

Human Genome news

Linking
Interdisciplinary
Contributors
and Users

Sponsored by the U.S. Department of Energy

www.ornl.gov/hgmis

ISSN: 1050-6101, Issue No. 45

Vol. 10, Nos. 3-4, October 1999

In This Issue

DOE '99 Oakland Highlights

DOE Genome Program Contractors and Grantees Present Progress at the Seventh Meeting, held in January 1999.

Reported by Denise Casey, HGMIS

Introduction	1
Reports on Progress, Challenges	3
Joint Genome Institute	3
Sequencing at Other Institutions	4
Functional Genomics	6
Informatics	8
Education and Bioethics	9
Microbial Genome Explorations	9

Genome Project

HGP Accelerated Timetable	1
HGP Sequencing Progress	2
Refitting at JGI's Sequencing Facility	2
DOE 2000 Santa Fe Meeting	10

Other Topics

In the News

Chromosomes 21, 22 Sequence	7
<i>Drosophila</i> Sequence	10
DOE Award Winners	11
Human MHC Sequence	11
New Mouse Probes	11
Genetic Testing Advisory Committee	11
Bioengineering at National Laboratories	12
Public SNP Resource	13
SBIR 1999 Human Genome Awards	13
HUGO: London Office, Editors	15
1999 Hollaender Winners	19
Protecting Privacy in Research	19

Microbial Genomics

EcoCyc Database for <i>E. coli</i>	14
Microbial Genome News, Resources	14

ELSI News

Education and Counseling Foundation	16
ELSI Resources	16

Informatics

Mouse Genome Informatics 2.2	16
MRC's Staden Package	17
Other Resources	17

Web, Other Resources	17, 19
----------------------	--------

Calendars	18
-----------	----

Funding	19
---------	----

Subscriptions, Acronyms	20
-------------------------	----

HGP Sequencing Progress: p. 2
February 2000 Santa Fe Meeting: p. 10

HGP Leaders Confirm Accelerated Timetable for Draft Sequence

New Sequencing Resources Aid Effort

In September, international leaders of Human Genome Project (HGP) sequencing confirmed a plan to complete a rough draft of the human genome by next spring, a year ahead of schedule. This accelerated pace is made possible by the commercialization of a new generation of automated capillary DNA sequencing machines and by BAC mapping resources generated from DOE-sponsored clone projects.

The rough draft will provide a scaffold of sequence across about 90% of the human genome. Remaining gaps will

be closed and accuracy improved over the following 3 years to achieve a complete, high-quality human DNA reference sequence by 2003 [see *HGN* 10(1-2), 1 (www.ornl.gov/hgmis/publicat/hgn/v10n1/01goals.html)]. So far, about 13% of human sequence has been finished, and another 12% is available in draft form (genome.ornl.gov/GCat; www.ncbi.nlm.nih.gov/genome/seq).

Sequencing Allocation

About 60% of the draft sequence will be produced by six major NIH-funded (see *HGP Timetable*, p. 2)

Oakland Hosts DOE Genome Program Contractor-Grantee Meeting '99



Enthusiasm ran high for the DOE Human Genome Program (HGP) in response to impressive gains reported at the Seventh DOE Contractor-Grantee meeting in January of this year. Convened every 12 to 18 months, this workshop provides an effective forum for all DOE HGP investigators and invited guests to discuss their research, initiate collaborations, and share new material resources and software capabilities.

Although traditionally held in Santa Fe, New Mexico, the 1999 meeting was moved to Oakland, California, so attendees could visit the new Production Sequencing Facility of DOE's Joint Genome Institute (JGI) in nearby Walnut Creek. Two years ago, JGI began operations under the direction of Elbert Branscomb to address the challenge of high-throughput

sequencing, which remains the major task facing the HGP today. An overview of JGI progress begins on p. 3.

Investigators representing other HGP-funded projects reported that exciting improvements in mapping and sequencing technologies resulted in higher throughput. New approaches to regulating gene expression are helping to assess whole-cell effects that are key to functional genomics. Attendees

Meeting '99
Highlights Begin
on p. 3

also discussed the application of genome sequencing information to real-life problems in medicine and waste cleanup. DOE is continuing activities to educate the general public about the HGP and its societal impact. Researchers in the Microbial Genome Program, a spinoff of HGP, reported impressive gains in whole-genome sequencing and analysis, proteomics, and comparative genomics.

HGP Timetable (from p. 1)

sequencing centers, including those previously established at Washington University, St. Louis; MIT-Whitehead Institute; and Baylor College of Medicine. In July, laboratories at the University of Washington, Seattle; Genome Therapeutics Corp.; and Stanford University joined the NIH production sequencing effort.

The DOE-funded Joint Genome Institute (JGI) and the Sanger Center (United Kingdom) will generate about 10% and 30%, respectively, of the draft sequence. (See box, below right, for details on JGI sequencing.) France, Germany, and Japan are contributing significant amounts of human sequence, and China recently joined the worldwide project. To avoid duplicated work, each laboratory focuses on particular genomic regions (see Human Genome Sequencing Index, www.ncbi.nlm.nih.gov/HUGO).

Sequencing group members reaffirmed the policy of placing all sequence in publicly accessible databases within 24 hours of obtaining a continuous 1000- to 2000-base assembly. They also agreed that sequencing should continue to be based on a proven strategy using BAC clones containing DNA from known genomic locations.

BAC-End Sequences: Prime Resource

Data from the DOE-funded BAC-end sequencing projects at The Institute for Genome Research (TIGR) and the University of Washington, Seattle (UWS), are critical for achieving the new HGP goals. These projects are generating single-sequence reads from both ends of the human DNA insert in 450,000 BAC clones [*HGN* 10(1-2), 4 (www.ornl.gov/hgmis/publicat/hgn/v10n1/04bacend.html) and BAC Web page (www.ornl.gov/meetings/bacpac)].

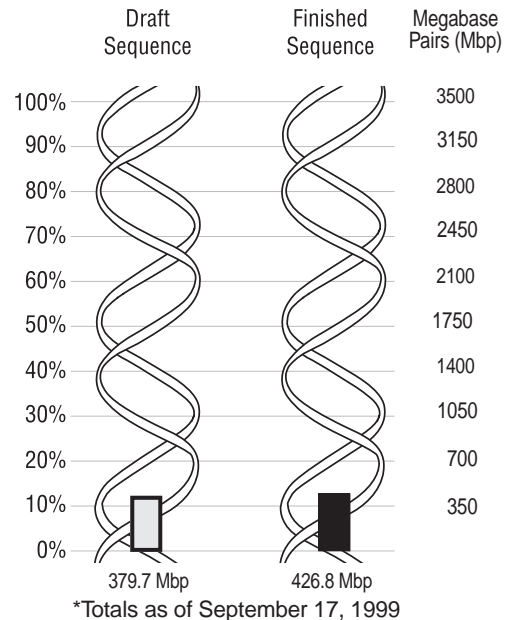
In March, DOE increased its support of BAC-end sequencing projects to accelerate their completion and usefulness for guiding production sequencing. The BAC-end sequences, called sequence tag connectors (STCs), are

FY 1999 HGP Budgets

- DOE: \$ 89.8M
- NIH: \$225.7M
- Total: \$315.5M

HGP Sequencing Progress*

Graph at right depicts progress toward generating a working draft of the human genome by spring 2000 and a complete, high-quality reference human DNA sequence by 2003. Sequences from roughly one-quarter of the human genome have been deposited into GenBank (www.ncbi.nlm.nih.gov/Genbank), and several individual chromosomes are expected to be completed before the end of this year. For updates on HGP sequencing, see Web site (www.ornl.gov/hgmis/project/progress.html).



JGI Refits Production Sequencing Facility

Accelerated goals are challenging the major Human Genome Project sequencing centers to restructure and streamline their processes. The Joint Genome Institute (JGI) Production Sequencing Facility is using 84 new high-throughput MegaBACE DNA sequencing machines acquired from Amersham Pharmacia Biotech. These machines use multiple thin "capillary" tubes instead of large flat gels to separate and "read" the DNA sequence. The technology was developed largely with funding from the DOE Human Genome Program and licensed to AP Biotech.



JGI's current sequencing targets are human chromosomes 5, 16, and 19, which together contain about 300 Mb of DNA or about 10% of the entire human genome. Sequencing director Trevor Hawkins reports that as of September 20, JGI had produced 31.2 Mb of finished sequence and about 41 Mb of draft sequence (see "HGP Sequencing Progress" box, above). The Genome Annotation Consortium is joining with JGI for sequence analysis (see p. 8 for details), and Stanford University will be part of the finishing effort.◊

valuable tools for eliminating redundant sequencing. They can direct researchers to particular BAC clones needed for extending a sequenced region along the chromosome and can identify clones representing genomic regions still not sequenced. STCs also can provide quality checks on sequence assemblies and are useful for spanning regions resisting standard sequencing biochemistry.

STC data will furnish researchers with markers spaced on average

every 3000 to 4000 bases across the entire human genome, a 100-fold improvement over other current human genome maps. Detailed data on BACs are available at TIGR (www.tigr.org/tdb/humgen/bac_end_search/bac_end_intro.html) and UWS (www.htsc.washington.edu). The STC data resource is complemented by several other types of mapping information, including FISH mapping of BAC clones.◊

DOE Genome Meeting Highlights

DOE
99

HGP Grantees Report Progress, Challenges

Some 360 researchers, program managers, and invited guests gathered in California on January 12-16 for the DOE Human Genome Program workshop. Plenary presentations and posters described a wide spectrum of accomplishments and activities, highlights of which are reported below.

Joint Genome Institute

Achievements, Future Plans Presented

A contingent of researchers from JGI reported on the challenges and triumphs involved in merging the sequencing capabilities of three national laboratories into a single, smoothly functioning unit while meeting much higher sequencing goals. Efforts were highly successful, and all objectives were reached as planned, with over 20 Mb of sequence submitted to GenBank in October 1998.

JGI converged the disparate sequencing processes of the three centers of Lawrence Berkeley, Lawrence Livermore, and Los Alamos national laboratories at the Production Sequencing Facility (PSF) in Walnut Creek, California. The team took up residence in one of the new buildings in December 1998, and Secretary of Energy Bill Richardson was keynote speaker at the April 19 formal PSF dedication (see box at right).

Some details of JGI presentations follow (for JGI update, see box, p. 2, and Web site: www.jgi.doe.gov).

Clone Resources

Jan-Fang Cheng reported that about 71.7 Mb of sequence-ready maps were generated in FY 1998 for target regions of chromosomes 5, 16, and 19, and some 3500 unique genes have been mapped to these regions. The goal for FY 1999 was to develop maps for over 250 Mb to allow production of a draft sequence of the three chromosomes by March 2000. The JGI mapping teams then plan to generate similar maps for the mouse genome in regions syntenic with these human chromosomes.

Clones containing target regions are isolated from 10x-coverage BAC libraries via a combination of colony hybridization and PCR approaches using STSs obtained mostly from

public databases. Minimally overlapping clones that represent contiguous regions (contigs) of the chromosome are selected for sequencing. Contigs, expanded by end-sequence STS walking, are oriented by STSs developed from known genes and ordered genetic and RH markers. All clones are sized by pulsed-field gel electrophoresis, and their chromosome map locations are confirmed by FISH. Detailed information on STS and restriction maps is available on the Web site (www-hgc.lbl.gov/human-maps.html).

Quality Control and Assurance

Norman Doggett described the robust system of quality control and assessment processes for sequence generated at JGI. He emphasized JGI's commitment to quality, with more than 95% of all finished bases having Phrap scores over 40 and at least 95% of all bases covered in reads from both

strands (or 2 chemistries). To ensure correct assembly, JGI validates the data before submitting the sequence to GenBank and performs postsubmission quality assessment. Validation includes visual inspection of all bases with Phrap values of 30 or less and comparison of the final assembled sequence to 3 independent high-resolution restriction fingerprints.

