Develop New or Improved Instruments and Technologies for Use in Research and Medicine

IN THIS SECTION:
SCIENCE ADVANCES | STORIES OF DISCOVERY

SCIENCE ADVANCES

New Protein Labels Are Visible in Both Light and Electron Microscopy
Automated Labeling of Neuroanatomical Structures in Human Brain
Development of New Technologies for Treatment of Cystic Fibrosis
Anopheles gambiae Genome Adds Firepower to Our Anti-Malaria Arsenal
Organizing Genetic Variability: Entrez SNP and dbMHC
Gene Expression Omnibus: Gene Expression and Hybridization Array Data Repository
Making New Connections: RefSeq and LocusLink
Mammalian Model Organisms: Mapping Diseases Onto the Human Genome
Profiles in Science
ClinicalTrials.gov: Linking Patients to Medical Research
National Library of Medicine Gateway
Multidimensional Approach to Web Evaluation
MIMCom.Net: A Malaria Research Network for Africa
Unraveling the Mysteries of Disease-Causing Germs
Microarray Technologies
Protein Profiling for Prostate Cancer Detection
Protein Chip Array Gives PSMA Clinical Use as a Prostate Cancer Biomarker
Burden of Disease in Mycosis Fungoides
Identification of Signature Profiles of Altered Gene Expression from Exposure to Specific Chemical Classes using Microarray Technology
System Developed for the Rapid Modification of Genes Within Cells
Heart Assist Device Found to Extend and Improve Live of Patients with End-Stage Heart Failure
New Noninvasive Imaging Technique Can Pinpoint Tissue Inflammation in Mice
Eyes are the Window to Organ Damage from Sickle Cell Disease
Wanted: A Bone Marrow Transplant Protocol that Does Not Eliminate a Recipient's Marrow, Does Not Require a Matching Donor, and Cures Sickle Cell Anemia
A New Test to Help Measure the Effectiveness of Novel HIV Vaccines
Sequencing Genomes: Progress in Model Organisms and Selection of New Target Genomes
Model Organism Databases
Bacterial Artificial Chromosome Resource Network
Yeast Gene Deletion Mutant Set - A Powerful Tool for Understanding Genes and Proteins
Hemochromatosis and Iron Overload Study Update (HEIRS)
User's Guide to the Human Genome
Direct Brain Control of Neuroprosthetic Devices
Tracing Brain Activation Using Manganese-enhanced Magnetic Resonance Imaging
Chemical from Sea Anemone Ameliorates Animal Model of Multiple Sclerosis

Microarray "Fingerprinting" of Disease
Spatial Language and the Development of Wayfinding Systems
New Approach to Identifying Disease Genes in the Visual System
Crucial Signaling Events are Visualized in Single Living Cells
Technique to Identify Shortened Cervix Predicts Premature Birth
Building Comfortable Sockets for Amputees
New Brain Imaging Techniques Allow Insight into Neuronal Functioning
Mouse Model Implicates Enzyme in Human Development and Disease
Important New Mouse Model for Ovarian Cancer
Delivering Dual Growth Factors to Engineer New Tissues
Critical Technical Advance in Extending Bone Forming Capacity of Adult Stem Cells

STORIES OF DISCOVERY

MEDLINE and MEDLINEplus: A Continuing Story of Discovery
First Analysis of the Mouse Genome Sequence
The Creation of a Haplotype Map
Early Cochlear Implantation Yields Best Results

New Protein Labels Are Visible in Both Light and Electron Microscopy

Background: Scientists use both optical and electron microscopy for visualizing proteins within cells. Optical microscopy can be used to observe living cells but is not capable of a high level of magnification. Electron microscopy can magnify objects considerably more than optical microscopy, but the cells are killed in the preparation process. To visualize specific cellular proteins in either type of microscopy, scientists attach molecular labels to the proteins. Labels exist for visualizing proteins using optical microscopy, and other labels exist for visualizing proteins using electron microscopy. Until now, no labels have been available for visualizing proteins in both types of microscopy.

Advance: Research teams at an NIH-supported Research Resource at the University of California, San Diego (UCSD) have developed a new technology that lets researchers sequentially label specific proteins as they are produced within a cell. The labels are visible in both optical and electron microscopy. The protein of interest is genetically engineered to contain a small binding domain, which is able to interact with a number of custom-designed chemical compounds. Two of these compounds, called FlAsH and ReAsH, can cross the plasma membrane of living cells and quickly bind to the protein. Upon binding, the compounds turn brightly fluorescent. FlAsH fluoresces green, whereas ReAsH glows red. To distinguish old from new proteins, FlAsH is applied first; then, after new proteins develop, ReAsH is applied. By administering FlAsH and ReAsH sequentially, the UCSD team was able to label two pools of a protein produced at different times. The first pool appeared green in the light microscope, while the second pool appeared red. After identifying the two pools within the living cell using light microscopy, the researchers used a chemical process called photo-oxidation to visualize the ReAsH-labeled proteins in electron microscopy. However, the FlAsH-labeled proteins could not be visualized through this process.

Implications: This technique could be used with any protein and will allow scientists to follow labeled proteins during the assembly and disassembly of intracellular structures. The scientists will be able to first visualize the proteins in living cells using light microscopy and then view the same scene at higher magnification using electron microscopy.

Gaietta G, Deerinck TJ, Adams SR, Bouwer J, Tour O, Laird DW, Sosinsky GE, Tsien RY, Ellisman MH: Multicolor and Electron Microscopic Imaging of Connexin Trafficking. Science 296: 503-508, 2002.

Automated Labeling of Neuroanatomical Structures in Human Brain

Background: Neurodegenerative disorders, psychiatric disorders, and healthy aging are all frequently associated with structural changes in the brain. These changes can alter the imaging properties of brain tissue, as well as the morphometric properties of brain structures. Morphometric changes may include variations in the volume or shape of subcortical brain regions, as well as alterations in the thickness, area, and folding pattern of the cortex. While surface-based analyses can provide an accurate assessment of cortical variability, volumetric techniques are needed to detect changes in noncortical brain structures. This type of analysis has commonly been accomplished by having a trained anatomist or technician manually label some or all of the structures in the brain, a procedure that can take up to a week for high-resolution images.

Advance: NIH-supported scientists at the Massachusetts General Hospital have developed a technique for automatically assigning a neuroanatomical label to each voxel in an MRI volume. In contrast to existing procedures that label only a small number of tissue classes, the current method assigns 1 of 37 labels to each voxel in many specific brain regions, including the thalamus, lateral ventricles, hippocampus, and amygdala. The technique is shown to be comparable in accuracy to manual labeling, and of sufficient sensitivity to detect changes in the volume of noncortical structures. The automated segmentation procedure requires approximately 30 minutes on current computer workstations, such as a 1 GHz Pentium III.

Implications: The additional capability to run multiple processes in parallel enables the labeling of thousands of brain images per day, which can be a major asset for clinical evaluation of neurological disorders and normal aging.

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, Van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM: Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33: 341-355, 2002.

Development of New Technologies for Treatment of Cystic Fibrosis

Background: Cystic fibrosis (CF), a recessive genetic disease, affects approximately 30,000 Americans and is diagnosed in approximately 2,000 children born each year. The gene responsible for CF, *CFTR*, produces a protein normally found in the membranes of secretory cells of the airway, pancreas, sweat glands, salivary glands, intestines, and reproductive organs. Although about 900 mutations have been identified, the delta F508 (_F508) mutation is most common, accounting for 70 percent of affected CF genes.

Advance: Research teams are exploring different ways to overcome the harmful effects of the _F508 mutation. One avenue is to develop safe, effective methods for correcting the defect by replacing the abnormal gene, in whole or in part, by a normal copy. However, the low efficiency of correction at the DNA level has proven to be an obstacle for gene transfer to become a viable treatment for humans. An alternative approach attacks the genetic problem at the messenger RNA (mRNA) level. Messenger RNA is the key intermediate in the process the cell uses to make a protein from the instructions encoded in the DNA. Researchers have used SMaRT (spliceosome mediated RNA trans-splicing) technology to replace the _F508 mutation at the mRNA level. Instead of incorporating the part of the CFTR mRNA with the mutation, the cell is provided with a partial mRNA containing a normal sequence that can be spliced into the mature mRNA molecule. This technique provides a partial restoration of the level of CFTR protein in airway cells obtained from CF patients. SMaRT offers several advantages over conventional gene therapy approaches since only cells that normally express the *CFTR* gene are targeted by this therapy. Also, since SMaRT only replaces the abnormal portion of the message, researchers can take advantage of gene delivery technologies that are unable to accommodate the entire *CFTR* gene.

Another therapeutic approach is to combat the effects of the _F508 mutation at the level of the CFTR protein. This mutation results in improper folding and retention of CFTR protein within the cell and inhibition of CFTR transport to the cell membrane, where it normally functions as a chloride channel to maintain the cellular electrolyte and fluid balance. Because the cellular concentration of calcium plays a role in the retention of abnormal CFTR within cells, depletion of cellular calcium stores may allow the misfolded _F508 CFTR protein to "escape" from within the cell and reach the cell membrane where it may function effectively and thereby correct the CF defect. Genetically engineered mice with _F508 CFTR treated with an inhaled calcium binding agent had decreased calcium concentrations in cells lining the airway thus, permitting the _F508 CFTR protein to be transported to the cell membrane where it functioned to increase chloride transport.

Implications: These results suggest that alternate approaches may be more effective in reversing CFTR abnormalities than conventional gene transfer. Development of non-toxic and effective small molecule therapies that can be delivered directly into the airway may allow production of a normal CFTR protein or improve function of an abnormal CFTR protein and correct the defect in

chloride transport seen in CF. Research to develop these technologies further may lead to new treatments for CF as well as other inherited genetic diseases.

Egan ME, Glöckner-Pagel J, Ambrose CA, Cahill PA, Pappoe L, Balamuth N, Cho E, Canny S, Wagner CA, Geibel J, Caplan MJ: Calcium-pump inhibitors induce functional surface expression of _F508-CFTR protein in cystic fibrosis epithelial cells. Nat Med 8(5): 485-492, 2002.

Liu X, Jiang Q, Mansfield SG, Puttaraju M, Zhang Y, Zhou W, Cohn JA, Garcia-Blanco MA, Mitchell LG, Engelhardt JF: Partial correction of endogenous _F508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA *trans*-splicing. Nat Biotechnol 20: 47-52, 2002.

Anopheles gambiae Genome Adds Firepower to Our Anti-Malaria Arsenal

Background: Anopheles gambiae (the mosquito) is the primary agent involved in the transmission of human malaria. This disease causes an estimated 200 million clinical cases and more than one million deaths annually. Pest control strategies that are designed to curb the local population of mosquitos have met with limited success because the overuse of insecticide compounds eventually leads to insecticide-resistant strains that continue to transmit the disease. In addition, various xenobiotic compounds pose a threat to the environment when not used in a controlled manner. Current preventative treatments for malaria include the administration of antiparasitics or specialized vaccines to humans. These treatments may also have limited success if the gene within the parasite that is targeted by the drug changes, thereby rendering the antimalarial treatment ineffective. Furthermore, not all individuals react the same way to such treatments, prompting researchers to develop new strategies to control malaria, especially in third world nations.

Advance: In an effort to better understand the biology of malaria and to serve the malaria genetics community, NIH has recently provided access to the genomic sequence data of *Anopheles* that complements growing genomic data from other organisms related to the malaria parasite or involved in the disease cycle. The *Anopheles gambiae* genome has recently been sequenced by Celera Genomics and deposited in GenBank®, the nucleotide sequence database at NCBI. The sequenced portion of the genome of *Plasmodium falciparum*, the parasite living in the gut of the female mosquito that is the causative agent of malaria in humans, is also available for searching at the NCBI website. Massive sequencing efforts are also underway for other Plasmodium species, namely those that cause malaria in rodents: *P. berghei*, *P. chabaudi*, *P. yoelli*, and *P. vinckei*, all of which are suitable model species for the study of malaria in humans. A graphical display of the *Anopheles* genome can be viewed with NCBI's MapViewer tool, which provides access to 7 different maps and allows users to examine their sequence in a full genomic context. MapViewer also provides protein matches to the *Drosophila melanogaster* genome, an ideal model genome for *Anopheles*, with well-defined genetics that is useful for the study of genes that may be conserved between the two organisms. NCBI's Malaria Genetics and Genomics site provides full and integrated access to a wide variety of resources both within and outside NCBI, including the assembled genomic data of *A. gambiae*, the genomic MapViewer, sequence similarity searches (BLAST®) against the *Plasmodium* databases, and access to the biomedical literature (PubMed®).

Implications: Coupled with the accumulating sequences of the human and rodent malaria species, and the complete genomic sequence of the host genome (human), NCBI's set of malaria-related genomic resources will contribute to the understanding of the mosquito genome structure and organization and will spur the development of more effective malaria control strategies and improved anti-malarial drugs and vaccines.

Malaria Genetics and Genomics: http://www.ncbi.nlm.nih.gov/Malaria/

Organizing Genetic Variability: Entrez SNP and dbMHC

Background: One of the most promising returns from the Human Genome Project is an increased ability to identify genes that are linked to the development of diseases. Each normal human contains a particular version, or allele, of every human gene. Some alleles of disease-linked genes result in normal, healthy function while others increase the likelihood of a particular disease. Thus, it is critical to identify not only those genes that are linked to a disease, but also those particular alleles that lead to an increased risk of disease. Because different alleles differ in their DNA sequence, identifying the alleles amounts to discovering their sequences. A database of the sequences of all known alleles of a given gene would be an invaluable asset in discovering the genetic basis of disease.

Advance: NIH maintains a database of genetic variation, dbSNP, as part of its text-based Entrez search system. A SNP, or single nucleotide polymorphism, is a site of variation between the genomic DNA of two individuals. The dbSNP database is a growing compendium of all known sequence differences between gene alleles. While a SNP database has been in development at NIH for over 3 years, it now is fully integrated with the other standard NIH databases, allowing users to search this data with more ease and power than ever before. For example, it is now possible to search for variations in a particular gene using simple text searches and then to view the result of this variation on the three-dimensional structure of the protein the gene produces. Alongside Entrez SNP, NIH has also released a related database named dbMHC that compiles all known alleles for the Major Histocompatibility Complex, a critical part of the immune system. The human MHC genes are currently unique in that they are exceptionally well studied with hundreds of known alleles, most of which have known sequence. The dbMHC database allows users to search these alleles by their DNA sequence and then to produce a graphic summarizing the similarities and differences to selected combinations of other MHC alleles. These data are used not only in the discovery of new MHC alleles, but also regularly to predict compatibility in procedures such as organ transplant surgery.

Implications: Integrating SNP into the Entrez system will accelerate the discovery and analysis of gene variations by giving researchers seamless access to the rapidly growing set of variation data. The integration of SNP data with other sequence data holds great promise in reaching the ultimate goal of discovering the genetic basis of many diseases. As a model of future allele databases, dbMHC anticipates the coming era when detailed sequence data of alleles will be available for many genes of medical importance, and serves as an example of how this data can be easily and directly disseminated to the clinic.

Entrez SNP: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=snp

dbMHC: http://www.ncbi.nlm.nih.gov/IEB/Research/GVWG/MHC/MHC.cgi

Gene Expression Omnibus:
Gene Expression and Hybridization Array Data Repository

Background: Over the past few years, molecular biology experiments using high-throughput hybridization array- and sequencing-based technologies have become commonplace. These techniques allow a scientist to analyze the expression of many genes in a single experiment, both quickly and efficiently. Today, proficiency in generating data has led to a deluge of expression information. Much of this information is scattered across the Internet, or is not yet available to the public. To make this avalanche of data accessible to, and understandable by, the biomedical community worldwide requires the standardization of storage, sharing, and publishing techniques.

Advance: In response to this need, NIH's National Center for Biotechnology Information has established the Gene Expression Omnibus (GEO), a public repository for high-throughput gene expression data accessible through the World Wide Web. GEO provides a flexible design that facilitates the submission, storage, and retrieval of heterogeneous data sets from high-throughput gene expression and genomic hybridization experiments. The primary goal in creating GEO was to attempt to cover the broadest spectrum of high-throughput experimental methods possible while remaining flexible and responsive to future experimental trends. As such, GEO is intended to serve as a central data distribution hub, not to replace in-house gene expression databases that benefit from coherent datasets and which are constructed to facilitate a particular analytic method.

Implications: By storing vast amounts of data on gene expression profiles derived from multiple experiments employing varied criteria and conditions, GEO will be critical in the study of functional genomics – the development and application of global experimental approaches to assess gene function. GEO will also facilitate the cross-validation of data obtained using different techniques and technologies and will help set benchmarks and standards for further gene expression studies. Data obtained from such experiments will enable researchers to understand fundamental aspects of growth and development as well as to explore the underlying genetic causes of many human diseases.

GEO: http://www.ncbi.nlm.nih.gov/geo/

Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30(1): 207-210, 2002.

Making New Connections: RefSeq and LocusLink

Background: GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences. NIH's National Center for Biotechnology Information (NCBI) is responsible for all phases of GenBank production, support, and distribution, including the timely and accurate processing of sequence records and biological review of both new sequence entries and updates to existing entries. Integrated retrieval tools have been built to search the sequence data housed in GenBank and to link the results of a search to other related sequences, as well as to bibliographic citations. Yet as an archival database, GenBank can be very redundant for some gene sequences. In addition, as GenBank sequence records are owned by the original submitter and cannot be altered by a third party, they may not always represent our current knowledge about a particular gene or gene sequence.

Advance: Two integrated resources have been developed by NCBI to address the problems of multiple sequences for single genes and sources of current information about specific genes. LocusLink organizes information around genes to generate a central hub for accessing gene-specific information for fruit fly, human, mouse, rat, and zebrafish. NCBI's Reference Sequence project (RefSeq), accessed through LocusLink, provides a reference sequence, or standard, for each principal class of molecule found in cells: chromosomes, mRNAs, pre-proteins, and mature proteins. RefSeq records are currently available for the following genomes and organelles: human, mouse, rat, fruit fly, nematode, yeast, *Arabidopsis*, *Plasmodium falciparum*, *Leishmania major*, bacteria, phages, viruses, viriods, mitochondria, chloroplasts, and plasmids. Together, RefSeq and LocusLink provide a non-redundant view of genes and other loci to support research on genes and gene families, variation, gene expression, and genome annotation. Many new features have been added to the databases over the past year, including GeneRIF – links to references into function. GeneRIF allows users to search and retrieve information via the sequence annotations and citations. As part of NLM's MEDLINE indexing process, new papers discussing the basic biology of a gene – function, structure, and genetics – are being linked to a LocusLink report. Concise text summarizes and describes the importance of the paper to the understanding of that gene and its products and function. In addition, using a PubMed function called LinkOut, users retrieving articles in PubMed will also be able to navigate to LocusLink to learn more about the genes being referenced.

Implications: Although the GeneRIF project is in its initial stages, NLM indexers have already found up to 3,000 new connections between locus identification numbers (IDs) and PubMed IDs. In this way, GeneRIF allows researchers to make new connections between sequence information and gene names; between gene sequence and gene function; between protein structure and function; and between previously unrelated proteins. For example, through information obtained from GeneRIF, NCBI investigators have been able to link the LocusID for a human taste receptor to a particular nucleotide sequence with a previously unknown function.

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/index.html

Mammalian Model Organisms: Mapping Diseases Onto the Human Genome

Background: Over the last century, research on a small number of organisms has played a pivotal role in advancing our understanding of numerous human biological processes. The mouse is a useful mammalian model organism for humans beause the gene sequences that code for numerous proteins responsible for carrying out vital biological processes in both the human and the mouse share a high degree of similarity. The rat, on the other hand, is the principal model organism used to link function to genes. In some instances, specific aspects of human disease are duplicated well only in the rat, making the rat an excellent organism for studying and identifying genetic pathways relevant to human disease. Efforts are now underway to develop public resources designed to collect, consolidate, and integrate the genomic data being generated from the mouse and rat genome sequencing efforts. The ultimate goals of these projects include the construction of a robust physical map and a high quality, finished genomic sequence as these data will provide an essential tool to identify and study the function of human genes.

Advance: NIH investigators, in collaboration with investigators from around the world, recently reported on the development of a radiation hybrid map of mouse genes. This gene-based map was constructed using genes from NIH's mouse UniGene collection and includes 3,658 genes that are homologous, or similar, to the human genome sequence. Human-mouse orthology maps – maps showing the correspondence between human and mouse conserved genomic segments – are publicly available on NIH's website. NIH investigators are also collaborating on the Rat Genome Database (RGD) project, an NIH-funded project whose mission is to collect, integrate, and disseminate the data being generated from ongoing rat genetic and genomic research efforts. To date, investigators have integrated large amounts of rat genetic and genomic resources in RGD and these resources are being expanded to include links to the literature, curated data on rat genes, quantitative trait loci, microsatellite markers, and rat strains commonly used for research purposes. In addition, RGD provides tools for homology searching, gene prediction, radiation hybrid mapping, and polymorphic marker selection.

Implications: The development of a radiation hybrid map of mouse genes provides not only a framework for sequencing the mouse genome, but for comparing and contrasting the human genome sequence to the mouse sequence. The RGD serves to facilitate the discovery of the genetic basis of disease in the rat, thereby providing a comparative view of the rat disease data with known human and mouse data. Together, these resources make it possible to conduct cross-organism studies for genes with shared locations, biological processes, and function.

Human-Mouse Homology Map: http://www.ncbi.nlm.nih.gov/Homology/index.html
UniGene: http://www.ncbi.nlm.nih.gov/UniGene/

Hudson TJ, Church DM, Greenaway S, Nguyen H, Cook A, Steen RG, Van Etten WJ, Castle AB, Strivens MA, Trickett P, Heuston C, Davison C, Southwell A, Hardisty R, Varela-Carver A, Haynes AR, Rodriguez-Tome P, Doi H, Ko MS, Pontius J, Schriml L, Wagner L, Maglott D, Brown SD, Lander ES, Schuler G, Denny P: A radiation hybrid map of mouse genes. Nat Genet 29(2): 201-205, 2001.

Twigger S, Lu J, Shimoyama M, Chen D, Pasko D, Long H, Ginster J, Chen CF, Nigam R, Kwitek A, Eppig J, Maltais L, Maglott D, Schuler G, Jacob H, Tonellato PJ: Rat Genome Database (RGD): mapping disease onto the genome. Nucleic Acids Research 30(1): 125-128, 2002.

Profiles in Science

Background: The NIH is making the archival collections of pioneering biomedical scientists available on its *Profiles in Science* Web site. The site was launched in September 1998 and promotes the use of the Internet for research and teaching in the history of biomedical science. Six of the collections have been donated to the NIH, and two were digitized in collaboration with their official repositories. The historical (pre-2001) Reports of the Surgeon General were made available in cooperation with the Office of the Surgeon General. The collections contain published and unpublished items, including books, journal volumes, pamphlets, diaries, letters, manuscripts, photographs, audiotapes, video clips, and other materials.

Advance: *Profiles in Science* currently provides access to the collections of eight scientists. In the last year the collections of two Nobel Laureates as well as the historical Reports of the Surgeon General were made available. In October 2001, in cooperation with the American Philosophical Society, the papers of Barbara McClintock (1902-1992) were added to the site. McClintock was an American geneticist who won the 1983 Nobel Prize in Physiology or Medicine for discovering the ability of genes to change position on the chromosome. In February 2002 the historical Reports of the Surgeon General were made available through *Profiles in Science* at http://sgreports.nlm.nih.gov/. In February 2002 the papers of Linus Pauling (1901-1994) were added, in collaboration with the Oregon State University Libraries. Pauling is the only person to win two unshared Nobel Prizes (1954 Nobel Prize in Chemistry and 1962 Nobel Peace Prize.) These papers have been added to the six collections already available through *Profiles in Science*. The site was launched with a selection of papers devoted to the life and work of Oswald T. Avery (1877-1955). Avery was a founder of modern molecular biology and the first person to establish genetic information is encoded in DNA. In March of 1999, the site was expanded with the release of a portion of the extensive Joshua Lederberg (b. 1925) collection. Lederberg received the Nobel Prize for his discoveries concerning bacterial genetics. In November 1999, a selection from the papers of Nobel Laureate Martin Rodbell (1925-1998) was posted, and in May 2000, the Julius Axelrod (b. 1912) collection was made available. Rodbell shared the Nobel Prize for his discovery of G-proteins and the principles of signal transduction in cellular communication, and Axelrod shared the Nobel Prize for his discovery of the actions of neurotransmitters in regulating the metabolism of the nervous system. In November 2000, the papers of Christian Anfinsen, Jr. (1916-1995) were added. Anfinsen was an American biochemist who shared the 1972 Nobel Prize for work that helped explain the structure and composition of proteins in living cells. In May 2001 the papers of Marshall W. Nirenberg (b. 1927) were added. Nirenberg is an American biochemist who shared the 1968 Nobel Prize for his work on deciphering the genetic code.

Implications: The *Profiles in Science* site is a growing resource for students, educators, and researchers. The site makes the successes of science readily accessible to anyone who is interested in the scientific process and in the people who have dedicated their lives to scientific discovery. The collections are particularly strong in the areas of cellular biology, genetics, and biochemistry, but also reflect issues in such areas as health and medical research policy, the application of computers in medicine, science education, and the search for extraterrestrial life.

http://profiles.nlm.nih.gov/
http://sgreports.nlm.nih.gov/

ClinicalTrials.gov: Linking Patients to Medical Research

Background: Clinical trials are medical research studies used to evaluate promising new treatments in people. Experimental drugs, medical procedures, and devices must be shown to be safe and effective in human participants before becoming available for widespread clinical use. Thus, clinical trials bridge basic biomedical research and clinical applications. Likewise, *ClinicalTrials.gov*, a freely available Web-based information resource, links members of the public to information about clinical trials. Established by Section 113 of the FDA Modernization Act of 1997 and launched by NIH on February 29, 2000, the site currently includes information on more than 6,350 trials for hundreds of diseases and conditions conducted in over 69,000 locations worldwide. Sponsors representing the Federal government, pharmaceutical industry, and non-profit organizations submit and maintain trial data using an NIH-developed Web-based Protocol Registration System (PRS). The site also contains numerous resources for promoting public awareness on the role of clinical trials, including risks and benefits to participants. *ClinicalTrials.gov* receives approximately 7,200 visitors daily and 2 million page views monthly.

Advance: *ClinicalTrials.gov* has been designed with an emphasis on ease of use and providing relevant medical information for non-health professionals, including patients, family members, and other health consumers. Visitors may access comprehensive, accurate, and timely information on clinical trials within a few mouse clicks. All study data are presented in a standard format with four sections: Purpose, Eligibility, Locations and Contact Information, and More Information. The condition, treatment or intervention, and phase of each study are highlighted in a table for easy reference. Additional context-sensitive resources for non-professionals are provided through numerous "just-in-time" hyperlinks to relevant documents in other NIH online biomedical resources. Thus, visitors who want to learn more about a condition being studied or the published biomedical literature related to a study simply click on hotlinks provided in *ClinicalTrials.gov* records. From within a record, they may access related health topics in MEDLINE*plus*, NIH's consumer health resource, and specific citations in PubMed/MEDLINE, NIH's bibliographic database. The overall design of *ClincialTrials.gov* emphasizes usability, accessibility (e.g., compliance with Section 508 of the ADA), and contextualization (e.g., just-in-time hotlinks). The site relies on natural language processing tools, terminological systems, and other medical informatics tools developed at NIH.

Implications: Due to the conceptual complexities and technical nature of biomedical information and clinical trials, information about experimental interventions is often inaccessible to non-professionals. As the national clinical trials data bank, *ClinicalTrials.gov* provides the public with access to ongoing clinical research by using the latest information technologies. Ongoing research into areas such as consumer health information needs and online seeking behaviors, methods for assessing reading ease in the health domain, and bilingual information retrieval, will yield additional mechanisms to help people learn about clinical trials in the context of general health care, basic and clinical biomedical research, and improving human health.

http://www.clinicaltrials.gov

National Library of Medicine Gateway

Background: The NIH National Library of Medicine (NLM) has at least 10 major Web sites providing important information to huge numbers of Internet visitors each day. These users find information that answers their questions in many databases and document collections on multiple retrieval systems at multiple NLM Web addresses. Some users, however, may not know exactly what is available at NLM or where and how best to search for it. For these users, NLM has recently developed and deployed a system called the NLM Gateway.

Advance: The NLM Gateway provides a single Web address with a single look and feel. The user enters one query that is translated and sent automatically to multiple NLM retrieval systems having different characteristics but potentially useful results. Results from those systems are presented in categories (journal article citations; books, serials and audiovisuals; consumer health information; meeting abstracts; other collections) rather than by database. In some categories, multiple collections are searched. The goal is to provide "first-stop shopping" as an overview to an increasing number of NLM's information resources. The NLM Gateway currently provides access to 11 databases or document collections in five NLM retrieval systems using six access methods. It also provides access to NLM's Unified Medical Language System (UMLS) Metathesaurus for query term selection and to NLM's Loansome Doc system for hardcopy document delivery.

Implications: The NLM Gateway site is a growing resource for students, educators and researchers. Users who were not sure where to look can be led to useful information they might not have known was present, in a database they weren't aware of. Some users will find what they need immediately. Others may find that one resource, such as PubMed or MEDLINEplus, has information they would like to know more about. They may then choose to go straight to that resource for a more focused search. This combination of the NLM Gateway for an initial scan, coupled with the native interfaces of other systems for carefully targeted searches, can be a powerful means of finding information.

http://gateway.nlm.nih.gov

Multidimensional Approach to Web Evaluation

Background: All NIH Institutes and Centers now depend on the Internet and World Wide Web for dissemination of biomedical and health information. This information revolution has taken place over the last five years. Additionally, NIH now has a mandate to provide information not only to researchers, scientists, medical librarians, and health care providers, but to the general public including patients, families, and students. The combination of web access and a broader user community has resulted in much higher usage levels than in the pre-web environment. For PubMed/MEDLINE, for example, the annual database search volume increased from about 7 million searches per year 7 years ago to over 400 million searches per year in FY02. For MEDLINEplus, usage increased from a cold start in 1998 to a current level of over one million unique users per month. PubMed and MEDLINEplus are the top two NIH web sites in terms of usage levels, and together account for between half and four-fifths of NIH usage (depending on the metric used). Overall, NIH web sites collectively rank number one in usage in the Federal Government health sector.

Advance: With this unprecedented growth and leadering position comes the need to monitor and evaluate our performance, and ask the question "How good a job are we doing?" Why, one might ask, should NIH take the trouble to do that? Why not just let the numbers and the marketplace speak for themselves? But what are those numbers? How are they generated? What should we be counting? Is there symmetry between our users' technical capacity and expectations – what we are able to deliver, and users receive, over a rapidly changing network infrastructure? And in the final analysis, determining whether NIH is really doing a good job is more than just counting numbers of page hits or eyeballs reached. Who are we serving? What content are users seeking? For what purposes? Is it understandable? What impact are we having on health care and the decisions people make for their families? From a societal standpoint, what role can the NIH information resources play in helping to reduce health disparities among minority and other medically underserved populations? And what enhancements should NIH plan the years ahead to help assure continued performance of the first-rank?

The challenges posed in answering these questions are considerable. But we are making good progress. NIH's web metrics and evaluation research takes many forms. They can be grouped into essentially five categories and question types, each with their own distinct methodologies:

1) End-to-End Internet Performance Testing – Asks the basic question, "How well are NIH's intended users able to use the Internet? " "Where are the bottlenecks – from the users' perspective?" Metrics include such objective measures as round trip time, bulk transfer capacity and similar tools for diagnosing network and local performance problems.

2) Broadband Infrastructure Capacity Assessment – Related to network performance, this approaches the question from the applications side, namely "What are the clinical and biomedical research applications that need broadband and more of it?" Our first exploratory effort will be launched in FY03 in collaboration with Internet2. It will employ in-depth

interviews that profile and document actual networked applications, and determine where available broadband (or lack thereof) has made a difference (good or bad) at various institutional settings – including health science libraries, hospitals and research centers.

3) Usability Testing – Asks the question, "Does the system and its functions perform as intended by the designers, when placed in the hands of actual users?" A developmental tool, these simulations are carried out under controlled experimental conditions using pre-production systems. They are typically small-scale observational studies of users' cognitive processes and actual behaviors. The MEDLINEplus development team has made extensive use of usability testing.

4) Audience Measurement – Asks the question, "Who's coming to NIH's web sites and what do they do when they get there?" This may involve the use of internal Web logs and and/or external independent user panels in which Web usage is monitored much like the Nielsen TV viewer ratings. System managers are beneficiaries of these data on a day-to-day basis. They also make their way into internal management reports on utilization. Some also are picked up for the public information displays, public usage reports, and related research and evaluations projects.

5) Information Seeking Behavior and Studies of User Satisfaction – Asks the question, "What do users want and are they getting it?" This is best accomplished with focus groups and large-scale randomized sample surveys of new and experienced users. This kind of data can really help put a human face on the numbers, and the techniques have been applied widely to NIH's key databases and web sites including PubMed/MEDLINE and MEDLINEplus.

Implications: Over the last several years, NIH has experimented with and applied evaluation techniques in all of these areas. What has evolved is what we call a multidimensional approach to web evaluation. We believe, taken as a whole, that this strategy constitutes an important advance in the science of evaluation. Moreover, it complements and leverages NIH's investment by translating the results of biomedical research, and provides a useful metric for gauging the impact of NIH's budget doubling by answering the broad question of how many more people today have rapid access to needed biomedical and health care information.

Wood FB, Cid VH, Siegel ER: "Evaluating Internet End-to-End Performance: Overview of Test Methodology and Results." J Am Med Inform Assoc 5(6): 528-545, 1998.

Wood FB, Cid VH, Siegel ER, Gill M J, Aronson J, Lindberg DAB: "Evaluating Internet End-to End Performance Part II." J Am Med Inform Assoc under revision.

Wood FB, Siegel ER, Lyon B, LaCroix EM, Benson D: "Multi-Dimensional Approach to Web Evaluation at NLM." in preparation.

MIMCom.Net: A Malaria Research Network for Africa

Background: NIH has led an international effort to provide malaria researchers in Africa with full access to the Internet and the resources of the World Wide Web. This project began with NIH's leadership in the Multilateral Initiative on Malaria in which African scientists identified electronic communication and access to scientific information as critical in the fight against the devastating and economically debilitating effects of malaria in developing countries.

Advance: The NIH, working in partnership with organizations in Africa, the United States, the United Kingdom and Europe, has created MIMCom.Net, the first electronic malaria research network in the world. The network provides full access to the Internet and the resources of the World Wide Web, as well as access to current medical literature, for scientists working in Africa. The African research sites are of recognized high quality, require improved communications to accomplish ongoing research, and have the necessary resources to purchase equipment and sustain the system.

Telecommunications Technology. The network technical hub is located at Redwing Satellite Solutions, Ltd. in the United Kingdom, where a large satellite dish, focused on a geo-stationary satellite 37,000 km above the Atlantic Ocean, is connected directly to the high-speed Internet backbone on the ground. At research sites with no local telecommunications service, a smaller ground station in the form of a Very Small Aperture Terminal (VSAT) is installed. The VSAT dish antenna connects, through a radio unit, to an existing local area network (LAN) used by the researchers. Some sites on the network operate a wireless connection to a local Internet Service Provider (ISP) or to another MIMCom site nearby.

The system provides an open link that allows researchers to send and receive email, search the literature and databases, or share files and images 24 hours a day 7 days a week. Permanent access to information is moving researchers in Africa toward a more efficient way of working with colleagues around the world.

Cost Savings through Sharing. Satellite systems are not subject to the problems and limitations of telephone wires or other more traditional means of obtaining an Internet connection and are, therefore, highly reliable. Although associated costs are high, MIMCom.Net is designed to allow hundreds of researchers in Africa to share satellite bandwidth, maximizing the usage of satellite capacity and minimizing cost per site.

Network Evaluation. NIH and the International Centre of Insect Physiology and Ecology (ICIPE) Mbita research site in Kenya are currently leading the first evaluation of MIMCom.Net. The evaluation covers network performance and efficient use of bandwidth, as well as information use and site growth, proposals funded, papers published, and numbers of collaborations carried out. Information is being gathered using baselines that were created before the network was installed.

Access to Medical Literature. NIH is using MIMCom.Net to launch two experimental programs that promote increased access to medical literature for malaria researchers in Africa. The Medical Library at the University of Zimbabwe, the Medical Research Council (MRC) in South Africa and NLM have established a pilot document delivery system for malaria researchers in Africa.

Training and New Tools for Research. NIH provides ongoing technical support and training to IT personnel at each site. Additional training opportunities, including individual course work and regular conference calls, are planned to build capacity among African IT specialists at the research sites. NIH is making follow-up visits to each site for updating and trouble shooting. Special training in the use of IT as it relates to specific research agendas will be offered to malaria researchers in Africa. This may include training in the use of various wireless communication devices as well as personal software agents.

Malaria Research Web Site. The web site, http://www.nlm.nih.gov/mimcom, comprises links to Medline, a variety of free on-line journals, databases, malaria-related sites, and general information. An NIH reference librarian serves as the webmaster and is expanding the site to include special news releases and articles of interest to researchers.

http://www.nlm.nih.gov/mimcom

Unraveling the Mysteries of Disease-Causing Germs

Background: Microscopic organisms cause a staggering amount of human suffering. In fact, infectious diseases caused by microbes are responsible for more deaths worldwide than any other single cause. Health experts estimate the annual cost of medical care for treating infectious diseases in the United States alone is about $120 billion.[1]

Advance: Computational biologists who were studying an unusual corkscrew-shaped fold within proteins, known as a beta helix, recently made a surprising find. The scientists' aim was to see if a computer program they developed, BETAWRAP, could accurately predict from amino acid sequences when a protein structure contains a beta helix fold. When they tested BETAWRAP on a large database of experimentally determined protein structures, they found that it did accurately predict the beta helices. However, the researchers were surprised when they noticed that a disproportionately high percentage of these beta helix proteins came from disease-causing organisms compared to the number found in harmless bacteria, human, or animal cells.

Implications: This finding suggests that the beta helix fold may be a key to the workings of proteins from organisms that cause diseases ranging from cholera to Lyme disease. The scientists say their findings imply that the beta helix fold may play a more extensive role in human disease than was previously known. The fold could also offer medical researchers a new way to identify pathogens as well as a new target for disabling them. For example, drug developers might find a new way to kill disease-causing germs with a compound that targets the beta helix shape in their proteins. The advantage of such a therapy is that it could thwart invading pathogens with few, if any, harmful side effects for human patients.

Bradley P, Cowen L, Menke M, King J, Berger B: Successful Prediction of Parallel â-helices from Primary Sequence Reveals an Association with Many Microbial Pathogens. Proc. Natl. Acad. Sci. 98(26): 14819-14824, 2001.

[1] *Microbes in Sickness and in Health,* National Institute of Allergy and Infectious Diseases, NIH Publication No. 01-4914, September 2001.

Microarray Technologies

Background: Brain researchers are eager to harness our understanding of the human genome for studies of mental disorders. One technology uses the structure of DNA itself to understand which genes are turned on in particular structures of the brain. Scientists have put minute quantities of very large numbers of genes onto "microarrays," regular arrays of identified spots of DNA on glass microscope slides. Researchers interested in gene expression in various parts of the brain can then grind up small quantities of tissue, extract the RNA from it (which represents the genes that are turned on) and allow the RNA to bind specifically to its copy on the slide. As information emerges from the Genome Project, researchers have worked to develop this technology to allow more and more genes to be included in a single array. At the same time, they are working to develop methods that will allow ever-smaller amounts of tissue to be used to conduct a reliable experiment. Ultimately, researchers would like to be able to do an experiment using RNA from a single cell.

Advance: Researchers at NIH have developed a microarray that includes 31,000 unique RNA copies of mouse genes and a second array with 21,000 unique RNA copies of human genes. An additional 15,000 human genes will be included in the human microarray within the next few months. In addition, the same researchers have developed methods for getting RNA from tissues that allow them to perform experiments with as little as 1 milligram of tissue, 20–fold less than was previously possible.

Implications: This technology enables greater qualitative analyses that are more sensitive and at a lower cost than other techniques currently available. The new, high-density microarrays are approaching inclusion of all of the expressed genes in mouse and human. These will be a critical resource to the entire biomedical research community as they work to understand the mass of data generated by the Genome Project. These microarray experiments on smaller numbers of genes have already begun to yield important insights into gene expression and misexpression in disease. Ultimately, identification of gene alteration in disease is expected to point toward novel targets for pharmaceutical interventions in disease processes.

Xiang CC, Kozhich OA, Chen M, Inman JM, Phan QN, Chen Y, Brownstein MJ: Amine-modified random primers to label probes for DNA microarrays. Nat Biotech 20: 738-742, 2002.

Protein Profiling for Prostate Cancer Detection

Background: The prostate-specific antigen (PSA) test has been a major factor in increasing awareness and improving patient management of prostate cancer, but it does not reliably distinguish prostate cancer from non-cancer conditions that can elevate PSA. Therefore, its use in diagnosis is limited and it is a weak tool for early detection of prostate cancer. Better biomarkers are needed to improve early detection accuracy and to distinguish slow-growing cancers from aggressive ones. However, because the molecular and cellular characteristics of prostate cancer are highly varied, it is likely that a combination or panel of biomarkers will be required to improve early detection of this disease.

Advance: The researchers coupled a technology known as a protein biochip with an artificial intelligence learning algorithm to look for 9 different proteins in serum samples from 167 prostate cancer patients, 77 patients with benign prostatic hyperplasia (BPH), and 82 age-matched healthy men. These samples were used to "train" the artificial intelligence program, which correctly classified 96 percent of the samples. Then the investigators tested the program on a second set of samples without knowing themselves which samples belonged to each of the three categories (i.e., prostate cancer, BPH, or healthy men). The sensitivity (true positive results) of the new test was 83 percent, and its specificity (true negative results) was found to be 97 percent. By comparison, the sensitivity of PSA testing is greater than 90 percent, but its specificity is only about 25 percent. The protein profiling test was found to have a positive predictive value of 96 percent in distinguishing men with prostate cancer from those with BPH or healthy men.

Implications: This high-throughput proteomic classification system could provide a highly accurate and innovative approach for the early detection and diagnosis of prostate cancer.

Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, Semmes OJ, Schellhammer PF, Tasui Y, Feng Z, Wright, GL Jr: Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. Cancer Research 62: 3609-3614, 2002.

Protein Chip Array Gives PSMA Clinical Use as a Prostate Cancer Biomarker

Background: The main screening method for prostate cancer is serum analysis for prostate-specific antigen (PSA), because an elevated PSA can indicate prostate cancer. However, PSA can also indicate benign conditions, including benign prostatic hyperplasia (BPH) – the abnormal growth of benign prostate cells – especially when levels are 4-10 ng/ml. In this "grey zone," biopsies have been the only way to distinguish between benign and malignant conditions.

Prostate-specific membrane antigen (PSMA) is a protein expressed in prostate tissue. Because it appears in low quantities in BPH and is overexpressed in prostate cancer, several studies have suggested its potential as a diagnostic or prognostic biomarker, but no one had been able to develop a test sensitive to quantify PSMA in body fluids.

Advance: NIH researchers used an innovative protein biochip immunoassay to measure and compare serum PSMA levels in healthy men and in patients with benign or malignant prostate disease. According to the assay, average PSMA values were greater than 600 ng/ml for men with prostate cancer, approximately 115 ng/ml for men with BPH, and about 270 ng/ml for healthy men.

Implications: Serum PSMA may be a more effective biomarker than PSA for differentiating between benign and malignant prostate disease, and warrants more research to evaluate clinical use of the protein biochip immunoassay.

Xioa Z, Adam B-L, Cazares LH, Clements MA, Davis JW, Schellhammer PF, Dalmasso EA, Wright GL Jr: Quantitation of serum prostate-specific membrane antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. Cancer Research 61: 6029-6033, 2001.

Burden of Disease in Mycosis Fungoides

Background: Mycosis fungoides is a slowly progressing lymphoma of the skin. Since these patients usually survive for twenty or more years with the disease but continue to undergo treatment, they are a good population for developing and quantifying a tool for a measurement of the impact of the disease on their activities.

Advance: A tool was developed that included measurements of the severity of the disease. It was termed a severity-weighted assessment tool (SWAT). Scores were determined by measuring body surface area involved with the various forms of lesions seen in mycosis fungoides and multiplied by a weighting factor determined by the impact of the particular type of skin lesion. Thus, the least severe lesion, the patch was multiplied by one while the most severe forms of the disease manifesting as tumors or ulcers were multiplied by three. For comparison, total body surface area measurements and physician global assessments were also recorded. The tool was shown to agree well with the physician global assessment and also correlated well with response to treatment. Thus, it was demonstrated to be a useful tool in assessing burden of disease in this disease.

Implications: There is a dearth of studies and measurement tools for determining burden of skin disease in general and the burden of specific skin diseases in particular. This article demonstrates the successful development of a tool for determining at least one aspect of burden of disease for this condition. The methodology used is one that could be generalized to other similar diseases and so may not be limited to use in this one relatively rare disease.

Stevent SR, