BIOLOGY CHEMISTRY PHYSICS ENGINEERING INFORMATICS

Human Genome news

A Forum for Genome Information Exchange

Sponsored by the U.S. Department of Energy

ISSN: 1050-6101

Vol. 8, Nos. 3 and 4, January–June 1997

Fear of Genetic Discrimination Drives Legislative Interest

Ownership, Predisposition Major Issues

By Philip R. Reilly, Shriver Center

Legislative interest in regulating the uses of genetic information is at an all-time high, after decades in which not one legislator at either the congressional or state level introduced a bill to control the use of genetic information. Since 1989, however, the number of such proposals, as well as the territory they seek to cover, has expanded steadily as genetic data has proliferated.

In the early 1990s, legislation focused almost exclusively on attempting to forbid insurers from using genetic information in health-insurance underwriting decisions. Currently, the social concern driving legislative interest is fear of genetic discrimination, defined here as discrimination against otherwise-healthy individuals on the basis of a genotypic variation.

As of April 1997, at least 15 states had enacted genetic privacy laws. More than 75 similar bills are pending in more than 30 states, according to a survey by the Biotechnology Industrial Organization, and several federal bills have been introduced into the 105th Congress.

Protection Under HIPAA

Significant legislation at the state level has been more than eclipsed by a new federal law, "The Health Insurance Portability and Accountability Act of 1996" (PL-104-191), which takes effect this year. HIPAA provides an important new protection for people who want to undergo genetic testing but fear discrimination by health insurers if their test results indicate an increased risk for developing a serious disease. Section 101 of HIPAA sharply curtails the right of group health insurers to limit coverage of new employees because of "preexisting conditions." As of August 1997, group health insurance plans may impose a preexisting-condition exclusion only when "medical advice, diagnosis, care, or treatment was recommended or received within the 6-month period before enrollment."

The new law also forbids group health insurance plans to apply the preexisting-condition rule to genetic information unless the person has been diagnosed with the illness predicted by the genetic test. For example, a woman who does not have cancer may not be denied coverage even if her test results indicate a

(see Genetic Privacy, p. 2)

In This Issue

Rapid Sequencing of Microbial Genomes Opens Door to Functional Genomics

"Microbial Genome Research and Its Applications" was the topic of the 35th Hanford Symposium on Health and the Environment held October 21–24, 1996, in Richland, Washington.[1] According to participants, the genomes of as many as 100 microorganisms are expected to be completely or partially sequenced by the turn of the century.[2] The overriding question during the meeting was, How do we deal with, interpret, and use all this new information?

The meeting served well as a status report on whole-genome microbial sequencing and on initial attempts to use the new information. The complete 3.6-Mb sequence for the cyanobacterium *Synechocystis* sp. PCC6803 was announced by Hirokazu Kotani (Kazusa DNA Institute, Japan).[3] Progress was reported for *Treponema pallidum*, the causative agent of syphilis (George Weinstock, University of Texas Medical School); the hyperthermophilic Archaeon *Pyrobaculum aerophilum* (Jeffrey Miller, University of California at Los Angeles); *Pyrococcus furiosus*, a thermophilic Archaeon (Robert Weiss, University of Utah); and the

(see Microbial Genomes, p. 3)

DOE Contractor-Grantee Workshop, p. 11

Genetic Privacy *(from p. 1)*

predisposing mutation for breast or ovarian cancer.

Given the scope of HIPAA, rapid proliferation of similar state laws, steady growth of managed care, and public mood, individuals who decide to undergo DNA-based predispositional testing may face relatively little risk of discrimination in health insurance. Nevertheless, a widespread concern—stimulated in part by the emergence of tests to identify persons at increased risk for cancer—is that test results can and will be used against people. Some are so concerned that even protective legislation is unlikely to reassure them.

The vast number of genetic privacy bills circulating in state legislatures has generated many questions. Four of the most important are (1) What is meant by a genetic test? (2) Is genetic information distinct from or merely one form of medical information? (3) Should a tissue sample and data derived from it be the "property" of the person from whom it was taken? (4) Do we need more stringent oversight of human genetic research?

Defining "Genetic Test"

Deciding what constitutes a genetic test is not so easy, and definitions vary widely. A current Texas bill (TX 75RSB 98) defines it as a test of an "individual's DNA, RNA, or chromosomes . . . associated with a predisposition for a clinically recognized disease or disorder." Note that this definition does not include proteins, so it excludes some newborn screening, prenatal tests for neural tube defects, and many tests currently used by geneticists to make diagnoses. A more-inclusive Vermont bill (H.89), on the other hand, defines genetic testing as analysis of a "chromosome, a gene, DNA, RNA, or protein encoded by a gene . . . ." Both bills also exclude certain standard medical tests from their reach. If each state bill has its own particular definition and many become law, these variations may someday haunt insurers, employers, and testing laboratories that seek to comply with laws in the individual states.

Protection for Other Clinical Data

Is genetic information so different from other clinical data that it deserves special protection? There is, admittedly, precedent for this. Our society traditionally accords a special level of protection to psychiatric records, and, to some extent, we have condoned a higher degree of protection to HIV test results. In essence, the argument that genetic data is different from regular medical information and deserves special protection is two pronged:

- genetic tests may predict future risks for healthy persons, and
- these tests may infer risk about relatives.

True enough—but the ability to treat genetic information with special care depends on how well it can be separated from other clinical information. If genetic testing permeates medical care, as it almost certainly will within the next 20 years, it will be very

Pending Legislation Could Impact Research

Two of the federal bills on genetics currently before Congress are the "Genetic Information and Nondiscrimination in Health Insurance Act of 1997" (H.R.306), introduced by Rep. Louise Slaughter (D-NY), and the "Genetic Confidentiality and Nondiscrimination Act of 1997" (S.422), introduced by Sen. Pete Domenici (R-NM).

H.R.306

H.R.306 addresses the genetics concern voiced most often by the American people: that test results could make health insurance more difficult or impossible to obtain. The bill goes beyond the scope of HIPAA in extending protection against this use of genetic information to virtually all Americans.

H.R.306 would prohibit group health plans from denying, canceling, refusing to renew, or changing the terms, premiums, or conditions of coverage based on genetic information. It would also prevent health insurers from requesting or requiring a genetic test as a condition of coverage and would require written informed consent before the health plan could disclose genetic information to a third party.

Contact:

- Hon. Louise Slaughter; 2347 Rayburn House Office Bldg.; Washington, DC 20515 (202/225-3615, *louiseny@hr.house.gov*)

Phillip Reilly suggests that scientists conducting genomic research read pending federal legislation, especially Title V of S.422, and consider commenting on potential impacts (access to H.R.306, S.422, and S.193: *http://thomas.loc.gov/home/c105query.html*).

S.422

The more-comprehensive S.422 attempts to regulate the use of genetic data by health insurers and employers, sets up a system by which the person who is the DNA sample source exercises ongoing control over release of the data it yields, and offers a plan to oversee genetic research. S.422 is believed to be the first bill to propose a specific scheme for regulating the conduct of genomic research.

Among its most important research provisions, S.422 extends Institutional Review Board (IRB) assessment to all genomic research regardless of funding source. Section 501(a)(2) sets up a balancing test for IRBs to use in assessing the value of a research proposal: Do societal benefits outweigh the risks to individual subjects? IRBs, already growing uneasy about informational risks associated with genomic research, have been admonished to safeguard individual subjects. This test would make IRBs take a hard look at such issues as the care with which investigators plan to protect the security of DNA samples and the data generated therefrom.

Section 501(a)(3)(E) of S.422 prohibits research information from being placed in clinical records, which many consider a positive step. Section (a)(3)(F) virtually creates a duty to disclose clinically relevant research findings to families of deceased persons who had been study subjects. This section tends to support the argument that researchers have a duty to recontact some individuals under certain conditions, an obligation that could be very difficult to fulfill.

S.422 strives to create a set of regulations under which genetic information generated by researchers will be used to help, not harm, individuals. At the same time, some are concerned that portions of the law could unfairly impede research.

Contact:

- Hon. Pete Domenici; 328 Dirksen Senate Office Bldg.; Washington, DC 20510 (202/224-6621, Fax: -7371, *senator_domenici@domenici.Senate.gov*) ◊

difficult to implement a law that requires separate treatment of portions of many people's medical records. Few bills have confronted this issue; in those that have sought to deal with it, genetic information is defined so narrowly that, if enacted, the bills would protect very little information about very few people. Too often lost in the discussion about genetic privacy is the point that nearly everyone would benefit from enactment of a general medical privacy law covering access to and use of all health information.

Individual Property Rights

In 1995, bills began to appear that were partly influenced by a model "Genetic Privacy Act" drafted by DOE ELSI grantee George Annas and his colleagues at Boston University. These bills asserted that tissue taken for genetic testing, as well as the test results, should remain the property of the individual tested. This principle challenges a century of practice in which discarded tissue samples have been used routinely by pathologists for research and teaching in a manner that respects the donor's privacy.

In 1995, Oregon enacted a law based on Annas's model, and the property proposal has since popped up in many other bills. In 1996, the New Jersey legislature passed a similar bill that was vetoed by Governor Whitman after 11th-hour protests from the pharmaceutical industry. Citing the bill's possible chilling effect on research, Whitman later signed an amended version that deleted the property provision but required that all genetic testing be preceded by written informed consent. A bill now before the New Jersey legislature seeks to reintroduce the property concept.

Bills that include provisions forever tying up tissue samples as property of persons from whom they were taken do raise questions about uses to which the academic and research community may put the DNA. Like it or not, given the importance of intellectual property concerns, scientists might not go forward with research if they don't have a clear right to use the samples.

IRB Oversight

The most recent trend in genetic privacy bills is the notion that genetic information is so sensitive and the threat of genetic discrimination so significant that our society needs a new level of oversight for research involving human genetics. Proponents of these provisions offer few concrete examples of studies in which genetic research has harmed human subjects; rather, they assert a potential future risk. Nevertheless, it would not be surprising to see state laws that define all research involving human gene mapping as constituting significant social risk to subjects. Any such research, regardless of its funding, thus would require the equivalent of an independent review by an institutional review board (IRB) or its equivalent. Although unlikely to become law this year, "The Human Research Subjects Protection Act of 1997" (S.193), introduced by Senator John Glenn (D-OH), proposes to extend IRB oversight to all U.S. research involving human subjects.

Federal Control

If the federal government decides to broaden its level of control over research, a comprehensive federal law could be enacted to preempt individual state regulations. Monitoring and abiding by a potpourri of varying state rules might be exceedingly expensive and time consuming for multicenter investigations.◊

Clinical geneticist and attorney Phillip Reilly, M.D., J.D., is executive director of the Shriver Center for Mental Retardation in Waltham, Massachusetts. He also serves on the board of directors of the American Society of Human Genetics. With his colleagues Dorothy Wertz and Robin J.R. Blatt, Reilly writes and edits *The Gene Letter*, a DOE-supported electronic newsletter about genetics and public policy. Six issues of the newsletter, published since July 1996, are now on the Web (*http://www.geneletter.org*).