Informatics

Meeting JGI's ambitious FY 1998 goals was something of a "miracle of elasticity" of the preexisting separate systems at member laboratories, according to Tom Slezak. He noted the complex efforts now under way as the laboratories work toward merging into a single PSF informatics system without interfering with production pressures. The long-term goal is to make information available in a single view through the entire process, from mapping to submission, while allowing for multiple viewing methods.

(See JGI, next page)

Richardson Attends Facility Opening

Secretary of Energy Bill Richardson (second from right) listens as JGI Director Elbert Branscomb explains equipment at the dedication of DOE's new JGI Production Sequencing Facility (PSF) in California. Others listening are, from left, DOE Office of Science Director Martha Krebs, DOE Office of Public Affairs Deputy Director Stu Nagurka, and Walnut Creek city officials Sue Rainey, Don Blubaugh, and Gwen Regalia. Located near both the Lawrence Livermore and Lawrence Berkeley national laboratories, PSF is equipped with flexible, modular laboratory spaces to accommodate changes in sequencing technologies.◊



Sequencing at Other Institutions

New Strategies, Resources Reported

The exciting potential of high-throughput sequencing was evident as presenters detailed progress in effective strategies for handling repeat telomeric regions, high-speed DNA analysis systems, and new vectors.

Verifying Sequence of Subtelomeric Regions

Producing correct maps and sequences in human DNA regions containing a high percentage of repetitive sequences is a challenge. Particularly troublesome are the chromosome ends (telomeres), which can have complex arrays of repeats. The subtelomeric region of human chromosome 7q was sequenced largely at Los Alamos National Laboratory while Bob Moyzis was director of the Center for Human Genome Studies. At the Oakland meeting, Moyzis [now at University of California (UC), Irvine] reported a quality-control analysis using the RARE (for RecA-Assisted Restriction Endonuclease) cleavage technology.

This technology allows restriction maps to be constructed directly from genomic DNA rather than from DNA clones that are more prone to rearrangements. Using DNA from PCR products generated on genomic DNA, the UC researchers also resequenced about 18% of the 0.13-Mb 7q telomere. These methods confirmed prior mapping and sequencing results. The group is completing the mapping and sequencing of two additional telomeres, 9q and 11q, and next will target telomeric regions of chromosomes 5, 16, and 19.

JGI (from p. 3)

"Data silos" in various locations currently are hard to connect, but eventually physical maps will become sequence maps as all marker and annotation data are merged. Querying, navigating, and displaying data soon will require seamless links to information on function, structure, and experimental results. Slezak emphasized the importance of implementing an infrastructure for functional genomics informatics.◇

Microfabricated Devices

Microfabrication of devices for DNA sample preparation, electrophoretic separation, and detection is advancing a new generation of high-speed DNA analysis systems. Microdevices require less sample input, present narrower sample zones, generate less heat during electrophoresis, and thus suffer less sample-diffusion broadening during shorter, more effective run times. Although some of these devices will be used to finish the Human Genome Project, most will be deployed into the next century as the number of sequencing projects increases.

Richard Mathies (UC, Berkeley) previously developed a capillary array electrophoresis device in which DNA was separated in individual glass capillaries, a prototype that has since been commercialized successfully. Sequencing throughput had been limited by the number of capillaries that could fit into an array and by the inefficiency of sample injection.

At the Oakland meeting, Mathies described a new microfabricated

sequencing device that promises higher throughput (see box, below). Detection in the new device is accomplished by the confocal fluorescence optical system used in Mathies' previous technologies, but the group is moving to other detection paradigms. Mathies and colleague Indu Kheterpal reviewed several approaches to using capillary arrays for high-throughput DNA sequencing [see *Anal. Chem.* **71**(1), 31-37A (1999)].

BAC Update: New Vectors, Sequencing Progress

Because of their stability and large insert size (average, 150 kb), BAC clones have played an increasingly important role in human and mammalian genomics in recent years. BAC libraries are being used or developed for almost every extensively characterized genome.

At the Oakland meeting, Melvin Simon (California Institute of Technology) talked about new BAC vectors for the genome project and for tagging BAC clones with gene information to enhance research into gene function in animals and plants. The number of human BAC clones generated is now close to one million. Simon's team has been annotating the "D" library [used for BAC-end sequencing at The Institute for Genomic Research (TIGR)

New Device Speeds Sequencing

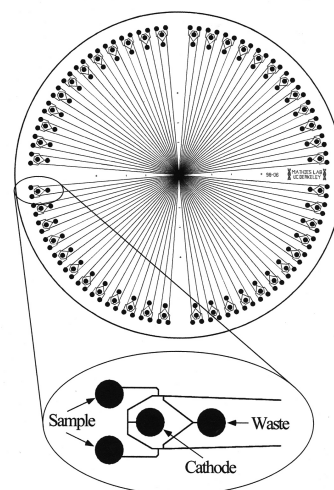


Richard Mathies
University of
California, Berkeley

The "micro-plate DNA analyzer" features 96 channels etched in a radial pattern into a 4- or 6-in. glass wafer disk. Sample injection occurs near the disk edge, and separated DNA frag-

ments are detected near the center. Separations of ds-DNA fragments are complete in <120 seconds, and DNA sequencing separations take only 20 minutes for 500 bp [see *Anal. Chem.* **71**(1), 566-573 (1999)].◇

Radial 96 Channel CAE Microplate



and the University of Washington (UW, see below)] with ESTs obtained from I.M.A.G.E. Consortium resources (www-bio.llnl.gov/image/image.html).

Simon also reported the development of a new BAC vector and an improved method of constructing BAC libraries. The group has begun constructing a series of BAC libraries with much larger insert sizes (182 to 202 kb) for sequencing projects on humans and other organisms, including *Arabidopsis*, maize, and rice. The larger-insert libraries will provide significant improvement, he noted, to applications in physical mapping, positional cloning, and DNA sequencing.

Pieter de Jong (now at Parke-Davis Laboratory) also described new approaches to library construction and the preparation of new BAC vectors for BAC cloning and transformation-associated recombination (TAR) cloning (<http://bacpac.med.buffalo.edu>). When comparisons of multiple gene alleles are desirable, scientists use the TAR strategy invented by Natalay Kouprina and Vladimir Larionov (both at NIH National Institute of Environmental Health Sciences). TAR provides for an economical, selective recloning of target genomic DNA segments. The TARBAC variant has been used to prepare BAC libraries for several less-complex genomes of unicellular eukaryotes to model future work with mammalian TARBAC libraries.

UW and TIGR are generating sequence reads from BAC ends. These sequence tagged connectors (STCs) are useful markers that speed human genome contig building and sequencing. A target of one STC for every 3 kb of the human genome will eliminate bottlenecks in contig development. Beyond use in human genome sequencing, the characterized clones provide excellent resources for biological studies. The STC resource is available on the Web for use by any researcher for clone or sequencing target selection, and the clones can be ordered from commercial resources.

Plasmid Vectors

John Dunn (Brookhaven National Laboratory) and his team have developed improved biochemistry and vectors for the "deletion-factory" approach to sequencing. This involves generating a set of nested deletions from a common parent plasmid and recognizing descendants with progressively shorter inserts. Those chosen for successive sequencing differ in length by less than one sequencing read, and assembly problems are negligible compared to those obtained with shotgun sequencing. The deletion factory approach is particularly useful for regions having numerous repeats that cause troublesome assembly ambiguities during shotgun sequencing.

BAC-End Project Web Sites

- Human BAC-End Sequencing: www.ornl.gov/bac
- University of Washington, Seattle: www.htsc.washington.edu
- TIGR: www.tigr.org/tdb/humgen/bac_end_search/bac_end_intro.html

The new "pZIP" plasmid vectors (named for the way the sequencing progresses along the insert) are maintained in a single-copy state that generally increases the stability of the foreign DNAs they carry and are amplified only when DNA is needed for sequencing. DNA segments up to at least 15 kb have been sequenced, and the reads have been assembled easily. Dunn's team currently is sequencing restriction fragments of human BACs with sizes in the 5- to 15-kb range.

Plasmid Unwinding For Sequencing

Sergei Kozyavkin (Fidelity Systems, Inc.) presented a poster describing applications of ThermoFidelase, a heat-stable enzyme that relaxes plasmid DNA supercoils. Upon relaxation at high temperature with a shift to lower temperature, the covalently closed, duplex DNA is forced into a partially denatured configuration. This enables the binding of primers for Sanger sequencing biochemistry. The contaminating fragments of host DNA retain their original duplex configuration and cannot bind primers. Useful Sanger reaction products thus can be generated selectively from templates of ThermoFidelase-treated plasmids, even in the presence of great excesses of fragmented host DNA. This demonstrates that even single-copy plasmids such as BACs can serve as templates to generate good-quality sequencing ladders within the total complement of excess host *Escherichia coli* DNA. Accordingly, this technology has substantial promise for simplifying several BAC-end sequencing protocols, including BAC-end sequencing.◊

Gene Riches of Chromosome 19 Revealed

Preliminary analysis of chromosome 19 sequence data, as reported by Jane Lamerdin (Lawrence Livermore National Laboratory, LLNL) at the Oakland meeting, reaffirms the chromosome's selection as a rich target for gene discovery through genomic sequencing. Lamerdin's group at JGI-LLNL has finished about 10 Mb of genomic sequence and has placed many large (1-Mb) contigs from well-mapped regions in its sequencing queue, including representative contigs from almost every cytogenetic band on the chromosome.

Several sequenced GC-rich areas exhibit a high gene density relative to the rest of the genome—on average 1 gene per 20 to 25 kb. To better understand evolution and subsequent functional diversification, the group currently is sequencing and making a detailed comparison of several clustered chromosome 19 gene families and their orthologs in the mouse. The genes coding for cytochrome P-450 proteins are one such family under study. These proteins are considered part of the detoxification pathway used by eukaryotes (animals, plants, algae, and fungi) to rid themselves of toxins. Other gene families being analyzed include the pregnancy-specific glycoprotein family, multiple zinc finger families, and olfactory receptors.◊

Functional Genomics

Efficient interpretation of the functions of human genes and other DNA sequences requires that resources and strategies be developed to enable large-scale investigations across genomes. Goals include studies into genome expression and control, creation of mutations that cause loss or alteration of function in such non-human organisms as the mouse, and development of experimental and computational methods for protein analyses. Some highlights of HGP functional genomics projects follow.

Deletion Studies

Eddy Rubin's group (Lawrence Berkeley National Laboratory, LBNL) uses the laboratory mouse to examine gene function, particularly for mouse genes similar to those found in human genomic regions sequenced at the Joint Genome Institute (JGI). Rubin described use of the Cre Lox system to create mice having several large gene deletions in a 4.5-Mb stretch of mouse chromosome 11 syntenic with human chromosome 5q31. The deleted region contains nine genes of unknown function. The majority of mice homozygous for the deletion exhibited triglyceride levels about tenfold greater than control animals and died prematurely. These deletion mice may prove useful in studying triglyceride metabolism, which is important for understanding atherosclerosis in humans.

Insertion of a human YAC containing about 120 kb of the deleted region successfully corrected the high-triglyceride condition and produced mice having normal lifespans. Three candidate genes are present in the YAC, including one with homology to a liver-

specific, transporter-like protein previously characterized in the rat.

Rubin's group also is investigating the functions of evolutionarily conserved noncoding sequences in the human 5q31 region. This region is biologically interesting because it carries a family of cytokine genes, which are important regulators of the immune response. Rubin's data suggests that a 400-bp conserved element in this region is involved in regulating the expression of the human (interleukin) *IL4* and *IL13* genes.

High-Throughput Mouse Mutagenesis, Phenotyping

Eugene Rinchik (Oak Ridge National Laboratory) discussed progress in creating a large mouse-mutation resource for function studies. The strategy, based on recovered recessive phenotypes, involves chemical mutagenesis of specific chromosomal regions, broad-based phenotype screening (in house and through a statewide consortium), and correlation of specific DNA sequences with phenotypes.

The group has been characterizing regions of mouse chromosome 7 while recovering recessive single-gene mutations induced by the powerful mutagen N-ethyl-N-nitrosourea (ENU) and mapped by two-generation hemizygosity screens with radiation-induced deletions. New work involves three-generation homozygosity strategies to induce mutations in proximal chromosome 7 (human 19q homology), mid-chromosome 7 (human 15q homology), and mid-to-distal chromosome 15 (human 8q, 22q, and 12q homologies). The emphasis is on developing genetic

reagents that enable any mutation to be maintained and used by a wide variety of investigators without the need for molecular genotyping.

Rinchik also described the high-throughput goals of the Tennessee Mouse Genome Consortium (TMGC). TMGC combines ORNL's resources and experience in mouse genetics with academic and clinical expertise across the state to induce and detect mouse-gene mutations. The goal is to create human-disease models and perform gene-function studies (<http://tmouse.org>).

Mouse-Human Comparative Sequence Analysis

Lisa Stubbs' team (JGI-LLNL) is comparing human chromosome 19 sequence (19q13.4), generated by the JGI, with a region of mouse chromosome 7 containing similar genes. The goal is to annotate the human sequence with information on gene function extrapolated from mouse genetics and biology.

The group is focusing on imprinted genes in these areas to find regulatory elements that control imprinting processes in humans and mice. Imprinted genes, which tend to be clustered and may share regulatory regions, are expressed differently depending on which parent contributed the allele. Stubbs reported identification of *Zim1*, a new maternally expressed gene located in both species next to the paternally expressed *Peg3* gene. About 100 human genes are thought to be imprinted, and 10 or more new imprinted genes may be located in the 19q13.4 mutation region.



From left, Darla Miller and Dabney Johnson (both of Oak Ridge National Laboratory) and Dave Smith (retired director of the DOE Human Genome Program).

► Acronyms

A list of acronyms is printed on the back page of this newsletter.

► Meeting Abstracts

More details of all HGP and MGP projects are in the DOE Contractor-Grantee meeting abstracts (www.ornl.gov/hgmis/publicat/99santa; print copies are available from HGMIS at the address on p. 14).

Phage Display Identification of Target Protein

A poster presentation by Andrew Bradbury's team (Los Alamos National Laboratory) described the use of phage display in functional genomics. Phage display offers the possibility of selecting single-chain antibodies (and the genes encoding them) from libraries of 1012 or more different polypeptides on the basis of their abilities to bind target proteins. This technology will enable derivation of ligands that recognize protein products of all human genes, and the ligands will be used to characterize proteins and protein complexes.

Ribozyme Inactivation of Target Gene Expression

Another method of determining gene function is to begin with an interesting cellular function and work back to the gene. Jack Barber's team at Immusol, Inc., has taken this approach using ribozymes (RZs), which are RNA molecules that can be engineered to cleave and inactivate other RNA molecules in a sequence-specific fashion. RZs can be designed to selectively inactivate the expression of any target gene ("gene knockdown") and therefore its corresponding protein. Immusol has developed a combinatorial library of RZ genes, delivered in viral vectors, to use as probes for finding and then cloning genes. The RZ

gene library is delivered into large numbers of tissue culture cells (one RZ gene per cell for each RZ gene in the library), followed by selection for individual cells that have lost a particular function. Barber reported cloning the first gene, a tumor suppressor, with no prior sequence information. The group now is identifying genes involved in regulating cancer gene expression and hepatitis C virus replication.

One Gene But Many Proteins

Complicating the study of gene function is the fact that multiple proteins can arise from a single gene. This can be the result of alternate splicing and editing of the mRNA, post-translational protein modification, or alternative translation. Nonstandard translation involves "recoding," in which ribosomes change the coding of an mRNA by frame shifting, changing a codon's meaning, or skipping over some message regions instead of translating them all. Ray Gesteland (University of Utah) described a pilot project using electrospray liquid chromatography mass spectrometry (MS) to catalog the number of each mRNA species' protein products from the 500 genes in yeast mitochondria. The project's ultimate goal is to develop ways to monitor all proteins expressed during physiological events such as development.

Proteome Dynamics

Richard Smith (Pacific Northwest National Laboratory, PNNL) further described how the complexities of the proteome—all the proteins expressed by an organism—cannot be accounted for by DNA sequences alone. He pointed out that, in contrast to an organism's virtually static and well-defined genome, the proteome continually changes in response to external and internal events.

News Update

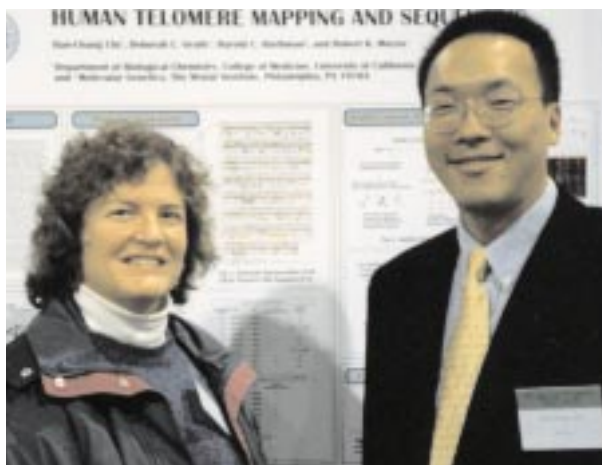
Chromosomes 21, 22 Near Sequencing Finish Line

This fall, international sequencing consortia are expected to announce the completion of chromosomes 21 and 22, the first to be declared "finished" according to Human Genome Project criteria. HGP standards include sequencing over 95% of the chromosome; determining the number, location, and size of remaining gaps; and closing all gaps over 150 kb.◊

Improved Sample Prep System

Although the last decade has brought many advances in analytical instrumentation for characterizing complex biological samples, the technology for sample preparation has lagged behind. In an article in *Analytical Chemistry* [70, 1797-1801 (1998)], Richard Smith (PNNL) described a sample-cleanup microdialysis system that was integrated into the measurement instrumentation, greatly reducing the amount of sample required for analysis. Smith observed that the device has potential for rapidly identifying microorganisms by using mass spectrometry to detect characteristic biomarker species and for analyzing DNA using MS or automated PCR instrumentation.◊

Smith's group is exploring ways to characterize the proteome directly by combining the speed of capillary isoelectric focusing electrophoresis, which separates proteins by charge, and the mass accuracy and sensitivity obtainable with Fourier transform ion cyclotron resonance MS. He described its initial application to several completely sequenced prokaryotes including *E. coli*, expressing confidence in the technique's applicability to larger genomes.◊



From left, Linda Ashworth (Lawrence Livermore National Laboratory) and Han-Chang Chi (University of California, Irvine).

Informatics

Oakland presentations emphasized that genome-sequencing projects are producing data at a rate exceeding current analytical and data-management capabilities. Additionally, some current computing problems are expected to scale up exponentially as the data increase.

Genome Annotation Consortium

Ed Uberbacher, Jay Snoddy, and Phil LoCasio (all at Oak Ridge National Laboratory) offered an update on progress at ORNL and the multi-institutional Genome Annotation Consortium (GAC), which was established to address massive computational and informational challenges.

The goals of this work are to develop a system for whole-genome annotation that (1) organizes various types of data around genome frameworks that can be cross-indexed, compared, and cross-navigated and (2) allows multiple analytical methods to be applied to the same data. Steps in the annotation process include the following:

- retrieving data and assembling genomes;
- computationally finding genes and other sequence-level features;
- computationally determining homology, function, and other relationships;
- genome-wide structural modeling of gene products;
- analyzing and modeling pathways and systems; and
- managing, accessing, and visualizing data.

Snoddy, Uberbacher, and LoCasio discussed the growing complexity and expense involved in biological computing for genome assembly and annotation. They noted that assembly problems will increase as billions of nucleotides are entered as draft sequences into the sequence databases by mid-2000, when the daily assembling of new data alone will require over 1600 workstations.

Other significant computational challenges include integrating the major community maps, which often have

inconsistencies and discrepancies, and performing comprehensive sequence analyses for gene modeling, which requires the time-consuming application of several algorithms. Furthermore, completing some desired analyses for protein classification currently could require about 70 days on a 1024-node processor. Challenges are similar for such other comparative processes as genome-to-genome alignment for studying mouse and human synteny. As sequence numbers and lengths increase, challenges become even greater for making phylogenetic gene and species trees. Meeting these and other high-performance biological computing needs, the speakers emphasized, demands a centralized approach with advanced infrastructure and specialized facilities.

Uberbacher gave an overview of GAC progress in developing tools, servers, and special data views to serve the community. Achievements include establishment of data-acquisition and semiautomated sequence-assembly components and modules that are integrated to allow comprehensive genome-wide analysis. He noted that the computation-based GRAIL-EXP is finding about 10 times more human genes than investigators had identified previously, as indicated in the GenBank annotation. All human and microbial gene-analysis tools are available to researchers. At present, the Oak Ridge group is focusing on urgent annotation challenges from the massive sequencing ramp-up

under way at the DOE Joint Genome Institute (JGI: www.jgi.doe.gov; Genome Catalog: <http://genome.ornl.gov/GCat/species.html>; Genome Channel: <http://grail.lsd.ornl.gov/tools/channel>).

Multiple Genome Analysis: WIT

A poster by Natalia Maltsev (Argonne National Laboratory, ANL) and colleagues described ANL's WIT system (<http://wit.mcs.anl.gov/WIT2>). WIT was designed and implemented to support genetic sequence and comparative analysis of sequenced genomes and metabolic reconstructions from sequence data. It now contains data from 34 genomes (some incomplete).

The authors believe that parallel analysis of a large number of phylogenetically diverse genomes can add much to the understanding of higher-level functional subsystems and major physiological designs. They reported a new method for using conserved clusters of genes from numerous genomes to predict functional coupling between genes. Although early results are encouraging, investigators believe the precision of prediction and the amount of accessible functional coupling will increase dramatically as more genomes are added. They emphasized that this class of data may well become a significant resource for establishing the function of hypothetical proteins, better understanding the functions of paralogous genes, and reconstructing connections in higher-level functional subsystems.◊



From left, Jeroo Kotval (University of Albany), Cynthia Needham (Microbial Literacy Collaborative), and Sara Tobin (Stanford University).

Education and Bioethics

The Ethical, Legal, and Social Issues (ELSI) components of the U.S. Human Genome Project represent the world's largest bioethics undertaking. DOE's ELSI Program focuses on genetic education, privacy, fair use of personal genetic information, and genetics and the workplace. Presentations of two ELSI grantees are reported below.

Revealing DNA Differences

David Micklos (Cold Spring Harbor Laboratory, CSHL) described a program to introduce high school biology teachers to a laboratory-based unit on human DNA polymorphisms, which can be useful in disease diagnosis, forensic identification, and other applications. In addition to teaching the science, the program provides a personal perspective that enhances discussion of the uses and potential abuses of genetic technology. [Ready-to-use teaching kits developed at CSHL to support this program are available through Carolina Biological Supply Company (800/331-5551, www.carosci.com).]

Other DOE-funded educational programs initiated at CSHL include a mitochondrial DNA sample-processing service that sequences student DNA samples and posts the data to an Internet-based sequence server (<http://vector.cshl.org/resources/bioserver.html>). Step-by-step directions allow students to use their own data to find related sequences in GenBank, compare themselves to other people, and test whether or not Neanderthal hominids were direct ancestors of modern humans. Micklos demonstrated the Bioform program (<http://vector.cshl.org/resources/bioforms>). With this program, students can analyze mitochondrial sequences to identify the remains of the last Russian czar and his family and determine whether Anna Anderson was the missing princess Anastasia, as she claimed.

Importance of Microbes

Microbes can teach people much, since all living things today evolved

from them and still share fundamental biological properties with them. Such is the message of *Intimate Strangers: Unseen Life on Earth*, a four-part science documentary developed for public television with the support of the DOE Human Genome and Microbial programs. Cynthia Needham described this and other educational efforts that make up the Microbial Literacy Collaborative, a partnership of organizations committed to advancing scientific literacy through a focus on the microbial world.

Needham reported on four components of the initiative: the science documentary, a set of 17 hands-on community-based microbial activities, a youth leadership training program for precollege students from traditionally underrepresented communities, and a 12-part telecourse for undergraduates. *Intimate Strangers* will air on many PBS stations for four consecutive Tuesdays beginning Nov. 9 at 8 p.m. (check local listings). Microbe World, an interactive Web site, was designed to build on the themes of all these activities (www.microbeworld.org).◇

Microbial Genome Explorations

Analysis of microbial genomes can provide clues to genome organization and evolution, contribute to a healthy citizenry, and offer potential solutions to long-standing challenges in renewable energy production, chemical and materials production, and environmental cleanup. To take advantage of these opportunities for fulfilling key missions, in 1994 DOE initiated its Microbial Genome Program (MGP), a spinoff of the Human Genome Program. At the Oakland Contractor-Grantee meeting, MGP researchers reported exciting progress.

“Superbug” Analysis

David Schwartz (now at University of Wisconsin, Madison), Owen White (The Institute for Genomic Research), and Kenneth Minton [now retired from Uniformed Services University of the Health Sciences (USUHS)] described mapping, sequencing, analyzing, and genetically engineering the 3-Mb genome of *D. radiodurans*. This microbe can survive radiation exposure thousands of times greater than doses that are lethal to humans. Although its chromosomes shatter into hundreds of fragments when hit with millions of rads of gamma radiation, the organism can stitch itself back together in about a day.

Scientists hope that analyzing *D. radiodurans*' genome will give some clues to its remarkable DNA-repair mechanisms and that the

microbe will be useful in cleaning up toxic mixed-waste sites around the globe. Schwartz's optical mapping of the *D. radiodurans* genome was critical to the discovery that it has four chromosomal elements rather than just one [see *Science* **285**(5433), 1558-62 (www.sciencemag.org/cgi/content/full/285/5433/1558); *HGN* **10**(1-2), 12 (www.ornl.gov/hgmis/publicat/hgn/v10n1/12deino.html)].

Schwartz described using a single-molecule approach to produce ordered restriction maps from individual DNA molecules. The technique, proven useful for producing high-resolution detailed maps of clones and entire genomes, is expected to facilitate large-scale sequencing projects. Optical mapping also generates an in situ picture of the entire genome's architecture, revealing the number of chromosomes and the existence of extrachromosomal elements. The team plans to use optical mapping to complete a human reference map that will include 10× to 15× coverage and link with other physical maps by aligning restriction-mapped BAC contigs.

White discussed results of an early survey of the *D. radiodurans* genome to determine how this organism withstands extraordinarily high levels of radiation and oxidative stress. But so far, analysis of the DNA-repair genes has turned up nothing unique that would account for the capability to knit double-stranded breaks back

together. "They're pretty much the garden variety of DNA-repair genes," White noted. A survival strategy unique to this organism is that the genome is partitioned into regions of genes that have specific functions.

Minton discussed efforts to annotate the *D. radiodurans* sequence, focusing on properties that render this organism resistant to radiation. This work has been taken over by Michael Daly (USUHS). Features noted to date include a novel enzyme that combines potential repair domains from three independent repair proteins. Minton reported on Daly's work to engineer this genome to enhance the potential for organopollutant degradation in radioactive mixed-waste environments; for example, Daly introduced genes to degrade toluene to less-dangerous substances and showed that the genes retain effectiveness even at high-radiation doses.

Minton also described Daly's engineering of heavy-metal resistance by transferring a mercury-resistance gene from *Escherichia coli* and the team's plans to transfer a uranium-reduction gene from the microbe *Shewanella putrefaciens* to *D. radiodurans*. In *Shewanella*, uranium acts as a final electron acceptor, being reduced from uranium⁺⁶ to uranium⁺⁴, which settles like a stone or

mineral and does not enter the groundwater.

Archaeal Proteomics

Carol Giometti (Argonne National Laboratory) described the Archaeal Proteomics Project, whose goal is to identify proteins and regulatory pathways relevant to bioremediation and energy technology. She explained that proteomics includes information on relative protein abundance, post-translational modifications, changes in stimuli-response kinetics, and subcellular location. Important proteomics tools include two-dimensional gel electrophoresis (2-DGE) and mass spectrometry.

Initial work is focused on the *Pyrococcus furiosus* and *Methanococcus jannaschii* proteomes (both genomes have been sequenced completely). Both are hyperthermophilic archaea with growth temperatures near 100°C and enzymatic capabilities that have promise for bioremediation, energy conversion, and chemical-processing systems. Investigators are using 2-DGE to purify and quantitate proteins expressed in archaea that are grown under a variety of conditions designed to modulate specific metabolic pathways. Giometti discussed preliminary results, which provide a foundation

DOE Sets Workshop For February 2000

The eighth DOE Human Genome Program Contractor-Grantee workshop will be held February 27– March 2, 2000, in Santa Fe, New Mexico. At least one investigator from each funded project is expected to attend the entire meeting and represent the project at poster sessions. Speakers for platform presentations will be notified by January 18. Abstracts should be submitted through the Web site (<http://lcbcg.lbl.gov>, click on Meetings).

- **Abstract deadline:** December 10
- **Contact:** Leonora Castro (510/486-5874, Fax: -5717, licastro@lbl.gov) ◇

for studying the *M. jannaschii* and *P. furiosus* proteomes.

Comparative Genomics

Projects have begun to leverage microbial sequence information to sequence other strains more rapidly. Gary Andersen (Lawrence Livermore National Laboratory, LLNL) spoke about using suppressive subtractive hybridization to identify genomic differences among enteropathogenic strains of *Yersinia enterocolitica* and *Y. pseudotuberculosis*. The technique uses PCR amplification to enrich for unique segments of restricted DNA and simultaneously limits nontarget amplification by suppression PCR. Of the two pathogens, *Y. enterocolitica* is more often associated with human infection.

These explorations are likely to reveal unique DNA regions that define the genetic basis for the underlying differences in their phenotypic variation. Streamlining, automating, and increasing this technique's throughput should enable large-scale genomic comparison among closely related strains and generation of strain-specific oligonucleotide probes for molecular epidemiology studies. Ultimately, this technology could enable the much more rapid determination of closely related microbial sequences based on completed reference strain sequences. ◇

In the News

Drosophila Sequencing Nears Completion

In September researchers at Celera Genomics announced that they had obtained the raw sequence for the 140-Mb euchromatic region of *Drosophila melanogaster's* 180-Mb genome (www.celera.com). The Celera group is collaborating with the Berkeley *Drosophila* Genome Project (BDGP) consortium, directed by Gerald Rubin, to deliver a completely finished sequence map of the *Drosophila* genome by the end of the year (www.fruitfly.org).

Celera plans to begin making the sequence data available to the public in October, and publication in collaboration with BDGP is expected early in 2000. Because the human and fruitfly genomes share many similarities, the finished sequence should provide an important key to understanding human biology.

Celera researchers used high-throughput machines and a whole-genome shotgun strategy (average coverage, 10×) to decode sequence from many small, random DNA fragments. BDGP provided 26.5 Mb of completed genome sequence and a low-coverage (~1.5×) "scaffold" shotgun sequence of each BAC and P1 clone containing the fragmented DNA. This scaffold map will assist researchers in assembling fragments and finishing gaps.

BDGP is a consortium of scientists at the University of California, Berkeley; Lawrence Berkeley National Laboratory; Baylor College of Medicine; and Carnegie Institution of Washington. It is funded by NIH, DOE, and the Howard Hughes Medical Institute. ◇

Human Genome Project Directors, Researchers Receive Awards

Patrinos, Others Honored by Smithsonian, Platinum Technology

Ari Patrinos, head of DOE's Human Genome Program, received the 21st Century Pioneer Award in June along with two of his predecessors, Charles DeLisi (Boston University) and David Galas (Keck Graduate Institute). Other recipients were Francis Collins, director of the NIH National Human Genome Research Institute, and representatives from the Wellcome Trust and the Institute of Medical Science at the University of Tokyo. This new award is presented by the Smithsonian Institution and Platinum Technology to individuals who have demonstrated vision and leadership as they strive to use information technology in innovative ways.

LBNL's Bissell Cited for Pioneering Cancer Research

In April, the American Association of Cancer Research (AACR) presented the G.H.A. Clowes Memorial Award to Mina Bissell, Life Sciences director at Lawrence Berkeley National Laboratory. Bissell was recognized for her research demonstrating that the extracellular matrix plays a vital role

in gene expression and thus bears significantly on cell growth, functional differentiation, programmed cell death, and cancer. The award was established by the Eli Lilly Company to honor one of its research directors who also was a founding member of AACR. [For more information on Bissell's work, go to www.lbl.gov/Science-Articles, search on "Bissell."]

ORNL's Vo-Dinh Receives Sixth R&D 100 Award

Researcher Tuan Vo-Dinh, group leader and Corporate Fellow at Oak Ridge National Laboratory, won a sixth prestigious R&D 100 Award and the Editors' Award for Most Promising New Technology for 1999. Vo-Dinh's multifunctional biochip for rapid screening and detection of pathogens and diseases was featured in this year's PBS series, *Frontiers of Medicine*.

Integrating microelectronics and biotechnology in a single system, the chip contains different types of bio-probes that allow diagnosis of multiple diseases. When fully developed, the technology will provide test results at the doctor's office for AIDS, cancer, tuberculosis, and other diseases and illnesses. ◊

Human MHC Region Sequenced

Key to Transplant Rejection, Autoimmune Disease

Sequencing of the 4-Mb human Major Histocompatibility Complex (MHC) region of chromosome 6 has been completed by the Sanger Centre, University of Washington, and Tokai University (www.sanger.ac.uk/HGP/Chr6/MHC.shtml).

Proteins encoded by genes residing in the MHC region are responsible for helping the body defend itself against microscopic invaders by distinguishing normal body constituents ("self") from everything else, which it then marks for extinction. Researchers hope that a better understanding of MHC proteins will lead to the development of new ways to minimize transplant rejection and fight such infectious and autoimmune diseases as arthritis and juvenile diabetes. ◊

Genetic Testing Advisory Committee

The Secretary's Advisory Committee on Genetic Testing (SACGT) of the Department of Health and Human Services (DHHS) held its first meeting in June. Chaired by Edward McCabe (University of California, Los Angeles), the 13-member committee's task is to help DHHS formulate policies on the development, validation, and regulation of genetic tests, particularly DNA-based diagnostics (www.nih.gov/od/orda/sacgt/docs.htm). SACGT was formed last year on the recommendation of the NIH-DOE Task Force on Genetic Testing.

At the meeting, the committee set up working groups to address the adequacy of governmental oversight of genetic tests and to plan how to gather public perspectives. Members also identified nine broad areas that will require the committee's consideration over the next 2 years. These include education, access to testing, diversity issues, stigmatization, rare disorders, introduction of tests into clinical practice, use of evidentiary-based models and outcome assessments, economic issues in genetic testing and oversight, and direct marketing of tests. ◊

New Mouse Probes Aid Gene Mapping

A team led by Julie Korenberg (Cedars-Sinai Medical Center, Los Angeles) has produced a family of DNA probes that "light up" mouse chromosome sites under fluorescence microscopy. The markers, spaced at an average 19 million bases throughout the 3 billion bases of the mouse genome, aid researchers who are hunting for mouse counterparts of human genes. The work is described in the May 1999 cover article of *Genome Research*. The mouse is a favorite research model for understanding human biology, and researchers are finding more and more corresponding genes on human and mouse chromosomes. Unlike human chromosomes, however, mouse chromosomes have few distinguishing features that help cytogeneticists hunt for particular disease genes. The new markers thus far have led researchers to the mouse counterpart of the human Down's syndrome region.

The DNA resource consists of 157 BAC clones, each an identifier of specific bands or band borders, with 42 linked to genetic markers from the centromeric and telomeric ends of the Whitehead-MIT recombinational maps. Inclusion of BAC clones containing markers from the ends of genetic maps is expected to facilitate development of an integrated view of mouse cytogenetic, genetic, and physical maps. ◊

Outstanding Research Makes DOE National Laboratories Major Asset in Biomedical Engineering

Biomedical Engineering Research at DOE National Labs, *DOE Office of Biological and Environmental Research (OBER)*. Report DOE/SC-1999-1, 106 pp., paperback. The text below was adapted by HGMIS from the report preface by Michael Viola, director of the DOE OBER Medical Sciences Division.

Biomedical engineering—the application of physics, chemistry, and engineering principles to problems of human health—has been a critical part of work at the DOE national laboratories since 1947.

The Atomic Energy Commission's first director of biology and medicine,

Shields Warren, had the vision to develop a formal research program outside the restricted scope of industrial health and safety. The national laboratories were recognized immediately as a source of medically important radioisotopes for use both as research reagents and as potential weapons against cancer. By the end of 1947,

almost 2000 radioisotope deliveries had been made to laboratories and hospitals.

During the next decade, cyclotrons and reactors generated radioisotopic tracers under controlled conditions. The new field of radiopharmacy (attaching radionuclides to biologically active molecules and studying their activity) was greatly advanced by work in the national laboratories, where instrumentation originated to detect the cellular and total body distribution of these new radiotracers. The results have been spectacular—all major hospitals in the world rely heavily on technologies developed, in part, in the DOE national laboratories.

Their talented scientists and extraordinary technologies have made the national laboratories a major asset for biomedical engineering, and their

specialized resources traditionally have led to multidisciplinary science projects and enabled science often not possible at universities or industry. During the past two decades, bioengineering programs have been beneficiaries of rapid advances in nuclear physics, nuclear engineering, nuclear

chemistry, and molecular biology. Research on the medical applications of such technologies as synchrotron light sources, lasers, mass spectrometry, high-field magnets, micro-fabricated machines, biosensors, and DNA chips is ongoing in many of the laboratories.

This January 1999 inventory of

biomedical engineering projects (see box, right, for examples), supported in the national laboratories by different offices of DOE and with discretionary lab funds, will be updated yearly. The extraordinary breadth and depth of projects, investigators, and areas of expertise should stimulate networking and collaboration among DOE scientists and those in universities and industry.

[Print copies: Sharon Betson (301/903-3213, sharon.betson@science.doe.gov); download from Web: <http://apollo.osti.gov/sc73/doe-sc-1999-1.pdf>]

[For more information on DOE's role in the development of medical imaging, see "A Healthy Citizenry," beginning on p. 20 in the downloadable PDF version of *A Vital Legacy* (www.ornl.gov/hgmis/publications.html#vital).]

Some Projects Described In OBER Report*

- Advanced Lower-Limb Prosthetics (Morton Lieberman, SNL)
- Alzheimer's Disease as a Systemic Disorder (Thomas Budinger, LBNL)
- Biochip for Medical Diagnosis (Tuan Vo-Dinh, ORNL)
- Carbon-Based Prosthetic Devices (David Devlin, LANL)
- Clot Removal by Laser (R.P. Godwin, LANL)
- Diagnostics for Lyme Borreliosis (John Dunn, BNL)
- Disinfection of Drinking Water (Paul Weaver, National Renewable Energy Laboratory)
- Dissolvable Stents (William Samuels, PNNL)
- Electronic Nose on a Chip (R.J. Warmack, ORNL)
- Genetic Reader (Edward Yeung, Ames)
- Human Skin Cancer Prevention (B. Sutherland, BNL)
- Noninvasive Biomedical Monitoring (David Haaland, SNL)
- Noninvasive Measurement of Intracranial Pressure (William Wray, LANL)
- Oxygen Generation for Medical Applications (Timothy Armstrong, PNNL)
- Positron Emission Tomography (PET) in Cocaine Abuse (N.D. Volkow, BNL)
- PET in Schizophrenia (S. Dewey, BNL)
- PET Imaging of Estrogen Metabolism in Breast Cancer (Y-S. Ding, BNL)
- Radiation Therapy for Pediatric Brain Tumors (F.A. Dilmanian, BNL)
- Rapid Identification of Microbial Species (Paul Jackson, LANL)
- Sex Differences of Dopamine Metabolism (Thomas Nordahl, LBNL)
- Silicon Microgripper for Surgeons (Abe Lee, LLNL)
- Tissue Welding Using Laser Energy (Duncan Maitland, LLNL)
- Ultrasound Mammography (Jim Candy, LLNL)

*An acronym list is on p. 20. ◇

Major Drug Firms Create Public SNP Resource

In April, ten large pharmaceutical companies and the U.K. Wellcome Trust philanthropy announced the establishment of a consortium headed by Arthur L. Holden to find and map 300,000 common DNA sequence variations. The goal is to generate a widely accepted, high-quality, extensive, publicly available map using single-nucleotide polymorphisms (SNPs) as markers evenly distributed throughout the human genome. SNPs can occur in both coding (gene) and noncoding regions.

SNPs may help scientists identify small genetic differences that could predispose people to disease or influence their response to a drug. A SNP map thus may be of great value for biomedical research and for developing pharmaceutical products or medical diagnostics. Also, the map is expected to simplify navigation of the much larger genome map being generated by researchers in the Human Genome Project (HGP).

Consortium Members and Laboratories

The international member companies, which together are committing at least \$30 million, are Bayer Group AG, Bristol-Myers Squibb Co., Glaxo Wellcome PLC, Hoechst Marion Roussel AG, Monsanto Co., Novartis AG, Pfizer Inc., Roche Holding Ltd., SmithKline Beecham PLC, and Zeneca Group PLC. The Wellcome Trust is contributing at least \$14 million.

The laboratories searching for SNPs are located at the Whitehead Institute, Sanger Centre, Washington University (St. Louis), and Stanford University. Data management and analysis will take place at Cold Spring Harbor Laboratory, which will scan for sequence matches in public databases to help determine a SNP's genomic location. Other research sites will use various processes. Expectations are to have 150,000 SNPs mapped by mid-2001.

Industry-Sponsored Public Map

The SNP consortium views its map as a way to make available an important, precompetitive, high-quality research tool that will spark innovative work throughout the research and

industrial communities. Several groups (including some in the HGP) are working to find SNPs, but the likelihood of duplication is small because of the estimated 3 million SNPs in the human genome, and the potential payoff is high.

Because the value of SNP technology is not yet proven for finding subtle genetic differences related to disease and pharmaceutical development, the creation of a joint map distributes the financial risk. If widely accepted, the joint map could serve as an important standard for the U.S. Food and Drug Administration and other regulatory agencies.

DNA Resources

The SNP consortium will use DNA resources from a pool of samples

obtained from 24 individuals representing several racial groups. This is a subset of the DNA reference panel for SNP identification being collected by the NIH National Human Genome Research Institute. The anonymous, voluntary DNA contributions are made with informed consent specifically for this use.

Data Availability

SNP data will be made available through a consortium Web site (<http://snp.cshl.org>) at quarterly intervals during the project's first year and at monthly intervals during the second year. SNPs also will be deposited in the public dbSNP database (www.ncbi.nlm.nih.gov/SNP). ◇

SBIR 1999 Human Genome Awards Announced

The DOE Office of Biological and Environmental Research has announced six Phase I and three Phase II awards for 1999 in genome, structural biology, and related technologies topics of the Small Business Innovation Research (SBIR) program. The highly competitive SBIR awards are designed to stimulate commercialization of federally funded research and development for the benefit of both the private and public sectors. SBIR emphasizes cutting-edge, high-risk research with potential for high payoff in hundreds of areas, including human genome research (contacts: p. 19).

Phase I Awards

- **Atom Sciences, Inc.** (Oak Ridge, Tennessee; Tom Whitaker): DNA Diagnostics Using Electrical Detection
- **CombiMatrix Corp.** (Burlingame, California; Francis Rossi): Microarrays of Affinity Probes for the Analysis of Gene Products
- **CyberConnect Corp.** (Storrs, Connecticut; Wally Grajewski): A Visual Data-Flow Editor Capable of Integrating Data Analysis and Database Querying
- **Genome Informatics Corp.** (Oak Ridge, Tennessee; Doug Hyatt): Commercialization of the GRAIL EXP Gene-Discovery System
- **Physical Optics Corp.** (Torrance, California; Tin Aye): Versatile Liquid Crystal Tunable Interference Filter for Chromosome Analysis
- **SymBiotech, Inc.** (Wallingford, Connecticut; Edward Davis): Fluorescent-Based, High-Throughput Protein Kinase and Phosphatase Assays

Phase II Awards

- **Atom Sciences, Inc.** (Oak Ridge, Tennessee; Tom Whitaker): A Quantitative Analytical Tool for Producing DNA-Based Diagnostic Arrays
- **Fidelity Systems, Inc.** (Gaithersburg, Maryland; Sergei Kozyavkin): D-Strap DNA Sequencing Chemistry
- **MacConnell Research Corp.** (San Diego, California; William MacConnell): Automated Purification of Blood or Bacterial Genomic DNA ◇

Articles on Genetics

A collection of articles on genetics appears in the July 24 issue of the British medical journal *The Lancet*, supplement on Molecular Medicine **354**(1). The full text is on the Web (www.thelancet.com/newlancet/sub/supplements/vol354s1/menu_NOD59.html). ◇

Microbial Genomics

EcoCyc Database for *E. coli*

The EcoCyc electronic database is a literature-derived resource that describes the genome and biochemical machinery of *Escherichia coli*. The database contains up-to-date annotations and the DNA sequences of all genes in *E. coli* and describes all known pathways of its small-molecule metabolism. Each pathway and its component reactions and enzymes are annotated in rich detail, with extensive references to the biomedical literature. For example, the detail provided for each *E. coli* enzyme includes its cofactors, activators, inhibitors, and subunit structure.

In July, the database included 159 metabolic pathways, 946 reactions, 629 enzymes, and 4390 genes. EcoCyc also is used by investigators to annotate other microbial genomes. Because *E. coli* has the largest fraction of gene products whose functions were determined experimentally, sequence-similarity matches to *E. coli* are less likely to result in incorrect function predictions than are matches to other microbial genomes. Those genomes may have a higher rate of annotation errors due to computational, and perhaps transitive, misannotation.

EcoCyc Features

EcoCyc's Pathway Tools software provides a powerful environment for creating, managing, and publishing pathway and genome databases (DBs) on the Web. One of the components is the Pathway Genome Navigator, which allows users to query these DBs and to visualize and compare the resulting pathways, genes, genome maps, reactions, and enzymes. The PathoLogic program computationally predicts an organism's metabolic-pathway complement and creates a DB describing that prediction. A set of graphical tools allows users to edit the annotations interactively.

Future Directions

The EcoCyc project is now moving beyond *E. coli* metabolic pathways to include signal-transduction pathways, transport proteins, regulation of gene expression, and tRNAs. Version 5.0, released in June, contains detailed annotations of *E. coli* phosphotransferase-system transporters authored by collaborators Milton Saier and Ian Paulsen (University of California, San Diego). In addition, Julio Collado (Universidad Nacional Autonoma de Mexico) is adding descriptions of *E. coli* gene regulation, including operons, promoters, and DNA-binding proteins. Once the regulatory mechanisms are added to EcoCyc, researchers will be able to compare known mechanisms of *E. coli* gene regulation with microarray-derived gene-expression data.

Pangea Systems, a bioinformatics company in Oakland, California, makes EcoCyc available free to the academic community and for a fee to commercial

Developing EcoCyc

The EcoCyc project was conceived in 1990 by Peter Karp (then at SRI International) and Monica Riley (Marine Biological Laboratory) to provide a central, integrative resource for the quickly changing knowledge about *E. coli*'s genes and metabolism. Karp's group developed Pathway Tools, a software environment that included but went beyond genome sequence data by integrating richly curated information on gene function and pathways. Riley's group searched the biomedical literature for material on *E. coli* enzymes and pathways. An early task was to determine what to gather about each enzyme and pathway and to design a database schema that would accommodate that information. Riley published a system for categorizing *E. coli* genes according to their function; this and other similar systems have since been used by genome project researchers to classify their annotated genes. Kenneth Rudd (National Center for Biotechnology Information) contributed his collection of all known *E. coli* DNA sequences and assembled them into an ordered set of maximal sequences that were pinned to the Kohara physical restriction map of the *E. coli* chromosome. After Fred Blattner's group at the University of Wisconsin completed *E. coli*'s DNA sequence, the data were integrated into EcoCyc.◇

organizations. The company is interested in collaborating with academic genome centers to use the Pathway Tools for curation and Web publishing of their genomes. [Peter Karp, pkarp@pangeasystems.com; Monica Riley, mriley@mbl.edu; EcoCyc, <http://ecocyc.pangeasystems.com/ecocyc>] ◇

Human
Genome
news

This newsletter is intended to facilitate communication and collaboration, help prevent duplication of research effort, and inform persons interested in genome research. Views expressed are not necessarily those of the Department of Energy Office of Biological and Environmental Research. Suggestions are invited.

Human Genome Management
Information System (HGMIS)

Oak Ridge National Laboratory
1060 Commerce Park, MS 6480
Oak Ridge, TN 37830
423/576-6669, Fax: /574-9888
www.ornl.gov/hgmis

Managing Editor
Betty K. Mansfield
BKQ@ornl.gov

Production Assistants
Marissa D. Mills
Laura N. Yust

**Editors/Writers/
Designers**

Anne E. Adamson
Denise K. Casey
Sheryl A. Martin
Judy M. Wyrick

**U.S. Department of Energy Office of
Biological and Environmental Research**

Ari Patrinos, Associate Director
www.er.doe.gov/production/ober/ober_top.html

Life Sciences Division, OBER

Marvin E. Frazier, Director
www.er.doe.gov/production/ober/hug_top.html

Contact: Daniel W. Drell, 301/903-6488, Fax: -8521
daniel.drell@science.doe.gov or genome@science.doe.gov

This newsletter is prepared at the request of the DOE Office of Biological and Environmental Research by the Toxicology and Risk Analysis Section of the Life Sciences Division at Oak Ridge National Laboratory, which is managed by Lockheed Martin Energy Research Corp. for the U.S. Department of Energy, under Contract DE-AC05-96OR22464.◇

Microbial Genome News
Thermotoga Sequence
PresentedEarly Analysis Suggests Gene
Exchange Between Kingdoms

In May, researchers led by Karen Nelson (The Institute for Genomic Research) reported obtaining the complete 1.8-Mb genomic sequence of the heat-loving bacterium *Thermotoga maritima*, first isolated from

Microbial Genome News

geothermally heated marine sediment in Vulcano, Italy. Early analysis reveals some unusual features that could affect our understanding of how earth's simplest life forms evolved.

The three major life groups or kingdoms are eubacteria and archaea, which include the simplest life forms lacking a central nucleus; and the more complex eucaryota, which include animals and plants. *T. maritima* has been considered one of the deepest and most slowly evolving lineages in the eubacteria kingdom.

Authors of the May 27 article in *Nature* (399, 323-29) reported that almost a quarter of *T. maritima*'s genes are similar to those found in archaea, with 81 archaeal-style genes clustered in 15 genomic regions. These results pose new questions about defining organisms that have mosaic-like genomes, with features shared across two domains. The authors note that these findings do not necessarily reflect a closely shared common ancestor but could point instead to lateral gene transfers.

Because of growing evidence for a high frequency of gene transfers and a lack of agreement in different phylogenetic (evolutionary) analyses on individual genes, the authors suggest that sequence comparisons of individual genes may be inaccurate indicators of organismal evolution. Relationships among the eubacteria and archaea will be understood better as other microbial genomes are fully sequenced and analyzed. [For data and requests for *Nature* reprints, see www.tigr.org/tdb/CMR/btm/htmls/SplashPage.html]

Genome of Historic Microbe Sequenced

A team headed by Douglas Smith (Genome Therapeutics, GTC) has finished sequencing the 4.1-Mb genome of *Clostridium acetobutylicum* and has placed the data in GenBank and on the Web (www.cric.com/genesequences/clostridium/clospage.html).

C. acetobutylicum, a nonpathogenic microbe that can convert starch into the solvents acetone and butanol, enjoys an unusual place in history. Discovered in 1915 by Chaim Weizmann, the microbe was used by Great Britain during World War I for generating

acetone to produce cordite for artillery shells. In gratitude, the government offered to honor Weizmann, but he asked instead for British support of a Jewish homeland in Palestine. This led to the Balfour Declaration of 1917, committing Britain to sanction what became in 1948 the state of Israel, with Weizmann as its first president.

GTC previously sequenced *Methanococcus thermoautotrophicum*, one of the first three microbes (along with *Pyrococcus furiosus* and *M. jannaschii*) targeted by the DOE Microbial Genome Program (MGP). As of September, ten microbial genomes had been sequenced completely with the support of MGP (www.er.doe.gov/production/ober/microbial.html).

The growing body of microbial sequences, accumulating rapidly in public databases, has advanced the field of comparative genomics and is leading to new insights into how genomes change over time. Scientists also are learning how microbes cause disease and contribute to natural environmental and biological processes. Studies may suggest ways to harness microbes' abilities for bioremediation, production of industrial chemicals and enzymes, development of antibiotics and other pharmaceuticals, and many other uses.◇

Microbial Gene Finder

Gene Locator and Interpolated Markov Modeler (Glimmer) finds genes quickly in microbial DNA, especially bacterial and archaeal genomes, by identifying and distinguishing coding from noncoding DNA regions. Glimmer is the primary gene finder at The Institute for Genomic Research (TIGR), where it has been used to annotate the complete genomes of both TIGR and non-TIGR projects. The home page contains statistics on Glimmer's accuracy (www.tigr.org/softlab/glimmer/glimmer.html).◇

Neisseria Sequence

The Sanger Centre has announced the completion of the genome sequence of the 2.23-Mb *Neisseria meningitidis* serogroup A strain Z2491 (www.sanger.ac.uk/Projects/N_meningitidis).◇

Microbial Web Sites

Genomes and Genome Projects

- <http://geta.life.uiuc.edu/~nikos/genomes.html>
- www.tigr.org/tdb/mdb/mdb.html
- www.beowulf.org.uk
- www.er.doe.gov/production/ober/EPR/mig_top.html
- www.sanger.ac.uk/Projects/Microbes
- <http://grail.lsd.ornl.gov/tools/channel>
- www3.ncbi.nlm.nih.gov/Entrez/Genome/org.html
- www3.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html
- www.ncgr.org/microbe (archive only)
- www.fp.mcs.anl.gov/~gaasterland/genomes.html (archive only)

Discussion Group

- www.medmicro.mds.qmw.ac.uk/microbial-genomes

Microbial Information Broker

- <http://mol.genes.nig.ac.jp/gib>

Metabolic Pathways

- **EcoCyc:** <http://ecocyc.pangeasystems.com/ecocyc>
- **KEGG** (Kyoto Encyclopedia of Genes and Genomes): www.genome.ad.jp/kegg
- **Metabolic Pathways on Internet:** <http://home.wxs.nl/~pvsanten/mmp/mmp.html>
- **WIT:** <http://wit.mcs.anl.gov/WIT2>

HUGO News



Moves London Office

The London office of the Human Genome Organisation (HUGO) has moved and changed its e-mail address. Telephone and fax numbers remain the same.

HUGO
144 Harley St.
London W1N 1AH, U.K.
+44/171-935-8085, Fax: -8341
hugo@hugo-international.org
www.hugo-international.org/hugo

Recruits Editors

HUGO is continuing to recruit chromosome editors for the Genome Database, now at the Hospital for Sick Children in Toronto, Canada. Editors are elected by HUGO's Human Gene Mapping Committee. [Contact: David Callen (dcallen@medicine.adelaide.edu.au)] ◇

Education and Counseling Foundation Established

The newly established Foundation for Genetic Education and Counseling will develop and promote genetic literacy about complex human diseases for the general public and healthcare professionals. The first educational programs will target individuals affected by or at risk for schizophrenia or bipolar disorder, and genetic counseling services will be developed for those who have participated in genetic-research protocols.

Ann Pulver (Johns Hopkins School of Medicine) is foundation president,

and initial funding is from Genset, a French biotechnology company engaged in human genome analysis.

Daily operations of the foundation are the responsibility of Joseph McInerney, a longtime grantee of the DOE Human Genome Program and former director of the Biological Sciences Curriculum Study (BSCS) in Colorado Springs, Colorado. Roger Bybee succeeded McInerney as BSCS director.

[Contact: McInerney, 410/955-0455, joemcinerney@genetic-medicine.org] ◇

¶ ELSI Retrospective

DOE ELSI Program Emphasizes Education, Privacy: A Retrospective (1990-99), Daniel Drell of DOE and Anne Adamson of HGMIS. This comprehensive review of research projects funded by the ELSI component of the DOE Human Genome Program is available in print and on the Web (www.ornl.gov/hgmis/resource/elsi.html#products). The retrospective discusses goals and provides such specifics as resource availability, Web sites, and contacts. It is accompanied by a list of products generated by DOE ELSI projects for the same period. For a print copy, see HGMIS address on p. 14.◇

✎ Collection Typifies Ethnic Populations

The Human DNA Repository at the National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University has a large DNA collection from healthy unrelated individuals and small families. The collection represents Israel's Jewish and Arab populations. Of more than 1400 cell lines established, over 3000 DNA samples have been distributed to researchers in the United States, Canada, Europe, and Japan. [More information: www.tau.ac.il/medicinel/NLGIPI/nlgip.htm] ◇

¶ ELSI Studies

The following two research studies, supported by the Ethical, Legal, and Social Issues component of the DOE Human Genome Program, are available on the Web.

Report on Family Attitudes

Pathways and Barriers to Genetic Testing and Screening: Molecular Genetics Meets the High-Risk Family (1997), Troy Duster (University of California, Berkeley) and Diane Beeson (California State University, Hayward). Now available to the public, the study's major finding was that all high-risk families, regardless of their cultural backgrounds, adapt genetic information they are given to fit the divergent values and priorities of family life. Sickle cell disease, cystic fibrosis, and thalassemia were

Informatics

MGI 2.2 Released

The Mouse Genome Informatics (MGI) Web site is the home of the Mouse Genome Database (MGD), Gene Expression Database (GXD), and other resources at the Jackson Laboratory (www.informatics.jax.org). MGI provides a comprehensive and growing source for information on the biology and genetics of the laboratory mouse. Release 2.2 includes the following changes and additions.

- With the exception of the Home Page, the names and locations of all files in the MGI Web site have been reorganized to facilitate mirror-site installation and maintenance.
- On the Home Page, "Quick Gene Search" enables searching for mouse genes using a single or partial gene symbol or name.
- All MGI pages contain a "Search Forms" menu for quick access to query forms.
- The marker-detail record now lists sequence database accession IDs with hyperlinks to DDBJ, EMBL, GSDB, and GenBank.
- New "Mouse Facts" pages present data from various sources, including a mouse sequencing-progress summary; genome-length estimates; MGD chromosomal length and gene distribution; mouse physiology information; and links to other sites of interest to researchers.
- The MGI Searches, Data, and Reports section incorporates such information as Strains and Polymorphisms, Gene Expression, and Maps and Mapping Data.
- Several data sets contributed by various authors are available via ftp (<ftp://informatics.jax.org/pub/informatics/datasets/index.html>).

Submitting Data to MGI

MGI encourages contributions of electronic data sets from the scientific community. For data-submission guidelines, visit the Web site (www.informatics.jax.org/support/submissions/submit.shtml) or send an e-mail message (submissions@informatics.jax.org). Use the following site to submit a proposed locus or gene symbol (www.informatics.jax.org/support/nomen).◇

chosen for investigation because they are found primarily in different ethnic and racial groups. To obtain a copy of the report, contact Janice Tanigawa; Institute for the Study of Social Change; 2420 Bowditch St.; University of California; Berkeley, CA 94720 (510/642-0813, Fax: -8674).◇

Trottier Report on Web

Public Sector Genetic Services in Florida and Georgia: Current Status and Potential Issues Raised by the Human Genome Project (1996), Ralph Trottier (Morehouse University) and Lee Crandall (University of Illinois, Champagne). Now on the Web, the study explores the effects of changing science and technology on public genetic services in two states (www.ornl.gov/hgmis/resource/trottier.html).◇

Resources

New Staden Package Released

Release 1999.0 of the Staden Package from the MRC Laboratory of Molecular Biology (Cambridge, U.K.) incorporates the new PREGAP4 program, which has a graphical user interface and a batch mode (www.mrc-lmb.cam.ac.uk/pubseq). Used to prepare sequences for assembly, PREGAP4 includes trace file conversion, vector and quality clipping, contaminant screening, repeat masking, and batch-mode assembly via such programs as Phrap, FAKII, CAP3, and GAP4. New algorithms increase VECTOR_CLIP's sensitivity and flexibility, and the new SCREEN_SEQ program filters readings contaminated with *Escherichia coli* or yeast.

Besides bug fixes, the GAP4 program has many additions. The trace display includes a multicolumn format, which by use of a vertical scrollbar can show any number of traces. The "auto-display-traces" mode works for more search types, and the search for contig joins is more sensitive. The maximum reading length is increased from 4096 to 30,000 bases, and a "notes" data type stores notebook-style comments about readings, contigs, and databases. The contig editor displays the confidence for the bases in the readings and in the consensus. "Sort Matches" is applied automatically for several commands, and matches are processed in order of their significance. GAP4 includes interfaces to Phrap, FAKII, and CAP3. NIP4 and SIP4 provide improved sequence library browsing using SRS indexes.

The CD-ROM containing package versions for Solaris, Digital Unix, IRIX, and Linux is free to academic sites. Courses on using the package are announced on the Web site.◊

Some Biotechnology Business Web Sites

- <http://ldnapatent.com>—DNA patenting and biotechnology for lay audiences
- www.signalsmag.com—Online magazine of biotechnology industry analysis
- www.pharminfo.com—Pharmaceutical information network
- <http://biomedicine.about.com>—Biotech and biomedical industry
- http://biz.yahoo.com/research/indgrp/med_biomed_gene.html—Medical and biomedical research by industry ◊

PDB Newsletter

Initiated in January of this year, *PDB Newsletter* is published quarterly by the Research Collaboratory for Structural Bioinformatics (RCSB), which operates the Protein Data Bank (www.rcsb.org). The newsletter, available on the Web, highlights the RCSB system and its features as well as future plans, collaborations, and projects (www.rcsb.org/pdb/newsletter).◊

Human Polymorphism Database

HGBASE is a public database of human intragenic polymorphisms. It is designed to contain all types of sequence variations, especially SNPs, found in normal individuals (<http://hgbase.interactiva.de>). Because many of these polymorphisms are likely to influence phenotypes, HGBASE is expected to be useful in the design of association studies and similar analyses.◊

Bacterial Protein Sequence Database

The Homologous Bacterial Genes Database (HOBACGEN) contains all the protein sequences of bacteria organized into families (<http://pbil.univ-lyon1.fr/databases/hobacgen.html>). Users can select sets of homologous genes from bacterial species and visualize multiple alignments and phylogenetic trees. These capabilities make HOBACGEN useful for comparative genomics, phylogeny, and molecular evolution studies on bacteria.◊

DOE BER Program, Symposium Publications

Fiftieth Anniversary Symposium

Serving Science and Society into the New Millennium was a 1997 symposium sponsored by the U.S. Department of Energy and the National Research Council to celebrate the 50th anniversary of DOE's Biological and Environmental Research (BER) program. Three publications growing out of this meeting are available.

- Color booklet profiling 13 scientists recognized with Exceptional Service Awards (39 pp., paper and electronic). 1999. [Print copies: HGMIS*; online: www.ornl.gov/hgmis/publicat/berawards]
- Color brochure highlighting some major BER achievements (14 pp., paper and electronic). 1999. [Print copies: HGMIS*; online: www.er.doe.gov/production/ober/ber50.html]
- Meeting proceedings, including transcripts of plenary talks (156 pp., paper). 1998. [National Academy Press: 888/624-6242 or 202/334-3313, Fax: -2451, amerchan@nas.edu, www.nap.edu/bookstore (search on "Serving Science and Society")]

Research Program

A Vital Legacy, an attractive 50-year progress report on the revolutionary program that gave rise to the Human Genome Project, is in PDF format on the Web (www.ornl.gov/hgmis/publicat/publications.html#vital). Suitable for both lay and scientific audiences. [Print copies: HGMIS.*]

*See p. 14 for address.◊



Human Genome Project Information Web Site

This Web site is devoted to Human Genome Project information, related research, and societal implications.

Home Page

www.ornl.gov/hgmis

Benefits

www.ornl.gov/hgmis/project/benefits.html
(links to downloadable poster on Human Genome Project benefits)

Goals

www.ornl.gov/hgmis/hg5yp

Progress

www.ornl.gov/hgmis/project/progress.html

Frequently Asked Questions

www.ornl.gov/hgmis/faq/faqs1.html

Topical Fact Sheets

www.ornl.gov/hgmis/home.html#topics

Chromosome Launchpad

www.ornl.gov/hgmis/launchpad

Medicine and the New Genetics

www.ornl.gov/hgmis/resource/medicine.html

Education

www.ornl.gov/hgmis/resource/education.html

Ethical, Legal, and Social Issues

www.ornl.gov/hgmis/resource/elsi.html

Research

www.ornl.gov/hgmis/research.html

Meetings

www.ornl.gov/meetings

Publications

www.ornl.gov/hgmis/publicat/publications.html

Images

www.ornl.gov/hgmis/resource/images.html

Calendar of Genome and Biotechnology Meetings*

More comprehensive lists of genome-related meetings and organizations offering training are available on the Web (www.ornl.gov/hgmis) or from HGMIS (see p. 14 for contact information).

November 1999

1-7. DNA Repair and Mutagenesis; Hilton Head, SC [ASM, 202/942-9248, Fax: -9340; MeetingsInfo@asmusa.org; www.asmusa.org/mtgsrsrc/mtgs.htm]

2-5. Chips to Hits '99: Harnessing Power of Microtechnol.; Berkeley, CA [IBC, 508/481-6400, Fax: -7911; reg@ibcusa.com; www.ibcusa.com]

7-9. Gene Therapy: Delivering the Medicines of the 21st Century; Washington, DC [BioEdge, 402/996-9185, info@bioedge.net; www.bioedge.net]

8-9. Disease Biomarkers: Genetic and Proteomic Approaches; Baltimore [CHI, 617/630-1300, Fax: -1325; chi@healthtech.com; www.healthtech.com]

10. From Mad Cows to "Psychotic" Yeast: A New Paradigm in Genetics; TIGR/NRC/DOE/NASA Speaker Series: Susan Lindquist (Univ. Chicago); Washington, DC [D. Hawkins, 301/838-3501, Fax: -0209; dhawkins@tigr.org; www.tigr.org]

11-14. In Silico Biology: Sequence, Structure, and Function. 2nd Georgia Tech Intl. Conf. on Bioinformatics; Atlanta [M. Borodovsky, 404/894-2400, Fax: -8925; register@conted.swann.gatech.edu; <http://exon.biology.gatech.edu/conference>]

12-15. Genetics in the New Millennium: Meeting the Challenge; Arlington, VA [AGSG, 202/966-5557 ext. 201; www.geneticalliance.org/99cnfrnc/conf99.html]

13-17. Seventh Conf. on Small Genomes; Arlington, VA [K. Smith, 423/576-4860, Fax: -8646; smithky@ornl.gov; www.esd.ornl.gov/programs/microbes/genome99.html]

15-16. Protein Structure; Washington, DC [see contact: Nov. 8-9]

15-17. NIST Advanced Technol. Program 1999 Meeting; San Jose, CA [L. Buckland, 301/975-4513, Fax: /948-2067; lori.phillips@nist.gov; www.atp.nist.gov/nationalmeeting]

18-21. Third Computational Genomics Conf.; Baltimore [TIGR, 301/610-5959, Fax: -0229; cg@tigr.org; www.tigr.org]

28-Dec. 1. German HGP: Implications, Progress, and Future; Munich, Germany [J. Maurer, +49-30/32639-171, Fax: -262; j.maurer@dhgp.de; www.dhgp.de/genome99.html]

30-Dec. 2. NanoTech '99: 3rd Annu. Conf. on Micro- and Nanoscale Technol. for the Biosciences; Montreux, Switzerland [Secretariat, +44-21/626-4630, Fax: /624-1549; symposia@nanotech99.com; www.nanotech99.com]

December 1999

1-3. Functional Proteomics VI; San Diego [see contact: Nov. 2-5]

6-8. Second Conf. on Genetics and Disease Prevention: Integrating Genetics into Public Health Policy, Research, Practice; Baltimore [J. Kelly, 301/530-1619 ext. 16, Fax: /571-1988; jenniferkelly@conferencemanagers.com; www.cdc.gov/genetics/events/meetings.htm#cdc]

9-12. Physiological Genomics and Rat Models; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845; meetings@cshl.org; www.cshl.org]

January 2000

4-9. Pacific Symp. on Biocomputing; Honolulu [K. Lauderdale, 650/725-0659, Fax: -7944; psb@cgl.ucsf.edu; www.cgl.ucsf.edu/psb]

6-12. Gene Therapy: The Next Millennium; Keystone, CO [Keystone Symp., 970/262-1230, Fax: -1525; keystone@symposia.com; www.symposia.com]

9-12. Plant and Animal Genome VII; San Diego [D. Scherago, 212/643-1750 ext. 20, Fax: -1758; pag@schicago.com; www.intl-pag.org/pag]

11-15. Molecular Toxicology, Toxicogenomics, and Associated Bioinformatics Applied to Drug Discovery; Santa Fe, NM [see contact: Jan. 6-12]

13-14. Agriculture Microbes Genome Conf.; San Diego [P. Johnson; pjohnson@reeusda.gov; www.ag-microbial.org/lam]

14-18. DNA Repair Defects; San Diego [AACR, 215/440-9300, Fax: -9313; meetings@aacr.org; www.aacr.org]

20. Life at the Ends of Your Chromosomes: How to Stay Young Forever; TIGR/NRC/DOE/NASA Speaker Series: Thomas Cech (Univ. of CO); Washington, DC [see contact: Nov. 10]

31-Feb. 1. Cardiovascular Genetics; Miami [see contact: Nov. 8-9]

February 2000

5-8. Advances in Genome Biology and Technology; Marco Island, FL [G Corp., 877/343-8895; info@gcorp.org; www.gcorp.org]

5-9. DNA, RNA, and Cancer; Miami [MNBWS, 305/243-3597, Fax: /324-5665; mnbws-biochem@miami.edu; www.med.miami.edu/MNBWS]

7-8. Gene Quantification: Clinical Applications and Drug Development; La Jolla, CA [see contact: Nov. 8-9]

12-15. TIGR Microbial Genomes Conf.; Chantilly, VA [see contact: Nov. 18-21]

17-22. AAAS Annu. Meeting; Washington, DC [M. Strauss, 202/326-6466; mstrauss@aaas.org; www.aaas.org/meetings]

26-27. Genomic Partnering: Emerging and Early-Stage Companies; San Francisco [see contact: Nov. 8-9]

27-Mar. 2. DOE HGP Contractor-Grantee Workshop VIII; Santa Fe, NM [S. Spengler, 510/486-5874, Fax: -5717; sjspengler@lbl.gov; <http://teamweb.lbl.gov/spengler>]

28-29. NIH NHGRI Advisory Council Meeting; Bethesda, MD [K. Malone, 301/402-2205, Fax: -0837; kimberly@od.nhgri.nih.gov]

28-Mar. 1. HGP: Commercial Implications; San Francisco [see contact: Nov. 8-9]

March 2000

2-3. Gene Functional Analysis; San Francisco [see contact: Nov. 8-9]

9-12. Joint ACMG-MOD Clinical Genetics Meeting; Palm Springs, CA [www.faseb.org/genetics/acmg/meet2000/meet2000.htm]

12-17. Fulfilling the Promises of Genomics Research. Keystone 2000; Taos, NM [see contact: Jan. 6-12]

22-26. Forty-First Annu. Drosophila Meeting; Pittsburgh [P. Geyer, 319/335-6953, Fax: -9570; pamela-geyer@uiowa.edu; <http://flybase.bio.indiana.edu:82/docs/news/announcements/meetings>]

26-28. Genes, Proteins, and Computers VI; Chester, U.K. [J. Ison, jison@hgmp.mrc.ac.uk; www.hgmp.mrc.ac.uk/CCP11/gpcvi/index.html]

26-30. BIO 2000; Boston [S. Spears, 202/857-0244, Fax: /331-8132; www.bio.org/meetings/bio2000/bio2000preview.html]

April 2000.....

9-12. ACEDB 2000: A *C. elegans* Database; Vancouver, B.C., Canada [R. Bruskiwicz; rbsk@sanger.ac.uk; www.sanger.ac.uk/Info/Events/Acedb2000]

9-12. HGM 2000; Vancouver, B.C., Canada [HUGO, +44-171/935-8085, Fax: -8341; hugo@hugo-international.org; www.gene.ucl.ac.uk/hugo/hgm2000.html]

11-15. Genomes 2000: Intl. Conf. on Microbial and Model Genomes; Paris [Inst. Pasteur, +33-145/688-869, Fax: -786; www.pasteur.fr/units/Gmp/Gmp_conference.html]

15-18. Experimental Biol. 2000; San Diego [AFMR, 202/429-5161, Fax: /857-1115; afmr@dc.sba.com; www.afmr.org]

May 2000.....

11-12. HGP: Science, Law, and Social Change in 21st Century; Boston [ASLME Conf., 617/262-4990, Fax: /437-7596; www.aslme.org]

22-23. NIH NHGRI Advisory Council Meeting; Bethesda, MD [K. Malone, 301/402-2205, Fax: -0837; kimberly@od.nhgri.nih.gov] ◊

Training Events*

November 1999

4-9. Computational Genomics; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845; meetings@cshl.org; www.cshl.org]

29-Dec. 3. Advanced Linkage Course; New York [K. Montague, 212/327-7979, Fax: -7996; montagk@rockefeller.edu; http://linkage.rockefeller.edu/suzanne/advance_course.html]

December 1999

1-11. Genome Sequencing and Differential Gene Expression Analysis; Heidelberg, Germany [W. Ansorge, +49-6221/387-355, Fax: -306; reunis@embl-heidelberg.de; www.embl-heidelberg.de/~saffrich/DNASEQ99/DNASEQ99.html]

April 2000.....

2-5. Genetic Analysis of Complex Human Diseases; Durham, NC [V. Roberts, 919/684-2470, Fax: -2275; vroberts@chg.mc.duke.edu; <http://phg.mc.vanderbilt.edu/gachd.htm>] ◊

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

¶ Behavioral Genetics *The Clash of Culture and Biology*

In this volume, editors Ronald A. Carson (University of Texas Medical Branch at Galveston) and Mark A. Rothstein (University of Houston) bring together well-known experts to address the cultural, legal, and biological underpinnings of behavioral genetics. Authors from the fields of genetics, ethics, neuroscience, psychiatry, sociology, and law discuss a broad range of topics. Throughout, they focus on two basic concerns: the quality of the science behind behavioral genetics claims and the need to formulate an appropriate, ethically defensible response when the science is valid. 1999, 224 pp., hardcover. [Johns Hopkins University Press, Baltimore (www.press.jhu.edu/press); also available through booksellers and electronic bookstores] ◇

 **Executive Summary of NIH Biomedical Information Science and Technology Initiative:** www.nih.gov/welcome/director/060399.htm

¶ DOE Office of Science Reports

The DOE Office of Science (OS), which oversees the DOE Human Genome Program, has completed and published *Strategic Plan*, *Science Portfolio*, and "Strategy at a Glance." These colorful companion reports can be downloaded from the Web (www.sc.doe.gov/scsplas.htm); print copies: Christine Chalk (202/586-7203, christine.chalk@science.doe.gov).

The OS mission is to advance basic research and the instruments of science that are the foundations for DOE's applied missions, a base for U.S. technology innovation, and a source of remarkable insights into our physical and biological world and the nature of matter and energy. ◇

ELSI News

Protecting Privacy in Genetic Research

Recommendations to protect privacy of genetic information in research, developed by the Privacy Workshop Planning Committee of the National Action Plan on Breast Cancer (NAPBC), were published in the August 27 issue of *Science* [285(5432), 1359-61; www.sciencemag.org/cgi/content/full/285/5432/1359; NAPBC: www.napbc.org]. ◇

1999 Hollaender Winners Announced*

DOE has announced the award of nine 1999 Alexander Hollaender Distinguished Postdoctoral Fellowships for up to 2 years of research at DOE laboratories having substantial programs supportive of the Office of Biological and Environmental Research's mission. The mission is to understand health and environmental effects associated with energy technologies and to develop and sustain research programs in life, biomedical, and environmental sciences.

Fellowship winners were chosen from a field of applicants who received their doctoral degrees after April 30, 1997. Listed below are each fellow's name, university and subject of doctoral degree, host laboratory and research mentor, and research topic.

Brian Bovard (Duke University, Botany); University of Michigan Biological Station, Peter Curtis (Ohio State University) and James Teeri (University of Michigan); Determining Plant and Soil Contributions to Ecosystem Respiration Using Stable Isotopes

Sophia Chernikova (Colorado State University, Cellular and Molecular Radiobiology); Lawrence Berkeley National Laboratory, Priscilla Cooper; Role in Base Excision Repair in Mammalian Cells of DNA Damage Signaling Usually Associated with Double-Strand Breaks

Jeffrey Dukes (Stanford University, Biology); University of Utah, James Ehleringer; Responses of Rangeland Invaders to Climate Change

Christine Goodale (University of New Hampshire, Natural Resources); Carnegie Institution of Washington, Chris Field; Effects of Land Use Change on Carbon Cycling

Len Pennacchio (Stanford University, Biological Sciences):

Lawrence Berkeley National Laboratory, Edward Rubin; A Genomic Approach to Identify Genes Involved in Neurological Disease

Derek Radisky (University of Utah, Pathology); Lawrence Berkeley National Laboratory, Mina Bissell; Molecular Characterization of the Mechanism of BCE-1 Activation

Karah Street (University of Texas Health Sciences Center, Cellular and Structural Biology); Lawrence Livermore National Laboratory, Lawrence Thompson; Role of RAD51B in Recombinational Repair of Radiation Damage

Xiling Wu (Mayo Clinic Medical School, Pharmacology); Los Alamos National Laboratory, David Chen; Identification of DNA-PK Substrates in DNA Double-Strand Break Repair

Xuqiong Wu (University of Virginia, Physiology); Lawrence Berkeley National Laboratory, Terumi Kohwi-Shigematsu; SATBI-Mediated Transcription Regulation Complex at Methylated DNA

*See contact information in box at right.

U.S. Genome Research Funding

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

DOE Office of Biological and Environmental Research Human Genome Program

- Funding information, inquiries: genome@science.doe.gov or 301/903-6488
- Relevant documents: www.er.doe.gov/production/ober/hug_top.html

Alexander Hollaender Distinguished Postdoctoral Fellowships

Research opportunities in energy-related life, biomedical, and environmental sciences, including human and microbial genomes, global change, and supporting disciplines.

- Next deadline: January 15, 2000
- Contact: Barbara Dorsey, Oak Ridge Institute for Science and Education (423/576-9975, Fax: /241-5220, dorseyb@ornl.gov, www.ornl.gov/ober/hollaend.htm)

Computational Molecular Biology Postdoctoral Fellowships

Topic: Support career transitions into computational molecular biology from other scientific fields. Funded by DOE and the Alfred P. Sloan Foundation to give young scientists an intensive 2-year postdoctoral opportunity in an appropriate molecular biology facility.

- Next deadline: February 1, 2000
- Contact: Melissa Stoudenheimer; Alfred P. Sloan Foundation (212/649-1649, Fax: /246-7585, stoudenheimer@sloan.org, www.sloan.org/main.htm)

NIH National Human Genome Research Institute

- NHGRI program: 301/496-7531, Fax: /480-2770, www.nhgri.nih.gov/About_NHGRI
- Program announcements: www.nhgri.nih.gov/Grant_info
- ELSI: 301/402-4997

Small Business Innovation Research Grants

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

Contacts:

- DOE SBIR/STTR Office: 301/903-1414 or -0569, Fax: -5488, sbir-sttr@science.doe.gov. SBIR/STTR applications due Feb. 29, 2000. SBIR: sbir.er.doe.gov/sbir; STTR: sttr.er.doe.gov/sttr
- Bettie Graham (see contact, NHGRI). NIH SBIR due April 15, August 15, and December 15. STTR, April 1, August 1, and December 1

National SBIR/STTR conferences: November 21-23, Las Vegas, NV, and May 5-7, 2000, Washington, DC. Contact: 360/683-5742, Fax: -5391, sbir@ztn.com. For regional conferences, see Web site (www.ztn.com/sbir). ◇



(see Informatics, p. 23)



