography published since 1992. Library collection, including copies of most materials, open to researchers; custom searches through Yesley (Fax: 505/665-4424, msy@lanl.gov).

Stanford Human Genome Center: Three high school curriculum units on DNA profiling and genetic disease testing, screening, and treatment; laboratory experiments integrated with social and ethical decision making. Teacher workshops, mobile laboratory kits, and school open houses. Lectures on genome science and human genetics for teachers and students, summer research internships, support for local public science museums, and genome center tours. Project partially supported by Perkin Elmer Corporation. [Contacts: Lane Conn (lconn@shgc.stanford.edu) and Cynthia Keleher (keleher@shgc.stanford.edu), Fax: 415/812-1916]

University of Michigan Medical Center, Diane Baker (Fax: 313/764-4133): One-week workshops (including resource material) for high school teachers. One-week course for genetic counselors; covers molecular diagnostics, genetic counseling, and the Human Genome Project.

Colleges, Universities, State Departments of Education

Air Academy High School, Colorado Springs, Doug Lundberg (Fax: 719/472-1413, lundberg@kadets.d20.co.edu): One-week summer course on genetic

engineering sponsored by the Colorado Department of Education for 8 to 12 new and 10 to 20 returning teachers; video tapes and student discussions via Internet during school year. GENTALK Listserv created and moderated by Lundberg for teachers, students, and researchers. Addresses laboratory protocols, technical questions, genetic engineering, and bioethical issues. No charge. To subscribe, send a message to listproc@usa.net with *subscribe gentalk firstname lastname* in the body.

Johns Hopkins University, Robert Robbins: WWW base page (<http://www.gdb.org/rjr/rjr.html>) under development; access to growing collection of technical and educational materials on molecular biology and informatics, presented in Adobe Acrobat format (which gives camera-ready quality). Site includes manuscripts, preprints, reprints, slides, transparencies, and data relevant to computational biology. Links to sources for downloadable free Acrobat Reader software.

Kingwood College Biotechnology Department, Kingwood, Texas, Brian Shmaefsky (Fax: 713/359-1612, bshmaefs@kc.nhmocd.cc.tx.us): Demonstrations at public schools about the impact of biotechnology and genomic information on daily life, career opportunities in biotechnology. Biotechnology workshops for teachers.

University of Stellenbosch (US) and Medical Research Council Centre for Molecular and Cellular Biology, Cape Town, South Africa, Valerie Corfield (Fax: +27/21-931-7810, vc1@maties.sun.ac.za): US TOO (Teacher Operation Outreach) program to train teachers through workshops in molecular genetics and ethical issues.

Washington University, St. Louis, Cynthia Moore (Fax: 314/935-4432): Development and field testing of a molecular genetics curriculum unit for all levels of high school students, emphasizing hands-on activities and bioethical decision making. Funded by the NIH Science Education Partnership Award Program.

Programs Affiliated with Corporations and Private Industry

Access Excellence (<http://www.gene.com:80/ae>): Genentech-sponsored initiative launched last year with a 5-day course for 105 biology teachers from all 50 states and Puerto Rico. Computer networking via America Online (AOL) to

connect teachers with researchers, computerized information services, and each other. Three-week Science Seminar online discussion group (to volunteer, send message to BioEditAE@aol.com describing topic chosen). AOL software, free diskette containing 50 favorite lesson activities of the AE fellows in Macintosh or DOS, and quarterly newsletter: 800/295-9881.

Edison BioTechnology Centers, Ohio: Science and Societal Issues Symposium of presentations by high school students, modeled on a Lawrence Hall of Science project. Examples of 1995 student research topics: genetic testing, gene therapy, fetal tissue research, and release of genetically engineered microorganisms. [Contact: Bettie Sogor, Symposium Coordinator (Fax: 216/229-7323)]

Human Genome Symposium for Teachers: One-day workshop conducted by Lynne Gordon (San Diego High School, Fax: 619/231-0973, Lgordon@ec.edcs.k12.ca.us) and Patricia Winter (General Atomics, Fax: 619/455-3379, winters@vaxd.gat.com). Hosted and underwritten by General Atomics, which also sponsors other teacher workshops.

Keys to Science Institute, Keystone, Colorado: Intensive 12-day summer teacher training in molecular and cellular biology with emphasis on the role of biotechnology in current medical and scientific research. Interested corporations sponsor one or more teachers. Teachers are matched with a resource liaison mentor in their area, often a practicing scientist from the sponsoring company, who helps to implement the institute curriculum in the classroom. [Mary Schwartz,

Wisconsin Biolissues

The June issue of *Wisconsin Biolissues*, published by the University of Wisconsin Biotechnology Center (UWBC), focuses on biotechnology education and resources. Subscriptions free. (UWBC Documents; 1710 University Ave.; Madison, WI 53705 (608/262-2604, Fax: -6748))

Teacher Institutes Coordinator, Keystone Science School (Fax: 970/468-7769, [tkck\\$\\$@keystone.org](mailto:tkck$$@keystone.org))

Woodrow Wilson National Fellowship Foundation (WWNFF): Biology institutes stressing hands-on teaching and learning strategies in selected themes (including bioethics, biotechnology, and evolution) relevant to middle- and secondary-level teachers. Administered by the foundation's National Leadership Program for Teachers (NLPT) and funded by the Howard Hughes Medical Institute. Through national competition, 50 teachers are chosen for each institute. Participants may apply for outreach grants; teacher teams lead 1-week workshops at selected sites nationwide. [NLPT; WWNFF; CN5281; Princeton, NJ 08543-5281 (Fax: 609/452-0066, dale@wwnff.org)] [Anne Adamson, HGMS, with introductory text by Sarah Elgin, Washington University, St. Louis] ♦

Students from Gunn High School in Palo Alto, California, put together a YAC map model during a Stanford Human Genome Center tour.



CSHL Mapping and Sequencing Meeting Held

The Eighth Annual Genome Mapping and Sequencing Meeting, held May 10–14 at Cold Spring Harbor Laboratory, was attended by more than 450 participants with a strong international representation. Over 300 abstracts covered a broad array of topics related to genome analysis of numerous organisms. The meeting was organized by David Bentley (Sanger Centre, U.K.), Eric Green [NIH National Center for Human Genome Research (NCHGR)], and Robert Waterston [Washington University (WU)].

Sessions covered a variety of areas, including gene discovery and transcript mapping, informatics, mapping methods and technologies, physical mapping of human chromosomes, DNA sequencing, model organism mapping and biology, and human genetics and biology.

Rapid progress was demonstrated in all areas, with particular emphasis on consolidating and integrating different approaches for human

genome mapping. The human genetic map based on short tandem repeat polymorphisms is approaching completion, and the Généthon genetic map is essentially finished (J. Morissette, Généthon) with over 5000 (CA)_n-type markers developed and mapped.

Similarly impressive progress was reported by the Cooperative Human Linkage Center (CHLC) group in the continued growth of its high-quality, well-integrated genetic map (K. Buetow, Fox Chase Cancer Center). Work by other groups will lead to further map refinement and mapping of additional genetic markers.

Construction of physical maps of human chromosomes, in both genome-wide and chromosome-specific efforts, is well advanced. High-resolution maps of chromosomes 16 and 19 are essentially complete and feature BACs, PACs, YACs, and cosmids, with integration of genes and genetic markers [see *HGN* 6(5), 2–3 (January–February 1995)]. YAC-based maps of many

Two Bacterial Genomes Sequenced

At the 95th meeting of the American Society of Microbiology held May 21–25 in Washington, D.C., Craig Venter [The Institute for Genomic Research (TIGR)] and Hamilton Smith (Johns Hopkins University) announced the complete sequencing of two bacterial genomes. With in-house TIGR support, the 1.9-Mb *Haemophilus influenzae* genome was finished and all gaps closed in less than a year; funding from the DOE microbial genome project, administered by Jay Grimes, allowed the 580-kb *Mycoplasma genitalium* genome to be completed in 3 months by Claire Fraser's team at TIGR.

"This is really an incredible moment in history," said Frederick Blattner (University of Wisconsin), who heads the NIH project to sequence the bacterium *Escherichia coli*. "It demonstrates the ability to take the whole sequence of an organism and work down from that to its genes, which is what geneticists have been dreaming of for a long time." Blattner noted that geneticists traditionally have studied genes by identifying functions that are impaired when a gene is mutated.

Whole-Genome Approach

The usual method of breaking DNA into overlapping segments (mapping), sequencing the pieces, and reassembling them is time-consuming. Smith (who won the Nobel prize for isolating restriction enzymes) and Venter developed a whole-genome shotgun sequencing approach that skipped the mapping stage. They employed ultrasonic waves to break the DNA into fragments, which were then sequenced and reordered with computer software developed at TIGR.

Free-Living Bacteria

H. influenzae and *M. genitalium* are free-living; that is, they contain all the genetic information needed for living independently. This contrasts with viruses, many already sequenced, which lack genes for independent living and replicate by using genetic information from the cells they infect.

M. genitalium, a eubacterium thought to be the simplest known self-replicating and free-living life form, has been used as a model for the minimum number of genes and protein products necessary for independent existence. As with other *Mycoplasma* species, it has been shown to have a very rapid rate of evolution.

Third Organism Under Way

Venter reported that his primary DOE project to sequence *Methanococcus*

jannaschi is ahead of schedule and should be completed by September—less than a year after the award was made. *M. jannaschi*, an extreme thermophile belonging to the ancient Archaeobacterial family, was isolated at the base of a Pacific Ocean thermal vent.

Sequence Analysis

Venter's group has started to analyze the full DNA sequence of *H. influenzae*, a common cause of ear infections in children. He said that although "it will take all of us months, if not years, to truly understand it," he already has predicted the biological role of most of the 1749 genes by comparing them with other genes of known function. After the paper is published in *Science* in late July, the entire *H. influenzae* sequence will be deposited in the Genome Sequence Data Base. Sequence data and a table of putative gene identifications and role categories will be available through the TIGR Home Page (<http://www.tigr.org>).

Venter predicted that TIGR could sequence ten or more microbrial genomes each year with high efficiency. Further research on these and other bacterial genomes will enable researchers to identify bacterial genes, including those responsible for causing disease, and help search out comparable human genes that may be involved in disease processes. ◊

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

Production Manager/Editor
Judy M. Wyrick

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

Special thanks to
Kathleen H. Mavoumin

HGMIS Address
Betty K. Mansfield
Oak Ridge National
Laboratory
1060 Commerce Park
MS 6480
Oak Ridge, TN 37830
615/576-6669
Fax: 574-9888
bkq@ornl.gov
[http://www.ornl.gov/
TechResources/Human_
Genome/home.html](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
Daniel.Drell@mailgw.
er.doe.gov

Leslie Fink
NIH National Center for
Human Genome Research
Bethesda, MD
301/402-0911, Fax: -2218
LeslieF@od.nchgr.nih.gov



chromosomes are progressing steadily, and YAC contigs covering chromosomes 22, X, and 12 are at or near completion (similar to those for 21 and Y). A great deal of progress also has been made for chromosomes 3, 4, 7, 10, 11, and 13. Although other chromosomes are mapped less extensively, work is proceeding with considerable momentum, particularly at Whitehead Institute—Massachusetts Institute of Technology (MIT) and CEPH-Génethon.

Remaining chromosome maps are likely to be constructed more quickly than those now approaching completion. While YACs continue to be the major source of clones for large-scale projects, the newer, large-insert bacterial cloning systems are emerging as important components for physical mapping of mammalian chromosomes. Contigs based on these bacterial clones have been constructed in a few specific regions of the human genome. P1s, PACs (P. de Jong, Roswell Park Cancer Institute), and BACs (H. Shizuya, California Institute of Technology; B. Birren, Whitehead—MIT) are contributing significantly. Such bacteria-based cloning systems will be important for converting lower-resolution framework maps into a more suitable sequencing form.

An important complementary technology for constructing highly integrated maps of human chromosomes is whole-genome radiation hybrid (RH) mapping [D. Cox, Stanford University (SU); G. Gyapay, Génethon]. This technique promises to provide a framework map of ordered markers; a subset of existing genetic markers already is being used to integrate evolving RH and genetic maps. RH mapping offers an alternative method for long-range ordering of landmarks, which will be particularly vital for genome-wide mapping.

Identification and mapping of human genes continues to be a major focus of attention. Over the last year, cataloging genes by "tag" (i.e., EST) sequencing has grown explosively: the work of The Institute for Genomic Research (M. Adams) is complemented by the Genexpress program (R. Houlgatte, Centre National de la Recherche Scientifique, France) and the recently established St. Louis—Merck initiative (L. Hillier, WU). The latter efforts are being incorporated rapidly into the RH mapping program by an international consortium of U.S. and European centers and mapped onto YAC clones (J. Sikela, University of Colorado) to build comprehensive and integrated transcript maps.

The challenge of finding mammalian genes in a more targeted fashion continues to stimulate new developments, including bacteriophage lambda (T. Boehm, German Cancer Center, Heidelberg) and cosmid-based exon-trapping systems (G. van Ommen, Leiden University, Netherlands). Other cDNA-based methods, such as development of chromosome-specific cDNA direct-selection libraries (M. Lovett, University of Texas), also are being refined.

The rapidly growing catalogs of gene tags derived from ESTs, direct-selection libraries, and CpG island libraries (S. Cross, Edinburgh University)

will provide better access to coding sequences and, in the long run, more powerful ways to identify genes within genomic sequence. Methods geared to studying gene expression, using refinements in technologies such as differential display (T. Ito, University of Tokyo), also will be critical if the wealth of new gene information is to be exploited fully.

The meeting produced reports of exciting advances in genome mapping technology, including glimpses of possible future technologies involving microfabricated chips (D. Burke, University of Michigan) and improved throughput for optimal mapping of single DNA molecules (D. Schwartz, New York University). Other advances were high-throughput physical mapping approaches incorporating multiple complete restriction digests [J. Yu, University of Washington, Seattle (UWS)], creating new yeast strains for facilitating YAC isolation (L. Borbye, NCHGR), using FISH to study duplicated genomic segments (B. Trask, UWS), and developing vectors for introducing large DNA segments into mammalian cells [C. Huxley, St. Mary's Hospital (SMH), London].

With the genome project rapidly approaching a critical mapping-to-sequencing transition, major accomplishments in various large-scale DNA sequencing projects were not surprising. Model-organism sequencing continues to lead with the completion of 16 Mb in *Caenorhabditis elegans* (M. Berks, Sanger Centre), 2.8 Mb in *Drosophila* (M. Palazzolo, Lawrence Berkeley Laboratory), and, most important, the entire *Saccharomyces cerevisiae* genome by year's end (M. Cherry, SU).

Significant progress was also reported in sequencing megabase-sized stretches of human DNA [D. Buck, Sanger Centre; B. Roe, University of Oklahoma; E. Chen, University of Houston; A. Rosenthal, Institute of Molecular Biology, Germany; and D. Nelson, Baylor College of Medicine (BCM)]. Although no single advance in DNA sequencing technology was described, a general consensus seemed to emerge that current approaches could be scaled cost-effectively for large-scale sequencing of human DNA. This optimism was based on impressive evolutionary advances in almost every step of typical large-scale sequencing projects, including development of improved algorithms that require less decision making by humans. The effect of these advances is a dramatic reduction in overall DNA sequencing cost.

To keep pace with the explosive accumulation of new mapping and sequencing data, genome informatics continues to grow and mature. Advances were showcased in a platform session and projection-style computer demonstrations. The platform session highlighted a variety of areas, including the establishment of databases [e.g., the Integrated Genome Database (IGD) for human mapping data (O. Ritter, DKFZ, Heidelberg)], a yeast database (Cherry), and a human gene database derived from the

wealth of new EST sequences [G. Schuler, National Center for Biotechnology Information (NCBI)].

Other talks reported development of informatics tools for organizing laboratory work in a large genome center (L. Stein, Whitehead-MIT), aligning and analyzing sequence (R. Smith, BCM), and automating high-throughput genotyping (T. Christenson, Marshfield Medical Research Foundation); improved software for image processing (D. States, WU); and algorithm refinement for analyzing RH mapping data (T. Matise, Columbia University). New to this year's meeting were daily projection-style computer demonstrations that allowed real-time viewing of various software packages. Participants found these demonstrations effective in introducing numerous informatics tools for managing and analyzing genome mapping and sequencing data.

Genomic studies in other organisms continue to be critical for research, not only as models but also for further biological and genetic studies. Sequence analysis in *Escherichia coli* (E. Koonin, NCBI) continues to play an important role in assigning function to individual gene products.

The usefulness of yeast genome analysis was demonstrated, both in providing the ability to compare genomes by cross-referencing EST sequences (F. Spencer, Johns Hopkins University) and allowing systematic studies of gene function by mutation analysis (P. Ross-MacDonald, Yale University). The pufferfish *Fugu*, with its apparently compact genome, continues to reveal interesting possibilities for identifying genes in mammalian DNA (M. Trower, Glaxo Research and Development Ltd.; R. Sandford, Addenbrookes Hospital, Cambridge). The mouse genetic map consisting of >6000 (CA)_n repeat-type markers is now complete (E. Lander, Whitehead-MIT), while efforts are being initiated to build genetic and physical maps of the rat genome (H. Jacob, Whitehead-MIT). The value of high-resolution genetic maps was illustrated by two successful mouse positional-cloning projects identifying the genes for the shaker 1 deafness locus (S. Brown, SMH) and the nude locus (Boehm).