¶ Genetic Testing Report

Promoting Safe and Effective Genetic Testing in the United States: Principles and Recommendations is now on the Web (*http://www.med.jhu.edu/tfgtelsi/promoting*). This report, by the Task Force on Genetic Testing of the NIH-DOE Joint Working Group on Ethical, Legal, and Social Implications of Human Genome Research, supersedes the proposed recommendations and interim principles released early this year.◊

Microbial Genomes

(from p. 1)

hyperthermophilic eubacterium *Aquifex VF5* (Ron Swanson, Recombinant BioCatalysis, Inc.).

Presentations and discussions also focused on techniques for generating clones for sequencing, different ways of identifying genes and gene functions, and databases.

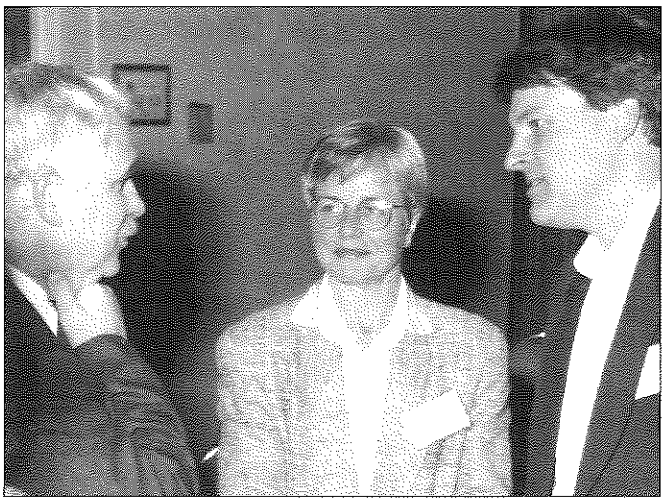
Identification of Gene Function

The explosion of information from the genome projects has spawned a number of efforts, some of them coordinated, to develop readily accessible databases that are truly useful to individual scientific investigators. The session chaired by Ross Overbeek (Argonne National Laboratory) covered some of these attempts and their underlying problems. A major challenge is to annotate completed genomes at a rate comparable to the generation of raw sequence data.

Owen White [The Institute for Genomic Research (TIGR)], describing specialized in-house software to help identify frame-shift errors, pointed out that tracking problems can result in lost information when data is moved between projects or among institutions. He described how, for some metabolic pathways, genes for key enzymatic proteins cannot be found. Is this because the microbe lacks the enzyme, or is it due to the inability to identify corresponding open reading frames (ORFs)?

Overbeek pointed out other problems in relying on sequence homology for gene identification. He cited one case in which enzymes shared sequence homology but belonged to unrelated metabolic pathways. To help identify these problems, Overbeek and Niels Larsen (Michigan State University) are linking these gene lists to known metabolic pathways on a Web site called WIT,[4] which replaces the former PUMA site.

Others are using different computer-assisted approaches to identify ORFS of unknown function. Monica Riley (Woods Hole Oceanographic Institute) suggested that genes encoding enzymes need to be broken into functional domains before they are assigned to databases. Enzymes can



Linda Lasure, president-elect of the Society for Industrial Microbiology, discusses microbial genome science with David Galas (left) and Jay Short at the 35th Hanford Symposium on Health and the Environment held in October 1996. Galas, now at Darwin Molecular Corporation, is former director of DOE OHER, and Short is with Recombinant BioCatalysis, Inc.

exhibit multiple functions by encoding multiple functional domains in a single protein or by complexing multiple protein subunits of different function. Domain order may be shuffled and still result in a similarly functioning enzyme that would go unrecognized if entire gene sequences were compared. Using Riley's domain approach could suggest additional functions for enzymes having only one known function (see p. 6 for more details on Riley's approach).

Keynote speaker J. Craig Venter (TIGR) was one of a number of participants who expressed concern about the consequences of poor control over the quality of data entering public databases. Another issue was how much time should pass between sequence determination and deposition in a public database for government-sponsored projects. Some attendees supported immediate deposit of new information to create and maintain a level playing field and to minimize the commercial advantage of firms paid by federal sponsors to acquire the sequence. Others felt that a greater risk of archiving bad data would accompany its deposition without some measure of preliminary quality assurance.

Laboratory Challenges

Innovative work in the laboratory clearly is needed to test hypotheses suggested by in silico methods and to search for functions associated with unidentified ORFs. A novel approach for identification was presented by George Church (Harvard Medical School), who described the development of Genomically Engineered Multiplex Selection (called GEMS) to measure simultaneously the survival effects of in-frame deletions in many genes in many environments. To determine whether specific genes are required under a given set of conditions, each gene is deleted in frame, mutants are pooled, and mutant loss is monitored during growth under various culture environments.

Some functionally equivalent enzymes have evolved independently and thus are coded for by genes with unrelated sequences. Jay Short (Recombinant BioCatalysis) described a high-throughput screening strategy in which total genomic DNA is extracted from the environment, cloned, and tested for expression of enzymatic functions of interest. The result is a pool of gene sequences (an "environmental library") for proteins capable of carrying out a single function. The collective genomes are archived in recombinant form using cloning vectors containing the *Escherichia coli* F-factor origin of replication, which permits high-fidelity replication of 40- to 300-kb cloned DNA fragments. As an example of this innovative concept's potential, Short reported the discovery of previously unknown Archaeal genes for RNA helicase, which is responsible for ATP-dependent alteration of RNA secondary structure; and glutamate semialdehyde aminotransferase, which is involved in initial steps of heme synthesis.

Others such as Gerben Zylstra (Rutgers University) are taking an alternative approach to identifying enzymes with unique sequences. Zylstra's interest is in pathways involved in degradation of aromatic compounds, but the same approach could be applied to other pathways. By selecting for lack of hybridization to a suite of probes for all representative enzymes known to be involved in a specific pathway, Zylstra has successfully identified unique enzymes from enrichments of aromatic hydrocarbon-degrading bacteria from polluted environments.

Future Concerns

Hans Peter Klenk (TIGR) discussed a report in *Science*[5] that the majority (62%) of predicted protein-coding genes in the *Methanococcus jannaschii* genome are of unknown function. The report has highlighted concerns that underlying assumptions used to assign probable function to ORFs need to be assessed carefully. Thus, approaches described above will lead to sequences of genes that share function but not necessarily sequence, thus underscoring a major point to emerge from the symposium—that care must be taken when interpreting results from the flood of sequence information being placed in the public domain.

The inverse has also been shown[6] for very different sequences encoding unrelated primary protein structures that result in similar functions. Gary Olsen's (University of Illinois) presentation on the use of sequence information (primarily rRNA gene sequences) to infer phylogeny demonstrated the need for careful consideration of sequence similarity for purposes other than predicting gene-product function.

Many scientists predict that the availability of dozens of complete microbial sequences is a major breakthrough in fundamental biology that surely will result in new ways of approaching industrial microbiology and fighting infectious diseases. Without adequate tools for managing information or plans for prioritizing ever-shrinking financial resources for science, however, much may go unrealized, and many of the benefits may be deferred longer than necessary. *[F. Blaine Metting (*fb_metting@pnl.gov*) and Margaret F. Romine (*mf_romine@pnl.gov*), Pacific Northwest National Laboratory]*

Conference on Small Genomes Held at Hilton Head

At the January 1997 Conference on Small Genomes: Sequencing, Functional Characterization, and Comparative Genomics, over 250 participants gathered on Hilton Head Island to discuss recent progress and future directions in this emerging and exciting area of research. As stated by Craig Venter [The Institute for Genomic Research (TIGR)] in the opening session, small-genome research is growing exponentially, and a new era of biological insight is emerging because of it. This first meeting on small genomes was sponsored by TIGR and organized by Claire Fraser (TIGR), Hamilton O. Smith (Johns Hopkins University Medical School), and E. Richard Moxon (Oxford University). Selected meeting highlights follow.

Plasmodium falciparum, the organism causing 300 million to 500 million new cases of malaria each year, contains over 5000 genes distributed among 14 chromosomes. As discussed by Stephen L. Hoffman (Naval Medical Research Institute), less than 5% of this organism's DNA is cloned at the present time. A consortium of sponsors including The Wellcome Trust, Burroughs Wellcome Fund, U.S. Department of Defense, and the NIH National Institute for Allergy and Infectious Diseases has joined to sequence this pathogen, many strains of which have become resistant to chloroquine, the only effective drug for treating malaria. The hope is that new targets for different drugs will be revealed by the complete sequence of the plasmodial genome. Cloning and sequencing the plasmodium DNA, however, is proving difficult because of the organism's high AT content (as much as 76% in coding regions and 90% to 100% in intergenic regions). Further, *P. falciparum's* DNA has proven unstable in *Escherichia coli*, and the organism itself is difficult to maintain in culture.

ORFs: Open reading frames, one method researchers use to identify genes in DNA sequencing.

*Attendees Discuss Small Genome Sequencing Progress**

Sequences Completed Recently

- *Helicobacter pylori*, Jean-Francois Tomb (TIGR)
- *Escherichia coli*, Fred Blattner (University of Wisconsin)
- *Archaeoglobus fulgidis*, Hans-Peter Klenk (TIGR)
- *Methanobacterium thermoautotrophicum*, Douglas Smith (Genome Therapeutics)

Sequences Nearing Completion

- *Aquifex aeolicus*, Ronald Swanson (Recombinant BioCatalysis)
- *Bacillus subtilis*, European and Japanese consortium
- *Deinococcus radiodurans*, Owen White (TIGR) and Kenneth Minton (Uniformed Services University of the Health Sciences)
- *Neisseria gonorrhoea* and *Streptococcus pyogenes*, Bruce Roe (University of Oklahoma)
- *Pyrobaculum aerophilum*, Sorel Fitz-Gibbon (UCLA) and Melvin Simon (Cal Tech)
- *Pyrococcus furiosus*, Robert Weiss (University of Utah)
- *Streptococcus pneumoniae*, Brian Dougherty (TIGR)
- *Sulfolobus solfataricus*, NRC, Canada
- *Treponema pallidum* and *Borrelia burgdorferi*, Claire Fraser (TIGR)

*See box, p. 6, for early history.

This article was adapted from the meeting synopsis in *Microbial and Comparative Genomics* **2**(1), 1997. Program and abstracts: **1**(4), 1996.

Several interesting characteristics of *Helicobacter pylori*, which resides in mucosa or apposed to epithelia and causes gastric and duodenal ulcers in humans, were reported by Jean-Francois Tomb (TIGR) and Douglas E. Berg (Washington University Medical School). *H. pylori* contains over 1500 open reading frames (ORFs), with many repetitive sequences. A large porin-like family of cell adhesion molecules may be involved in mucosal or epithelial cell adhesion. Most genes for flagellar structure and function are present. Dinucleotide repeats may affect the ORF for some surface proteins and provide a possible mechanism for eluding the host immune system. There is much variation among *Helicobacter* strains. Different variants predominate in members of the same host species, which may reflect spontaneous mutation, horizontal gene transfer, and selection by different hosts over a long interval of infection. A challenge for the future will be to identify the function of unknown or "orphan" genes in *H. pylori* and determine how different genes allow this organism to live in a relatively hostile acid environment, causing active disease in some cases but not others.

Bernard Dujon (Institut Pasteur) presented some highlights of the recently completed yeast genome sequence. Three signature features are apparent in the more than 6000 ORFs of this organism: A large number of genes have unknown function, the genome contains many redundant sequences, and about 10% to 15% of the ORFs are thought to be essential for life. Work is proceeding to assign functions to orphan ORFs by using deletion and other mutants to examine effects on expression in different vector systems. Two hybrid systems are being used to search for relationships and interactions with other yeast gene products.

Satoshi Tabata (Kazusa DNA Research Institute, Japan) reported that 3,573,470 bp of DNA specify

Microbial Genomes

(from p. 4)

References

1. See *http://www.pnl.gov:2080/health/attend.html* for a list of symposium participants.

2. See *http://www.mcs.anl.gov/home/gaasterl/genomes.html* and *http://www.tigr.org/tdb/mdb/mdb.html* for listings of microorganisms whose sequences have been or are being determined.

3. *http://www.kazusa.or.jp/cyano/cyano.html*

4. *http://www.cme.msu.edu/WIT* (viewable only with frames-capable browser).

5. C.J. Bult et al., *Science* **273**, 1058–73 (1996).

6. P.C. Babbitt et al., *Science* **267**, 1159–61 (1995).◊

more than 3100 potential ORFs in the cyanobacterium *Synechocystis*. This single-celled organism is photoautotrophic and capable of oxygenic photosynthesis. *Synechocystis* contains more than 126 genes related to photosynthesis, and about 90% of algal plastid genes appear to be conserved in *Synechocystis*.

Richard Hermann (University of Heidelberg) reported on the closely related bacteria *Mycoplasma genitalium* (580,070 bp) and *Mycoplasma pneumoniae* (816,394 bp). All the ORFs in *M. genitalium* are present in *M. pneumoniae*, whose genome consists of six segments. Segment order is not conserved between the two species, but gene order is conserved within each segment. Differences in genome size could have evolved by deletion of *M. genitalium* genes that are not essential for life outside the host and by gene amplification in *M. pneumoniae*.

Bacterial virulence factors are often encoded in extrachromosomal plasmids, phages, and transposons that can be transmitted horizontally between different species or strains. Pathogenic *Vibrio* cholera, for example, carries a phage kappa. The phage has a gene for Glo, a virulence factor very similar to a small eukaryotic G protein. Present on the chromosomal DNA of toxigenic strains are cholera toxin–encoding genes similar to the genome of a filamentous bacteriophage. The cholera toxin and pilus are regulated coordinately by a transcriptional regulator called Tox R. In fact, the pilus is the receptor for the phage (John Mekalonos, Harvard Medical School).

As reported by Brian G. Spratt (University of Sussex, U.K.), both inter- and intraspecies recombination can occur in bacteria, particularly when the bacteria coexist in an environment such as the nasopharynx. In these cases, Spratt noted, recombination occurs more frequently in housekeeping genes and in genes under strong selection.

Analysis of Gene-Product Function

Several groups are attempting to analyze the function of gene products specified by ORFs of completely sequenced organisms. Richard Moxon and his collaborators have identified and cloned 25 genes involved in the biosynthesis and regulation of the lipopolysaccharide (LPS) of *Haemophilus influenzae*. LPS is a major virulence determinant for this human pathogen. Analysis of these genes and their mutants with monoclonal antibodies, polyacrylamide gel electrophoresis, and mass spectrometry has, in fact, confirmed a role for most of these proteins in strain pathogenicity.

H. influenzae has an estimated 1700 ORFs. Only about 500 polypeptides, however, can be detected by Coomassie blue staining following separation of cell extracts by 2-D polyacrylamide gel electrophoresis. About 650 polypeptides can be detected by autoradiography after biosynthetic labeling of cells with ^{35}S-methionine before electrophoresis. The effects of protein inhibitors and RNA synthesis on these resolved polypeptides were discussed by Stefan Evers (Hoffmann-LaRoche). Responses to different inhibitors were similar for some resolved proteins. For example, up-regulation occurred for some transcription and translation bacterial components, including ribosomal proteins and RNA polymerase. Most puzzling is the observation that, in many cases, a change in gene transcriptional level did not correspond to a change in translation rate.

Ian Humphery-Smith (National Innovation Center, Australia) discussed limitations of the 2-D polyacrylamide gel approach to protein resolution and analysis. Improved methods are needed for extracting protein from cells, fractionating and enriching protein classes, and conducting 2-D polyacrylamide gel electrophoresis, especially when providing a larger separation area to obtain better resolution. Better methods to detect resolved polypeptides by using mass spectrometry and nanoelectrospray mass spectrometry are under development.

Both Monica Riley (Marine Biological Laboratory, Woods Hole) and Peter Karp (Artificial Intelligence Center, SRI International) spoke about progress and problems in assigning functional roles to ORFs identified by complete genome sequence analysis. Sequence-similarity analyses of amino acid residues aligning 100 to 200 *E. coli* residues showed that all the proteins can be grouped in families ranging in size from 2 to more than 60 members. Interestingly, not all proteins performing a similar function have sequence similarity, emphasizing the challenge of assigning a specific function to a protein from its deduced amino acid sequence. Current methods are being improved to place a protein sequence deduced from an ORF into a metabolic pathway by using databases that describe the genes and intermediary metabolism of *E. coli* and *H. influenzae*. Use of these databases should lead to a better and more reliable system for identifying biological function during annotation.

As discussed by Hamilton Smith, nongenic information between ORFs may be as interesting and important as the

Early History of Small Genome Sequencing

In 1977, bacteriophage phi-x174 (5386 bp) became the first organism to be sequenced completely, by Sanger and colleagues [*Nature* **246**, 687 (1977)]. In 1982, bacteriophage lambda (48,502 bp) was completed using a strategy based on sequencing random fragments of DNA, in this case produced by digesting the lambda genome with restriction enzymes, again by Sanger and colleagues [*J. Mol. Biol.* **162**, 729 (1982)].

Thirteen years elapsed before the first nonviral organism was sequenced completely, this time using whole-genome random sequencing and assembly, called shotgun sequencing. In July 1995, Fleischmann and colleagues reported the completion of *Haemophilus influenzae* (1,830,137 bp), the first free-living organism to be sequenced [*Science* **269**, 469 (1995)]. At the end of 1995, the complete DNA sequence of *Mycoplasma genitalium* (580,070 bp), another free-living organism, was published by Fraser and colleagues. *M. genitalium* DNA encoded only 470 predicted ORFs, providing an estimate for the minimal number of genes needed to support life [*Science* **270**, 349 (1995)].

Since 1995, complete genomic sequences have been published or made available for four more organisms: *Methanococcus jannaschii*, an Archaeon; *Synechosystis*, a cyanobacterium species; *Mycoplasma pneumoniae*, a eubacterium closely related to *M. genitalium*; and the yeast *Saccharomyces cerevisiae*, the first eukaryotic organism to be completely sequenced. [See box, p. 5, for more recent progress in small genome sequencing.]

genes themselves, but these sequences often would be overlooked during annotation. For example, a 34-amino acid ORF is oriented oppositely to two flanking *H. influenzae* genes and has a strong promoter and ribosome-binding site. Additional experimentation is needed to determine this gene's function and to identify similar genes.

Biological Diversity

Although a large number of microorganisms will be sequenced in the near future, they will represent a minute sample of biological diversity on earth. Only a very small fraction of living organisms can be cultivated, cloned, and grown in a defined laboratory environment, 0.001% to 0.1%, as estimated by Norman R. Pace (University of California, Berkeley). However, such newer techniques as PCR, gene cloning, sequencing of amplified products, and ribosomal RNA typing will allow a survey of different organisms in their natural environment, thus eliminating the need for laboratory cultivation. Similar techniques are being used to examine biological diversity among the Archaea (Edward F. DeLong, University of California, Santa Barbara). Initial results indicate that the microbial world has just begun to be appreciated and that much new information and many surprises in the biochemical, genetic, metabolic, physiological, and evolutionary realms will be forthcoming from studies of this most abundant and diverse group of organisms.

The second annual Small Genomes meeting will be held at Hilton Head, South Carolina, from January 31 to February 5, 1998. *[Darrell Doyle, TIGR]* ◊

Microbial Resources

☛ MAGPIE: Data on Gene Sequences

MAGPIE (multipurpose automated genome project investigation environment) provides genome sequencing projects with an automated platform for collecting computed data about an emerging or finished genome sequence. Once installed, two daemons run side by side: one collects data through a multitude of automated requests to remote and local servers, and the other synthesizes collected data into a "knowledge base" of queryable information. MAGPIE, which was created at Argonne National Laboratory by Terry Gaasterland, in collaboration with Christoph Sensen of the Canadian NRC Institute for Marine Biosciences, includes a prepackaged set of queries that generate hierarchically organized reports about the genome sequence data.

The MAGPIE site (*http://www.mcs.anl.gov/home/gaasterl/magpie.html*) also lists the status of genome sequencing projects and links to genome databases at other institutions and to metabolic and functional pathway resources.◊

☛ Microbial Database at TIGR

The Microbial Database (MDB, *http://www.tigr.org/tdb/mdb/mdb.html*), created by Anthony Kerlavage at The Institute for Genomic Research, lists the sequencing status of microbial genomes. Information for each completed or in-progress genome includes size, domain classification, sequencing institution, funding source, and anticipated date of publication. Links are provided to data and to papers published about completely sequenced genomes.◊

☛ Complete Microbial Genome Clones

The American Type Culture Collection and The Institute for Genomic Research have created a new special collection to store and distribute clones resulting from the complete sequences of three microbial genomes:

- *Haemophilus influenzae* Rd, ATCC 51907 (*Science* **269**, 496–512, 1995);
- *Methanococcus jannaschii* (*Science* **273**, 1058–73, 1996); and
- *Mycoplasma genitalium* G37, ATCC 33530 (*Science* **270**, 370–403, 1995).

Purified DNA from these organisms also will be available from ATCC. [Sequence and specific clone information: *http://www.tigr.org/tdb/mdb/mdb.html*; ordering: *http://www.atcc.org/hilights/microbe_gene.html* or 301/881-2600] ◊

President's Bill Would Prohibit Human Cloning

Acting on a recommendation by the National Bioethics Advisory Commission (NBAC), President Clinton announced on June 9 that he will send Congress a bill to outlaw the cloning of humans. Biomedical and agricultural research that could develop new medical therapies would be unaffected. Prompted by the cloning of the lamb "Dolly" in a U.K. laboratory, NBAC's report noted that attempting to clone a human would be premature and unacceptably risky.◊

HGMIS Web Sites Win Awards

The HIGMS Web sites:

- *http://www.ornl.gov/hgmis/home.html*,
- *http://www.ornl.gov/hgmis/publicat/publications.html*, and
- *http://www.ornl.gov/hgmis/vl.html*

have been granted the Editor's Choice Award by the LookSmart Internet directory for providing a useful and high-quality Web site. LookSmart International, a subsidiary of *Reader's Digest*, is committed to quality and superior Web experience for users. Previous awards include the highest ratings in the Magellan and MEDSITE directories.◊

Want Notification of *HGN* Web Version?

Readers wishing to be notified when the latest issue of *HGN* is placed on the Web should send their request and e-mail address to *yustln@ornl.gov*. A table of contents and URL will be e-mailed to requesters when each issue is posted. This electronic version usually is available at least 2 weeks before readers receive the printed copy.◊

Planning a Meeting?

To avoid conflicts, see genome calendars on pp. 14–15 and a more extensive listing on the HGMIS Web site (*http://www.ornl.gov/hgmis*). For inclusion in the calendars, submit information to HGMIS (see contact information, p. 10) or to the Web site (*http://www.ornl.gov/hgmis/forms/form.html*). Announcements may also be submitted to the meetings Website (*http://www.ornl.gov/meetings*).◊

Genetics in the News

NCHGR Becomes NIH Institute

In January, the NIH National Center for Human Genome Research (NCHGR) was granted institute status and a new name—National Human Genome Research Institute (NHGRI). The new status will facilitate collaboration with other institutes, give NHGRI's director equal standing with other directors, and allow NHGRI to operate under the same legislative authorities as other NIH research institutes. The new name is thought to reflect more accurately the growth and accomplishments of the former NCHGR, which was established 7 years ago to carry out the NIH role in the Human Genome Project.◊

Chromosome X Map Completed

On March 14, researchers at Washington University School of Medicine in St. Louis announced the completion of a high-resolution map of chromosome X. This chromosome, which determines gender, is associated with many inherited disorders. Researchers also located hot spots for genes and detected a large region where the DNA remains intact as it passes from one generation to another.

Published in the March issue of *Genome Research*, the map has 2100 unique landmarks—three times as many as any previous X chromosome map. If it were a road map from St. Louis to San Francisco, it would show a marker every mile. On average, the new map has a landmark every 75,000 bp. The national goal for chromosome mapping is one landmark every 100,000 bp.

David Schlessinger, director of the Center for Genetics in Medicine and principal investigator for the NIH-funded X project, said, "Completion of a map with this level of detail has made X one of the earliest chromosomes for DNA sequencing—the next phase of the Human Genome Project."

X chromosome's completion has permitted a refined comparison between a physical map and a genetic map, which is constructed by studying the passage of traits from one generation to another. When researchers compared the X genetic map, with its few hundred markers, to the physical map, they found an area in the middle that corresponds with a much longer stretch (17 Mb) of the physical map. "So this region is uneventful on the genetic map, whereas it contains a whole bunch of markers on the physical map," Schlessinger said. "But we don't know why the X chromosome should have this large area of poor recombination." He speculates that the answer may involve the X-inactivation locus, which in women turns off most of the genes on one copy of X, leaving the other to direct biological activities. The region of low recombination is on X's long arm, beginning near the X-inactivation locus and ending at a distinctive region that also is seen on the Y chromosome.

As YACs containing the relevant regions of X became available, Schlessinger and colleagues located genes for several diseases. These included an overgrowth disorder called Simpson-Golabi-Behmel syndrome and ectodermal dysplasia, which impairs the development of hair follicles, teeth, and sweat glands.◊

HUGO Calls for Patent Policy Changes

The Human Genome Organisation (HUGO) released a statement in May appealing for a change in international patent policies to encourage fast release and free availability of DNA sequencing data in the international Human Genome Project. The statement asks for a grace period to file for a patent after announcing a discovery, a policy already in effect in the United States.

HUGO also urged the U.S. Patent and Trademark Office (PTO) to rescind its position on granting patents for gene fragments called ESTs. In February, PTO Deputy Director Lawrence Goffney was quoted as saying that PTO decided to grant patents on ESTs based on their usefulness as probes, even through the biological function of a gene fragment may be unknown. Researchers who merely identify the fragment may thus have a prior claim when other uses of the fragment—based on its biological function—are identified in the future.◊

Cancer Gene Web Site Set to Debut

The Cancer Genome Anatomy Project (CGAP) of the National Cancer Institute (NCI) is set to go online (*http://www.ncbi.nlm.nih.gov/NCICGAP*) with the first installment in its cancer gene catalog. The goal of CGAP, which began last year, is to develop new diagnostic tools based on understanding molecular changes that underlie all cancers. These tools eventually will help doctors develop and select treatments designed to fight specific cancers. Europe is establishing a similar program called the Cancer Gene Expression Program.

One of CGAP's short-term goals is to compile an index of all genes that are turned on during the cancer process. Such an index would allow scientists, for the first time, to create complete genetic profiles differentiating among normal, precancerous, and malignant cells. The index will feature EST sequences from 45 cDNA libraries of lung, colon, prostate, ovarian, and breast tumors. Data about the tumors and source libraries also will be presented.

DOE researchers are generating the bacterial clones needed to hold ESTs and longer cDNA fragments, and the Integrated Molecular Analysis of Gene Expression consortium will make them available to all researchers. DOE support stands at about $1 million. NCI is putting forward $4 million for the index, $6 million to develop gene-analysis technologies and generate long cDNAs, and $10 million to develop clinical applications for this research data. [*Contact: robert_strausberg@nih.gov*] ◊

Homologs of Human Disease Genes Found in Model Organisms

Researchers at the National Center for Biotechnology Information (NCBI) used known DNA sequences of 70 human genes linked to such disorders as colon cancer and obesity to search public sequence databases for counterparts in yeast, bacteria, and roundworm. They found the highest number of matches in worm databases (36%) and expect to find more as the other half of the worm's genes are sequenced. Another 10% to 20% of the human genes had counterparts in bacteria and yeast. New functions were predicted for a number of disease genes. A paper reporting this work by Eugene Koonin and colleagues (NCBI) is published in *Proc. Natl. Acad. Sci. USA* **94**, 5831–36 (May 1997).◊

¶ Yeast Genome Directory

Published as a separate supplement to the May 29th issue of *Nature*, the Yeast Genome Directory contains papers on the sequence of all unpublished *Saccharomyces cerevisiae* chromosomes (IV, V, VII, IX, XII, XIII, XIV, XV, and XVI), along with gene maps and tables summarizing the yeast genome's structure and gene properties. The yeast genome sequence was released to the public last year on the Web (*http://genome-www.stanford.edu/Saccharomyces*). The directory also contains a "News and Views" article by Craig Venter (The Institute for Genomic Research) on the implications of yeast genome research (*Nature* Web site, *http://www.nature.com*).◊

Whitehead-MIT Teams With Consortium

A new consortium of companies has signed a 5-year arrangement with Eric Lander's group at Whitehead Institute–Massachusetts Institute of Technology (WIMIT) to develop automated systems for analyzing gene and protein activities in normal and diseased cells. Bristol-Myers Squibb, Affymetrix, and Millennium Pharmaceuticals are contributing $40 million in cash and equipment in return for commercial rights to technologies developed under the program.

WIMIT Resources

☛ **Mouse Data Release 13.** In March, the WIMIT Center for Genome Research (CGR) announced Data Release 13 of the Mouse Genomic Mapping Project (*http://www-genome.wi.mit.edu/cgi-bin/mouse/index*). CGR is constructing a mouse physical map to consist of 10,000 markers, many of which are single-sequence length polymorphisms from the mouse genetic map; random STSs are being added to reach the 10,000-marker total. In Release 13, data are included for 6018 markers successfully screened against a mouse YAC library. Some 4917 markers have been placed into singly linked physical map contigs, and 4191 are in doubly linked contigs.

☛ **Completed Mouse Genetic Map.** The July 1996 data release was WIMIT CGR's final release of the mouse genetic map. It reflects data represented in the March 14, 1996, issue of *Nature*; also see *Genetics* **131**, 4213–47 (1992) for descriptions of materials and methods used to construct both maps.

☛ **Human Physical Map Update.** The updated 11th release from the Human Physical Mapping Project is available via ftp (*ftp-genome.wi.mit.edu*) and the Web (*http://www.genome.wi.mit.edu*). This release contains radiation hybrid data for 14,308 STSs screened on the Genebridge4 Panel, representing an additional 1700 mapped ESTs since the October 1996 release. The current integrated map contains 11,722 ESTs and 24,230 total markers.◊

☛ Stanford Center's High-Resolution Map

A new STS-based radiation hybrid map of the human genome appears in the May issue of *Genome Research* [**7**(5), 422–33]. The new map, from the Stanford Human Genome Center, contains over 10,000 loci covering most of the human genome.◊

Merck Genomics Institute Established

On April 9, Merck & Co., Inc., announced the establishment of the Merck Genome Research Institute, Inc. (MGRI) to support development of scientific technology for linking human genetic traits and resolving biological function of disease genes. This not-for-profit institute will promote and sponsor projects for broadly applicable assays and methodologies to improve the accuracy and speed with which function can be associated with sequences of genetic information.

"We believe this institute's mission meets a current scientific need to translate our knowledge of gene sequence into function," said MGRI President C. Thomas Caskey. "In the spirit of the Merck Gene Index Project, the institute will ensure that such genetic technology is available to the entire biomedical community."

Awarding one of its first grants, MGRI announced in April that it plans to give $8 million to Lexicon Genetics, Inc. (The Woodlands, Texas) to create 150 new strains of "knockout" mice that have had particular genes disrupted. To help researchers determine the functions of the disabled genes, the mice will be made available at low cost to academic researchers.

Grant Research Areas

MGRI will fund grants or agreements to support research on gene function in the following general areas:

- **Merck Gene Index Expansion:** Identify tissues with disease associations and increase the usefulness of sequence data by targeting complete gene sequences.
- **Informatics:** Develop new algorithms for predicting gene function based on sequence content.
- **Disease Models:** Develop methods to create gene-targeted mutations for studying gene function in specific models such as bacteria, *Drosophila*, yeast, and mice.
- **Human Genetics:** Archive cells from informative families with common heritable diseases.

Grant applications should propose 1- to 2-year research projects that broadly address program objectives. Grants are expected to range from about $100,000 to $150,000 per project year and, in some cases, may be renewable. Large research projects and small pilot programs will be considered. [Contact for information, applications: Finley Austin, Administrative Director; MGRI; P.O. Box 4, WP 42-300; Sumneytown Pike; West Point, PA 19486; *mgri@merck.com*] ◊

☛ BSCS Genome Educational Modules

The Biological Sciences Curriculum Study (BSCS), with support from the DOE Ethical, Legal, and Social Issues program, is offering single copies of three educational modules at $4 each for shipping and handling. Designed for use by high school and college teachers and students, these modules provide background information and classroom activities and explore ethical, legal, and social issues related to the module's topic.

Currently Available Modules

- *Mapping and Sequencing the Human Genome: Science, Ethics, and Public Policy* (1992): Objectives and technology behind the Human Genome Project.
- *The Human Genome Project: Biology, Computers, and Privacy* (1996): Information technology as it relates to the genome project. Includes structures, uses, limitations, and implications of genomic databases through both print materials and software.
- *The Puzzle of Inheritance: Genetics and the Methods of Science* (1997): Uses non-traditional inheritance to demonstrate the nature and methods of science.

Orders: BSCS; Attn: Publications; 5415 Mark Dabling Blvd.; Colorado Springs, CO 80918-3842 (719/531-5550, Fax: -9104)

Module Under Development

Under a 24-month grant from DOE, BSCS is developing a fourth module, tentatively titled *Genes, Environment, and Human Behavior*. It will be distributed free of charge to all interested high school biology teachers and other educators in late 1998 or early 1999.◊

☛ Mapping Panel from Coriell

Coriell Institute is distributing the Human-Rodent Somatic Cell Hybrid Mapping Panel #2, Version 3, as cell cultures or DNA from the Human Genetic Mutant Cell Repository of the National Institute of General Medical Sciences. Version 3 consists of 24 human-rodent somatic cell hybrids, each retaining a single intact human chromosome. The new monochromosomal hybrids for chromosomes 6 (GM/NA11580) and 14 (GM/NA11535) retain the human chromosome in 100% and 92% of the cells, respectively. The panel has been characterized by (1) G-banded chromosome analysis, (2) in situ hybridization using biotinylated total human DNA, (3) Southern blot hybridization, and (4) PCR analysis. [Information: 800/752-3805 or 609/757-4848; Fax: 609/757-9737; *ccr@arginine.umdnj.edu*. Online catalog: *http://arginine.umdnj.edu/coriell/nigms.htm*] ◊

Resources and Publications

☛ FlyBase Updated

In its most recent update of May 1997, the *Drosophila* database FlyBase contains information about more than 46,000 alleles of some 14,000 genes. Many gene reports now link to reports about expression patterns and other features for associated proteins and transcripts. FlyBase presents descriptions of over 15,000 chromosomal aberrations as well as molecular maps and information about more than 1000 molecular constructs and 1000 transposons. The bibliography includes some 83,000 listings, many with links to associated genes and aberrations, and an address book lists over 5300 *Drosophila* researchers. Genotypes and ordering information for more than 13,000 *Drosophila* stocks are available.

Reports retrieved from gene searches have been enhanced by dividing alleles for each gene into "classical" and "in vitro" and references into "primary," "review," and "abstract." In its gene reports, FlyBase has extensive hyperlinks to other databases, most notably to sequence databank records, gene homologs from other organisms, and Medline citation records.

FlyBase Access

- WWW: *http://flybase.bio.indiana.edu*
- Gopher: *flybase.bio.indiana.edu*
- Ftp: *flybase.bio.indiana.edu* (type *anonymous* when prompted for login name; enter full e-mail address as password)

Inferences from automatically generated gene-location maps in the map directory are incorporated in individual gene reports and used for other map-based searches. Hardcopy versions of gene-order and annotated maps have been published as part of the FlyBase-edited *Drosophila* Information Service, volumes 78 and 79.

Interactive Fly

Flybase servers now provide access to the Interactive Fly, a database developed by Tom and Judy Brody (free-lance biologist and writer/editor). The Interactive Fly, which is designed to make *Drosophila* data more accessible to the nonspecialist, provides overviews of various *Drosophila* developmental and cellular processes integrated with information on vertebrate systems. The Interactive Fly is housed in the *Allied Data* section of FlyBase. Gene records are maintained with links to FlyBase, along with lists of genes participating in the described developmental and cellular processes.

FlyBase has developed a hierarchy of the Interactive Fly that links to specific pages, and gene lists link to individual gene records in both FlyBase and the Interactive Fly. Accessible in the FlyBase *Allied Data* and *Genes* directories, this hierarchy permits searches for genes grouped according to developmental and cellular pathways and functions.

Encyclopaedia of Drosophila

Some FlyBase data is merged with data from the Berkeley *Drosophila* Genome Project (BDGP) to form the *Encyclopaedia of Drosophila*, a collaboration between BDGP and FlyBase. [Version 3.0 (BDGP server), *http://shoofly.bdgp.berkeley.edu*; Version 2.0 CD-ROM, *eofd-sales@morgan.harvard.edu*]

Supported by grants from NIH and the British Medical Research Council, FlyBase is a consortium of *Drosophila* biologists and computer scientists at Harvard University, Cambridge University (U.K.), and Indiana University, where the main FlyBase server is located. Several mirror sites are located around the world. ◊

☛ ProDom Release 34.1

Release 34.1 of ProDom, the Protein Domain database, was constructed by clustering homologous segments derived from 65,376 nonfragmentary sequences present in SWISS-PROT 34 as of May 21. Some 18,086 multiple alignments and consensus sequences for homologous domain families are provided. The enhanced Web user interface links to and from PROSITE and PDB. Domain families can be searched by keyword, and graphical representations of domain arrangements facilitate structural interpretation of large protein families.

- **ProDom access:** *http://protein.toulouse.inra.fr/prodom.html* or *ftp://ftp.toulouse.inra.fr/pub/prodom/prodom34*
- **Requests:** *proquest@toulouse.inra.fr*
- **MultAlin program:** *http://www.toulouse.inra.fr/multalin.html*

All multiple alignments have been recalculated ab initio using the MultAlin program, and a new expertise procedure validates some domain boundaries. A sensitive homology search procedure scans all domain sequences in ProDom and retrieves matches with only one sequence for each domain family, thus drastically reducing output redundancy. The most significant matches are visualized graphically to assist with interpretation. For long queries, the former less-sensitive but faster search on consensus sequences is also provided. Users may choose between the classical NCBI BLAST 1.4.9 and the new WU-BLAST 2.0a13, which allows for gapped output (*http://blast.wustl.edu*). *[Jerome Gouzy and Daniel Kahn (Centre National de la Recherche Scientifique–Institut National de la Recherche Agronomique [INRA]) and Florence Corpet (INRA)]* ◊

¶ Gene Patenting

The *Bulletin of Medical Ethics* devoted its January issue to the patenting of human genes. Articles cover such topics as the ethics of patenting DNA; a Christian view of patenting; the legal protection of biotechnological inventions; patents, ethics, and improving health care; and the principle of nonownership of the human body.

Contact: Publications Subscription Dept.; Royal Society of Medicine Press, Ltd.; 1 Wimpole St.; London WIM 8AE (+44-171/290-2927 or -2928, Fax: -2929, *106223.3241@CompuServe.com*) ◊

Human Genome news

This newsletter is intended to facilitate communication, help prevent duplication of research effort, and inform persons interested in genome research. Suggestions are invited.

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

Managing Editor
Betty K. Mansfield
bkq@ornl.gov

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

Production Assistants
Murray Browne
Marissa D. Mills
Laura N. Yust

U.S. Department of Energy
Office of Health and Environmental Research
Ari Patrinos, Associate Director

Health Effects and Life Sciences Research Division, OHER
Marvin E. Frazier, Director
http://www.er.doe.gov/production/oher/hug_top.html

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

This newsletter is prepared at the request of the DOE Office of Health 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.◊

☛ TIGR Gene Index

Researchers now have free Internet access to the Human Gene Index (HGI) database released by The Institute for Genomic Research (TIGR). Designed to integrate research from human genome projects worldwide, the massive HGI holds full-length gene sequences, more than 600,000 ESTs, and 63,000 tentative human consensus sequences. The ultimate goal is to represent a nonredundant view of all human genes with data on their expression patterns, cellular roles, functions, and evolutionary relationships. HGI will also include links to genomic sequences, mapping data, 3-D structures, and literature references (*http://www.tigr.org/tdb/hgi/hgi.html*).

Anthony Kerlavage, director of the Department of Bioinformatics at TIGR, said of HGI, "Rather than searching for and locating bits and pieces of information in various places and databases where researchers often have to spend valuable time playing detective, the TIGR Human Gene Index brings together all available information in one location with various means of pinpointing specific areas of inquiry. It opens a vast universe of data to help fill in the blanks for scientists."

Researchers can search data using their own protein or nucleic acid sequences, search gene-product names assigned to many sequences, and examine detailed expression data associated with the sequences. Complex cross-checking of information is permitted through database queries. Each sequence is also tagged to clones in the American Type Culture Collection (ATCC), allowing researchers to obtain clones through the TIGR-ATCC Special Collection.

Until HGI's release, much of the TIGR information was available only to academic researchers who signed written agreements with SmithKline Beecham and Human Genome Sciences in an arrangement that has since expired.◊

¶ AAAS Book

Exploring Public Policy Issues in Genetics, a 224-page paperback book edited by Mark Frankel (American Association for the Advancement of Science), is a compilation of presentations and resource material prepared for the AAAS Congressional Seminar Series on the Human Genome Project. Funded in part by the DOE Human Genome Program, the seminars were held in May and June of 1996.

Material in the book is organized into seven sections: Introduction; Scope of the Human Genome Project; Patenting Our Genes; Privacy, Confidentiality, and Genetic Information; Genetics and Behavior; Resources; and Seminar Speakers.

Individual copies: Kamla Butaney; Directorate for Science and Policy; AAAS; 1200 New York Ave. NW; Washington, DC 20005 (202/326-6792) ◊

¶ JAMWA Genetics Issue

The Winter 1997 issue of *Journal of the American Medical Women's Association* (*JAMWA*) is devoted to genetics and women's health. The issue includes articles on the Human Genome Project from NIH and DOE perspectives, clinical molecular genetic testing, genetic identification of children of the disappeared in Argentina, and genetic-susceptibility testing for breast and ovarian cancer. Other articles are on communicating about chromosomes, genetic counseling and prenatal diagnosis, women's ethics in genetics, deficiencies in obtaining consent for genetic testing, genetic discrimination and health insurance, and the revival of eugenics in American popular culture.

Authors include Francis Collins and Leslie Fink [both at NIH National Human Genome Research Institute (NHGRI)], Ari Patrinos and Daniel Drell (both of the DOE Human Genome Program), Linda Brzustowicz (Rutgers University) and Bernice Allitto (Genzyme Genetics), Victor Penchaszadeh (Albert Einstein College of Medicine), Barbara Bowles Biesecker and Lawrence Brody (both at NHGRI), Rayna Rapp (New School for Social Research), Diana Puñales-Morejon (Beth Israel Medical Center), Dorothy Wertz (Shriver Center), Lori Andrews (Chicago-Kent College of Law), Karen Rothenberg (University of Maryland School of Law), and Dorothy Nelkin (New York University) and Susan Lindee (University of Pennsylvania, Philadelphia).

Back issues: JAMWA; 801 North Fairfax St.; Alexandria, VA 22314 (703/838-0500, Fax: /549-3864) ◊

¶ Booklet on Genetics

Your Genes, Your Choices: Exploring the Issues Raised by Genetic Research is a color-illustrated, 90-page booklet designed for adults with low literacy skills. Funded by the DOE Human Genome Program and written on the 6th- to 8th-grade level by Catherine Baker, the booklet describes the Human Genome Project; the science behind it; and selected ethical, legal, and social issues that can affect individuals in their daily lives.

As part of the Science + Literacy for Health Project directed by Maria Sosa for the American Association for the Advancement of Science, the booklet is accompanied by a curriculum, implementation framework, short video on genetics, database of resources, and fact sheets to assist others in preparing easy-to-read materials about the Human Genome Project. The goal is to disseminate project materials to libraries and community organizations carrying out literacy programs throughout the United States.

The booklet is on the Web (*http://www.nextwave.org/ehr/books/index.html*), and individual copies are available for a nominal fee from Sosa (202/326-6453, Fax: /371-9849, *msosa@aaas.org*). ◊

Notice to DOE Contractors, Grantees

The sixth DOE Human Genome Program Contractor-Grantee workshop will be held November 9–13 in Santa Fe, New Mexico. At least one investigator from each funded project is expected to attend the entire meeting and represent the project at poster sessions. Some projects also will be represented in platform presentations. More information on registration and abstracts will be forthcoming from Sylvia Spengler; Human Genome Program Coordination; 459 Donner Laboratory; Lawrence Berkeley National Laboratory; Berkeley, CA 94720 (510/486-4879, Fax: -5717, *sjspengler@lbl.gov*).◊

☛ CBIL Web Site

The Computational Biology and Informatics Laboratory (CBIL) at the University of Pennsylvania creates and maintains software and databases for computational biology. The CBIL Web site (*http://agave.humgen.upenn.edu*) includes the following:

- bioWidgets for Java, a collection of classes for rapid development of graphics applications (applets) in the molecular biology domain.
- Genome Annotation and Information Analysis (GAIA) storage system for genomic sequence and its annotation.
- Database on the specific regulation of a number of genes in muscle tissue (MTIR).
- Human Genome Center for Chromosome 22 interface.
- Databases developed by CBIL for public use.
- Tools for computational biology.◊

¶ NFCR Newsletter

The quarterly *National Flow Cytometry Resource Newsletter*, edited by Carolyn Bell-Prince, began publication at Los Alamos National Laboratory in July 1996. Available electronically and in hard copy, the newsletter contains a Director's Chair section by James Jett, articles on flow cytometry, meetings and course announcements, and notices of open positions. Current and back issues are accessible on the Web (*http://telomere.lanl.gov/NFCR/newsletter.html*). [Contact: Carolyn Bell-Prince (505/667-2836, Fax: /665-3024, *cbp@telomere.lanl.gov*)] ◊

Resources

Genetics Networks

To aid people who are seeking information about specific genetic conditions, contact with others impacted by a genetic condition, or professionals providing genetic services, *HGN* is making available the following list of resources.

All the entities below receive some funding from the Genetics Services Branch of the Maternal and Child Health Bureau, U.S. Department of Health and Human Services.

Alliance of Genetic Support Groups

35 Wisconsin Circle, Ste. 440; Chevy Chase, MD 20815

800/336-4363 or 301/652-5554, Fax: /654-0171, *alliance@capaccess.org*

http://medhelp.org/www/agsg.htm

The Alliance was established in 1986 as a nonprofit organization dedicated to helping individuals and families affected by genetic disorders. The Alliance Web site lists publications (many of them free), upcoming events such as meetings and conferences, and some 301 support groups in the *Directory of National Genetic Voluntary Organizations*. This directory, which gives full contact information for each support group and links to other Web sites, is also available in hard copy.

The Council of Regional Networks (CORN) for Genetic Services

Emory University; Pediatrics-Medical Genetics; 2040 Ridgewood Drive; Atlanta, GA 30322

404/727-1475, Fax: -1827, *cfh@rw.ped.emory.edu*

http://www.cc.emory.edu/PEDIATRICS/corn/corn.htm

CORN was formed in 1985 to coordinate 10 regional networks representing all 50 states; District of Columbia, Puerto Rico, and the Virgin Islands; national sickle cell disease programs; and the Alliance of Genetic Support Groups. CORN brings together network representatives to facilitate communication and planning for genetic services and to address national public health priorities in genetics.

Regional Genetics Networks

Within their regions, the ten networks coordinate genetic services; promote communication among genetic professionals and consumers through CORN membership, network newsletters, and meetings and other events; share resources; and promote education and awareness of genetic disorders.

Each network will answer inquiries about genetic evaluation and counseling services.

Genetics Network of New York, Puerto Rico, and the Virgin Islands (GENES)

Genetic Services Program; Wadsworth Center; Empire State Plaza, Room E299; P.O. Box 509; Albany, NY 12201-0509

518/486-2215, Fax: /473-1733, *kbh02@health.state.ny.us* or *kxg03@health.state.ny.us*

Great Lakes Regional Genetics Group (GLaRGG)

328 Waisman Center; 1500 Highland Avenue; Madison, WI 53705-2280

608/265-2907, Fax: /263-3496, *elbaum@waisman.wisc.edu*

http://www.waisman.wisc.edu/glargg

States served: Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin

Great Plains Genetics Service Network (GPGSN)

Pediatrics-Medical Genetics; University of Iowa; Iowa City, IA 52242-1083

319/356-4860, Fax: -3347, *dolores-nesbitt@uiowa.edu*

http://www.unmc.edu/mrimedia/bradweb/gpgsn.html

States served: Arkansas, Iowa, Kansas, Missouri, North Dakota, Nebraska, Oklahoma, South Dakota

Mid-Atlantic Regional Human Genetics Network (MARHGN)

260 South Broad Street, Ste. 1000; Philadelphia, PA 19102-3865

215/985-6759, Fax: -6763, *marhgn1@aol.com*

States Served: District of Columbia, Delaware, Maryland, New Jersey, Pennsylvania, Virginia, West Virginia

Mountain States Regional Genetic Services Network (MSRGSN)

Colorado Department of Health; FCHS-MA-A4; 4300 Cherry Creek Drive South; Denver, CO 80222-1530

303/692-2423, Fax: /782-5576, *joyce.hooker@state.co.us*

http://www.ahsc.arizona.edu/~msrgsn/msrgsnhp.htm

States Served: Arizona, Colorado, Montana, New Mexico, Utah, Wyoming

New England Regional Genetics Group (NERGG)

P.O. Box 670; Mt. Desert, ME 04660

207/288-2704, Fax: -2705, *76363.3114@compuserve.com*

States Served: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont

Pacific Northwest Regional Genetics Group (PacNoRGG)

CDRC-Clinical Services Building; 901 E. 18th Avenue; Eugene, OR 97403-5254

541/346-2610, Fax: -2624, *kerry_silvey@ccmail.uoregon.edu*

States served: Alaska, Idaho, Oregon, Washington

Pacific Southwest Regional Genetics Network (PSRGN)

Genetic Disease Branch; California Department of Health Services; 2151 Berkeley Way, Annex 4; Berkeley, CA 94704

510/540-2852, Fax: -2095, *hkuliopu@genetic.dhs.cahwnet.gov*

States served: California, Hawaii, Nevada

Southeastern Regional Genetics Group (SERGG)

Emory University; Pediatrics-Medical Genetics; 2040 Ridgewood Drive; Atlanta, GA 30322

404/727-5844, Fax: -5783, *mrl@rw.ped.emory.edu*

http://www.cc.emory.edu/PEDIATRICS/sergg/sergg.htm

States served: Alabama, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee

Texas Genetics Network (TEXGENE)

Bureau of Children's Health; Texas Department of Health; 1100 West 49th Street; Austin, TX 78756-3199

512/458-7111 Ext. 2129, Fax: -7421, *jlivingston@wc1.tdh.state.tx.us*

Other Groups

National Health Information Center

http://nhic-nt.health.org

National Organization of Rare Disorders

http://www.NORD-RDB.com/~orphan

Genetic Syndrome Support Group Listings

http://members.aol.com/dnacutter/sgroup.htm

Compiled by Laura Yust, June 1997

Web Resources

From time to time *HGN* will list pertinent uniform resource locators (URLS) for sites of interest to our readers. Although some of these sites are not related directly to molecular biology or biology, their subject matter is important to many investigators in the Human Genome Project and those using its data.

Biology

Life Science Dictionary of 3700 terms: *http://biotech.chem.indiana.edu/search/dict-search.html*

Database of Genome Sizes: Scientific and common names, classifications, and references: *http://www.cbs.dtu.dk/databases/DOGS/index.html*

Data Analysis and Databases

ABIView V.1.0 for displaying ABI trace and basecall data in Windows: *http://www.paranoia.com/~dhk/abiinfo.html*

Bioinformatics and computational molecular biology, *GENE-COMBIS* journal: *http://www.elsevier.nl/journals/genecombis/Menu.html*

CEPH Genotype Database V8.1 (over 2.5 million family genotypes): *http://www.cephb.fr/cephdb* or *ftp://ftp.cephb.fr/pub/ceph_genotype_db*

Chromosome 11 Sequencing Project (data, resources, maps): *http://mcdermott.swmed.edu:80/datapage*

GENET database (gene network organization): *http://www.iephb.ru/~spirov/genet00.html*

Homeobox gene-classification issues, information: *http://copan.bioz.unibas.ch/homeo.html*

Institute of Cytology and Genetics, Siberian Regional Center for Gene Informatics: *http://benpc.bionet.nsc.ru/SRCG*

Marshfield Clinic, Wisconsin, human genetics project: *http://www.marshmed.org/genetics*

Molecular Bioinformatics of Gene Regulation project: *http://transfac.gbf-braunschweig.de*

RESTSITE v.1.2 to process restriction-site or restriction-fragment text files for evolutionary analysis: *http://www.genome.wi.mit.edu/~jmiller/restsite.htm*

SEQIO for reading and writing sequence files: *ftp://ftp.cs.ucdavis.edu/pub/strings/seqio.tar.gz* or *http://wwwcsif.cs.ucdavis.edu/~knight/seqio.html*

Sequence Analysis and Comparison Bibliography of 1700 theory and methodology papers: *http://www.pitt.edu/~csna*

Sequence tagged connector (BAC end-sequencing) approach to genomic sequencing: *http://serac.mbt.washington.edu/index.html*

TREMBL protein sequence database supplementing SWISS-PROT: *http://srs.ebi.ac.uk:5000*

Webcutter for restriction mapping nucleotide sequences: *http://www.medkem.gu.se/cutter*

ELSI and Education

Bioethics news and trends, *Eubios Journal of Asian and International Bioethics*: *http://www.biol.tsukuba.ac.jp/~macer/NBB.html*

DNA Learning Center, including Student Allele Database: *http://darwin.cshl.org*

Genetic Diseases

Canadian Genetic Diseases Network: *http://www.cgdn.generes.ca*

Duchenne/Becker Muscular Dystrophy (dystrophin gene data, physical map, mutations): *http://ruly70.MedFac.LeidenUniv.nl/~duchenne*

