Human Genome news

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

ISSN: 1050-6101

Vol. 7, Nos. 3 & 4, September–December 1995

Genome Project Finishes Fifth Year Ahead of Schedule

Project History, Progress, Challenges, and Impact Highlighted in This Issue

October 1 marked the fifth anniversary of the Human Genome Project in the United States. This issue of *HGN* provides an overview of progress toward the initial short-term goals set out by NIH and DOE in October 1990 and the long-term goal of developing scientific resources and technologies for the DNA-based biology of the 21st century.

Impressive progress has been made in all targeted areas: genetic and physical mapping of the human genome; DNA sequencing; analysis of the genomes of several important nonhuman organisms; informatics; resource and technology development; and ethical, legal, and social implications (ELSI) of genetic research for individuals and society. In 1993, 3 years into the project, advances in technology allowed the original 5-year goals to be updated with more-ambitious approaches. However, a broad range of technological advances is still needed to achieve extended goals by 1998 and also to reach the ultimate objectives—the complete sequence of the 3 billion human DNA base pairs and identification of all human genes—by 2005.

Advances in genome research have come about through the efforts of a large number of investigators in the United States and abroad. The U.S. Human Genome Project is supported by NIH and DOE at 22 specialized human genome research centers (see genome centers box, p. 9) and in many university, national, and private-sector laboratories. At least 14 other countries now have programs for analyzing the genomes of a variety of organisms ranging from microbes to economically important plants and animals to humans (see partial list below). A remarkable spirit of cooperation has facilitated human genome research around the world.

Human Genome Research Programs Established Worldwide

Brazil	Italy
Canada	Japan
China	Mexico
Denmark	Netherlands
European Union	Russia
France	Sweden
Germany	United Kingdom
Israel	United States

One purpose of the Human Genome Project is to provide new "infrastructure," including data and material resources and technology that will improve the ability of investigators to do biological research rapidly, efficiently, and cost-effectively. To achieve this purpose, the accomplishments of genome researchers must be easily accessible. Guidelines describing reasonable sharing practices were developed by advisors to NIH and DOE in 1992, and the genome research community has responded admirably. Information sharing and services have been facilitated by the physics community's gift of WWW protocols, which have greatly simplified data presentation and transfer across the Internet. Additionally, materials and reagents produced by genome research investigators have been made widely available, often through commercial sources.

NCHGR URL
http://www.nchgr.nih.gov

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

Acronyms, p. 19.

As the scientific community hoped, the genome infrastructure has rapidly accelerated the study of inherited human disease. A wide range of other genome-research applications also indicates that the promise of the genome project is being realized.

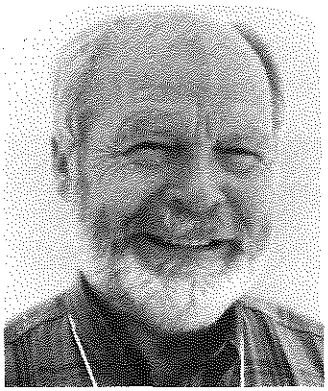
David Smith (DOE) and Francis Collins (NIH), Directors of the U.S. Human Genome Project, offer their perspectives on goals, benefits, and challenges beginning on the next page. Research highlights begin on p. 4.◊

In This Issue

Evolution of a Vision: Genome Project Origins, Presen

For David Smith, the impressive early achievements and spin-off benefits of the Human Genome Project offer more than mere vindication for project founders. They also provide a tantalizing glimpse into the future where, he observes, "scientists will be empowered to study biology and make connections in ways undreamt of before."

David Smith is a founder and current Director of the DOE Human Genome Program. In September, at the seventh annual Genome Sequencing and Analysis meeting in Hilton Head, South Carolina, Smith reflected on principles guiding the establishment and management of the genome project and proffered some insights into where it all may lead. Along the way, he answered one of the questions asked most frequently of DOE genome program staffers: Why is DOE involved in the genome project?



David Smith, Director DOE Human Genome Program

Smith became intrigued by DNA studies in the mid-1950s at a small college in Nebraska where, as he explains it, "An embryology teacher opened my eyes to the significance of DNA, and from there, my main scientific motivation was almost philosophical."

After doctoral studies in biochemistry at the University of Southern California, Smith began his DOE career with an appointment at Los Alamos National Laboratory. In 1977 he began managing the molecular biology programs at the Energy Research and Development Administration, a forerunner of DOE, in Germantown, Maryland. Here Smith's ideas about the importance of understanding DNA eventually took shape and became a rationale guiding DOE toward implementation of what has become the largest project to be carried out in biology.

Detecting Heritable Mutations

Smith views establishment of the DOE Human Genome Program as a natural outgrowth of that agency's long-term mission to develop better technologies for measuring health effects, particularly induced mutations. As he explains it, "DOE had been supporting mutation studies in Japan, where no heritable mutations could be detected in the offspring of populations exposed to the atomic blasts at Hiroshima and Nagasaki. The program really grew out of a need to characterize DNA differences between parents and children more efficiently. DOE led the development of many mutation tests, and we were interested in developing even more-sensitive detection methods. Mortimer Mendelsohn of Lawrence Livermore National Laboratory, a member of the International Commission for Protection Against Environmental Mutagens and Carcinogens, and I decided to hold a workshop to discuss DNA-based methods. (See Human Genome Project chronology, beginning on p. 4.)

"Ray White (University of Utah) organized the meeting, which took place in Alta, Utah, in December 1984. It was a small meeting but very stimulating intellectually. We concluded the obvious—that if you really wanted to use DNA-based technologies, you had to come up with more-efficient ways to characterize the DNA of much larger regions. And the ultimate sensitivity would be if you could compare the complete DNA sequences of parents and their offspring."

Project Gets Started

Smith recalls reaction to the first public statement that DOE was starting a program with the aim of sequencing the human genome. "I announced it at the Cold Spring Harbor meeting in May 1986, and there was a big hullabaloo." After a year-long review, a National Academy of Sciences–National Research Council panel endorsed the project and the basic strategy proposed.

Smith points out that NIH and others were also having discussions on the feasibility of sequencing the human genome. "Once NIH got interested, many more people became involved. DOE and NIH signed a Memorandum of Understanding in October 1988 to coordinate our activities aimed at characterizing the human genome." But, he observes, it wasn't all smooth sailing. The nascent project had many naysayers.

Responding to Critics

Many scientists, prominent biologists among them, thought having the sequence would be a misuse of scarce resources. Smith, laughing now, recalls one scientist complaining, "Even if I had the sequence, I wouldn't know what to do with it."

Other critics worried that the genome project would siphon shrinking research funds away from individual investigator-initiated research projects. Smith takes the opposite view. "In fact, individual investigators can do things they would never be able to do otherwise. We're beginning to see that demonstrated at this Hilton Head meeting. For the first time, we're finding people exploring systematic ways of looking at gene function in organisms. The genome project opens up enormous new research fields to be mined. Cottage-industry biologists won't need a lot of robots, but they will have to be computer literate to put it all together."

The genome project is also providing enabling technologies essential to the future of the emerging biotechnology industry, catalyzing its tremendous growth. According to Smith, the technologies are capable of more than elucidating the human genome. "We're developing an infrastructure for future research. These technologies will allow us to efficiently characterize any of the organisms out there that pertain to various DOE missions, with applications such as better fuels from biomass, bioremediation, and waste control. They will also lead to a greater understanding of global cycles, such as the carbon cycle, and the identification of potential biological interventions. Look at the ocean; an

(see Smith, p. 17)

and Future Challenges, and Far-Reaching Benefits

FY 1996 marks the end of the first 5 years and the beginning of a new era in the Human Genome Project. It is very rewarding to note that we have met or exceeded most of our ambitious goals—some ahead of time and all under budget. The genetic map is complete, the chromosome physical maps are within 18 months of completion, and pilot programs to begin sequencing the entire human genome are under consideration.

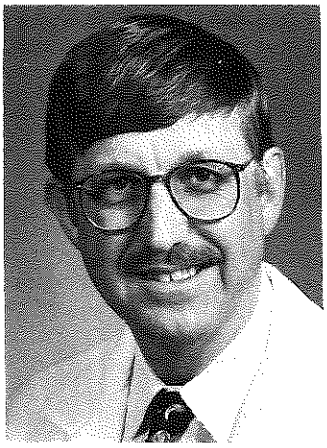
Fortunately, we do not have to wait until the end of the project to reap its benefits; information generated by the Human Genome Project is quickly disseminated through public electronic databases and used by researchers across the United States and the rest of the world. Already this information is changing the way biomedical research and the practice of medicine are being conducted. The benefits are widely apparent, as significant discoveries of the genetic basis of a vast number of genetic diseases—from polycystic kidney disease and Alzheimer's to breast cancer, colon cancer, and diabetes—are happening at a staggering pace.

New Maps Aid Gene Hunters

Detailed maps coming out of the Human Genome Project are critical to understanding the basis of many common diseases—complex disorders resulting from the effects of multiple genes and environmental influences. Not long ago, for example, a researcher used genetic maps to discover at least five different chromosome regions that appear to play a role in insulin-dependent (type 1) diabetes. This kind of genome scan would not have been possible without high-resolution genome maps. The maps will be invaluable in teasing apart the contributions of genes to other complex disorders, including heart disease, asthma, cancers, and psychiatric disorders. In conjunction with the detailed maps, genome technologies also played a key role in the isolation of 13 disease genes in FY 1995, including the *BRCA*1 gene for hereditary breast cancer.

Francis Collins was appointed Director of the NIH National Center for Human Genome Research in 1993, following the resignation of James Watson. Collins was formerly at the University of Michigan, where he was a Howard Hughes Medical Institute investigator, professor, and Director of the NCHGR-supported human genome center. In this article, Collins shares his thoughts on the Human Genome Project and its growing impact on the practice of medicine.

He received the B.S. degree from the University of Virginia, M.S. and Ph.D. degrees in physical chemistry from Yale University, and M.D. from the University of North Carolina School of Medicine. Collins pioneered the development of a powerful gene-finding method known as "positional cloning," in which investigators localize a disease gene to a chromosomal subregion by studying the inheritance pattern of the disease within families. His group and others used this technique to isolate the genes for cystic fibrosis, neurofibromatosis type 1, Huntington's disease, and ataxia telangiectasia.



Francis S. Collins, Director NIH NCHGR

Meeting Sequencing Challenges

Having met or exceeded our original goals for genetic mapping, we now turn our vision for the next phase of the Human Genome Project to the most challenging technological undertaking of all: determining the sequence of DNA bases in the entire 3-billion base-pair length of human DNA. Knowing the order of DNA bases tells an investigator where genes are located, as well as what instructions are carried in a piece of DNA. This information is critical to understanding the function of genes and how they cause disease. Technology development for this work primarily has been carried out experimentally on the DNA of important model organisms. Researchers sequencing the genome of the roundworm have amassed almost 28 million bases of DNA from that organism—over one-quarter of the animal's genome—and have increased their annual production rate from 2 million to 14 million DNA bases. These investigators expect to complete the sequence of the roundworm genome by the end of 1998.

Thus far, technology development in sequencing DNA has aimed at reducing cost and increasing rate. This past year, NCHGR began a new initiative to reduce the scale of sequencing instrumentation and increase sequencing speed. We will also explore new strategies for minimizing time-consuming bottlenecks by developing integrated, matched components throughout the sequencing process. The urgency of pushing these advances now is considerable because reducing the cost by only $.01 per base pair will save $30 million in the production phase of human DNA sequencing.