The importance of genome research in studying complex human diseases was illustrated by descriptions of novel genetic mapping strategies through analysis of isolated populations (A. Chakravarti, Case Western Reserve University) and DNA pooling approaches (V. Sheffield, CHLC). More detailed accounts of specific human DNA regions associated with human disease were also presented [G. Landes, Integrated Genetics; S. Glucksmann, Millennium Pharmaceuticals; E. Eichler, BCM; V. Van Heyningen and A. Brookes, both at Medical Research Council, Edinburgh].

By many accounts, the meeting highlight was the newly added keynote speaker session, which this year featured presentations by Maynard Olson (UWS) and John Sulston (Sanger Centre). These

talks drew together major themes emerging from the meeting while highlighting the current state of the genome project. Central to both talks were issues surrounding completion of the human physical map and initiation of large-scale human DNA sequencing.

The major emphasis of Olson's talk was that construction of high-quality human DNA maps must not be overlooked when plans are formulated for large-scale DNA sequencing. High-resolution maps will serve an important role in assembling DNA sequence contigs. Sulston's discussion focused on a proposal to begin large-scale sequencing in a coordinated, international effort that would yield an initial human genome sequence around the end of 2001. [Written for HGN by David Bentley (Sanger Centre), Eric Green (NCHGR), and Robert Waterston (WU)] ♦



Human Gene Map Workshops Held

The following article was adapted by HGN staff from two reports prepared for the Human Genome Organisation (HUGO) by freelance writer and editor Alison Stewart and published in Human Genome Digest 2 (2), 1-4 (April 1995) and 2(3), 6-9 (July 1995).

Substantial progress in developing a public-domain human transcript (gene) map was reported May 9-10 by researchers at Human Gene Map Workshop (HGMW) II in Cold Spring Harbor, N.Y., the second such workshop organized by HUGO. The Human Gene Map Initiative—an international effort to find and map expressed human genes and deposit the information in public databases—began last October in Washington, D.C., at a meeting sponsored by The Wellcome Trust; strategies were developed January 24-25 at HGMW I in London. This article summarizes highlights from the London and Cold Spring Harbor meetings.

Cooperation for mutual advantage was the "take-home message" of the January HGMW, which brought together academic and industrial scientists with representatives from pharmaceutical companies and institutions funding public research. At the meeting, an overall picture emerged of several loose associations or laboratory consortia that are sharing materials, coordinating activities to minimize overlap, and contributing their information to public databases. Various consortia reported using complementary strategies and exchanging information.

The strategy of choice for building up the gene map is based on ESTs—short, identifying sequences obtained by partially sequencing cDNAs. ESTs are obtained from cDNAs represented in arrayed libraries from various tissues. ➔

Contacts

ATCC

800/638-6597, Fax: 301/816-4361
help@atcc.org, Gopher: culture.atcc.org
<http://www.atcc.org/>

dbEST

<ftp://ftp.ncbi.nlm.nih.gov>
<http://www.ncbi.nlm.nih.gov/dbEST/index.html>

GDB

410/955-9705, Fax: /614-0434
help@gdb.org
<http://gdbwww.gdb.org>

IMAGE Consortium

info@image.llnl.gov
<http://www-bio.llnl.gov/bbrp/image/image.html>

Research Genetics

800/533-4363, Fax: 205/536-9016
<http://www.resgen.com>

WU-Merck EST Project

<http://genome.wustl.edu/est/esthmpg.html>

Genome News

Chromosome Web Activities

Chromosome 9 Workshop

The full report and abstracts from the 1995 Chromosome 9 Workshop are available in hypertext form via the Chromosome 9 Home Page (<http://www.gene.ucl.ac.uk/chr9home.html>) or by ftp from <ftp.gene.ucl.ac.uk> as .txt files in the <pub/c9workshop/1995> directory. [Comments or electronic problems: John Attwood (john@mrc-hbgu.ucl.ac.uk)]

Chromosome X Virtual Workshop

To accompany the June 15-18 X chromosome workshop in Banff, Alberta, a "virtual" workshop is being assembled by Bob Cottingham and David Nelson (both at Baylor College of Medicine) at <http://gc.bcm.tmc.edu:8088/chrX/home.html>. Through this page, users can participate in online dialog about X chromosome mapping and access workshop details and a list of participants and abstracts.

If suitable primers for an EST are designed, PCR can be used to amplify the corresponding sequence from genomic DNA. The EST is thus converted to an STS that can be mapped to a genomic location using radiation hybrids (RHs) or genomic clones such as YACs and BACs.

IMAGE

Charles Auffray (Centre National de la Recherche Scientifique and Généthon) and Greg Lennon (Lawrence Livermore National Laboratory (LLNL)), two founding members of the Integrated Molecular Analysis of Gene Expression (IMAGE) Consortium, reported on the use of arrayed cDNA libraries for gene sequencing, mapping, and expression studies [*HGN* 6(6), 3 (March 1995)]. All IMAGE collaborators deposit their data into public databases; as of June 27, 154,669 of the 206,654 human clone-derived EST sequences in dbEST were from IMAGE clones (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>). Records for over 147,000 IMAGE clones have been submitted to the Genome Data Base (GDB) as well. The LLNL group, which as part of the IMAGE consortium supplies cDNA clones to both the Washington University (WU)—Merck & Co. and Généthon sequencing groups, will construct master arrays (perhaps 10,000 at a time) of clones representing unique genes, regardless of the library of origin. A running update of unique genes resulting from the Merck Initiative is available from the IMAGE Home Page (see box, p. 7).

At the May meeting, Lennon set out some remaining tasks; for example, preparing master arrays of definitive cDNAs, sequencing full-insert cDNAs to complement genomic sequencing projects, and finding the remaining human genes.

Merck-WU Initiative

The Merck-WU Initiative, involving groups at Merck, WU, and LLNL, aims to isolate one cDNA clone for each expressed human gene. This collection, which will be distributed by nonprofit and commercial concerns at reasonable cost, will be called the Merck Index. By March 1996, according to Rick Wilson (WU) and Keith Elliston (Merck), the initiative plans to obtain ESTs from 200,000 clones from a variety of primary and normalized cDNA libraries. Since February 10, new ESTs obtained by Merck collaborators have been added to dbEST at the National Center for Biotechnology Information (NCBI) at the rate of 4000 to 6000 each week. In the first phase of the project, highly and moderately expressed genes are likely to be overrepresented. In the second phase, such techniques as oligonucleotide fingerprinting and subtractive hybridization procedures may be used to identify clones representing more rarely expressed genes. Perhaps 70 to 80% of the 50,000 to 100,000 genes will be represented in dbEST by the scheduled completion date for the Merck initiative.

International RH Mapping Consortium

David Cox (Stanford University) outlined the organization of an international RH consortium that includes laboratories at the Stanford Human

Genome Center and Whitehead-MIT Genome Center in the United States; University of Cambridge (UC), Sanger Centre, and Oxford University in the United Kingdom; and Généthon in France. The consortium employs two different panels of whole-genome RHs to generate high-resolution maps integrating human cDNAs with meiotic linkage markers. The Stanford RH panel consists of 83 hybrids and results in maps of 0.5-Mb resolution, and the Genebridge panel (Généthon-Cambridge) produces lower-resolution maps but allows markers to be linked more readily. Consortium laboratories are using RH DNAs distributed by Research Genetics Inc.

STSs generated from the 3' ends of cDNAs are being used in conjunction with RH DNAs to map most human genes over the next 2 to 3 years. A database at the Sanger Centre (RHALLOC) helps to coordinate efforts by listing markers being mapped in consortium laboratories. Raw mapping data and finished maps are being deposited in a variety of public databases, including European Bioinformatics Institute (EBI) and GDB.

ESTs and Genes

In early May, dbEST contained 132,674 human EST sequences and accounted for almost half the number of sequence entries in GenBank®. Of these sequences, 83,602 had been submitted by the Merck sequencing group. According to Mark Adams (The Institute for Genomic Research (TIGR)), over 100,000 additional ESTs generated by TIGR will be made available through the TIGR database.

The question of redundancy in the set of ESTs is an important one: Just how many different genes are represented? Groups at Merck, TIGR, Généthon, and NCBI are tackling this question by grouping sequences into overlapping clusters that are likely to originate from the same gene. In a preliminary version of the Merck Gene Index, 38,149 ESTs reduced to 17,743 different genes. Auffray estimated that Généthon's collection of nearly 27,000 EST sequences represents about 9000 to 10,000 different genes. Adams described development of a TIGR database of tentative human consensus sequences based on 270,000 private and publicly available ESTs.

At NCBI, a UniGene-UniEST database is being built to supply high-quality, nonredundant sequences to mapping groups. The UniGene set, compiled from all gene sequences in GenBank that have a bona fide 3' untranslated region, contains 3125 sequences. UniEST, with 13,900 sequences, represents a set of unique ESTs based on 3' sequence reads from the Merck sequencing group. Mark Boguski (NCBI) estimated that UniEST and UniGene provide sequence data for at least 15,000 different human genes, perhaps 15 to 25% of all human genes, and urged that mapping groups use sequences from both sets. Boguski and Greg Schuler (NCBI) will coordinate the supply

of gene-based mapping reagents to major mapping consortia and other interested groups.

As the number of ESTs grows, assessing the accrual rate of new sequences and developing strategies for finding unrepresented expressed sequences are important. The normalized cDNA libraries prepared by M. Bento Soares (Columbia University) continue to be the richest and cleanest source of novel ESTs; for example, the infant brain was still yielding substantial numbers of new sequences after about 20,000 clones had been sampled. To track down rare transcripts that may be represented only in specialized tissues, new libraries will be prepared from more specific cell and tissue types, for example, adult retina, spinal cord, pineal gland, and multiple sclerosis lesions (Soares); and inner ear, hair roots, and glomerulus (Kousako Okubo (Osaka University)). Some remaining problems in library construction include apparent differences in mRNA synthesis efficiency from different genes and the presence of clones without a poly A tail, for example, resulting from internal priming events—now less than 10% in Soares' libraries.

Building the Gene Map

Building a complete gene map requires information integration from different mapping approaches. Expressed sequences can be mapped at different levels including chromosome assignment (50- to 250-Mb resolution) and mapping to a DNA clone or contig or an RH (less than 1-Mb resolution). This year has seen a large increase in the number of ESTs and other markers mapped to the highest-resolution levels. Of more than 2000 IMAGE transcripts assigned to chromosomes by the time of the January meeting, perhaps half have now been mapped to higher resolution by the worldwide IMAGE community of over 100 researchers.

Mihael Polymeropoulos [National Center for Human Genome Research (NCHGR)] described progress in pilot experiments to map chromosome 8 cDNAs to YACs. Tom Hudson (Whitehead-MIT) reported screening 2300 gene-derived markers, of which 1500 were ESTs, against the CEPH mega-YAC library. This is part of a larger project in which 10,000 chromosomally assigned STSs have been screened against the YAC library, giving an average marker spacing of 300 kb. Currently, 70% of the contigs are anchored by genetic markers. Pilot experiments are under way to compare and integrate the map with RH maps. In general, good concordance exists among genetic, RH, and YAC maps.

Two sets of RHs are being used to construct framework maps with a resolution of around 500 kb that can be used to locate unknown ESTs. The Stanford hybrids are generated with a higher dose of X rays and have about twice as many breaks, thus allowing mapping at higher resolution, but the Cambridge-Généthon set allows significant linkage to be obtained with a higher fraction of cDNAs. These sets were

described by Cox, Karin Schmitt (UC), and Jean Weissenbach (Généthon). Pilot EST mapping studies were explained by Cox and David Bentley (Sanger Centre) for the Stanford set and by Weissenbach for the UC-Généthon set. Cox also described the development of a new set of hybrids that can be used to construct maps at around 100-kb resolution; about 30,000 markers will be needed for this panel.

Ken Buetow (Cooperative Human Linkage Center) described progress toward generating, maintaining, and distributing integrated high-resolution genetic maps highly enriched for user-friendly, PCR-based markers. ESTs are proving to be a source of useful genetic markers that can be linked explicitly to the genetic map.

As the gene map is built up, it will provide reagents and information that can be applied to genome sequencing. BACs, with an average insert size of 150 kb, could be an important mapping resource in establishing a sequence-level map. In a pilot study described by Hiroaki Shizuya [California Institute of Technology (Caltech)], 78 of 80 STSs from chromosome 22 were positive on the Caltech BAC library of 100,000 clones, while only 3 of 60 chromosome 22-specific cDNAs found no corresponding BAC clone. Research Genetics, in collaboration with Caltech, has prepared mouse and human genomic BAC libraries for distribution (see article below).

Construction of suitable PCR primers allowing EST conversions to genomic STSs is, for many researchers, the most expensive component →

Genome News

HGM 96

HGM (Human Genome Meeting) 96: March 22-24 in Heidelberg, Germany. [Conference Secretariat; HUGO Europe; One Park Square West; London NW1 4LJ, U.K. (+44/171-935-8085, Fax: -8341)] ◊

Biological Resources

Monochromosomal Somatic Cell Hybrid Panel

A new monochromosomal somatic cell hybrid panel from BIOS Laboratories consists of 24 distinct hybrids, of which 21 contain one and 3 contain two chromosomes. Most of the hybrids, developed and characterized by Raghbir Athwal (Temple University Medical School) in collaboration with BIOS, are constructed in the A9 mouse cell line, with four remaining in a Chinese hamster ovary background. The human donor genome is derived from the normal human fibroblast cell line GM0634. [BIOS Laboratories, Inc.; New Haven, Conn. (800/678-9487 or 203/773-1450, Fax: 800/315-7435)] ◊

BAC Libraries Available

In collaboration with researchers led by Melvin Simon (California Institute of Technology), Research Genetics has constructed mouse and human genomic BAC libraries from which DNA pools, high-density membranes, and individual clones are now available. Library clones are arrayed in 384-well microtiter plates. For each library, large genomic DNA is prepared from cultured mouse 129/sv or human 978SK cell lines. Average insert size is 100 to 150 kb. [Research Genetics, Inc.; Huntsville, Ala. (800/533-4363, Fax: 205/536-9016, <http://www.resgen.com>)] ◊

Coriell Offers Trinucleotide Samples

DNA samples with characterized trinucleotide expansions are available from Huntington's disease and myotonic dystrophy patients and from fragile X syndrome patients and unaffected carriers. Printed catalog. [Human Genetic Mutant Cell Repository; Coriell Institute for Medical Research; Camden, N.J. 08103 (800/752-3805 or 609/757-4849, Fax: -9737)] ◊

Genome News

of a mapping project. Adams described progress in TIGR's undertaking to provide 10,000 STS primer pairs for ESTs in public databases. By May, 4014 primer pairs had been ordered from Applied Biosystems. Of 2700 received, about 900 were checked and over 700 distributed to members of the RH mapping consortium. Primers developed by TIGR will be sold via American *Type Culture* Collection (ATCC) on condition that information resulting from their use is placed in public databases as quickly as possible.

Mapping groups at the May meeting expected to map around 37,500 ESTs within the next year; this probably represents a realistic total of 10,000 to 20,000, a huge increase over the nearly 1000 mapped at the end of 1994. Direct cost estimates for mapping an EST ranged from around \$170 to \$240, taking into account some 30% that fail; costs could fall if this percentage can be lowered.

At the January meeting, Robert Waterston (WU) outlined his scenario for completing a sequence-level map of the entire human genome by 2001 with 99% coverage and 99.99% accuracy in coding regions. Waterston estimated that this ambitious project could be completed by three sequencing centers at a cost around \$0.10 per base. In the view of Waterston and John Sulston (Sanger Centre), mapping and sequencing are part of a continuum; transcript mapping should be seen as an intermediate stage in progress toward the ultimate sequence-level map, which will reveal the remaining genes. Then the "real work" of functional characterization will begin. In this context, supporting complementary work on other mammals such as the mouse is seen as vital in helping to identify true transcripts, aid mapping by taking advantage of syntenic relationships between genomes, and provide a system for functional studies.

Informatics and Data Exchange

To be usable by the scientific community, all information generated by public-domain gene-mapping projects must be accessible. Boguski, Graham Cameron (EBI), Keith Elioston (Merck), and Chris Fields [National Center for Genome Resources (NCGR)] outlined some aspects of the informatics network being set up to meet these requirements.

All public-domain ESTs are deposited in dbEST, which, as the result of an international data agreement,

is mirrored at EBI. Boguski (NCBI) described the development of Chromoscope, which allows sequence retrieval from dbEST by map location. Chromoscope acts as a "front end" to ENTREZ, which interfaces with other data sources on protein sequences and structures, for example. dbEST links to the Genome Sequence Data Base (GSDB) at NCGR. Release 2.2 of GSDB, due in August, will feature an alignment representation of sequence information. Its coordinate system will enable representation of multiple overlapping or allele sequences such as those for "disease" genes. Ken Fasman (Johns Hopkins University) described work under way at GDB, the major public repository for human map data. GDB release 6.0, due in early fall, will contain linkage, physical, and RH map data and information on reagents such as clones and ESTs. An enhanced WWW interface for GDB 6.0 will also be ready at the same time.

Many major laboratories involved in the Human Gene Map Initiative have developed their own WWW servers. Although these are probably the best way for inquirers to access the most up-to-date data, attendees argued forcefully that such sites are not an alternative to public databases. All laboratories have a responsibility to provide their data to at least one public database. Database curators must ensure that the information is distributed among different database sites and that the "average user" can gain access to information. Considerable confusion could be avoided if every reagent, such as a cDNA clone, carried a unique identifier (e.g., an IMAGE identification number and associated GDB accession number) used always by anyone submitting information to a database.

Identification systems by various consortia should ensure that all materials, such as clones and arrayed libraries, are traceable to their source. As part of the Merck initiative, the Computational Biology and Informatics Laboratory (University of Pennsylvania) is developing software for tracking clones, libraries, and all information accumulated about them. Many clones and libraries, such as those in Lehrach's reference library system, are available directly from laboratories where they are maintained [*HGN* 6(3), 7 (September 1994)]. Others, such as the Caltech BAC library, clones generated by the IMAGE consortium, and Cambridge-Généthon and Stanford RH panels, are being made available to individual users at reasonable prices via such sources as Research Genetics and ATCC.

New types of "downstream" public databases are also being developed for information such as gene expression patterns. Although the idea of a single, integrated database sounds appealing, this is probably not feasible. Attendees argued that a set of databases supporting different types of query represents a richer

(see *HGMW*, p. 11)

Harvard Releases Encyclopaedia of *Drosophila*

Release 1.0 of the *Encyclopaedia of Drosophila* (EofD) Macintosh CD-ROM version is a graphical interface based on ACaDB software developed for the *Caenorhabditis elegans* database, now a powerful browsing and querying tool for many genome databases. Release 1.0 contains much FlyBase data and all Berkeley *Drosophila* Genome Project data except DNA sequences. The long-term goal is to provide different views of the same comprehensive data set on both EofD and the FlyBase Home Page (<http://morgan.harvard.edu>).

EofD runs in native mode on higher-end Macintosh computers equipped with 68040 or PowerPC processors and a CD-ROM drive. It requires System 7.1 or higher (System 7.5 to use Guide help facilities), a minimum of 16 MB of RAM, and virtual memory of at least 32 MB. Minimal cost. [Orders: EofD; c/o Dawn Palmer, Biological Laboratories, Harvard University; 16 Divinity Ave.; Cambridge, MA 02138 (617/495-2906, Fax: -9300, eofd-sales@morgan.harvard.edu). EofD structure questions: eofd@fly2.berkeley.edu] ♦

NSF Publishes *Arabidopsis* Booklet

Multinational Coordinated Arabidopsis thaliana Genome Research Project. Progress Report: Year Four (1995, NSF 95-43) summarizes project status and achievements and makes recommendations for the future. The 48-page booklet includes traditional scientific reports and a collection of brief articles for a more general audience. The text is available electronically via the Internet on the National Science Foundation Home Page (<http://www.nsf.gov>), and a hard copy may be ordered at pubs@nsf.gov. [Contact: Machi Dilworth (703/306-1422, mdilwort@nsf.gov)] ♦

Advisory Council Notes Progress

The National Advisory Council for Human Genome Research convened for its 13th meeting on January 30 in Washington, D.C. Francis Collins, Director of the NIH National Center for Human Genome Research (NCHGR), presided. Noting that 1994 was a banner year for scientific achievements in the Human Genome Project, Collins highlighted progress in physical and genetic mapping and in sequencing.

Report of the Center Directors' Meeting

Lloyd Smith (University of Michigan) reported that the December 15, 1994, meeting of NCHGR directors of genome science and technology centers (GESTECs) focused on transition to large-scale sequencing. At the meeting, three concerns related to the scientific community's desire for transcription maps were cost, funding balance between production and technology development, and portfolio division among technologies.

Smith pointed out that the NCHGR budget proposed at the directors' meeting includes a seven-fold funding increase for sequencing in a few

years, representing a reorientation of project goals. After Robert Waterston (Washington University) outlined for the directors a concrete plan for accomplishing human genome sequencing (see "Scenario" below), meeting attendees agreed on the need to determine how sequencing methods will scale, where substrate cosmids will originate, and how to assess accuracy. Smith also described a proposal by Leroy Hood (University of Washington, Seattle) to regroup genome project goals into high resolution (sequence) and low resolution (genetic and STS maps).

David Botstein (Stanford University School of Medicine) recommended that political considerations not deter the council from their sequencing emphasis and that mouse sequencing, while critical, be added only when completion of the human sequence is within sight. He asserted that mouse geneticists can readily study sequences conserved in human and mouse. Collins stated that the current 5-year plan includes limited mouse sequencing in parallel with human DNA sequencing.

A Scenario for Sequencing the Human Genome

Waterston reviewed the status of *Caenorhabditis elegans* sequencing for the council and described his ambitious scenario, devised with John Sulston (Sanger Centre), for human genome sequencing. Drawing upon their experience in

Genome News

Web Site

Jurg Ott's group (Columbia University) recently set up a WWW site that includes announcements on genetic linkage courses, newsletters, software, and publications (<http://linkage.cpmc.columbia.edu>). Linkage-analysis packages are also available via ftp (<ftp://linkage.cpmc.columbia.edu>).

Videotapes

Stories from the Scientists is a set of classroom videotapes and teachers' guides growing out of the 1992 symposium *Winding Your Way Through DNA*. The 30-min. documentary describes the collaborations between James Watson and Francis Crick, who discovered the structure of DNA; and Herbert Boyer and Stanley Cohen, who pioneered recombinant DNA techniques. Entire symposium also available on videotape. [Cold Spring Harbor Laboratory Press (800/843-4388 or 516/349-1930, Fax: -1946)]

HGMW (from p. 10)

overall resource, as long as connections are transparent to the user.

Funding and Future Meetings

At the January meeting, Francis Collins (NCHGR), David Owen (U.K. Medical Research Council), David Smith (DOE), and Michael Morgan (Wellcome Trust) pledged continuing support for projects integral to all gene-mapping consortia. As outlined by Manuel Hallen [European Commission (EC)], a welcome injection of new money is coming from the EC's new Biomedicine and Health Research Programme (Biomed 2). In 1995 and 1996, it will allocate 40 million ECUs to transnational genome research projects in the European region.

At the close of the London meeting, Peter Goodfellow (UC) appealed for increased cooperation between public and private sectors. He pointed out that academic institutions provide "seed corn" for private industry in the form of both people and ideas; companies willing to invest in the public sector can expect a rich return. Merck's continuing involvement is seen as essential to the rapid assembly of a public-domain gene map. Other companies represented at the meeting also signaled an interest in playing a role.

Two more meetings for fall 1995 and spring 1996 will be organized by HUGO to report progress in the Human Gene Map Initiative and usher in the era of the sequence-level map.

Database Workshops Held

Two workshops have been held entitled Interconnection of Molecular Biology Databases. The first was at Stanford University on August 9-12, 1994, with a 1995 follow-up on July 20-22 in Cambridge, England. Several workshop-related documents are available now via WWW (<http://www.ai.sri.com/~pkarp/mimbd.html>). These include the 1994 final report; attendee abstracts and contact information; bibliography on database interoperation; summary of biological databases and WWW pointers; and a list of presentations for the 1995 conference.

Some 55 bioinformatics researchers, computer scientists, and biologists from 9 countries attended the 1994 meeting, which surveyed the roughly 100 existing databases and requirements for integrating the diverse information they contain. Computer scientists presented an overview on the need for database interoperation and suggested techniques for solving the problem. Participants described a wide range of approaches that are presently generating practical results, such as systems allowing multi-database queries to the sequence databases, Genome Data Base, and Protein Data Bank. These systems and approaches vary according to their capability to handle complex queries, implementation difficulty, required user expertise, and scalability.

The workshop identified a number of barriers to interoperation, including resistance to standards, database inaccessibility to structured query via Internet, and inadequate documentation of many databases. Attendees felt, however, that interoperation is proceeding at a rapid pace and will soon enable researchers to answer questions that are laborious or impossible today. [Peter Karp, SRI International Artificial Intelligence Center (pkarp@ai.sri.com)]

Genome News

coordinating the *C. elegans* project, Waterston proposed that a human genome map should be produced at sequence level by 2001 to locate all human genes from various sources in context. Sequencing would be done chromosome by

chromosome, permitting methodological flexibility; and mouse and human data comparison could be included at some stage. Waterston's strategy is to carry out low-pass shotgun sequencing with fully automated data analysis. Results expected are 99% coverage with 0 to 2 gaps per cosmid, oriented contigs, estimated gap sites, and 99.9% accuracy at an estimated \$.10 per base. For each center doing 200 Mb annually, the cost will be \$20 million to \$25 million each year.

Jane Peterson (Chief, NCHGR Mammalian Genomics Branch) presented for discussion a draft RFA, arising from the December GESTEC meeting, for pilot projects for large-scale sequencing of human DNA. After suggested revisions have been made, the draft will be brought to the council again.

A second RFA proposed for concept clearance involved sequencing the *Escherichia coli* genome. Discussion leader Robert Strausberg (Chief, NCHGR Sequencing Technology Branch) observed that cost-effectiveness must be demonstrated and time lines and milestones made clear. He suggested that applicants establish goals for accuracy and that peer reviews evaluate them. The goal is to complete these projects in 2 years at a cost of no more than \$2 million, which the group agreed seemed realistic for research groups already doing sequencing. The council endorsed the RFA as presented.

ELSI Education Portfolio

Elizabeth Thomson (Acting Chief, NCHGR ELSI Branch) reviewed ELSI education grants for FY 1990-94 and reported that 122 applications were received for FY 1995, of which 22 were education applications. She added that the NCHGR ELSI program has funded about two to four education programs each year and that a group of advisors will be invited to discuss priorities for funding ELSI education grants.

Reporting on the DOE ELSI program, Daniel Drell stated that DOE concentrates on genetic privacy research and scientific and ELSI education. Applications have stabilized at 35 per year with one review cycle. The previous 2:1 ratio of research to education applications has reversed because the education applications competed more favorably in peer review.

Commenting on recent activities of the ELSI Working Group at which she represents the council, Anita Allen (Georgetown University Law Center) stated that the group collaborated with the NIH Office of Protection from Research Risks (OPRR) to issue informed-consent guidelines for which dissemination mechanisms are being explored. Joan Porter (OPRR) reported that participants at the first meeting of the NCHGR-National Cancer Institute cancer studies consortium had pointed out generic and genetics-specific problems with informed consent, some of which may be addressed in a limited way

(see Council, p. 13)

Informatics Resources

Baylor WWW Tools Available for Database Searching, Multiple Alignment

Several new WWW database-search services developed by Randall Smith's group [Baylor College of Medicine (BCM)] are available through the BCM Search Launcher (<http://gc.bcm.tmc.edu:8088/search-launcher/launcher.html>). The Search Launcher organizes searches related to molecular biology by function and provides a single point of entry for related searches. New tools and services include the following.

- BEAUTY (BLAST Enhanced Alignment Utility) incorporates information on the locations of conserved and annotated domains and sites (drawn from a local database and the Entrez, PIR, SWISS-PROT, PROSITE, BLOCKS, and PRINTS databases) directly into BLAST search results. These enhancements greatly facilitate detection of weak but functionally significant matches in BLAST database searches. (A more detailed description is available at <http://dot.imgen.bcm.tmc.edu:9331/seq-search/Help/beauty.html>).
- Interface to the National Center for Biotechnology Information's BLAST server adds links to the Entrez and SRS (Sequence Retrieval System) databases to BLAST search results. These links provide easy access to related information such as Medline abstracts.
- Multiple Sequence Alignment Server: CLUSTAL-W, MAP, and Pattern-Induced Multiple Alignment (PIMA) alignments can be run remotely on the BCM server, which uses the readseq program to input sequences in one of several formats (e.g., FASTA, MSF, NBRF, EMBL). Both DNA and protein multiple alignments can be performed.
- FASTAPAT and BLASTPAT Pattern Database Search Tools: Modified versions of FASTA and BLAST rapidly and sensitively search the PIMA Pattern Database for matches to protein sequences not identified by standard searches. A more detailed description of these programs is available (<http://dot.imgen.bcm.tmc.edu:9331/seq-search/Help/fastpat.html>).

Gene-Server Offers Internet Services

The University of Houston (UH) Gene-Server, managed by Dan Davison and funded by the National Science Foundation, the National Library of Medicine, and DOE, offers a number of Internet services for the genomics and molecular biology communities. These services include electronic access to software and data via ftp, Gopher, WWW, e-mail, and WAIS [see addresses below]. The Protein Information Resource (PIR), in release format and formatted for use in the GCG sequence analysis system, is available via ftp.

Another service from Gene-Server is Gene-Finder, to which a series of servers for human gene identification and protein secondary structure analysis were added recently in collaboration with Victor Solovyev (BCM). These servers are *texh* (find exons in human genes), *fgeneh* (human gene modeling), *hexon* (find internal exons), *hspl* (splice-site prediction for human genes), *nnspp* (nearest-neighbor-based protein secondary structure prediction), *rnaspl* (RNA splice-site prediction), *cdsb* (bacterial coding-region prediction) and *ssp* (segment-oriented protein secondary structure prediction). Full details on each service can be obtained at service@bchs.uh.edu with *man* <servicename> in the subject line (e.g., *man hexon*).

Another server allows recognition of human and bacterial sequences (HBR) to test a library for *Escherichia coli* contamination by sequencing sample clones. The program calculates the probability of each sequence being human (P) or *E.coli* (1-P) and the total percentage of human and bacterial sequences in the set. The method is based on linear discriminant functions [Solovyev et al., *Nucleic Acids Res.* 22(24), 5156-63 (1994)].

These servers have been added to the Bioinformatics Information Engine at the Weizmann Institute of Science in Rehovot, Israel. In the future, Gene-Server will focus on providing multiple sequence alignment via WWW, e-mail, and client-server programs. Some multiple sequence alignment is now available from the BCM Sequence Launcher in collaboration with Randall Smith's group (<http://kiwi.bcm.tmc.edu:8080/launcher.html>).

- Gene-Server questions (davison@uh.edu)
- Gopher (gopher.bchs.uh.edu or [gopher://gopher.bchs.uh.edu](http://gopher.bchs.uh.edu)): PIR sequence-retrieval pointers to other Bio-Gophers and some local information
- WWW (<http://www.bchs.uh.edu>): ftp site, software descriptions, and yellow pages of molecular biology software
- WAIS (www.bchs.uh.edu or gopher.bchs.uh.edu): indexed software descriptions
- Ftp (ftp.bchs.uh.edu): PIR releases and Macintosh, DOS, UNIX, and VMS molecular biology software
- E-mail (gene-server@bchs.uh.edu): molecular biology software and data



GDB Web Server Mirror Sites Available Worldwide

Mirror sites of the GDB Web server (<http://gdbwww.gdb.org>) are now available from many GDB nodes worldwide (see URLs in box at right). Users can access a GDB Web server via the network connection that provides the quickest response time for their location. Links to these sites are also available from the main GDB Web server in Baltimore on the GDB Home Page under "Links to Related Resources." New GDB Web sites will be included as they become available. GDB data accessed from all GDB Web servers are updated daily.

Genetic Maps on Web

A set of integrated genetic maps of human chromosomes has been generated from maps stored in the Genome Data Base using an experimental map construction and integration program called CPROP. These graphical maps are available in PostScript format on WWW (<http://gdbdoc.gdb.org/letovsky/cprop/human/maps.html>). Stanley Letovsky (letovsky@gdb.org) would like feedback on the accuracy and usefulness of these maps from knowledgeable persons in the chromosome-mapping community.◊

Council (from p. 12)

through institutional review boards [*HGN* 6(4), 15 (November 1994)].

Other Business

Noting that the ELSI Working Group has never had a formal mission statement outlining goals and delegation of authority, Collins submitted a mission statement to the council for review.

In response to concerns expressed at a previous council meeting, Bettie Graham reported that she had analyzed the NCHGR portfolio to determine the funding status of young investigators. After comparing NCHGR's R01, R26, and GESTEC statistics historically with those of NIH, she concluded that statistics for the two groups are very similar.

Smith recommended that the council discuss software modularity and interchangeability at its next meeting. He urged that investigators reduce duplication of effort by cooperatively establishing standards and following guidelines. Jay Snoddy (DOE) stated that DOE is addressing these issues through informatics meetings, and David Benton (NCHGR Office of Scientific Data Management) reported that the NCHGR GESTECs planned two joint informatics meetings.

A report from the DOE genome program will be a regular feature of future council meetings.

Application Review

The council reviewed 154 applications requesting \$40,841,162, including 103 regular research, 5 pilot-project, 19 ethics, 6 center, 4 conference, 10 small business innovation research, 1 research career development, and 3 continuing-education training grants. A total of 128

GDB Access Via WWW

The GDB Web server is available directly at the following URLs:

- United States <http://gdbwww.gdb.org/>
- Australia <http://morgan.angis.su.oz.au/gdb/docs/gdbhome.html>
- France <http://www.infobiogen.fr/gdbwww/>
- Germany <http://gdbwww.dkfz-heidelberg.de/>
- Israel <http://inherit1.weizmann.ac.il/gdb/docs/gdbhome.html>
- Netherlands <http://www-gdb.caos.kun.nl/gdb/docs/gdbhome.html>
- United Kingdom <http://www.hgmp.mrc.ac.uk/gdb/docs/gdbhome.html>

Other GDB nodes with WWW servers in English:

- Sweden <http://www.bmc.uu.se/Computing-Dept.html>
- Japan <http://www.jicst.go.jp/>

GDB User Support, Registration, Training

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. The Help Line in Baltimore is staffed from 9 a.m. to 5 p.m. EDT 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 and Genomic Data on the Internet" classes will be held in Baltimore on September 25-26 and December 4-5. Courses related to GDB and OMIM are also offered at some of the nodes. Contact the appropriate User Support office for details.

User Support Offices

UNITED STATES
GDB User Support
Genome Data Base
Baltimore, Maryland
410/955-9705
Fax: /614-0434
help@gdb.org

AUSTRALIA
Carolyn Bucholtz
ANGIS
Sydney, Australia
+ 61/2-692-2948
Fax: -3847
bucholtz@angis.su.oz.au

FRANCE
Phillippe Dessen
INFOBIOGEN
Villejuif, France
+33/1 4559-5241
Fax: -5250
gdb@infobiogen.fr

GERMANY
Molecular Biophysics Dept.
DKFZ
Heidelberg, Germany
+ 49/6221-42-2349
Fax: -2333
gdb@dkfz-heidelberg.de

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

JAPAN
Mika Hirakawa
JICST
Tokyo, Japan
+81/3-5214-8491
Fax: -8470
mika@gdb.gdbnet.ad.jp

NETHERLANDS
GDB User Support
CAOS/CAMM Center
Nijmegen, Netherlands
+ 31/80-653391
Fax: -652977
post@caos.caos.kun.nl

SWEDEN
GDB User Support
Biomedical Center
Uppsala, Sweden
+ 46/18-174057
Fax: -524869
help@gdb.embnet.se

UNITED KINGDOM
Administration
HGMP Resource Centre
Cambridge, U.K.
+ 44/1223-494511
Fax: -494512
admin@hgmp.mrc.ac.uk

applications requesting \$29,781,878 were recommended for approval.

Future Meetings

The council agreed to meet May 22-23 and September 11-12, 1995, and January 29-30, 1996.◊

Calendar of Genome Events*

September 1995

- 6-10.** Eukaryotic DNA Replication; [CSHL, 516/367-8346, Fax: -8845, meetings@csih.org, <http://www.cshl.org>]
- 7.** Marty Rodbell; Gaithersburg, MD [TIGR/NIST Distinguished Speakers Series, D. Hawkins, 301/869-9056, Fax: -9056, dhawkins@tigr.org, <http://www.tigr.org>]
- 7-8.** **2nd European Conf. on Gene Therapy of Cancer; London [Conf. Secretariat, +44-171/346-3126, Fax: /733-3877, mkuiper@hgmp.mrc.ac.uk]
- 11-12.** Intellectual Property & Patent Protection for Pharmaceuticals & Biologics; Washington, DC [IBC, 508/481-6400, Fax: -7911]
- 11-12.** NIH Natl. Advisory Council for Human Genome Res.; Washington, DC [J. Ades, 301/402-2205, Fax: -2218, ja51b@nih.gov]
- 13.** Biochemical Society Colloquium: Navigating the Genome; Dublin [K. Wolfe, +353-1/608-1253, Fax: /679-8558, khwolfe@otto.tcd.ie]
- 13-14.** European Biotechnol. Symp.; Amsterdam [BioConf. Intl., 301/652-3072, Fax: -4951]
- 13-17.** Biol. & Genetics of Complex Mammalian Traits; Bar Harbor, ME [S. Serreze, 207/288-23371 ext. 1378, Fax: -5079, complex95@aretha.jax.org]
- 14-15.** **4th DIMACS Intl. Algorithm Implementation Challenge: Two Problems in Computational Biology; Piscataway, NJ [M. Farach, 908/445-4580, Fax: -5932, special@dimacs.rutgers.edu, <http://dimacs.rutgers.edu>]
- 16-20.** 7th Intl. Genome Sequencing & Analysis Conf.; Hilton Head, SC (abs. deadline: May 1) [Conf. Office, 301/869-9056, Fax: -9423, seqconf@tigr.org]
- 17-20.** Intl. Conf. on Molecular Structural Biol.; ACS, Vienna [A. Kungl, +43-1/587-249, Fax: -966, msb95@helix.mdy.univie.ac.at]
- 17-20.** **1995 Molecular Biology Conf.; San Diego [R. Dana, 619/945-2321, Fax: /724-8171, rodana@greengene.com]
- 17-20.** 2nd Intl. Workshop on Human Y Chromosome; Pacific Grove, CA [C. Lau, 415/476-8839, Fax: /502-1613, clau@itsa.ucsf.edu or N. Affara, +44-223/333-700, Fax: -346, na@mbuc.bio.cam.ac.uk]
- 19-22.** Data Banks & Computer Support of HGP; Moscow [V. Tsitovich, +7-095/135-2311, Fax: -1405, imb@imb.msk.su]
- 22-24.** 2nd Single Chromosome Workshop on Human Chromosome 1; Vienna [A. Weith, +43-1/797-30-625, Fax: /798-7153, weith@aimp.una.ac.at]
- 28-29.** 2nd Annu. Self-Assembling Nanostructures for Gene Transfer; Wakefield, MA [CHI, 617/487-7989, Fax: -7937, chi@world.std.com, <http://www.healthtech.com/conferences>]
- 28-29.** Microfabrication Technology for Research and Diagnostics; CHI, San Francisco [see contact: Sept. 28-29, above]
- 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 or N. Moschonas, +30-81/210-079, Fax: /230-469, moschon@victor.imbb.forth.gr]

October 1995

- 5-8.** Conf. on Hist., Philos., & Ethical Implications of HGP; Notre Dame, IN [P. Sloan, 219/631-5015, Fax: -8209, phillip.r.sloan.1@nd.edu]
- 10-11.** 2nd Annu. Gene Therapy Technol.; CHI, Washington, DC [see contact: Sept. 28-29]
- 12.** Bruce Alberts; TIGR/NIST, Gaithersburg, MD [see contact: Sept. 7]
- 13-14.** **Promoting & Managing the Application of Genome Research; Concord, NH [T. Field, tfield@fplc.edu or C. Ruh, 603/228-1541, Fax: /224-3342, cruh@fplc.edu, <http://www.fplc.edu/tfield/genpro.html>]
- 13-14.** Workshop on Gene-Finding & Gene Struct. Prediction; Philadelphia (abs. deadline: May 1) [D. Searis, 215/573-3107, Fax: -3111, dsearis@cbil.humgen.upenn.edu]
- 22-24.** 3rd Intl. Karger Symp.: GeneTherapy; Basel, Switzerland (early reg.: Aug. 31) [S. Karger, +41-61/306-1111, Fax: -1434]
- 23-24.** 2nd Intl. Conf. on Strategic Business Info. in Biotechnol.; Res. Triangle Park, NC [IBI, 919/544-5111, Fax: -5401, nurban@biotechinfo.pdail.interpath.net, <http://www.bio.com>]
- 23-24.** Chromosome 3 Workshop; Minneapolis [S. Naylor, 210/567-3842, Fax: -6781, naylor@uhscsa.edu]
- 23-25.** BioWest '95; San Jose, CA [BioConf. Intl., 301/652-3072, Fax: -4951]
- 24-28.** ASHG 95; Minneapolis [FASEB, 301/530-7010, Fax: -7014]
- 24-28.** 3rd Intl. Conf. on Nanometer/Scale Sci. & Technol.; Denver [C. Marrian, 202/767-3150, Fax: -4998]
- 29-31.** 3rd Intl. Chromosome 13 Workshop; Tarrytown, NY [D. Warburton, 212/305-7143, Fax: -7436, cuh@cuccfa.ccc.columbia.edu]
- 29-Nov. 1.** NSGC 14th Annu. Educational Conf.; Minneapolis [B. Leopold, 610/872-7608 voice mailbox #6, Fax: -1192]
- 31-Nov. 1.** Managing the Privacy Revolution '95; Washington, DC [L. Vidal, 201/996-1154, Fax: -1883, CTRSRLR@aol.com]

November 1995

- 5-7.** 5th Intl. Workshop on Identification of Transcribed Sequences; Marseilles, France (abs. deadline: Aug. 15) [N. Matthews, 303/333-4515, Fax: -8423, nanm@druid.hsc.colorado.edu]
- 5-8.** 3rd Intl. Conf. on Automation in Mapping & DNA Sequencing; Berkeley, CA (poster deadline: Aug. 18) [M. Field, 510/486-6386, Fax: -5548, MOField@LBL.gov]
- 9-11.** 4th Conf. on Molec. Nanotechnology; Palo Alto, CA [Foresight Institutes, 415/324-2490, Fax: -2497, foresight@cup.portal.com, <http://ftp.parc.xerox.com/pub/nano/nano4.html>]
- 12-16.** 9th Intl. Mouse Genome Conf.; Ann Arbor, MI (abs. deadline: Aug. 1) [D. Miller, 716/845-4390, Fax: -8169, dmiller@mcbio.med.buffalo.edu]
- 13-14.** 4th Intl. Workshop on Human Chromosome 16; Leiden, Netherlands [M. Breuning, +31-71/276-048 or -293, Fax: -048, breuning@rullf2.LeidenUniv.nl]
- 16-17.** 3rd Intl. Workshop on Human Chromosome 12 Mapping; Leuven, Belgium

[P. Marynen, +32-16/34-5891, Fax: -5997, peter.marynen@med.kuleuven.ac.be]

18. Chromosome 12 Genes In Human Cancer; Leuven, Belgium [see contact: Nov. 16-17]

December 1995

- 2-6.** Molecular Basis of Gene Transcription; San Diego [AACR, J. Ruben, 215/440-9300, Fax: -9313]
- 3-6.** **HUGO Comparative Genome Organisation Workshop; Queensland, Australia [HUGO Americas, 301/654-1477, Fax: /652-3368, hugo@gdb.org]
- 3-7.** 3rd UNESCO Human Genome Conf.; New Delhi [S. Matsui, +33-1/45-68-3887, Fax: -67-2639]
- 8-9.** DNA Databanks & Repositories; Birmingham, AL [AFIP/ARP, 800/577-3749, Fax: 301/427-5001, came@email.afip.osd.mil]
- 10-12.** Genome Informatics 1995; Yokohama, JP (abs.: Oct. 13) [N. Tomioka, Fax: +81-3/5449-5434, workshop@ims.u-tokyo.ac.jp]

January 1996

- 3-6.** Pacific Symp. on Biocomputing; Kohala Coast, Hawaii (early reg.: Oct. 2) [L. Hunter, 301/496-9300, Fax: -0673, hunter@nlm.nih.gov or T. Klein, 415/476-0663, Fax: /502-1755, klein@cgl.ucsf.edu, <http://cgl.ucsf.edu/psb/psb.html>]
- 14-18.** Plant Genome IV; San Diego (PG I, II, and III abs.: <http://probe.nalusda.gov:8000>) [Scherago Intl., 212/643-1750, Fax: -1758, scherago@biotechnet.com]
- 15-18.** BioEast '96; Washington, DC [BioConf. Intl., 301/652-3072, Fax: -4951]
- 28-Feb. 1.** 5th DOE HGP Contractor-Grantee Workshop; Santa Fe, NM [S. Spengler, 510/486-4879, Fax: -5717, sylviaj@ux5.lbl.gov]
- 29-30.** NIH Natl. Advisory Council for Human Genome Res.; Washington, DC [see contact: Sept. 11-12]

February 1996

- 5-6.** Intellectual Property Issues: Critical Challenges for Biomedicine and Genomics; CHI, Santa Fe, NM [see contact: Sept. 28-29] ♠

Training Calendar*

September 1995

- 15.** 2nd Annu. NCHGR Sci. Writers Workshop; Bethesda, MD [P. Gregory, 301/496-3978, Fax: /402-4929, edcore@nchgr.nih.gov]
- 25-26.** GDB/OMIM and Genomic Data on the Internet; Baltimore [GDB/OMIM, 410/955-9705, Fax: /614-0434, help@gdb.org, <http://gdbwww.gdb.org>]

October 1995

- 2-6.** Advanced Linkage Course; Zurich, Switzerland [K. Montague, 212/960-2507, Fax: /568-2750, jurg.ott@columbia.edu]
- 8-17.** Frontiers of Protein Structure Prediction; Rome (reg. deadline: July 1) [T. Gobbl, +39-6/910-93-201, Fax: -654]

November 1995

- 15-25.** DNA Sequencing: Adv. Approaches, Automated Methods and Analysis; EMBL, Heidelberg, Germany [W. Ansorge, +49-6221/387-355, Fax: -306, ansorge@embl-heidelberg.de] ♠

*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.

Ongoing Training Courses

Courses are being held in the following selected subject areas later this year. Check with contact person for specific course titles, places, and times.

ACS (P. Orton, 202/872-4508, Fax: -6336). Molecular modeling and computational chemistry, molecular biology and recombinant DNA technology, analytical methods for proteins.

AFIP/ARP (J. Centeno, 202/782-2839, Fax: -9215, CENTENO@email.afip.osd.mil). Analytical and molecular biology techniques in environmental toxicology and forensic science.

ATCC (M. Miller, 202/319-6161, Fax: -4467, millerm@cua.edu, <http://www.atcc.org/workshops/workshop.html>). Recombinant DNA technology and sequencing, PCR techniques, molecular approaches to understanding and diagnosis of genetic and infectious diseases.

BTP (S. Chance, 800/821-4861, Fax: 603/267-1993, biotraining@delphi.com). PCR and clinical applications, basic cloning and hybridization, quantitative RNA-PCR.

Carolina Workshop (W. Litaker, 919/962-8920, Fax: /966-6821). Gene targeting in ES cells and transgenic mice.

CATCMB/CUA (M. Miller, 202/319-6161, Fax: -4467, millerm@cua.edu). Recombinant DNA methodology.

CSHL (516/367-8346, Fax: -8845, meetings@cshl.org or <http://www.cshl.org>). *Arabidopsis* molecular genetics, molecular cloning of neural genes, expression of eukaryotic genes, yeast genetics, bacterial genetics.

Exon-intron (800/407-6546, Fax: 410/730-3983). PCR understanding and methodologies, RNA isolation and gene expression, rDNA, tissue culture and baculovirus, in situ hybridization, chemiluminescence principles.

GRC (401/783-4011, Fax: -7644, grc@grcmail.grc.uri.edu). Human molecular genetics, mechanisms of toxicity, quantitative structure-activity relationships.

LTI (L. Kerwin, 800/952-9166, Fax: 301/258-8212). In situ hybridization; gene expression systems; analysis of gene expression; cDNA library, PCR, recombinant baculovirus, recombinant DNA, and cell culture techniques.

MBL (508/548-3705 ext. 401, Fax: /457-1924, admissions@mbi.edu or <http://www.mbl.edu>). Cellular and molecular biology, molecular evolution.

PSC (N. Blankenstein, 412/268-4960, Fax: -8200, biomed@psc.edu). Methods of molecular mechanics and dynamics of biopolymers.

Oncor (800/556-6267, Fax: 301/926-6129). Introduction to molecular cytogenetics.

UCLA (310/825-3344, Fax: /206-2815). Computational chemistry for materials and drug design.

UMBC (C. Harriger, 410/455-2336, Fax: -1074, Carolyn_Harriger@umbcadmn.bitnet). Recombinant DNA.

UMDS (P. Falk, +44-171/403-6998, Fax: /407-5281, wss@umds.ac.uk). Genetic analysis from YAC to gene, analysis of multifactorial diseases.

UWS (M. Bamard, 206/616-1864, Fax: /685-7515, mbarnard@u.washington.edu). Genomic information and its ethical implications.◊

Extended calendars are available at http://www.ornl.gov/TechResources/Human_Genome/home.html or from HGMIS (see p. 6 for contact information).

For Your Information

U.S. Genome Research Funding Guidelines

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

NIH National Center for Human Genome Research (NCHGR)

Program announcements listed in *NIH Guide for Grants and Contracts* (gopher.nih.gov and <http://www.nih.gov> or 301/496-0844). Bracketed numbers below refer to application due dates.

- [1] February 1, June 1, and October 1;
- [2] April 5, August 5, and December 5;
- [3] May 10; [4] on a continuous basis; and
- [5] May 1 and November 15.

Program Categories 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 for predoctoral and postdoctoral trainees (PA 94-085) [3].
 - Individual predoctoral student fellowships for disabled (PA 95-028) [5] and minorities (PA 95-029) [5].
- Special emphasis research career awards in genomic research (PA 91-89) [1].

Special Programs

- Minority institution travel awards (PA 91-17) [4].
- Research supplements for underrepresented minorities and disabled [4].

NCHGR: 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 or Jane_Peterson@nih.gov
- Physical mapping technology, training, and special programs: Betty_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/oerhp/hug_top.html

Alexander Hollaender Distinguished Postdoctoral Fellowships

Research opportunities in energy-related life, biomedical, and environmental sciences, including human genome, global change, and supporting disciplines. Next deadline: January 15, 1996

- Contact: Barbara Dorsey, Oak Ridge Institute for Science and Education (615/576-9975, 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.
- Betty Graham (see contact, NCHGR). NIH SBIR due April 15, August 15, and December 15. STTR, December 1.

National SBIR/STTR conferences: 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.◊

¶ **Mouse Genetics Book:** *Mouse Genetics: Concepts and Applications* by Lee M. Silver (Princeton University) is designed for both advanced students and practicing scientists interested in understanding and using the mouse as a model system for genomic analysis. The first half of the book provides a comprehensive introduction to the mouse, including a history of the field; mouse phylogenetic relationships; standard crosses, nomenclature, and specialized strains; reproduction and breeding; genome organization and evolution; and an overview of mutagenesis and transgenesis. The second half is aimed at the practicing geneticist, with detailed approaches to breeding experiments and interpreting mapping data. A discussion of statistical analysis as it applies to well-defined experimental crosses is provided, along with graphs and tables for interpreting linkage data derived from recombinant inbred strain and backcross studies. Although focused on the mouse as a model system, much of the material applies equally to mapping studies in other experimental mammals. 1995, 376 pages. [Oxford University Press (800/451-7556)] ◊