Rare Diseases in Children (resources, references, contacts, support): *http://mcrcr4.med.nyu.edu/~murphp01/homenew.htm*

Genomic Applications

Biotechnology Information Institute, *Federal Bio-Technology Transfer Directory Database, Antiviral Agents Bulletin*, Biopharm Registry and Reference Project: *http://www.bioinfo.com/biotech*

Canadian Institute of Biotechnology (information dissemination, technology transfer): *http://www.biotech.ca*

Discussion group on bioremediation for reducing environmental contamination: *http://biogroup.gzea.com/about.htm*

DNA Vaccine (progress review of clinical trials): *http://www.genweb.com/Dnavax/dnavax.html*

HotMolecBase database (proteins and other medically promising, biologically active molecules): *http://bioinfo.Weizmann.ac.il/hotmolecbase*

National Center for Infectious Diseases, *Emerging Infectious Diseases* journal: *http://www.cdc.gov/ncidod/EID/eid.htm*

Laboratory Tools

PCR Jump Station, theory and applications: *http://www.apollo.co.uk/a/pcr*

Plasmid Processor tool for plasmid presentation: *http://www.hytti.uku.fi/~kiviraum/plasmid/plasmid.html*

Nonhuman Genomes

Bugs 'N Stuff microbial genome site: *http://www.ncgr.org/microbe*

DOE Microbial Genome Project: *http://www.er.doe.gov/production/oher/mig_top.html*

Flybase (*Drosophila* genetics, biology): *http://flybase.bio.indiana.edu*

Genome sequence status, including microbial, eukaryotic, organelle, phage, and virus genomes: *http://www.mcs.anl.gov/home/gaasterl/genomes.html*

Microbial genomes, completed and in progress: *http://www.tigr.org/tdb/mdb/mdb.html*

MAGPIE (automated platform for collecting computed genome sequence data): *http://www.mcs.anl.gov/home/gaasterl/magpie.html*

Sacch3D (extension of the *Saccharomyces* Genome Database), presenting structure information for proteins in the yeast genome: *http://genome-www.stanford.edu/Sacch3D*

Seldin-DeBry Human-Mouse Homology Map of genes from human and mouse sources: *http://www.ncbi.nlm.nih.gov/Omim/Homology*

Structure and Function

Chromatin structure and function: *http://rampages.onramp.net/~jrbone/chrom.html*

Newsgroup on genome and chromatin structure and function: *http://www.bio.net/hypermail/GENSTRUCTURE*

Newsgroup to study 3-D macromolecule structure using nuclear magnetic resonance spectroscopy: *http://www.bio.net/hypermail/STRUCTURAL-NMR*

Newsgroup to survey structural biology and protein engineering in Europe: *http://www.cryst.bbk.ac.uk/CEC/eupage.html*

Protein Structure Prediction Center (Lawrence Livermore National Laboratory) to advance methods of identifying protein structure from sequence: *http://PredictionCenter.llnl.gov/Center.html*

ProAnWin (Protein Analyst for Windows) program for analyzing protein sequences and structure, studying structure-activity relationships in protein families, and designing protein-engineering experiments: *ftp://ftp.ebi.ac.uk/pub/software/dos/proanwin*

2-D PAGE Databases for Proteome Analysis in Health and Disease: *http://biobase.dk/cgi-bin/celis*

Compiled by Anne Adamson, June 1997

Calendar of Genome and Biotechnology Meetings*

More comprehensive lists of genome-related meetings and organizations offering training are available at *http://www.ornl.gov/hgmis* or from HGMIS (see p. 10 for contact information). A list of training organizations is on the next page.

August 1997

3–8. Society for Industrial Microbiology Annu. Meeting; Reno, NV [SIM, 703/691-3357, Fax: -7991]

7–8. Financing and Strategic Alliances for Emerging Genomic Companies; Boston [J. Feigenbaum, 212/366-3249, Fax: /645-4490; *globalbr@ix.netcom.com*]

7–8. Microfabrication and Microfluidic Technol.: Advances in Miniaturization of Bioanalytical Devices; San Francisco [IBC, 508/481-6400, Fax: -7911; *reg@ibcusa.com*; *http://www.ibcusa.com*]

10–15. Human Molecular Genetics; Newport, RI [GRC, 401/783-4011, Fax: -7644; *grc@grcmail.grc.uri.edu*; *http://www.grc.uri.edu*]

12–17. Yeast Cell Biology; Cold Spring Harbor, NY [CSHL, 516/367-8346, Fax: -8845; *meetings@cshl.org*; *http://www.cshl.org*]

18–19. Post-Genomic Analysis of Therapeutic Targets; La Jolla, CA [NMHCC, 617/505-8000, Fax: /270-6004; *register@nmhcc.com*]

24–29. 17th Intl. Congress of Biochemistry and Molecular Biol. in conj. with ASBMB 1997 Annu. Meeting; San Francisco [FASEB, 301/530-7010, Fax: -7014; *kmirabal@faseb.org*; *http://www.faseb.org*]

27–31. EMBL Mouse Molecular Genetics Meeting; Heidelberg, Germany [I. Fatscher, +49-6224/929-025, Fax: -026; *fatscher@embl-heidelberg.de*; *http://www.embl-heidelberg.de/CourseInfo/index.html*]

September 1997

13–16. 9th Intl. Genome Sequencing and Analysis Conf.; Hilton Head, SC [TIGR, 301/838-3515, Fax: -0229; *debbieg@tigr.org*; *http://www.tigr.org*]

22–23. Seizing Opportunities in Emerging Biochip Technologies—Scientific Breakthroughs and Commercial Applications; San Diego [IBC, see contact: Aug 7–8]

29–30. Cancer and Genetic Screening; Washington, DC [CHI, 617/630-1300, Fax: -1325; *chi@healthtech.com*; *http://www.healthtech.com/conferences*]

29–Oct. 1. Bioinformatics: Tools for Drug Discovery and Design; San Diego [IBC, see contact: Aug. 7–8]

October 1997

12–15. 5th Intl. *E. Coli* and Small Genomes Conf.; Snowbird, UT [ASM, 202/942-9248, Fax: -9340; *meetingsinfo@asmusa.org*; *http://www.asmusa.org*]

12–16. 11th Intl. Mouse Genome Conf.; St. Petersburg, FL [IMGS, D. Miller, 716/845-4390, Fax: -8169; *dmiller@mcbio.med.buffalo.edu*; *http://mcbio.med.buffalo.edu/11imgc*]

16–19. Gene Therapy; Hilton Head, SC [see contact: Sept. 13–16]

27–28. 4th Intl. Mutation Database Meeting; Baltimore [HUGO, Fax: +44-171/935-8341; *hugo@hugo-europe.org.uk*]

28–Nov. 1. ASHG; Baltimore [M. Ryan, 301/571-1825, Fax: /530-7079]

November 1997

1–4. Conf. on Computational Genomics; Herndon, VA [see contact: Sept. 13–16]

2–5. Functional Genomics: Application of Genomic Technologies to the Understanding of Biological Systems; Newport, RI [see contact: Sept. 29–30]

9–13. 6th DOE Human Genome Program Contractor-Grantee Workshop; Santa Fe, NM [S. Spengler, 510/486-4879, Fax: -5717; *sjspengler@lbl.gov*]

16–17. Gene Mutational Analysis; Bermuda [see contact: Sept. 29–30]

21–22. Science Policy Symp.; Washington, DC [see contact: Sept. 13–16]

December 1997

5–7. Genetic Information: Acquisition, Access, and Control; Blackpool, U.K. [L. Smith, +44-1772/892-255, Fax: -938; *lsmith1@uclan.ac.uk*]

15–17. Conf. on Scientific and Technical Data Exchange and Integration; Bethesda, MD (abs. deadline: Aug. 1) [P. Uhlir, 202/334-2421, Fax: -1684; *codataco@nas.edu*; *http://www.nas.edu/cpsma/codata.htm*]

January 1998

5–8. Pacific Symp. on Biocomputing; Kapalua, HI [L. Hunter, 301/496-9303, Fax: -0673, *hunter@nlm.nih.gov*; *http://www.cgl.ucsf.edu/psb*]

12–13. Gene Quantification; San Diego [see contact: Sept. 29–30]

31–Feb 4. Microbial Genomes II: Sequencing, Functional Analysis, and Comparative Genomics; Hilton Head, SC [see contact: Sept. 13–16]

February 1998

7–12. Bacterial Chromosomes; Santa Fe, NM [Keystone Symp., 800/253-0685 or 970/262-1230, Fax: -1525; *keystone@symposia.com*; *http://www.colorado.net/symposia*]

9–11. Human Genome Project: Commercial Implications; San Francisco [see contact: Sept. 29–30]

11–13. FISH: New Technologies and Clinical Applications; Steamboat Springs, CO (abs. deadline: Dec. 1) [B. Wolf-Ledbetter, 773/702-1375, Fax: /834-0659; *fish98@genetics.uchicago.edu*; *http://www.genes.uchicago.edu*]

12–13. Gene Functional Analysis; San Francisco [see contact: Sept. 29–30]

12–17. Genome Seminar at AAAS Annu. Meeting; Philadelphia [see contact: Sept. 13–16]

27–Mar. 1. 5th Joint Clinical Genetics Meeting: 29th Annu. MOD Clinical Genetics Conf. and 4th Annu. ACMG Meeting; Los Angeles [M. Greenfield, 301/530-7127, Fax: /571-1895; *mgross@genetics.faseb.org*]

May 1998

1–4. Biomedicine '98; AAP/AFMR/ASCI; Washington, DC [K. Greenwood, 202/429-5161, Fax: /223-4579; *karen_greewood@sba.com*] ◊

*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.

Training Events*

August 1997

3–15. MBL Workshop on Molecular Evolution; Woods Hole, MA [C. Hamel, 508/548-3705, Fax: /457-1924; *admissions@mbl.edu*; *http://www.mbl.edu*]

17–21. Symp. at 8th European Congress on Biotechnol.: Databases in Molecular Biology and Biotechnol.; Budapest [S. Pongor, +39-40/375-731, Fax: /266-555; *pongor@icgeb.trieste.it*; *http://www.icgeb.trieste.it*]

18–20. 3rd Nordic Workshop on Genetic Algorithms; Helsinki [J. Alander, +358-6/3248-444, Fax: -467; *Jarmo.Alander@uwasa.fi*]

September 1997

15–16. 15th Annu. Biotech Patent Forum; Rockville, MD [ATCC, 301/231-5566, Fax: /816-4364; *workshops@atcc.org*; *http://www.atcc.org:80/workshops/workshop.html*]

17–19. Cytogenetics and FISH; Rockville, MD [see contact: Sept. 15–16]

October 1997

12–25. Genetic Approaches in Complex Heart, Lung, and Blood Diseases; Bar Harbor, ME [Jackson Laboratory, 207/288-6262; *education@jax.org*; *http://www.jax.org*]

14–27. Positional Cloning: Contig to Candidate Gene; Cold Spring Harbor, NY [CSHL, 516/367-8345, Fax: -8845; *meetings@cshl.org*; *http://www.cshl.org*]

24–28. The New Frontier (incl. Short Course: Cancer Genetic Counseling: A New Era Unfolds); Baltimore [NSGC, B. Leopold, 610/872-7608, Fax: -1192; *nsgc@aol.com*; *http://members.aol.com/nsgcweb/ec16.htm*]

28. Intl. Workshop on Chromosome 15; Baltimore [C. Morton, 617/732-7980; *ccmorton@bics.bwh.harvard.edu* or R. Nicholls, 216/368-3331; *rxn19@po.cwru.edu*]

31–Nov. 1. Advanced Methods in Fluorescent Genotyping; Foster City, CA [Perkin-Elmer, 800/874-9868 ext 5015; *http://www2.perkin-elmer.com*]

November 1997

4–7. Microbial DNA Fingerprinting; Rockville, MD [see contact: Sept. 15–16]

4–17. Molecular and Cell Biology of *S. Pombe* and Other Yeasts; Cold Spring Harbor, NY [see contact: Oct. 14–27]

6–11. Computational Genomics; Cold Spring Harbor, NY [see contact: Oct. 14–27]

10–14. Advanced Recombinant DNA: Techniques and Applications; Rockville, MD [see contact: Sept. 15–16]

12–22. DNA Sequencing: Advanced Approaches Automated Methods and Analysis; Heidelberg, Germany [W. Ansorge, +49-6221/387-355, Fax: -306; *ansorge@embl-heidelberg.de*; *http://www.embl-heidelberg.de/CourseInfo/index.html*]

12–25. Genetic Approaches to Heart, Lung, and Blood Disease; Rockville, MD [see contact: Sept. 15–16] ➨

Training continued

17–20. PCR Applications/Cycle DNA Sequencing; Rockville, MD [see contact: Sept. 15–16]

20–22. Ethics and Genetics: Advanced European Bioethics Course; Nijmegen, Netherlands [B Gordijn, +31-24/361-5320, Fax:/354-0254; *b.gordijn@efg.kun.nl*] ◊

Sponsors of Genome Related Training

American Medical Informatics Association (301/657-1291, *http://amia2.amia.org*)

American Type Culture Collection (800/359-7370, Fax: 301/816-4369, *http://www.atcc.org*)

Barr Enterprises (301/898-3772, Fax: -5596)

BioConferences International, Inc. (914/834-3100, Fax: -4329)

Cambridge Healthtech Institute (617/630-1300, Fax: -1325, *http://www.healthtech.com/conferences*)

Catholic University of America (202/319-6161, Fax: -4467, *http://www.cua.edu/www/catc*)

Cold Spring Harbor Laboratory (516/367-8346, Fax: -8845, *http://www.cshl.org*)

Discrete Mathematics and Computational Science (908-445-5932, Fax: -5932, *http://dimacs.rutgers.edu/Workshops/index.html*)

Exon-Intron, Inc. (800/407-6546 or 410/730-3983, *http://www.dnatech.com/exonhm.htm*)

Federation of American Societies for Experimental Biology (301/530-7010, Fax: -0650, *http://www.faseb.org*)

Gordon Research Conferences (401/783-4011, Fax: -7644, *http://www.grc.uri.edu*)

International Business Communications (508/481-6400, Fax: -7911, *http://www.ibcusa.com*)

International Centre for Genetic Engineering and Biotechnology (+39-40/37-571, Fax: /226-555, *http://www.icgeb.trieste.it*)

Jackson Laboratory (207/288-6262, Fax: -6051, *http://www.jax.org/training/documents*)

Keystone Symposia (800/952-9166, Fax: 970/262-1525, *http://www.colorado.net/symposia*)

Life Technologies, Inc. (800/952-9166, Fax: 301/610-8011, *http://www.lifetech.com*)

Marine Biological Laboratory (508/548-3705, *http://www.mbl.edu*)

Oncor, Inc. (800/776-6267, *http://www.oncor.com/courses.htm*)

Perkin-Elmer (800/762-4000, Fax: 203/762-6000, *http://www2.perkin-elmer.com*)

Pittsburgh Super Computing (412/268-5206, Fax: -8200, *http://www.psc.edu/biomed/workshops.html*)

The Institute for Genomic Research (301/838-0200, Fax: -0208, *http://www.tigr.org*)

Wellcome Summer Schools (+44-171/403-6998, Fax: /407-5281, *http://www.umds.ac.uk/wlmg*) ◊

Funding Opportunities

NIH NHGRI

National Research Service Award Fellowships

Topic: To engage in research relevant to the Human Genome Project. Postdoctoral, senior postdoctoral, and minority predoctoral fellowships are available to U.S. citizens or permanent residents; research in ethical, legal, and social issues (ELSI) is not open to predoctoral students through this program.

- **Applications for postdoctoral and senior postdoctoral:** December 5, April 5, and August 5
- **Applications for minority predoctoral:** May 1 and November 15
- **Contacts:** ELSI topics, Eric Meslin (301/402-4997, *eric_meslin@nih.gov*); all other topics, Bettie Graham (301/496-7531, *bettie_graham@nih.gov*) ◊

Mapping, Sequencing, Analysis Technologies

PA-97-044

Topic: Develop new or significantly improved technologies to facilitate and accelerate genome mapping, sequencing, and analysis. Resources produced will be used to further the study of diseases and other biological phenomena. (Supported by R01, R29, and P01 grants.)

Pilot Projects, PA-97-045

Topic: Pilot projects or feasibility studies to develop technologies as in PA-947-044. The purpose of this program is to encourage high-risk, high-potential technologies that are not yet developed fully enough to compete successfully for a standard R01 grant. (Supported by the exploratory R21 mechanism.)

- **Application receipt dates for both topics:** February 1, June 1, and October 1
- **Contact for both topics:** Bettie Graham (see contact above) ◊

Low-Cost, High-Accuracy DNA Sequencing Technologies

RFA HG-97-002

Topic: To stimulate research on next-generation technologies that have the potential to reduce significantly the cost of high-accuracy genomic DNA sequencing.

- **Letter of intent:** August 1, 1997
- **Application receipt:** October 16, 1997
- **Contact:** Jeffrey Schloss (301/496-7531, *jeff_schloss@nih.gov*) ◊

Calendar Submissions

Items for the *Human Genome News* (*HGN*) "Calendar of Genome and Biotechnology Meetings" or "Training Events" should be submitted to HGMIS by mail, e-mail, or fax as soon as information is finalized (see page 10 for HGMIS contact information). *HGN* is published quarterly.◊

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 Health and Environmental Research Human Genome Program

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

Alexander Hollaender Distinguished Postdoctoral Fellowships

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

- Next deadline: January 1998
- Contact: Barbara Dorsey, Oak Ridge Institute for Science and Education (423/576-9975, Fax: /241-5219, 800/569-7749; *alexpgm@orau.gov*)

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 laboratory.

- Contact: Christine Trance; Alfred P. Sloan Foundation; 630 Fifth Ave., Ste. 2550; New York, NY 10111 (212/649-1649, Fax: /757-5117, *trance@sloan.org*)

NIH National Human Genome Research Institute

- NHGRI program: 301/496-7531, Fax: /480-2770, *http://www.nhgri.nih.Gov/About_NHGRI*
- Program announcements: *http://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 program to foster transfers between research institutions and small businesses.

Contacts:

- Kay Etzler; DOE SBIR Program, ER-16; DOE; Washington, DC 20585 (301/903-5867, Fax: -5488, *kay.etzler@oer.doe.gov*); SBIR, *http://sbir.er.doe.gov/sbir.htm*; STTR, *http://sttr.er.doe.gov/sttr.htm*
- Bettie Graham (see contact, NHGRI). NIH SBIR due April 15, August 15, and December 15. STTR, April 1, August 1, and December 1

SBIR/STTR conferences: Washington, DC (Oct. 14–16); Phoenix, AZ (Oct. 27–29). Conference information: 360/683-5742 ◊

Human Genome Management Information System Subscription/Document Request* (Vol. 8, Nos. 3 and 4)

Name _____
(First) (MI) (Last)

Affiliation _____

Department/Division _____

Street/P.O. Box/Building _____

City/State/Zip Code _____

Country _____ **Area of Interest** _____

Phone _____ **Fax** _____

E-Mail Address (important to list if you have one) _____

1. ___ *Human Genome News* ___New Subscriber ___Change of Name/Affiliation/Address (circle all that apply) ___Drop Subscription
2. ___ *To Know Ourselves* (DOE) ___DOE *Primer on Molecular Genetics* (also see *http://www.ornl.gov/hgmis/publicat/publications.html*)
3. ___ Reprint of "A New Five-Year Plan for the U.S. Human Genome Project" (Science, October 1, 1993) by Francis Collins and David Galas

***Please type, print carefully, or enclose a business card to ensure efficient shipping. To change name/address/affiliation or drop your subscription to *Human Genome News*, enclose your current *HGN* address label. Send to HGMIS address shown below and on p. 10.**

SELECTED ACRONYMS

AAAS Am. Assoc. for the Advancement of Sci.
AAP Assoc. of Am. Physicians
ACMG Am. Col. of Med. Genetics
AFMR Am. Federation for Medical Research
ASBMB Am. Soc. for Biochem. and Mol. Biol.
ASCI Am. Soc. for Clinical Investigation
ASHG Am. Soc. for Hum. Genet.
ASM Am. Soc. for Microbiol.
ATCC Am. Type Culture Collection
BAC bacterial artificial chromosome
bp base pair
CHI Cambridge Healthtech Inst.
cM centimorgan
CSHL Cold Spring Harbor Lab.
DOE Dept. of Energy
ELSI ethical, legal, and social issues
EMBL European Mol. Biol. Lab.
EMBO European Mol. Biol. Org.
EST expressed sequence tag
FASEB Fed. of Am. Soc. for Exp. Biol.
FISH fluorescence in situ hybridization
GRC Gordon Res. Conf.
HGMIS Human Genome Management Information System
HUGO Human Genome Organisation
IBC Intl. Business Communications
ICGEB Intl. Centre for Genetic Engineering and Biotech.
IMGS Intl. Mammalian Genome Society
MBL Marine Biological Laboratory
MOD March of Dimes
NIH Natl. Institutes of Health
NMHCC National Managed Health Care Congress
NSGC Natl. Soc. of Genetic Counselors
ORF open reading frame
PAC P1 artificial chromosome
PCR polymerase chain reaction
SIM Soc. for Indus. Microbiol.
STS sequence tagged site
TIGR The Inst. for Genomic Res.
WWW World Wide Web
YAC yeast artificial chromosome

OAK RIDGE NATIONAL LABORATORY • MANAGED BY LOCKHEED MARTIN ENERGY RESEARCH CORP. • FOR THE U.S. DEPARTMENT OF ENERGY

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

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

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