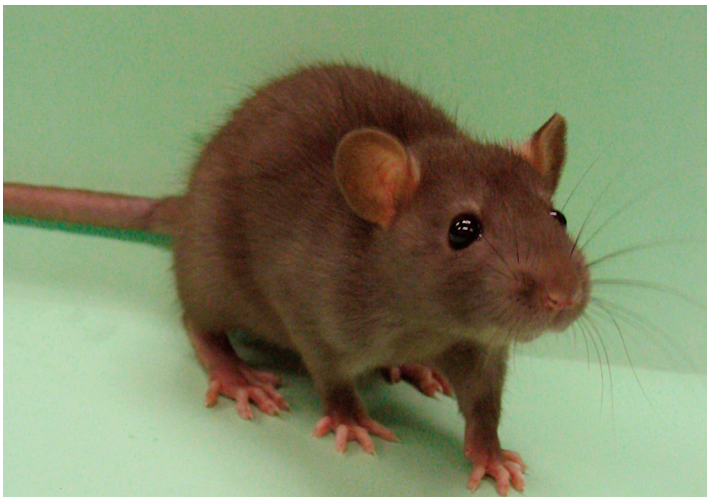
In science, the important thing is to modify and change one's ideas as science advances.
Herbert Spencer (1820–1903)

GENOMICS

The Year of the Rat

After two years of intensive efforts by an international consortium of researchers, the Brown Norway rat (*Rattus norvegicus*) joins the human and the mouse as the third mammalian genomic sequence to be completed. The achievement is expected to yield important new knowledge about mammalian evolution and human disease processes, and should also contribute significantly to progress in toxicogenomics.

The project, funded primarily by the National Human Genome Research Institute (NHGRI) and the National Heart, Lung, and Blood Institute (NHLBI), was conducted by the Rat Genome Sequencing Project Consortium. The Human Genome Sequencing Center at Baylor College of Medicine led the collaboration, assembled the genome, and coordinated the data and resource contributions of a large network of academic and private research centers. Next, an international team comprising more than 20 groups in six countries analyzed the results vis-à-vis the human and mouse genomes. The sequence was published in the 1 April 2004 issue of *Nature*, along with more than 30 papers analyzing the results relative to the human and mouse genomes published simultaneously in the April 2004 issue of *Genome Research*.



Three's company. The Brown Norway rat joins the human and the mouse as the third mammal to have its genome fully sequenced.

The Brown Norway rat has long been one of the primary models employed in biomedical, toxicological, and pharmaceutical research. "A large number of human diseases are mimicked in the rat model, and having the genome sequence lets us easily walk between what we understand in the physiology and biology in the rat, and translate that to a better understanding of human biology and disease processes," says Susan Old, associate director of the Clinical and Molecular Disease Program in the NHLBI Division of Heart and Vascular Diseases and a project officer for the sequencing initiative. She adds, "Our hope is to improve the health of the individual by better understanding of the mechanism of disease, and to develop better therapeutics and diagnostics. We are going to be able to do that a lot more effectively than we had been able to previously."

The availability of the rat genome sequence should also have a profound impact on toxicogenomics. "At the present time, when we generate expression profiles that are associated with a particular disease or genotype, or in response to a toxicant, we still have difficulty putting together the story of what genes and what pathways are being affected, because a lot of the [genes] represented by features on the DNA microarrays have still not been identified," says Helmut Zarbl, a toxicologist at the Fred Hutchinson Cancer Research Center in Seattle. Knowing the locations and identities of all of the genes in the rat genome will aid toxicogenomicists in their efforts to accurately characterize their functions—"to actually put the story together and come up with the predictive toxicology we're looking for," says Zarbl.

Zarbl's group uses rat models to search for genes associated with human breast cancer. He says having the sequence in hand will advance his work as well as toxicogenomics studies of many other complex diseases strongly suspected to be linked to gene–environment interactions. "By being able to map some of these complex diseases in the rat model and find the causal genes," he says, "we can then very quickly go to human studies using comparative sequence analysis to formulate hypotheses about what genes are involved in human disease."