NCHGR recently solicited proposals for pilot projects to test strategies for full-scale production sequencing of mammalian DNA. Applicants must expect to sequence at least one million bases of contiguous DNA over the 3-year period; they must also explain how their approach will help improve sequencing technology and capability to achieve the complete, accurate, finished human DNA sequence by 2005. This proposal request is justified by the technological progress and project-management experience already achieved and by the need to gain more experience in large-scale human DNA sequencing, which scientists anticipate will present different issues from those encountered in sequencing nonhuman DNA.

Impact of Genome Research

Our knowledge about gene function will continue to grow as researchers analyze the genetic causes of disease at the molecular level and associate specific gene alterations with an individual's risk for disease. An immediate

(see Collins, p. 16)

Five Years of Progress in the Human Genome Project

This article describes Human Genome Project accomplishments and progress toward short and long-term goals. Topics include genetic and physical mapping of the human genome; DNA sequencing; gene identification; analysis of model organism genomes; informatics; and explorations of ethical, legal, and social implications (ELSI) arising from genome research.

GENETIC MAPPING

Genetic linkage maps are critical for mapping genes underlying identifiable phenotypes including diseases. In late 1994 the first major initial genome project goal, a 2- to 5-cM human genetic map, was reached when an international group of investigators published a comprehensive map comprising 5840 loci covering 4000 cM. Of those markers, 970 are ordered with high confidence (odds of >1000:1) and provide a framework map. The comprehensive map can be said to represent an average marker density of 0.7 cM, with the more highly reliable framework map subset having a resolution of about 4 cM.

Progress toward this goal was extremely rapid; the map was, in fact, completed a year ahead of schedule. This accomplishment resulted, in part, from the use of a new type of genetic marker known variously as a microsatellite repeat, STRP, or SSLP. Advantages of microsatellites include a high level of variation from individual to individual (polymorphism), an abundant and relatively even distribution throughout the genome, and the ability to be assayed by PCR.

Although the initial genetic-mapping goal has been attained, the 1993 extended 5-year plan recognized the importance of continued improvement in genetic-mapping technology. Easier, automatable, and more cost-effective genotyping methods remain a priority. Such methods probably will require the development of new types of genetic markers, novel genotyping technology, and new analytical tools. Maximizing the usefulness of the genetic map will be particularly important for dissecting the genetics of such complex traits as susceptibilities to heart disease, hypertension, and diabetes.

PHYSICAL MAPPING

Physical maps are used to isolate and characterize individual genes and other DNA regions of interest and provide the substrate for DNA sequencing. As stated in the 1993 extended 5-year plan [*Science* **262**, 43–46 (1993)], a current Human Genome Project goal for physical mapping is to complete an STS-based map of the human genome with markers spaced every 100 kb on average. Investigators are generating STS maps using both chromosome-specific and genome-wide strategies, and collective progress has been impressive.

Constructing a 100-kb resolution map will require generating and ordering some 30,000 STS markers. A number of different strategies are being applied on a genome-wide basis to build such a map; these strategies include STS-content mapping using large-insert YAC clones, radiation hybrid mapping, and clone fingerprinting. Adoption of a whole-genome approach for map building has been important in the rapid progress of the past 2 to 3 years. Investigators also plan to map a common subset of STS markers on the different maps currently under construction, resulting in well-integrated maps with many more mapped STSs than any one laboratory could produce.

For example, efforts already under way will produce a radiation hybrid map in which a sufficient number of markers will be ordered at very high confidence (1000:1 odds) to provide a resolution higher than 200 kb. Additional STSs will be mapped, albeit with order established at lower confidence levels, with overall map resolution higher than 100 kb. When the map is completed, investigators will be able to select markers from any of the contributing maps, confident that the markers will fall in either the same or adjacent defined regions (or bins) on the chromosome.

Investigators are also placing polymorphic markers within physical maps to allow integration of physical and genetic mapping data across chromosomal regions. These maps will facilitate finer-scale mapping, sequencing, and disease-gene identification. Large-scale efforts to map YACs and BACs onto metaphase chromosomes are linking cytogenetic and sequence/cosmid-based maps.

Major Events in the Human Genome Project (Acronym List, p. 19)

1984

DOE OHER and ICPEMC cosponsor Alta, Utah, conference highlighting the growing role of recombinant DNA technologies. OTA incorporates Alta proceedings into report acknowledging value of human genome reference sequence.

1985

Robert Sinsheimer holds meeting on human genome sequencing at University of California, Santa Cruz.

Charles DeLisi and David Smith develop plans for DOE Human Genome Initiative.

1986

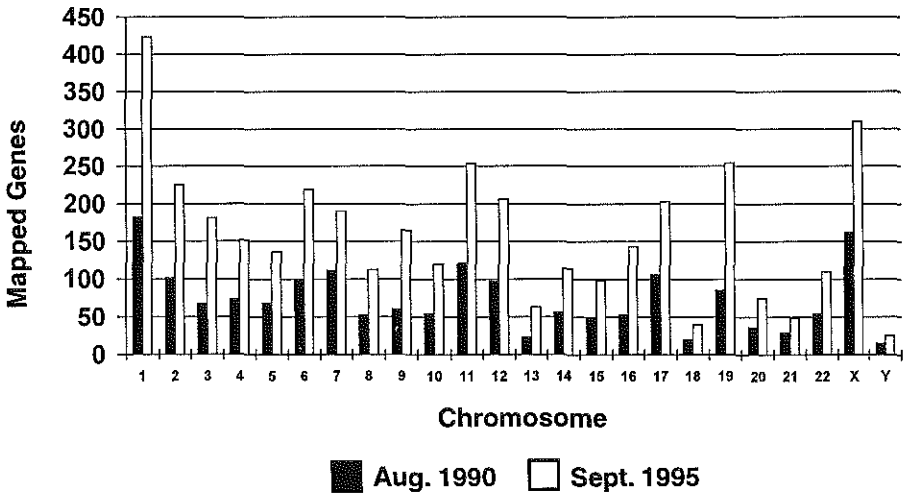
DOE OHER announces Human Genome Initiative after meeting in Santa Fe, New Mexico, to explore its feasibility. Pilot mapping and informatics projects are pursued at DOE national laboratories to develop critical resources and technologies for genomic analyses.

1987

Funding for NIH genome research begins, with funds administered by NIGMS for 2 years.

DOE HERAC publishes recommendations for 15-year, multidisciplinary, scientific and technological undertaking to map and sequence the human genome. DOE designates specialized human genome centers. DOE management publishes first

Initial physical-mapping goals included construction of contig maps (overlapping clone sets) of human chromosomes. Long-range clone contiguity has been achieved for several individual chromosomes in a number of laboratories. Clone-STS maps of entire euchromatic regions of chromosomes 21 and Y were published in 1992. YAC-based clone-STS maps of chromosomes 3, 11, 12, and 22 were finished more recently, and similar maps of chromosomes 4, 5, 7, and X are nearing completion. Maps principally based on cosmid contigs were published recently for chromosomes 16 and 19, and a cosmid-based chromosome 13 map is almost finished.



Growth of Mapped Genes, 1990–1995. The number of mapped genes has risen sharply over the past 5 years, from 1772 genes at the inception of the Human Genome Project in 1990 to 3695 genes by September 15, 1995. Gene distribution depicted here may reflect mapping activity per chromosome rather than relative gene density, which will remain unknown until the majority of genes are mapped. Numbers in this graph do not include genes not yet assigned to chromosomes. (Source: GDB, 1995)

None of the first-generation physical maps is error free. Errors come from at least two sources: rearrangement of clones relative to the native genome and map-assembly procedures that do not always produce the correct order. Some problems with the initial physical maps will resolve themselves. As marker density increases, internal inconsistencies will become evident and will be corrected upon data reexamination. The use of multiple, independent mapping methods also will contribute significantly to map validation, and using the same markers in different mapping projects will promote quality control. Criteria for assessing and reporting map quality and mapping progress were proposed recently by an international group of scientists.

In spite of these impressive advances, further improvements in mapping technology are essential. New host-vector systems may be required, for example, to capture regions not represented well in current maps and for particular map applications such as sequencing.

DNA SEQUENCING

The most technologically challenging goal of the Human Genome Project remains the complete sequencing of the human genome within the projected 15 years. In the past 5 years, significant progress has been made toward developing the capability for large-scale DNA sequencing. When the genome project began, the longest DNA sequence obtained was the 250,000-bp cytomegalovirus sequence, which took several years to complete. Now, several laboratories have each generated at least 1 Mb; some have determined more than 10 Mb of DNA sequence, mostly from model organisms. The longest contiguous human sequence is 685 kb from the human T-cell receptor β locus, a chromosomal region involved in immune responses.

Substantial technical, strategic, and organizational experience in managing large data-production projects has been gained through recent efforts to sequence the genomes of several nonhuman organisms (see Model Organisms section). The capacity of automated sequencing instruments has increased, and newer, higher-throughput instruments are almost ready for introduction into a large-

program plans with Program Director Benjamin Barnhart.

1988

OTA and NRC's Committee on Mapping and Sequencing the Human Genome publish highly influential reports recommending concerted genome research program and laying out scientific strategy. NRC recommends budget reaching $200 million a year.

NIH Director James Wyngaarden assembles scientists, administrators, science-policy experts to develop NIH plan for HGP.

HUGO founded by scientists to coordinate international collaboration.

First annual Cold Spring Harbor Laboratory meeting on human genome mapping and sequencing.

NIH creates Office for Human Genome Research within the Office of the Director. James D. Watson named Associate Director for Human Genome Research.

NIH establishes PACHG to advise on all aspects of genomic-analysis research.

DOE and NIH sign MOU outlining plans for cooperation on genome research.

DOE forms HGCC to provide external expertise and facilitate development and implementation of genome technology and informatics.

1989

NIH and DOE hold first planning retreat to develop joint 5-year goals.

DHHS Secretary Louis Sullivan establishes NCHGR at NIH. James Watson named first Director.

STSs proposed as common mapping language.

First genome sequence and analysis conference held at Wolf Trap, Virginia.

Five-Years of Progress (continued)

scale sequencing environment. As a result of these and other developments, confidence is growing that continued incremental improvements to current DNA sequencing approaches can be scaled up cost-effectively and probably will enable completion of the first-generation human DNA sequence by 2005.

Continued improvement in sequencing technology will be essential to meet the demands of sequence-based approaches to biological analysis. Achieving the capability for inexpensive sequencing at high-throughput levels will require technology far beyond that available today.

GENE IDENTIFICATION

One of the long-range genome project objectives is to identify all genes and other functional elements in genomic DNA, although understanding their functions will extend far beyond the project. With steady improvements in physical-map resolution and increased sequence data, an attendant need is for robust, high-throughput, and cost-effective methods to identify, map, and study functional elements in the genomes of humans and other organisms.

One method for tabulating genes on a genome-wide basis involves the determination and mapping of unique tags (ESTs) for cDNAs. Identification and initial analysis of large sets of ESTs have been published, and over the next year an even larger number of ESTs are expected to become available. The cDNA clones from which ESTs are derived are also available through the IMAGE Consortium [*HGN* **6**(6), 3 (March–April 1995)], repositories, and industry. Another international consortium is mapping a large number of publicly available ESTs on both radiation hybrids and YACs. By providing information on the chromosomal locations of genes represented by ESTs, this gene map will increase the value of the EST set for investigators engaged in gene hunting and other analytical activities.

A significant fraction of all human genes is expected to be represented ultimately in the EST and clone sets, but this approach is unlikely to reveal all human genes. Additionally, the amount of sequence and structural information about a gene identified by an EST will be limited. An optimal technology or combination of technologies that will allow high-throughput, cost-effective gene identification remains an important goal of the Human Genome Project.

Disease-Gene Identification

The speed with which human genes are being identified, particularly those responsible for genetic diseases, continues to increase rapidly because of improved genetic and physical maps. As a result, new disease genes are being discovered at a rate of several per month, compared with a few per year not so long ago.

For the past several years, improved maps have increased the efficiency with which investigators use the powerful positional-cloning approach to isolate human disease genes. Positional cloning is essential for identifying genes underlying a particular condition or trait when no prior knowledge of gene function is available.

As genome maps have improved and become increasingly enriched with gene sequences, a new strategy known as positional-candidate cloning has emerged. This approach begins with mapping the disease gene to a small interval on a chromosome. All genes previously identified for that genomic region can then be tested, starting with any whose product suggests possible involvement. Now a gene can become a candidate for disease involvement by virtue of its properties and its map location.

Microbial Genomes Sequenced

This year, sequencing projects on the genomes of the bacteria *Haemophilus influenzae* (1.8 Mb) and *Mycoplasma genitalium* (0.58 Mb) were completed in record time, the latter with funding from the DOE Microbial Genome Initiative (MGI). Because these are the first free-living organisms whose genomes have been completely sequenced (with *M. genitalium* having the smallest genome of any independent organism), these data also provide scientists for the first time with a model of all the genetic information needed for independent existence. Another MGI sequencing project soon to be completed focuses on *Methanococcus jannaschi*, a bacterium that thrives in extremely hot environments.

DOE holds first contractor-grantee workshop in Santa Fe, New Mexico.

NIH and DOE establish Joint ELSI Working Group.

1990

David Galas named OHER Associate Director responsible for DOE Human Genome Program.

DOE and NIH release joint 5-year plan with specific goals for U.S. HGP.

Secretary Sullivan establishes NACHGR at NIH.

U.S. Human Genome Project formally begins in October.

1991

DOE and NIH begin GDB support.

1992

Watson resigns as NCHGR Director; Michael M. Gottesman (from NIH NCI) named NCHGR Acting Director.

First PCR/STS-based genetic linkage map of entire human genome published.

First physical maps presented, for human chromosomes Y and 21.

DOE and NIH release data- and resource-sharing guidelines.

1993

International IMAGE cDNA Consortium initiated to construct and characterize arrayed cDNA libraries, integrate data, and make clones and data publicly available.

Francis S. Collins appointed NCHGR Director.

Joint ELSI Working Group's Task Force on Genetic and Insurance Information releases recommendations.

U.S. HGP revises 5-year goals through September 1998.

Aristides Patrinos named DOE OHER Acting Associate Director when Galas resigns; David Smith heads DOE Human Genome Program.

MODEL ORGANISMS

Initial Human Genome Project goals included the characterization of the genomes of such important research organisms as the bacterium *Escherichia coli*, yeast *Saccharomyces cerevisiae*, roundworm *Caenorhabditis elegans*, fruit fly *Drosophila melanogaster*, and laboratory mouse. These well-studied organisms, which serve as useful, more cost-effective testing grounds for developing large-scale DNA sequencing technology, provide another approach to interpreting human genomic information.

Sequencing

- Progress toward completing the sequence of *S. cerevisiae* has been remarkable, and determination of its 15-Mb genome is expected to be completed within 6 to 9 months, with significant contributions coming both from European and U.S. laboratories.
- Almost 28 Mb of the 100-Mb *C. elegans* genome has also been determined, an accomplishment representing the largest amount of DNA sequence available from any single organism. Investigators involved in this coordinated, international effort are moving quickly toward the 1998 goal for completion.
- Using a different, directed-sequencing strategy that is gaining in popularity, researchers sequencing the 120-Mb euchromatic portion of the *D. melanogaster* genome have now completed over 2.5 Mb.
- Over 2 Mb of sequence has been determined for the *E. coli* genome, with the complete 4-Mb sequence expected within the next 2 years.

Mouse Maps and Human-Mouse Sequence Comparisons

The mouse genome is about the same size as the human genome. Many genes are conserved between the two species, as is gene order along some chromosomes. Mouse genome maps are thus extremely valuable tools for finding human genes and understanding their functions. This year, investigators completed a genetic map of the mouse genome containing over 6500 microsatellite markers among a total of 7300 genetic markers. Work has begun on a physical map of the mouse genome.

Other investigators are sequencing homologous regions in mouse and human genomes. One example is the region containing the T-cell receptor (TCR) genes that specify cell-surface receptors and play an important role in immune responses. Comparative analysis of this stretch of contiguous sequence from the two species has revealed important and interesting genomic features. These studies are expected to lead to insights into the biological function of TCRs that, in turn, may lead to new ways to counteract transplant rejection, infectious and autoimmune diseases, and allergies.

Other Projects

Human Genome Project sequencing successes have facilitated genome analysis of other interesting and important organisms in the United States and abroad. Examples include the DOE Microbial Genome Initiative for studying organisms of environmental or industrial importance (see box, p. 6); a privately funded effort that has generated the first complete sequence of the free-living organism *Haemophilus influenzae*; a project jointly supported by the National Science Foundation, the U.S. Department of Agriculture (USDA), and DOE to map and sequence the genome of the plant *Arabidopsis thaliana*; and projects focused on mapping the genomes of plants and animals of agricultural importance, organized by USDA and by agencies in other countries.

INFORMATICS

From the beginning of the genome project, informatics has been recognized as essential to the project's success. Much progress has been

Training Programs

Scientists with interdisciplinary skills are needed to achieve project goals and apply tools created for the genome project to a wide variety of scientific problems, including the biology underlying genetic diseases, the development of therapeutic and preventive strategies, new insights into human evolution, and agricultural and industrial applications. Active pre- and postdoctoral training programs have been established (see related article on p. 11), and efforts are being made to recruit talented investigators from other fields. In addition, short courses are supported for rapidly disseminating developments in genomic sciences to the general scientific and ELSI communities.

First human genome mega-YAC physical map published.

GSDB established.

IOM releases HGP ELSI-funded report, "Assessing Genetic Risks."

1994

Genetic-mapping 5-year goal achieved 1 year ahead of schedule.

PACHG functions absorbed into NACHGR.

NLGLP completes partial-digest libraries in lambda and cosmid vectors for each human chromosome.

Genetic Privacy Act, first HGP legislative product, proposed to regulate collection, analysis, storage, and use of DNA samples and genetic information obtained from them; endorsed by ELSI Working Group.

1995

EEOC guidelines extend ADA employment protection to individuals experiencing discrimination based on genetic information related to illness, disease, or other conditions.

HGP achieves goal of high-resolution mouse genetic map.

Second-generation mega-YAC physical map covers 75% of human genome.

Chromosomes 3, 11, 12, and 22 YAC physical maps published.

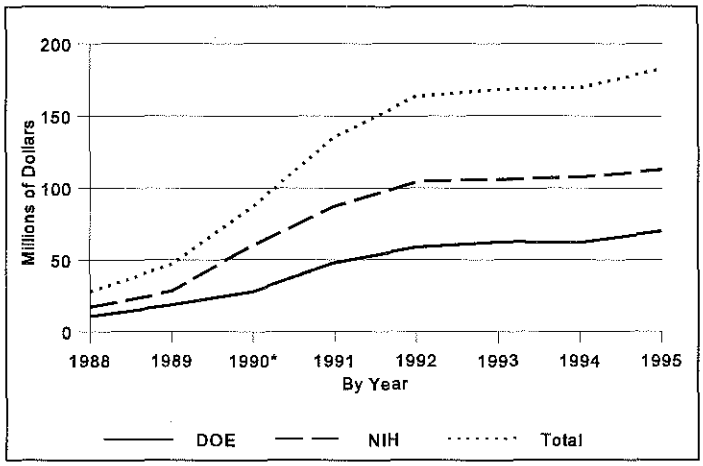
Chromosome 16 map comprising mega-YACs and high-resolution sequence-ready map of cosmid contigs and mini-YACs published.

Metric chromosome 19 map in YACs, BACs, PACs, and cosmids published.

Joint ELSI Working Group and NAPBC present recommendations for state and federal policymakers to protect against discrimination by insurance providers.◊

Five-Years of Progress (continued)

made in developing computer-based systems for automating the acquisition, management, analysis, and distribution of experimental data. Improvements in laboratory-systems integration and information-management systems have promoted large-scale genomics and other biology programs in academia and industry. A number of new databases have been created, and existing databases have been expanded to allow rapid distribution of genome data. In fact, the number of data sources and programs of interest is too large to summarize in this article, but information about many may be obtained from the NIH and DOE WWW sites listed on pp. 1 and 12.



Human Genome Project Funding. As recommended in reports by NRC and DOE HERAC, the original goals assumed a funding level of $200 million annually, adjusted for inflation. This funding level has not yet been achieved.

*HGP funding formally began in October 1990.

Improved software is critical to maximizing automated data acquisition and analysis in genetic and physical map construction, base calling, sequence-contig assembly and editing, project management, and feature recognition and annotation.

Beyond the development of these new tools, several other important informatics problems must be solved. The large number of informatics tools and data resources already available or still being developed is not fully integrated and coordinated. Research, development, and coordination efforts are under way to allow easier access to genome research data. With improved computer infrastructure, analyzing information for further and broader biological research will be easier.

Another major challenge is to integrate genome and genome-related databases. Some approaches under discussion include designing common interfaces, implementing "minimonolithic" databases that contain subsets of relevant data extracted from a set of larger public databases, improving database-query tools, and developing a new category of "middleware" to facilitate the construction of federated databases.

Ethical, Legal, and Social Implications

From the outset of the Human Genome Project, researchers recognized that the resulting increase in knowledge about human biology and personal genetic information would raise complex ethical and policy issues for individuals and society. Accordingly, ELSI investigations have been an integral element of genome programs around the world. In the first few years of the U.S. ELSI programs, NIH and DOE have taken two approaches.

The first approach is a research and education grant program supported by 3% to 5% of funds from each agency's budget. The research program has focused on identifying and addressing ethical issues arising from genetic research, responsible clinical integration of new genetic technologies, privacy and the fair use of genetic information, and professional and public education about ELSI issues. Progress in these areas is discussed in separate sections below.

The second approach involves the NIH-DOE Joint Working Group on ELSI of Human Genome Research. This group is charged with exploring and proposing options for sound professional and public policies related to human genome research and its applications and with identifying gaps in the current state of knowledge about ELSI issues.

Ethical Issues Surrounding the Conduct of Genetic Research

The NIH Office of Protection from Research Risks has developed guidelines for protecting the privacy, autonomy, and welfare of individuals and families involved in human genetic research. These recommendations grew out of a series of meetings and studies supported by the NCHGR ELSI program, which has worked with the National Centers for Disease Control and Prevention to develop recommendations for using stored tissue samples in genetic research.

Responsible Clinical Integration of New Genetic Technologies

Rapid development of new testing techniques and DNA-based diagnostic tests raises questions about their appropriate use beyond the research setting. The NCHGR ELSI program has supported a number of studies to identify issues and develop policy recommendations regarding the delivery of genetic tests into clinical practice.

One set of studies examined issues surrounding genetic testing and counseling for cystic fibrosis (CF) mutations. Results from this consortium led to proposals about preferred methods for providing CF testing to those who desire it. On the basis of these and other study results, clinical policy recommendations are expected to emerge from appropriate professional societies.

Last year, a second major effort in introducing genetic tests was initiated with a set of projects to examine testing and counseling for heritable breast, ovarian, and colon cancer risks. Issues include interest in, demand for, and impact of testing as well as alternative ways to provide the service.

In another approach to the use and regulation of new genetic tests, the ELSI Working Group created a Genetic Testing Task Force. This task force is reviewing genetic testing and examining strengths and weaknesses of current practices and policies. If needed, the task force will recommend changes to ensure that only necessary genetic tests are done and that they are conducted by qualified laboratories.

Finally, in 1994, the Institute of Medicine published a study of the clinical integration of new genetic tests. This report offered a number of recommendations for laboratory quality control of DNA diagnostics and for genetic testing in the clinical setting.

Privacy and Fair Use of Genetic Information

Information obtained from genetic testing potentially can serve the individual well by opening the door to therapeutic or preventive intervention. However, this information may also have such unwelcome effects as increased anxiety, altered family relationships, stigmatization, and discrimination on the basis of genotype. Concerns about stigmatization and discrimination are particularly troubling, especially regarding employability and insurability. In 1993 the ELSI Working Group established the Task Force on Genetic Information and Insurance to assess the potential impact of human genetic advances on U.S. health care and to make recommendations for managing that impact within a reformed health-care system.

A Genetic Privacy Act has been drafted with support from the DOE ELSI program. This act, a model for privacy legislation, covers the collection, analysis, storage, and use of DNA samples and the genetic information derived from them. This first legislative product of the ELSI component has been introduced into several state legislatures and was incorporated into a recently passed measure in Oregon. In November, a similar bill was introduced in the U.S. Senate.

Earlier this year, the U.S. Equal Employment Opportunity Commission ruled that genetic discrimination in employment decisions is illegal.

Designated U.S. Human Genome Project Research Centers
(Directors in Italics)

Affymetrix
Stephen Fodor

Albert Einstein College of Medicine
Raju Kucherlapati

Baylor College of Medicine
David L. Nelson

Children's Hospital of Philadelphia
Beverly Emanuel

Columbia University College of Physicians and Surgeons
Argiris Efstratiadis

Genome Therapeutics Corporation (Collaborative Research Division) Genome Sequencing Center
Jen-i Mao

*Lawrence Berkeley National Laboratory
Mohandas Narla, Acting Director

*Lawrence Livermore National Laboratory
Anthony Carrano

*Los Alamos National Laboratory
Robert Moyzis

Stanford Human Genome Center
Richard Myers

Stanford University DNA Sequence and Technology Center
Ronald Davis

University of California, Berkeley, *Drosophila* Genome Center
Gerald Rubin

University of California, Irvine
John Wasmuth

University of Iowa Cooperative Human Linkage Center
Jeffrey Murray

University of Michigan Medical Center
Miriam Meisler

University of Texas Health Science Center at San Antonio
Susan Naylor

University of Texas Southwestern Medical Center at Dallas
Glen Evans

University of Utah
Raymond Gesteland, Robert Weiss

University of Wisconsin, Madison, *E. coli* Genome Center
Frederick Blattner

Washington University School of Medicine
David Schlessinger

Washington University School of Medicine Genome Sequencing Center
Robert Waterston

Whitehead Institute for Biomedical Research and Massachusetts Institute of Technology
Eric Lander

*Supported by DOE. All others are funded by NIH.

[For more information on centers, see *HGN* **6**(4), 1–13 (November 1994).]

Professional and Public Education

ELSI programs have funded educational projects to increase understanding of the nature and appropriate use of genetic information by health-care professionals, policymakers, and the public. These projects include a reference work to assist federal and state judges in understanding genetic evidence; curriculum modules for middle and high schools; teacher-training workshops; short courses on genome science; radio and television programs on science and ethical issues of the genome project; and the development of educational materials.

REAPING THE BENEFITS

The beginning phase of the Human Genome Project has been remarkably successful. Public data describing human DNA and the DNA of other organisms has expanded enormously, and the information is being used at an increasing rate. Genome project contributions to the study of inherited disease and other biological phenomena are now widely recognized by the scientific community. Investigators are no longer arguing about whether the genome project is a good idea but are debating the most effective ways to reap its rewards. In the commercial sector, a burgeoning body of resources is providing a new base for a wide range of technology industries (see "Technology Transfer" article, p. 15).

Products of the Human Genome Project—including maps, DNA sequences, and improved technology for genomic analysis—will soon enable the era of sequence-based biological investigation to begin in earnest. *[Denise Casey, HGMIS, and NCHGR and DOE OHER program staff]* ◊

[For more detailed information on the Human Genome Project, contact HGMIS, DOE, or NCHGR (see p. 12) or access the WWW Home Pages (see pp. 1 and 12).]

Nature and *Science* Publish Special Genome Issues

Milestones in human genome mapping and sequencing are among the topics covered in special fall issues of Nature *and* Science.

Nature

The *Genome Directory*, a 379-page supplement to the September 28 issue of *Nature*, features a compendium by Craig Venter [The Institute for Genomic Research (TIGR)] of 88,000 unique ESTs from cDNA sequences expressed in 37 human tissues at various developmental stages.

The issue also features a description of the latest Généthon YAC contig map of the human genome by Daniel Cohen and Ilya Chumakov (CEPH) and colleagues. The map covers 75% of the human genome in 225 contigs having an average size of 10 Mb.

Detailed physical maps of four human chromosomes are also presented.

Chromosomes 3 and 12. More-detailed, second-generation YAC contig maps based on the original Généthon YAC physical map. The chromosome 3 map by Robert Gemmill and Harry Drabkin (Eleanor Roosevelt Institute) and colleagues covers 80% of the chromosome in 24 contigs and incorporates physical and genetic map data. Raju Kucherlapati (Albert Einstein College of Medicine) and his collaborators describe a map covering about 75% of human chromosome 12 in 13 contigs and incorporating genetic, physical, and cytogenetic map data.

Chromosome 22. High-density YAC contig map by Ian Dunham (Sanger Centre) and colleagues representing the practical limits of currently available YAC resources and comprising available physical and genetic data.

Chromosome 16. Map incorporating a low-resolution map of mega-YACs and a high-resolution sequence-ready map of cosmid contigs and mini-YACs. The map, by Robert Moyzis and Norman Doggett (Los Alamos National Laboratory) and associates, includes physical, genetic, and cytogenetic data and provides almost complete coverage of the chromosome 16 euchromatic arms.

The *Nature* Home Page is at *http://www.nature.com/*

Science

The annual Genome Issue of *Science* (October 20) features articles on technological developments in genome research, clinical applications, and concerns regarding the social impact of rapidly accumulating genetic information. The issue includes the following:

Genetic Discrimination and Health Insurance. Kathy Hudson [National Center for Human Genome Research (NCHGR)], Karen Rothenberg (University of Maryland School of Law), Lori Andrews (Chicago-Kent College of Law), Mary Jo Ellis Kahn (National Breast Cancer Coalition), and Francis Collins (NCHGR) present a series of recommendations for state and federal policymakers in "Genetic Discrimination and Health Insurance: An Urgent Need for Reform." Drafted by the NIH-DOE Joint Ethical, Legal, and Social Implications Working Group and the National Action Plan on Breast Cancer, the recommendations and definitions suggest that genetic information, including family histories, not be used to establish insurance premiums or eligibility.

Human Genome Project Sequencing Progress. Maynard Olson (University of Washington, Seattle) assesses technical progress in "A Time to Sequence." He argues for an early move to large-scale sequencing of human DNA.

[*Science* offers electronic forums on the above articles (*http://www.aaas.org/science/beyond.html*).]

***Mycoplasma genitalium* genome.** Claire Fraser (TIGR) and collaborators report the complete sequencing of the bacterium with the smallest known genome of any self-replicating organism in "The Minimal Gene Complement of *Mycoplasma genitalium*." In "Life with 482 Genes," André Goffeau (Université Catholique de Louvain, Belgium) discusses this achievement.

Gene Therapy. Ronald Crystal (New York Hospital–Cornell Medical Center) provides an overview of relevant clinical trials and concludes that the therapeutic transfer of genes into humans is feasible and should be pursued further in "Transfer of Genes to Humans: Early Lessons and Obstacles to Success." Two separate reports also in this issue describe efforts to apply gene therapy to people with ADA-SCID, a hereditary, usually fatal disease resulting in a nonfunctioning immune system.

***Caenorhabditis elegans* progress.** Jonathan Hodgkin (Medical Research Council, U.K.), Ronald Plasterk (Netherlands Cancer Institute), and Robert Waterston (Washington University School of Medicine, St. Louis) present a wall chart summarizing progress in the project to characterize the genome of the nematode *C. elegans*. A significant portion of the complete *C. elegans* DNA sequence has been determined, and its potential for yielding clues to understanding developmental, cell, and neurobiology is already unfolding. The chart is also accessible electronically (*http://www.aaas.org/science/science.html 2*).

Other Relevant Articles. Other genome-related articles include a story on a new strategy with the potential to analyze proteins directly and see how they change with disease, a report on a chromosome 4 physical map of the flowering plant *Arabidopsis thaliana*, and two reports describing new approaches to monitoring gene expression. Current and some back issues of *Science* are available electronically (*http://www.aaas.org/science/science.html*).◊

New HGMIS Telephone, Fax Numbers

The HGMIS area code has changed from 615 to 423:

- 423/576-6669, Fax: /574-9888

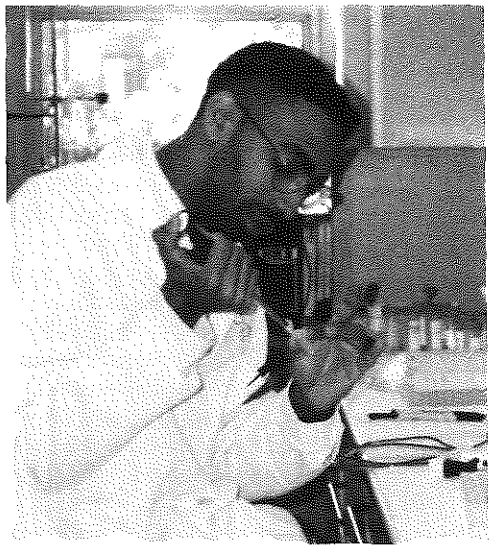
✔ Correction

CENSOR software was developed at Linus Pauling Institute, not at Stanford University as implied in 7(2), 11 (July–August 1995).

Genome News

Two NCHGR Awardees Share More Than Research Interests

Once a brother and sister have moved away from home and embarked on their own careers, their telephone conversations are usually sporadic and limited to an occasional birthday greeting, plans for Christmas, and maybe an update on the parents. However, long-distance chats between siblings Levi "Alec" and Isla Garraway sometimes take on a more serious tone as they discuss their shared dreams of careers in medicine and research. The pair are recipients of National Research Service Awards from the NIH National Center for Human Genome Research (NCHGR).



Siblings Alec (left) and Isla Garraway are recipients of NCHGR National Research Service Award Fellowships.

"Isla and I have always been close, and it is great to have her as a sister, friend, and scientific colleague," says Alec, who is currently an M.D.-Ph.D. candidate at Harvard Medical School and Harvard Graduate School of Arts and Sciences. "We have often spent hours on the phone discussing such diverse topics as apoptosis, protein purification, and neurosurgery, as well as more personal topics."

Isla Garraway, an M.D.-Ph.D. candidate at the Molecular Biology Institute of the University of California, Los Angeles (UCLA), shares her brother's sentiments. "I feel very fortunate to have a brother involved in research as well. Alec and I have always been supportive of each other, and we help each other thrive along the rigorous academic paths we have chosen for ourselves."

Although their career paths have converged in the field of molecular biology and genetics, the road traveled has not been the same. In 1991, Isla went directly to medical school after completing her B.S. in biochemistry and molecular biology at Brown University. Alec's studies took a different course. After completing his B.A. in biochemical sciences at Harvard, he chose to pursue graduate study in science for 2 more years before starting medical school. It wasn't the first time Alec had taken a more circuitous route.

"For most of my youth I assumed I would study music," says Alec, who plays the violin, viola, and piano, "but during high school, my direction shifted."

In contrast, Isla says she is one of the lucky people who has always known what she wanted to do in life—be a doctor. "Some of my earliest childhood memories are of my dad, a professor of plant pathology at Ohio State University, taking my brother and me to his laboratory. We were fascinated by the test tubes full of bugs that he would collect for various experiments."

The two students readily attribute their decision to pursue medical careers to their parents, Michael and Marie Garraway of Worthington,

NRSA Provides Financial Support to Qualified Minority Students and Students With Disabilities

Through its National Research Service Award (NRSA) program, NCHGR currently supports 11 minority students who are in Ph.D. or combined M.D. and Ph.D. programs. NRSA was established in 1974 to help ensure that highly trained scientists are available in adequate numbers and appropriate research areas to carry out the nation's biomedical and behavioral research agenda.

NCHGR is committed to training scientists who will use the tools and information generated by the genome project to address problems relevant to the health of the U.S. population. Because predicting future research areas is difficult, this program funds outstanding graduate students involved in a broad array of projects.

To be eligible for a predoctoral fellowship award for minority students, individuals must be U.S. citizens, noncitizen nationals, or permanent residents and from ethnic or racial groups that are underrepresented in biomedical or behavioral research. Annual stipends include $10,008 for living expenses, a tuition and fee allowance, and a $2000 institutional allowance for travel to scientific meetings and for laboratory and other training expenses. [Training opportunities, enrollment requirements, deadlines: Bettie J. Graham (301/496-7531, Fax: /480-2770, *Bettie_Graham@nih.gov*)] ◊

Human Genome news

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

Human Genome Management Information System
Oak Ridge National Laboratory
1060 Commerce Park, MS 6480
Oak Ridge,TN 37830
423/576-6669, Fax: /574-9888
http://www.ornl.gov/TechResources/Human_Genome/home.html

Managing Editor
Betty K. Mansfield *bkq@ornl.gov*

Editors/Writers/Designers
Anne E. Adamson
Denise K. Casey
Judy M. Wyrick

Production Assistants
Murray Browne
Larry W. Davis
Sheryl A. Martin
Marissa D. Mills
Laura N. Yust

Newsletter Sponsors

Health Effects and Life Sciences Research Division, OHER
David A. Smith, Director
http://www.er.doe.gov/production/oher/hug_top.html
Contact: Daniel W. Drell
301/903-6488, Fax: -8521
Daniel.Drell@oer.doe.gov
or
genome@er.doe.gov

National Center for Human Genome Research
Francis S. Collins, Director
http://www.nchgr.nih.gov
Contact: Leslie Fink
301/402-0911, Fax: -2218
LeslieF@od.nchgr.nih.gov

Ohio, a suburb of Columbus. Both elder Garraways hold Ph.D.'s from the University of California, Berkeley—his in plant pathology, hers in mathematics. "Our parents instilled in us the importance of education. They will always be my biggest inspiration," notes Isla.

Growing List of Credentials

In 1990, Alec attended a seminar by Stephen M. Beverley (Harvard Medical School) on the genetics of the *Leishmania* parasite. Inspired by the professor's rigorous standards of excellence, he later joined Beverley's laboratory, where he began exploring the genetic aspects of *Leishmania* pathogenesis to gain insights into human infectious diseases. Alec has since developed innovations that facilitate a more sophisticated approach toward future genetic studies of the parasite. Two reports he coauthored about *Leishmania* have appeared in the *Proceedings of the National Academy of Sciences.*

Isla says her college junior year was a turning point in her life. Her work as a summer research associate in Christopher Walsh's laboratory at Harvard Medical School introduced her to molecular biology and sparked an interest in research as a career. Currently in her third year of research at UCLA, Isla works in the laboratory of Stephen Smale, investigating the role of the conserved transcriptional elements of lymphocyte-specific genes. In 1993, Smale and Garraway coauthored an article about the mechanism of initiator-mediated transcription, which appeared in *Molecular Cellular Biology*, and Isla has recently submitted a second article for publication. She says her research area is important because "it has much relevance to diseases such as leukemia and immunodeficiency."

In addition to their research and academic responsibilities, the Garraways try to stay well-rounded as members of the larger academic community. Alec is a member of Harvard's Black Health Organization and Brookline secretary of the Minority Biomedical Scientists of Harvard. He is also currently involved with the African-American Achievement Program, which introduces black middle-school students from disadvantaged backgrounds to topics in science and medicine.

Isla is similarly involved. During her medical school days, she was active in the Student National Medical Association, an organization of African-American medical students. Currently she interviews prospective students for the Medical Scientist Training Program.

An Investment in the Future

Both Garraways agree that the NCHGR fellowships (his begun in 1992 and hers a year later) have relieved the financial burden of their education and allowed them to enrich their knowledge. The awards provide assistance for living expenses, tuition and fees, travel to scientific meetings, and laboratory and other training expenses.

"I can't imagine how I could have accomplished this level of scientific and medical training without the support of NIH," says Alec.

"The fellowship from NIH has been invaluable to my training," says Isla. "I hope one day to have my own laboratory that will help find solutions to medically relevant problems."
[Murray Browne, HGMIS] ◊

BAC Clone Web Pages

Information, protocols, and selected references on constructing and using BAC clones and libraries are accessible via the Home Page (*http://www.tree.caltech.edu*) of the Genome Research Laboratory at California Institute of Technology. Melvin Simon's group at Caltech developed BACs to clone and stably maintain large DNA fragments [Shizuya et al., *Proc. Natl. Acad. Sci. USA* **89,** 8794–97 (1992) and Kim et al., *Nucleic Acids Res.* **23**(10), 1838–39 (1995)]. Cytogenetic BAC maps are also available via WWW (*http://www.csmc.edu/genetics/korenberg/korenberg.html*).◊

GDB Forum

The Genome Data Base, Then and Now

Version 6.0 Release Expected in January 1996

When the Genome Data Base began public operation at Johns Hopkins University in 1990, researchers had been mapping the human genome for several decades. In keeping with the Human Genome Project's mission to make mapping data freely accessible to the scientific community, GDB set about curating the existing data and making the information available to those engaged in the mapping effort. The first release, GDB 1.0, debuted in September 1990 at HGM 10.5, offering a character-based interface with pull-down menus. It provided editorial functions, text-based genetic maps, a means for managing literature references, and data on probes and polymorphisms.

Since its inception as the official repository for mapping data in support of the genome project, GDB's efforts always have been dedicated to improving access and creating a data model that would accommodate different mapping methodologies and the data they generate. Allowing the user to view the data in a more meaningful way—one that reflected the richness of the biological data—became an important focus for GDB. This resulted in a series of releases aimed at making the task easier for the user to view and submit data. A statement in the report of the informatics committee at the August 1991 HGM 11 (when GDB 4.1 was in use) reads, "Graphical tools to display, analyze, and compare this information will be developed in the future." The future has finally arrived.

GDB has been aware of the need to provide a graphical representation of stored maps. An associated need is the logical imperative to provide an easier way for the user to edit stored maps and data and create new maps. With these two goals in mind, developers set out to create a set of applications that allow users to create, view, and edit maps off line, submit data directly, and retrieve data using more-flexible query tools.

GDB Version 6

The GDB 6 release will represent a fundamental change in both database structure and user interface. Unlike past major GDB releases that were issued infrequently and included many changes, the new series will be released incrementally to shorten the time between releases and add features as they become available. GDB 6.0 will take advantage of the user's familiarity with WWW to introduce a graphical interface. Subsequent releases will add additional query and editing functions.

A generic Web browser based on Stan Letovsky's Genera technology will allow users to query the database as before. Furthermore, this enhanced interface will enable customized map queries based on order and distance; this was not possible previously. Entering data by hand, including map data, will be accomplished through the Web interface. Once entered, data can be edited and displayed in a graphical format through the GDB map viewer.

Later GDB Releases

Later releases in the GDB 6 series will allow users to create and edit maps through the map editor, which has evolved from the map-viewer component of an applications suite. All users will be able to submit data directly, edit their own previous submissions, and annotate (but not change) the submissions of others. The ability to prepare submissions off line reduces dependence on the Internet, a plus as net traffic steadily increases. Map-viewer software will provide an export function allowing map customization with graphical drawing programs.

Maps that currently reside in GDB will be viewable through the map viewer. Cytogenetic, linkage, content-contig, and radiation-hybrid maps will be supported initially; other types will be added later. Consensus and observed maps will be available, providing both an authoritative source for the novice as well as the most-current mapping data to produce an integrated map for the experienced user. Clicking on an object in the map will display auto-

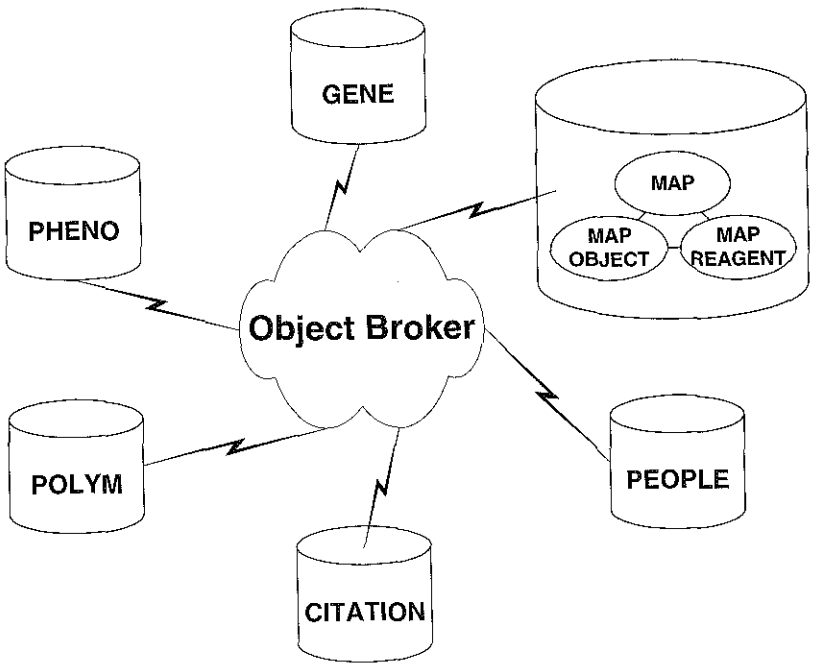


Fig. 1. GDB modular database architecture allows modules to be added without altering the underlying schema. The object broker provides a uniform interface across multiple databases.

GDB Forum

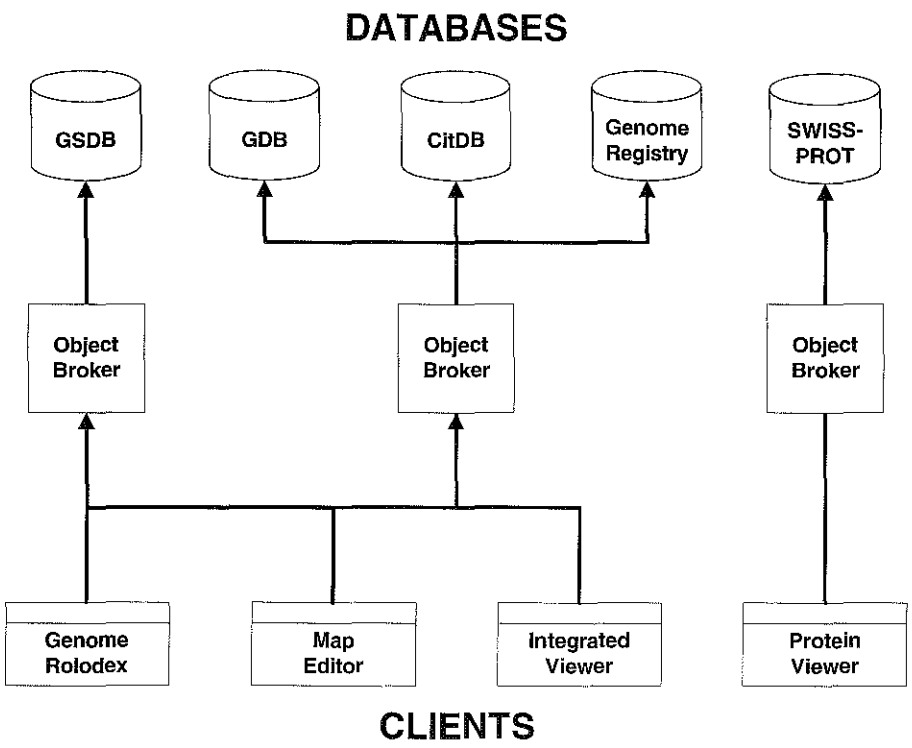


Fig. 2. Client-server relationships in GDB 6.0 and in a database federation.

matically the details of that object in the Netscape Web browser.

Because of the enhanced editorial features of GDB 6, the ability to identify data modifications or altered links to other data objects becomes an extremely important feature. Users will be able to specify and view map versions and display an edit trail of specified versions.

Behind the scenes, the database architecture has been significantly altered, moving from a monolithic to a modular structure. Previously, one large database of objects was stored in tables and linked in various relationships; now, data are stored in several smaller databases. The "object broker" at the core facilitates relationships among objects contained within the modular databases. This model allows other modules to be added without altering the underlying schema (Fig. 1, p. 13).

Database Federation

The use of an object broker opens up opportunities for a possible federation of genomic databases, and the new architecture allows a client to access information contained in multiple databases.

Figure 2 (above) depicts various client-server relationships as they will function conceptually in GDB 6.0 and a database federation.

GDB began assigning accession numbers to all database objects in July 1993. These external identifiers were a first step in linking objects in one database with relevant genomic data stored in other databases, regardless of format or location. The object broker will provide a means of retrieving linked data from multiple databases participating in the federation. The current version of GDB includes links through its Web server to such other databases as the Genome Sequence Data Base, Mouse Genome Database, and SWISS-PROT. This capability will be expanded in GDB 6.

Mapping data continues to increase as the genome project progresses. The release of GDB 6 will provide a new and improved method of navigating and viewing data and relating it to other new genomic insights.◊

GDB Seeks Comments on Schema

GDB is seeking feedback on the schema for version 6.0 of the database. GDB V6 is a major redesign that includes an enhanced map representation, support for direct community submission and curation of data, and other improvements based on an object-oriented data model.

All interested members of the genome community are encouraged to view the schema, which is available in Postscript format on GDB's WWW server (*http://gdbwww.gdb.org/*) under GDB Future Plans. Suggested changes will be implemented with the 6.1 release or later, and a public comment period will be a regular feature of future GDB development schedules. [Schema comments: *schema@gdb.org*, general GDB inquiries: *help@gdb.org*] ◊

Gene Family Database on WWW

The Gene Family Database is a prototype database for integrating biological data generated by the Human Genome Project. Currently, it is a WWW interface that includes descriptive text contributed by research collaborators as well as hypertext links to various biological databases. Entries include gene family definitions, member descriptions, and information on gene structure and sequence, RNA transcripts and protein products, gene expression and function, model organisms, and human genetic disease. Linked databases include GDB, GSDB, SWISS-PROT, PROSITE, BLOCKS, PRODOM, MGD, TBASE, FlyBase, OMIM, and Medline Entrez. [Gene Family Database access (*http://info.gdb.org/~avoltz/home.html*); comments and suggestions, Amy Voltz (*avoltz@gdb.org*)] ◊

GDB Makes New Appointments

Operations Director

Bob Cottingham, formerly of the Human Genome Center at the Baylor College of Medicine in Houston, has joined the Genome Data Base as Operations Director. The announcement was made by Managing Director David Kingsbury and Informatics Director Ken Fasman. "In addition to his administrative responsibilities," Fasman said, "Cottingham will oversee day-to-day operations, including user support and data curation and dissemination."

"GDB will continue to play a key role in the Human Genome Project and must be responsive to the changing needs of a very diverse user community," remarked Cottingham. "I'm looking forward to the challenge and to working with our creative and dedicated staff."

Cottingham's appointment became effective September 1. Said Kingsbury, "We hope the community shares our enthusiasm and excitement and will join us in welcoming Bob to GDB."

Acting Director, Data Acquisition and Curation

GDB also appointed Michael A. Chipperfield as Acting Director of Data Acquisition and Curation, effective October 1. Chipperfield assumed the position held for the past 4 years by A. Jamie Cuticchia, who recently joined the newly expanded genome informatics group at MITRE Corporation of McLean, Virginia. Cuticchia was instrumental in promoting GDB resources and establishing procedures for data collection and management.

Chipperfield, a former member of the Human Gene Mapping Library staff at Yale, continued to serve the community as Deputy Director of the data group when GDB was established at Johns Hopkins University. As Acting Director, his responsibilities include oversight of data group activities and continued interaction with Single Chromosome Workshops and Human Genome Organisation editors. When GDB 6.0 is released, the data group will play an important role in assisting the community with the transition to direct editorial and curatorial capabilities. ◊

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*
- **Japan** *http://gdb.gdbnet.ad.jp/gdb/docs/gdbhome.html*
- **Netherlands** *http://www-gdb.caos.kun.nl/gdb/docs/gdbhome.html*
- **Sweden** *http://gdb.embnet.se:443/gdb/docs/gdbhome.html*
- **United Kingdom** *http://www.hgmp.mrc.ac.uk/gdb/docs/gdbhome.html*

GDB User Support Offices

UNITED STATES
Baltimore, Maryland
help@gdb.org

AUSTRALIA
Sydney
bucholtz@angis.su.oz.au

FRANCE
Villejuif
gdb@infobiogen.fr

GERMANY
Heidelberg
gdb@dkfz-heidelberg.de

ISRAEL
Rehovot
lsprilus@weizmann.weizmann.ac.il

JAPAN
Tokyo
mika@gdb.gdbnet.ad.jp

NETHERLANDS
Nijmegen
post@caos.caos.kun.nl

SWEDEN
Uppsala
help@gdb.embnet.se

UNITED KINGDOM
Cambridge
admin@hgmp.mrc.ac.uk

Technology Transfer—Commercializing Genome Resources

When the genome project formally began in late 1990, available technologies were not nearly efficient or cheap enough to accomplish the goals for mapping, sequencing, and managing and analyzing data. Now, 5 years later, a burgeoning body of resources is providing a new base for a wide range of technology industries involving instrumentation, diagnostics, therapeutics, software and DNA chip development, bioengineering, and agriculture.

Instrumentation, Biological Resources for Future Research

The necessity for large-scale approaches to genome research has pushed technology development toward increasing capacity and decreasing size. Demand for high-throughput DNA sequencing methods, for example, has given rise to gel multiplexing and automated sequence-detection machines and gel readers. A new, multiplexed fluorescence detector for capillary electrophoresis, developed by genome project researchers and licensed this year to the private sector, will form the basis for a new sequencing instrument projected to increase significantly the DNA sequencing rate. Another resource transferred to the private sector this year is a new heat-stable enzyme for replicating DNA that promises to make sequencing faster and more accurate. In September, the "chromosome painting" technology developed by genome project researchers was licensed to a private company that will offer it for disease detection to both the research and clinical communities. The technology is used to detect many chromosomal abnormalities, including Down's syndrome and cancers. Other technology transfer is carried out through software licensing and library distribution. ➠

Technology Transfer (from p. 15)

New Diagnostics, Therapeutics

A government-sponsored plan to accelerate payoffs from genome research is the Tools for DNA Diagnostics component of the Advanced Technology Program, National Institute of Standards and Technology. Funded companies are engaged in such activities as developing diagnostic DNA arrays and adapting fluorescent mapping techniques for analyzing human tissues. Other projects focus on applications of DNA "superchip" technology that have high potential for disease diagnosis and treatment, extremely rapid sequencing, and industrial and environmental monitoring.

Cheap, rapid, and relatively easy-to-use tools for DNA analysis have increased dramatically the number of disease genes isolated during the past few years, providing the raw material for new strategies to diagnose, prevent, and treat disease. Almost 250 gene-derived products are in clinical development, and over 100 companies currently have DNA-based therapies in human clinical trials. Additionally, the top U.S. public biotechnology companies have an estimated 2000 therapeutics in early developmental stages, including monoclonal antibodies, clotting factors, growth factors and hormones, interleukins and interferons, and a variety of other protein or peptide molecules. Since 1988, more than 100 human gene-therapy or gene-transfer protocols have been approved by the NIH Recombinant DNA Advisory Committee.

Challenges

Clinical tests to detect disease-associated mutations offer powerful new tools for disease identification and management and are proving to be the most immediate commercial applications of gene discovery. These tests, however, also pose several medical and technical challenges. Key questions in determining whether a gene discovery will translate into a clinically useful diagnostic test include the following: How often does the test pick up disease-linked mutations? Are these mutations associated with disease development? Is the disease treatable or preventable? Does testing reduce medical cost or improve quality of life?

The NIH-DOE Joint ELSI Working Group is addressing some of these issues and has recently launched the Task Force on Genetic Testing to perform a comprehensive, 2-year evaluation of the current state of U.S. genetic-testing technologies. The task force will examine safety, accuracy, predictablity, quality assurance, and counseling strategies for the responsible delivery of genetic tests.◊

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

Collins (from p. 3)

spin-off of disease-gene discovery is the development of genetic tests that may indicate an individual's predisposition to disease. With the ability to test for disease genes, we can design medical programs for individuals that include lifestyle, diet, and medical surveillance to alleviate or prevent disease. Eventually, new treatments will be developed for many diseases that result from gene malfunctions, but there will be a lag time between our ability to offer a genetic test and the ability to understand the disease sufficiently to develop new treatments and therapies.

Genetic testing also can be potentially harmful if other people use test results to deny jobs or take away insurance. All of us carry probably four or five really fouled-up genes and another couple of dozen that are not so great and place us at some risk for something. People don't get to pick their genes, so their genes shouldn't be held against them. Using personal genetic information to discriminate would severely limit the anticipated medical benefits of human genetic research. In addition to justified concerns over health insurance and jobs, the fear of such misuse will make people unwilling to impart this information to doctors or even family members. As new preventive and treatment methods arise, protecting individuals from discrimination and stigmatization based on their genetic makeup will become increasingly important.

Transformation of Medicine

Beyond the development of new genetic tests and treatment strategies, my long-term dream is for scientists to figure out how diseases work and cure them in advance. In two or three decades, we hope to be able to find out what genetic disease a person is at risk for and fix it by putting in a gene that has the appropriate sequence. This dream has already started to come true in the case of cystic fibrosis. Finding the CF gene enabled researchers to identify people with two copies of the mutated gene and to begin gene-therapy trials soon after.

Unfortunately, medical training in genetics is lagging far behind these scientific advances. Most medical schools are not yet emphasizing the genetics courses so vital to medical education, and many physicians in practice today have had no genetics training at all. This situation will have to change dramatically if medical science is to keep up with and make use of the rapid and extremely valuable discoveries that can affect so many lives.

Genetics is going through a golden era right now. A century from now, people will look back and talk about how exciting it must have been to work in this field at a time when everything was breaking wide open. It's also very gratifying because science is moving so fast to help people suffering from disease. That's what I'm excited about—the chance to do something about diseases that have been completely untreatable in the past and now are beginning to yield their secrets.◊

Smith (from p. 2)

amazing number of microbes are in there, but we don't know how to use them to influence cycles to control some of the harmful things that might be happening. Up to now, biotechnology has been nearly all health oriented, but applications of genome research to modern biology really go beyond health. That's one of the things motivating our office—to try to develop some of these other biotechnological applications."

Responding to criticism about not researching gene function early in the project, Smith reasserts that the purpose of the Human Genome Project is to build technologies and resources that will enable researchers to learn about biology in a much more efficient way. "The genome budget is devoted to very specific goals, and we are tight about being sure that projects contribute toward reaching them."

International Scope

Smith credits the international community with contributing to many project successes. "The initial planning was for a U.S. project, but the outcome, of course, is that it is truly international, and we would not be nearly as far as we are today without those contributions. Also, there's been a fair amount of money from private companies, and support by the Muscular Dystrophy Association in France and The Wellcome Trust in the United Kingdom has been extremely important."

Technology Advances

While noting enormous advances across the board, Smith cites automation progress and observes that tremendously powerful robots and automated processes are changing the way molecular biology is done. "A lot of novel technologies probably won't be useful for initial sequencing but will be very valuable for comparing sequences of different people and for polymorphism studies. One of the most gratifying recent successes is the DNA polymerase engineering project. Researchers made a fairly simple change, but it resulted in a thermosequenase that may answer a lot of problems, reduce the cost of sequencing, and give us better data."

Progress in genome research requires the use of maturing technologies in other fields. "The combination of technologies that are coming together has been fortuitous; for example, advances in informatics and data-handling technologies have had a tremendous impact on the genome project. We would be in deep trouble if they were at a less-mature stage of development. They have been an important DOE focus."

ELSI

Smith describes tangible progress toward goals associated with programs on the ethical, legal, and social issues (ELSI) related to data produced by the genome project. "ELSI programs have done a lot to educate the thinkers, and it has produced a higher level of discourse in the country about these issues. DOE is spending a large fraction of its ELSI money on informing special populations who can reach others. Educating judges has been especially well received because they realize the potential impact of DNA technology on the courts."

According to Smith, more people and groups need to be involved in ELSI matters. "We have some ELSI products: the joint DOE-NIH ELSI working group has an insurance task force report, and a DOE ELSI grantee has produced draft privacy legislation. Now it's time for others to come and translate ELSI efforts into policy. Perhaps the new National Bioethics Advisory Commission can do some of this."

New Paradigm for Biological Research

Smith speaks of a changing paradigm guiding DOE-supported biology. "Some years ago, the central idea or dogma in molecular biology research was that information in DNA directs RNA, and RNA directs proteins. Today, I think there is a new paradigm to guide us: Sequence implies structure, and structure implies function. The word 'implies' in our new paradigm means there are rules," continues Smith, "but these are rules we don't understand today. With the aid of structural information, algorithms, and computers we will be able to relate sequence to structure and eventually relate structure to function. Our effort focuses on developing the technologies and tools that will allow us to do this better.

"That's how I think about what we do in DOE," he says. "We're working a lot on technology and projects aimed at human and microbial genome sequencing. For understanding sequence implications, we are making major, increasing investments in synchrotrons, synchrotron-user facilities, neutron-user facilities, and big nuclear magnetic resonance machines. These are all aimed at rapid structure determination."

Smith explains that now we are seeing the beginnings of the biotechnology revolution implied by the sequence-to-structure-to-function paradigm. "If you really understand the relationship between sequence and function, you can begin to design sequences for particular purposes. We don't yet know that much about the world around us, but there are capabilities out there in the biological world, and if we can understand them, we can put those capabilities to use."

"Comparative genomics," he continues, "will teach us a tremendous amount about human evolution. The current phylogenetic tree is based on ribosomal RNA sequences, but when we have determined whole genomic sequences of different microbes, they will probably give us different ideas about relationships among archaebacteria, eukaryotes, and prokaryotes."

Feeling good about progress over the past 5 years, Smith sums it up succinctly: "Genomics has come of age, and it is opening the door to entirely new approaches to biology." ◊

David Smith is retiring at the end of January 1996. Taking responsibility for the DOE Human Genome Program will be Aristides Patrinos, who is also Associate Director of the DOE Office of Health and Environmental Research. Marvin Frazier will become Acting Director of the Health Effects and Life Sciences Research Division.

Calendar of Genome Events*

January 1996

18. Michael Boehnke: Mapping Genes for Complex Human Diseases; Bethesda, MD [NCHGR Lect. Series, E. Feingold, 301/496-7531, Fax: /480-2770, *fey@cu.nih.gov*]

28–Feb. 1. 5th DOE Human Genome Program Contractor-Grantee Workshop; Santa Fe, NM (abs. deadline past) [S. Spengler, 510/486-4879, Fax: -5717, *sylviaj@ux5.lbl.gov*]

February 1996

4–8. 21st Annual Lorne Conference on Protein Structure and Function; Melbourne [R.J. Simpson, +61-3/9347-3155, Fax: /9348-1925, *simpson@licre.ludwig.unimelb.edu, http://grimwade.biochem.unimelb.edu.au/lorne/protein.htm*]

5–6. NIH Natl. Advisory Council for Human Genome Res.; Washington, DC [J. Ades, 301/402-2205, Fax: -2218, *ja51b@nih.gov*]

5–6. Intellectual Property Issues: Critical Challenges for Biomedicine and Genomics; Santa Fe, NM [CHI, 617/630-1300, Fax: -1325, *chi@healthtech.com, http://www.healthtech.com/conferences*]

10–14. **MBWS-96: Advances in Gene Technol.-Therapeutic Strategies & Molecular Medicine; Ft. Lauderdale, FL (abs. deadline past) [MBWS, 800/miagene, Fax: 305/324-5665, *mbws@mednet.med.miami.edu*]

10–16. Molecular Mechanisms in DNA Replication and Recombination; Taos, NM [Keystone Symp., 303/262-1230, Fax: -1525]

12–14. 1996 World Summit on Molecular Toxicology; IBC, Lake Buena Vista, FL [IBC, 508/481-6400, Fax: -7911, *inq@ibcusa.com*]

12–16. Lorne Genome Conference; Melbourne [J. Timmis, +61-8/303-4661, Fax: /303-4399, *jtimmis@genetics.adelaide.edu.au, http://grimwade.biochem.unimelb.edu.au/lorne/genome.htm*]

15. James M. Wilson; Rockville, MD [TIGR/NIST Distinguished Speakers Series, D. Hawkins, 301/838-3501, Fax: -0209, *dhawkins@tigr.org, http://www.tigr.org*]

15–16. Intellectual Property and Research Tools in Molecular Biology; NRC/IOM, Washington, DC [J. Peck, 202/334-2483, Fax: -1687, *jpeck@nas.edu*]

15–16. Oligonucleotide-Based and Gene Therapy–Based Antisense Therapeutics; IBC, Coronado, CA [see contact: Feb. 12–14]

19–25. Cancer Susceptibility Genes and Molecular Carcinogenesis; Keystone, CO [AACR, 215/440-9300, Fax: -9313, *aacr@aol.com*]

26–27. Gene Quantitation: Diagnostics, Monitoring, and Drug Development; CHI, San Diego [see contact: Feb. 5–6]

29. Philip Green: Genome Map and Sequence Assembly; NCHGR, Bethesda, MD [see contact: Jan. 18]

March 1996

4–6. 3rd Annual HGP: Commercial Implications; CHI, San Francisco [see contact: Feb. 5–6]

7–8. 2nd Annual Genetic Screening and Diagnosis of Human Diseases; CHI, San Francisco [see contact: Feb. 5–6]

10–14. 11th Symp. Intl. Soc. for Genetic Eye Disease and 8th Symp. Retinoblastoma Soc.; Hobart, Tasmania, Australia [Secretariat ISGED'96, *isged96@ariel.ucs.unimelb.edu.au*]

11–14. 3rd ACMG and 27th MOD; San Antonio, TX [S. Robinson, 301/571-1825, Fax: -1895, *srobinson@acmg.faseb.org*]

14. Bruce Stillman; TIGR/NIST, Rockville, MD [see contact: Feb.15]

15–17. Chromosome 5 Workshop; Manchester, U.K. [M. Dixon, Tel/Fax: +44-161/275-5620, *mdixon@fs2.scg.man.ac.uk*]

18–19. 4th Intl. Chromosome 4 Workshop and FSHD Group Workshop; Bochum, Germany [O. Riess, +49-234/700-3831, Fax: /709-4196, *riessoby@rz.ruhr-uni-bochum.de*]

21. Patrick Brown: Scanning a Genome; NCHGR, Bethesda, MD [see contact: Jan. 18]

21–23. Genetics Revolution: A Catalyst for Educ. and Public Policy; Dallas [M. Mays, 214/659-5328, Fax: -5171, *memays@dcccd.edu*]

22–24. HGM '96; Heidelberg, Germany [HUGO Europe Secretariat, +44-171/935-8085, Fax: -8341]

24–27. Electrophoresis '96; Atlanta (abs. deadline past) [D. Wiley, 800/627-0629, Fax: 913/843-1274, *dwiley@allenpress.com*]

April 1996

10. Single Chromosome 2 Workshop; London [N. Spurr, +44-171/269-3846, Fax: /269-3802, *spurr@mahler.clh.icnet.uk*]

13. Conversations: Personal, Prof., and Ethical Challenges in Treatment of Breast Cancer; Minneapolis [U.Minn. Ctr. for Biomedical Ethics, 612/626-9756, Fax: -9786]

18. Richard A. Mathies: Developing New Tools for the Genetic Revolution; NCHGR, Bethesda, MD [see contact: Jan. 18]

21–26. Molecular Cytogenetics; Barga, Italy [GRC, 401/783-4011, Fax: -7644, *grc@grcmail.grc.uri.edu*]

27–May 1. 37th Annual *Drosophila* Res. Conf.; San Diego [S. Bernstein, 619/594-5629, Fax: -5676, *sbernst@sunstroke.sdsu.edu, http://morgan.harvard.edu/dros-conf.html*]

May 1996

3–6. Biomedicine '96; AAP/AFCR/ASCI, Washington, DC [M. Stallings, 609/848-1000 ext. 264, Fax: -5274, *mstallings@slackinc.com*]

6. Samuel Broder; TIGR/NIST, Rockville, MD [see contact: Feb.15]

8–12. 1996 Genome Mapping and Sequencing Meeting; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845, *meetings@cshl.org, http://www.cshl.org*]

12–16. Chromosome 11 Workshop; Buffalo [T. Shows, 716/845-3108, Fax: -8449, *tbs@shows.med.buffalo.edu*]

16. David Burke: Microfabricated Structures for Integrated DNA Analysis; NCHGR, Bethesda, MD [see contact: Jan. 18]

20–21. NIH Natl. Advisory Council for Human Genome Res.; Washington, DC [see contact: Feb. 5–6]

23. Melvin Simon; TIGR/NIST, Rockville, MD [see contact: Feb.15]

June 1996

2–6. ASBMB/AAI/ASIP; New Orleans [FASEB, 301/530-7010, Fax: -7014]

2–6. **First Intl. Congress on Extremophiles; Estoril, Portugal; [G. Antranikian, +49-40/7718-3117, Fax: -2909, *Antranikian@tu-harburg.d400.de*]

2–7. 10th Intl. Workshop on Molecular Genetics of the Mouse; Spa, Belgium [A. Goffinet, +32-81/724-277, Fax: -280, *agoffinet@cc.fundp.ac.be*]

6–7. Intellectual Property Issues; CHI, Philadelphia [see contact: Feb. 5–6]

9–12. Genomic Information: Ethical Implications; Seattle [M. Barnard, 206/616-1864, Fax: /685-7515, *mbarnard@u.washington.edu*]

9–13. BIO '96 Intl. Biotechnol. Meeting and Exhibition; Philadelphia [BIO, 202/857-0244, Fax: -0237]

9–14. Frontiers of Sci.: Nucleic Acids; GRC, New Hampton, NH [see contact: April 21–26]

10–11. Fifth Intl. Bioinformatics and Genome Res. Conf.; CHI, Baltimore [see contact: Feb. 5–6]

12–15. 4th Intl. Conf. on Computational Biol.: Intelligent Systems for Molecular Biol. '96; St. Louis (papers due: Feb. 1, poster abs.: Apr. 1) [D. States, 314/362-2135, Fax: -0234, *states@ibc.wustl.edu, http://ibc.wustl.edu/ismb96*]

14–17. 5th Nordic Genome Workshop; Laugavatn, Iceland [S. Ingvarsson, +354/5601-906, Fax: -943, *siguring@rsp.is*]

14–25. 1st EL.B.A. Foundation Conference on Genome Structure and Function; Isle of Elba, Italy (abs. deadline: Feb. 28) [E. Cristiano, +39/565/901280, Fax: /901283, *Elba_Foundation_Conference@pst-elba.it, http://www.pst-elba.it/EF/conference.html*]

16–21. Diffraction Meth. in Molecular Biol.; GRC, Andover, NH [see contact: April 21–26]

19–20. Molecular Genetic Profiling; CHI, McLean, VA [see contact: Feb. 5–6]

20. Richard Myers: Mapping and Sequencing at the Stanford Human Genome Center; NCHGR, Bethesda, MD [see contact: Jan. 18]

July 1996

21–26. Plant Molecular Biology; GRC, New Hampton, NH [see contact: April 21–26]

28–Aug. 2. Molecular Genetics; GRC, Newport, RI [see contact: April 21–26]

September 1996

16–17. NIH Natl. Advisory Council for Human Genome Res.; Washington, DC [J. Ades, 301/402-2205, Fax: -2218, *ja51b@nih.gov*]

October 1996

29–Nov. 2. ASHG; San Francisco [M. Ryan, 301/571-1825, Fax: /530-7079]

November 1996

7–8. Natl. Conf. on Preparing Schools for the Genetic Revolution; Lincoln, NE (abs. deadline: April 1) [G. Wright, 402/472-8881, Fax: -8412, *gwright@unl.edu*]

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

For Your Information

Training Calendar*

February 1996

12–15. Computer-Aided Molecular Design; Los Angeles [M. Hennessy, 310/825-1047, Fax: /206-2815, *mhenness@unex.ucla.edu*]

26–March 1. **Basic Linkage Course; Zurich [K. Montague, 212/960-2507, Fax: /568-2750, *km165@columbia.edu, http://linkage.cpmc.columbia.edu*]

March 1996

11–12. Differential mRNA Display; New Haven, CT [M. Rossi, 203/932-7107, *rossimj@charger.newhaven.edu, http://www.newhaven.edu*]

11–13. MEMS for Medical and Biotechnol. Applications; Los Angeles [UCLA, 310/825-3344, Fax: /206-2815]

13–15. Computer Analysis of DNA Sequence Data; New Haven, CT [see contact: March 11–12]

April 1996

14–17. Genetic Analysis Methods for Medical Researchers; Boston (appl. deadline: January 15) [N. Powers, 919/684-6274, Fax: -6514, *genclass@genemap.mc.duke.edu, http://www.mc.duke.edu/depts/genetics/courses/index.html*]

May 1996

19–23. Biotechnology for Business: Training for Nonscientists; Durham, NC [B. Maciunas, 919/660-1570, Fax: -1591, *biotech@chem.duke.edu, http://www.chem.duke.edu/special/biotech/*]

29–June 5. Medical Informatics; NLM, Woods Hole, MA [Admissions Coordinator, 508/289-7401, Fax: /457-1924, *admissions@mbl.edu, http://www.mbl.edu*]

June 1996

6–18. Summer Inst. in Statistical Genetics; NCSU, Raleigh, NC [B. Weir, 919/515-3574, Fax: -7315, *weir@stat.ncsu.edu, http://ww2.ncsu.edu/ncsu/CIL/stat_genetics*]

10–14, 24-28. **Basic Linkage Course; New York [see contact: Feb. 26–March 1]

July 1996

7–Aug. 2. Scientific, Ethical, and Social Challenges of Contemporary Genetic Technology; NEH/NSF, Tacoma, WA; (appls. invited from U.S. college and univ. faculty; deadline: Mar. 1) [D. Magnus, 206/756-3508, Fax: -3500, *dmagnus@ups.edu*]

8–12. Introduction to Medical Cloning; New Haven, CT (also offered Aug. 12–16) [see contact: March 11–12]

11–19. 16th Wellcome Summer School. Human Genome Analysis: From YAC to Gene; UMDS, London (appl. deadline: March 29) [P. Faik, +44-171/403-6998, Fax: /407-5281, *wss@umds.ac.uk*] ◊

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

Request for *Arabidopsis* Sequencing Proposals

NSF, DOE, and USDA have been soliciting proposals to begin systematically sequencing the *Arabidopsis thaliana* genome. The project's ultimate goal is to complete the entire sequence within a reasonable time, by 2004 at the latest. To minimize duplication of effort and maximize efficient use of available resources, the project will be coordinated with other ongoing U.S. efforts, including NIH human genome and USDA plant genome research, and with other international programs.

Up to three 3-year awards are expected to be made in FY 1996, contingent on proposal quality and fund availability. Applications due to NSF by January 16, 1996. Potential applicants are strongly encouraged to discuss their plans with a project officer. [NSF: Machi Dilworth (703/306-1422, *mdilwort@nsf.gov*); USDA: Edward Kaleikau (202/401-1901, *ekaleikau@reeusda.gov*); or DOE: Gregory Dilworth (301/903-2873, *greg.dilworth@mailgw.er.doe.gov*)] ◊

SELECTED ACRONYMS

AACR Am. Assoc. for Cancer Research

AAI Am. Assoc. of Immunologists

AAP Assoc. of Am. Physicians

ACMG Am. Coll. of Medical Genet.

ADA Americans with Disabilities Act

AFCR Am. Federation for Clinical Research

ASBMB Am. Soc. for Biochem. and Mol. Biol.

ASCI Am. Soc. for Clinical Investigation

ASHG Am. Soc. for Hum. Genet.

ASIP Am. Soc. for Investigative Pathologists

BAC bacterial artificial chromosome

BIO Biotechnology Industry Organization

bp base pair

CEPH Centre d'Étude du Polymorphisme Humain

CF cystic fibrosis

CHI Cambridge Healthtech Inst.

cM centimorgan

CSHL Cold Spring Harbor Lab.

DHHS Dept. of Health and Human Services

DOE Dept. of Energy

EEOC Equal Employment Opportunity Commission

EL.B.A. Electronics Biotechnology Advanced

ELSI ethical, legal, and social issues

EST expressed sequence tag

FSHD facioscapulohumeral muscular dystrophy

GDB Genome Data Base

GRC Gordon Res. Conf.

GSDB Genome Sequence Data Base

HERAC Health and Environmental Research Advisory Committee

HGCC Human Genome Coordinating Committee

HGM Human Genome Meeting

HGMIS Human Genome Management Information System

HGP Human Genome Project

HUGO Hum. Genome Org.

IBC Intl. Bus. Communications

ICPEMC Intl. Comm. on Protection Against Environmental Mutagens and Carcinogens

IMAGE Integrated Molecular Analysis of Gene Expression

IOM Institute of Medicine

MBWS Miami Bio/Technol. Winter Symp.

MEMS Microelectromechanical Systems

MGD Mouse Genome Database

MGI Microbial Genome Initiative

MOD March of Dimes

MOU Memorandum of Understanding

NACHGR Natl. Advisory Council for Human Genome Research

NAPBC Natl. Action Plan on Breast Cancer

NCHGR Natl. Ctr. for Human Genome Research

NCI Natl. Cancer Institute

NCSU North Carolina State University

NEH Natl. Endowment for the Humanities

NIGMS Natl. Institute of General Medical Sciences

NIH Natl. Institutes of Health

NLGLP Natl. Laboratory Gene Library Project

NLM Natl. Library of Medicine

NRC Natl. Research Council

NRSA Natl. Research Service Award

NSF Natl. Sci. Foundation

OHER Office of Health and Environmental Research

OMIM Online *Mendelian Inheritance in Man*

OTA Office of Technology Assessment

PAC P1 artificial chromosome

PACHG Program Advisory Committee on the Human Genome

PG Plant Genome

RH radiation hybrid

SSLP single-sequence length polymorphism

STRP short tandem repeat polymorphism

STS sequence tagged site

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

UCLA Univ. of California at Los Angeles

UMDS United Medical and Dental Schools (Univ. of London)

USDA U.S. Department of Agriculture

WWW World Wide Web

YAC yeast artificial chromosome

Human Genome Management Information System Subscription/Document Request* (Vol. 7, Nos. 3 and 4)

Name _____
(First) (MI) (Last)

Affiliation _____

_____

Department/Division _____

_____

Street/P.O. Box/Building _____

_____

City/State/Zip Code _____

_____

Country _____ Area of Interest _____

Phone _____ Fax _____

E-Mail Address _____

1. ___ *Human Genome News* ___New Subscriber ___Change of Name/Affiliation/Address (circle all that apply) ___Drop Subscription
2. ___ Reprint of "A New Five-Year Plan for the U.S. Human Genome Project" *(Science*, October 1, 1993) by Francis Collins and David Galas
3. ___ DOE *Human Genome 1993 Program Report* DOE *Primer on Molecular Genetics* (see *http://www.gdb.org/Dan/DOE/intro.html*)

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

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

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

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

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

Printed on Recycled paper