According to Michael Waters, assistant director of database development at the NIEHS-based National Center for Toxicogenomics (NCT), having the sequence will enhance work being conducted in a variety of "omics" areas. "We at the NCT are using microarray technology, but we're also using proteomics, and we want to be able to use metabolomics to understand how toxicants operate, what their mode of action is, what some of the biomarkers are that would indicate when a toxic outcome is likely to occur, [and] . . . predict those effects at earlier times and lower doses," he says. Waters says the rat genome information provides the genetic scaffold for scientists to link expression profiles to a genome that is relevant to toxicology.

With the rat, mouse, and human genome sequences now completed, comparative genomics—identifying the essential functional and structural components of the human genome by comparing it with the genomes of other organisms—is positioned for rapid acceleration. "Every model organism has its advantages and disadvantages, and the more of them that we have to do experiments in, the more quickly we'll be able to find genes and genotypes associated with specific phenotypes, relate these genotypes back to the human genome, and find causes of human diseases," says Zarbl.

In August 2004, NHGRI announced that it has added 18 new model organisms to its sequencing pipeline, including the orangutan, the African savannah elephant, the rabbit, and the domestic cat. Other groups are working on sequencing the dog, the cow, the macaque, and several nonmammalian species. Many of these projects are expected to be completed within a few years.

Waters anticipates that the flood of additional genomes will be extremely valuable to toxicogenomics in two specific ways. First, he says, the value of existing databases will be enhanced, with the cross-species genomic information contributing to chemical risk assessment in the ecological domain, as well as in human health. Second, we could possibly learn far more about basic biological function in terms of phylogenetic relationships. **–Ernie Hood**

NHGRI

MOLECULAR BIOLOGY

Getting to the Core of Antimicrobials

Much research on host defenses against infection has concentrated on the amino acid sequences of antimicrobial peptides in the belief that the order of the acids and their replication reflect how they work against aberrant cells. Now researchers at the University of California, Los Angeles, (UCLA) suggest that the shape the sequences are arranged in may be a critical part of how these peptides work. A new report indicates that host defense systems across the spectrum of life rely on a universal core structure integral to many natural antimicrobial peptides. This core motif may play a key role in preventing or limiting infection, an insight that could accelerate a major advancement in antimicrobial drug development.

"It has been generally accepted that there is a wide diversity in amino acid sequences and sources of antimicrobial peptides," says study co-investigator Michael Yeaman, a professor of medicine at the David Geffen School of Medicine at UCLA. "But there hasn't been as much insight into the similarities that might exist that link all of these diverse groups of molecules."

The gamma (γ)-core motif—so called because it resembles the Greek letter—may be that missing link, providing a key ingredient in the signature of antimicrobial peptides. Yeaman and coauthor Nannette Yount, a molecular biologist at the Los Angeles Biomedical Research Institute, say the γ-core alone can have antimicrobial activity, but also appears to provide a scaffold on which critical modules are configured to create molecules that hunt down microbial pathogens and destroy them in diverse tissue contexts without injury to the host.

The duo studied the amino acid sequences and three-dimensional structures of over 500 antimicrobial peptides, and found the γ-core structure in molecules as diverse as pea defensins, fruit fly drosomycin, pig protegrin, and human hepcidin. Such molecules share the multidimensional signature of antimicrobial peptides. In a paper published 11 May 2004 in *Proceedings of the National Academy of Sciences*, the authors wrote, "This striking multidimensional signature is conserved among disulfide-containing antimicrobial peptides spanning biological kingdoms, and it transcends motifs previously limited to defined peptide subclasses."

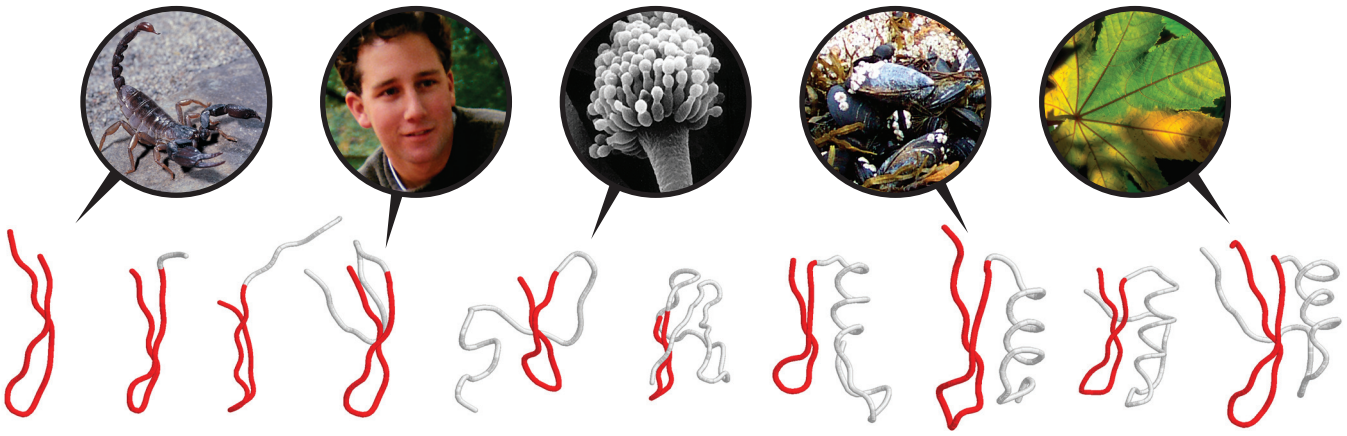
But the sequence, composition, and biochemistry of the amino acids that make up the signature still play a major role, says Yeaman. "We feel that some of the universality identified here may have been missed previously because to identify this signature, we had to look at amino acid sequences in both forward and reverse orientation, and that is not typically done," he says. "The broad conservation of the multidimensional signature identified may have been missed if we only performed amino acid sequence searches and alignments in a conventional way."

There are other critical aspects of the γ-core motif as well, Yeaman says. "The amino acid sequence is configured in three-dimensional space so that the γ-core has certain characteristics. For example, electrostatic charge tends to be placed in one part of this motif and hydrophobicity in another; disulfide linkages are also conserved. These hallmark features of the γ-core motif rely on both composition and three-dimensional structure."

Yeaman and Yount are now translating the motif into peptide mimetics and small molecules, and are designing so-called modular anti-infectives with customized payloads of drugs that attach to the γ-core motif. These compounds are at different stages of development—some are in the design phase, while some have been tested and proven to have antimicrobial efficacy. Still others are being optimized based on data generated in the lab as well as in initial *ex vivo* studies. "We are trying to develop entirely new types of 'smart' antibiotics that recognize and act against harmful microbes, particularly those that have become resistant to most all conventional drugs," Yeaman says.

The work has captured the attention of researchers in the drug development industry. "It's the structure that defines the signature," says Steve Projan, vice president of biological technologies at Wyeth Research in Cambridge, Massachusetts, and a consultant to the American Society for Microbiology, based in Washington, D.C. "That structure may be more important than sequence of amino acids. Even if the amino acids are different, it is the overall structure that defines the activity of the molecule." However, Projan admits, "I'll be skeptical about the impact of this work until we have a molecule that works by [these] rules and a molecule that also works in an infection model."

Yeaman suggests that learning how nature has evolved antimicrobial agents may allow scientists to use the γ-core motif or mimetics thereof as the scaffold that will guide the right peptide or molecule to the right target. "Nature has done much of the designing," he says. "We are capitalizing on the experiments that nature has performed over millions of years [and] trying to integrate the results of that process in new antibiotics." **–Ed Susman**



Common threads. The γ-core motif, visualized in red above, is seen in antimicrobial peptides from a breadth of organisms, including (left to right) the scorpion, the human, *Aspergillus*, the mussel, and the buckeye tree. The motif appears to provide a scaffold upon which disease-fighting molecules are configured.

Left to right: Digital Vision; Photodisc; National Institute of Allergy and Infectious Diseases; Albert E. Theberge; Corbis; molecular images: Nannette Y. Yount and Michael Yeaman

INFECTIOUS DISEASE

Tackling Innate Immunity

According to the 1999 World Health Organization report *Removing Obstacles to Healthy Development*, infectious diseases cause one-third of all human deaths worldwide. These diseases also cost the livestock industry billions of dollars yearly, according to figures from the U.S. Department of Agriculture National Center for Animal Health Surveillance. Infectious diseases are currently fought largely with vaccines (which generate so-called adaptive immunity) and antibiotics. But adaptive immunity can take months to acquire, and overuse of antibiotics may promote resistance in bacteria. If researchers with a Canadian project called Functional Pathogenomics of Mucosal Immunity (FPMI) can unlock the genetic mechanisms behind another branch of immunity—innate immunity—they may have the key to faster-acting, more effective medicines by harnessing the body's rapid-response agents. Indeed, project scientists recently identified a highly promising peptide candidate for future immunotherapies.

FPMI is funded by Genome Canada, a nonprofit corporation dedicated to advancing genomics and proteomics to improve human and animal health. The three-year project involves groups at the University of Saskatchewan, the University of British Columbia, Simon Fraser University, and the Vancouver firm Inimex Pharmaceuticals.

"The unique strength of FMPI is the application of animal and human models of infection to study evolutionarily conserved host responses," says microbiologist Vivek Kapur, co-director of the Biomedical Genomics Center at the University of Minnesota. Because the mechanisms of the innate immune system are not well understood, this comparative genomics approach to study host–pathogen interactions may lead to new immunotherapeutics to prevent infections, he adds.

Innate immunity appears highly conserved in evolution, suggesting that similar events occur in different species. Innate immunity is relatively nonspecific and acts rapidly to block pathogens at the point that they enter the body: the mucous membranes of the respiratory, digestive, and reproductive tracts. Agents produced by the innate immune system—such as cytokines, chemokines, and natural host defense peptides—act immediately in response to infection.

The researchers use microarrays to watch gene activity in humans and animals following exposure to six bacteria and three viruses associated with hospital-acquired infections, food poisoning, and livestock illnesses. "If we can show that the same genetic processes happen in cows, chickens, and humans, that gives us a great deal of confidence that we're on the right path [to understanding the mechanism involved]," says project co-leader Lorne Babiuk, director of the Vaccine and Infectious Disease Organization at the University of Saskatchewan.

The data generated by the thousands of microarray experiments are processed by bioinformaticists headed by Fiona Brinkman, an assistant professor of molecular biology and biochemistry at Simon Fraser University. The team's sophisticated software system, called ArrayPipe, "allows researchers from distant geographic regions to work together and view each others' analyses," says Brinkman. The software is available in an "open source" format that makes it very flexible and easy to customize. ArrayPipe can be downloaded for free at **http://www.pathogenomics.ca/arraypipe/**.

The genes related to innate immunity encode disease-fighting substances, which not only kill pathogens, but also produce inflammation. Although some inflammation is necessary to kill pathogens, it can escalate to undesirable conditions such as septic shock. One goal of the FPMI researchers is to find ways to induce desirable disease-fighting responses, yet quell undesirable ones related to inflammation.

A major breakthrough came when researchers in the laboratory of FPMI co-leader Bob Hancock, who is director of the Centre for Microbial Diseases and Immunity Research at the University of British Columbia, showed that the natural host defense peptide LL-37 cures infections as it suppresses inflammation. In a report published 15 March 2004 in *The Journal of Immunology*, Hancock and colleagues write that LL-37 up-regulates genes linked with the inflammation that kills microbes, but down-regulates those linked with the inflammation that promotes septic shock, suggesting that LL-37 serves as a watchdog to control inflammatory processes. "This . . . indicates that you can get the good aspects of innate immunity without the bad," says Hancock.

Scientists at Inimex are designing new drug compounds based on LL-37. The new strategy will encourage the body's innate immune system to attack foreign invaders, rather than bombard bacteria with antibiotics—an approach that increasingly leads to antibiotic-resistant strains. "It's a new perspective that's desperately needed to counteract antibiotic-resistant bacteria," says Hancock. **–Carol Potera**

Everybody's got one. The ability of innate immunity to block pathogens at the mucous membranes appears highly conserved across species.

Left to right: Pixtal; Image Source; ThinkStock

INNOVATIVE TECHNOLOGIES

Cellular Jigsaw Puzzles

Although scientists have great understanding of individual molecules, the limitations of modern technology have restricted the study of molecular groupings, or "machines," within cells. Now, however, scientists with the Structural and Computational Biology Programme at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, have developed a method to predict and gain understanding of how molecules assemble into machines—an advancement with significant potential application to toxicogenomics.

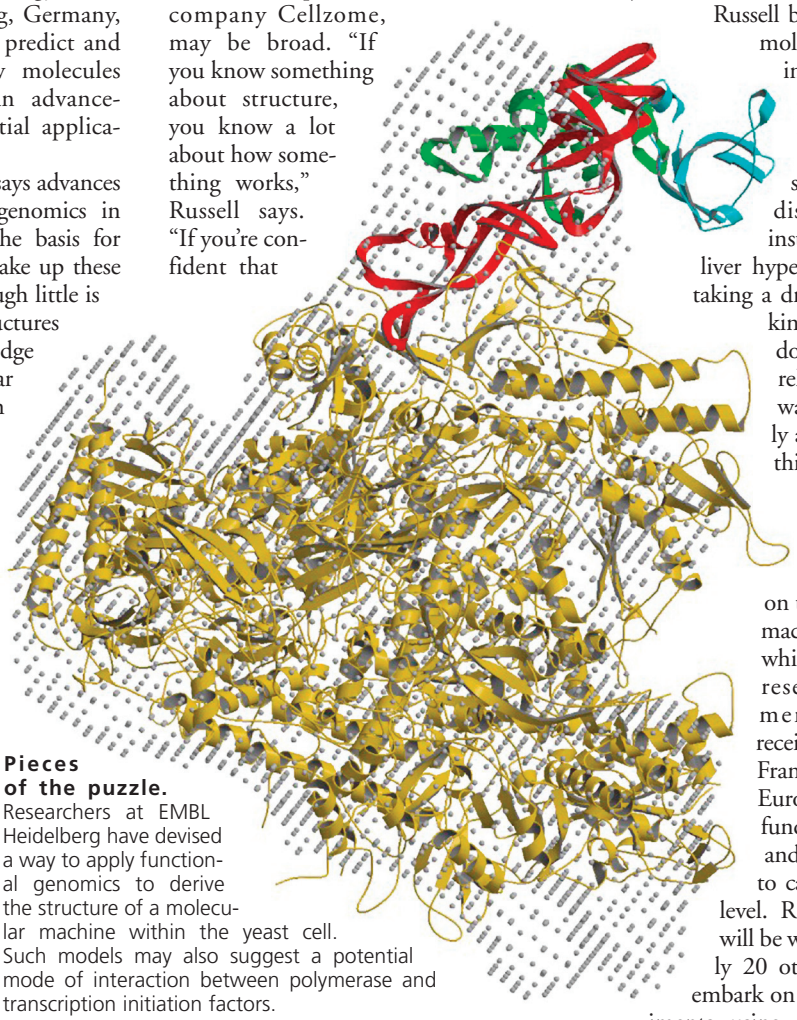
Group leader Rob Russell says advances in the study of functional genomics in recent years have provided the basis for knowing what components make up these molecular machines, even though little is known about what the structures look like. This new knowledge about the makeup of molecular complexes, combined with electron microscopy technology and computer methods developed by Russell and computational biologist Patrick Aloy, provided the framework for the project.

Russell and Aloy studied yeast proteins, identifying the components of hundreds of molecular machines in these cells. Using the "tandem affinity purification" method developed at EMBL Heidelberg, they attached molecular tags to selected proteins and "went fishing" for other proteins in the yeast that would interact with the bait. These interactive complexes form the basis of protein networks.

They liken the process, from that point on, to that of assembling a jigsaw puzzle, where the pieces are individual components of particular machines in yeast cells. They first divided the components into groups containing structural similarities, then proceeded to look for recurring patterns of molecular interactions. For example, if two similar molecules in one machine were also found in another, they were considered likely to fit together in the same way.

The scientists looked for those kinds of relationships and built upward, using knowledge about how various protein molecules fit together in one machine to predict the structure of other machines. In some cases, they were able to draw three-dimensional images of machines on their computer screens. Aloy cautions, however, that these images are predictions—not depictions—of structure.

The potential application of the research at EMBL Heidelberg, which was conducted in conjunction with the private German biopharmaceutical company Cellzome, may be broad. "If you know something about structure, you know a lot about how something works," Russell says. "If you're confident that the structure is right, you could conceivably design chemicals to target particular types of machines."



Pieces of the puzzle. Researchers at EMBL Heidelberg have devised a way to apply functional genomics to derive the structure of a molecular machine within the yeast cell. Such models may also suggest a potential mode of interaction between polymerase and transcription initiation factors.

Rob Russell and Patrick Aloy/EMBL Heidelberg

Andrej Sali, a professor at the University of California, San Francisco, departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, says that scientists working on structural genomics have become keenly interested in how protein assemblies function. "The general point is that structures of assemblies are informative about what the function of the assemblies is and how that function is performed mechanistically—how one might want to control that function, or modify it, and perhaps eventually how one could design new functions," Sali says. "So, for these purposes, knowledge of structure is very helpful."

Sali says Russell and Aloy's report of their research, which appeared in the 26 March 2004 issue of *Science*, has been widely read because it presents a new way to envision molecular structures as systems that appear in three dimensions, and not just as individual proteins.

Russell believes that knowledge of molecular machines is useful in toxicogenomics because so much in this science relies on being able to understand the relationship between often highly disparate processes. "For instance," he says, "how does liver hyperplasia arise when one is taking a drug acting on a particular kinase? This essentially boils down to understanding the relationship between pathways in the cell, and certainly a structural perspective on this can be a great boon."

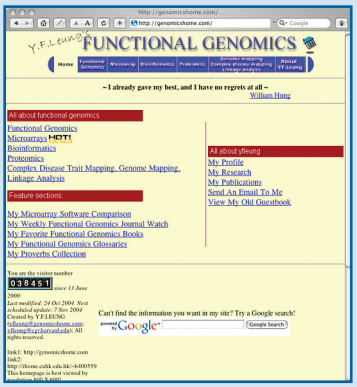
Russell says he and his colleagues have only begun to scratch the surface with their work on the functions of molecular machines. EMBL Heidelberg, which is funded by public research monies from 17 member states, recently received a grant from the Sixth Framework Programme of the European Community (which funds research, development, and demonstration activities) to carry the work to the next level. Russell says his laboratory will be working with approximately 20 other groups in Europe to embark on a variety of further experiments using tools including electron microscopy and X-ray crystallography. "The hope," he says, "is to do this in more detail than what we were able to do in the original paper."

Although there is obviously much work left to do to realize the potential of this new method, the possibilities appear wide open. "It's an exciting area," Russell says. "Our ultimate goal is to have a kind of dynamically updated view of the cell. . . . Ultimately, we want a complete picture of the cell." **–Richard Dahl**

txgnet Y.F. Leung's Functional Genomics

With its sequencing completed in 2003, scientists set their sights on determining the basic structure and inner workings of the human genome. This movement has spawned numerous new scientific specialties that have been supported by the growth of data-generating technologies. One of these interdisciplinary fields, functional genomics, is devoted to linking gene expression to function (or dysfunction) in cells, organs, and tissue. On his website titled Y.F. Leung's Functional Genomics, located at http://genomicshome.com/, Harvard researcher Yuk Fai Leung sketches out the current state of this new field of study.

The homepage of the site is divided into three central sections. The main section, titled All About Functional Genomics, is an assemblage of links to relevant outside resources such as "omics" glossaries and the Department of Energy Genomes to Life program. Also in this section are pages of resources for related fields including bioinformatics and proteomics. Leung has also brought together resources on the use of chaos and nonlinear dynamics in genomics, and on the ethical, legal, and social issues surrounding genomics. A group of links to institutes and core facilities conducting functional genomics work is also provided.



Microarray technology has been crucial to the development of functional genomics. The Microarrays subsection gives an overview of what exactly these tools are, as described through videos, technology reviews, even cartoons, and provides descriptions of all of the various equipment and technology required to perform this sort of analysis. The Language & Standard subpage lists links to resources on communicating results with others within the discipline. Listings of relevant courses, video seminars, conferences, and workshops are also available.

The Bioinformatics subsection of the website contains links to more than 50 databases. This subsection, like the Microarrays subsection, also has pages devoted to the language and algorithms used in bioinformatics as well as data standardization. The Ontology page has links to resources on the efforts to develop a standard, universal vocabulary that can be used across the "omics" fields for all organisms.

Other subsections are devoted to proteomics and to genome mapping, complex disease mapping, and linkage analysis. They are populated much as the other two subsections, with pages of glossaries, educational opportunities, calendars, and the like.

This website also has a novel Feature Sections element that bench scientists will find useful. Leung has put together a microarray software comparison featuring 13 primary types of software, including programs for data preprocessing, analysis, and annotation. The page for each software type has a definition of the software's use, suggested readings, and lists of the products available in each category. Also on offer is a compendium of peer-reviewed journal articles related to functional genomics, including a list of biotech business articles. Leung also provides a reading list of books on functional genomics topics and an in-depth functional genomics glossary. **–Erin E. Dooley**

Health Effects Institute

Requests for Applications

HEI announces its Fall 2004 Research Agenda. The new Requests for Applications booklet is now available at ***www.healtheffects.org/funding.htm.***

RFA 04-4 Measuring the Health Impact of Actions to Improve Air Quality solicits applications for studies designed to assess the health impacts of action taken to improve air quality, and to develop methods required for, and specifically suited to conducting such research. Preference will be given to studies that evaluate actions taken with the express intent of improving air quality and thereby benefiting public health. Up to $3.5 million over a 3-year period is available for these studies. Detailed Letters of Intent are due **November 10** 2004, applications are due **April 12** 2005.

RFA 04-5 Walter A. Rosenblith New Investigator Award has the purpose to bring new, creative investigators into active research on the health effects of air pollution. It provides three years of funding for a small project relevant to HEI's research interests to a new investigator with outstanding promise at the Assistant Professor or equivalent level. Please refer to the HEI website for eligibility criteria. Letters of Intent are due **November 23** 2004, applications are due **February 1** 2005.

RFPA 04-6 Health Effects of Air Pollution provides an application mechanism for investigators whose area of interest falls outside the topics targeted in other current research requests, but is relevant to HEI's current priorities. Applications can be submitted any time.

For more information please refer to the HEI website at *www.healtheffects.org/funding.htm.* If you have questions please contact Terésa Fasulo at +1-617-886-9330 ext 345 or *tfasulo@healtheffects.org*.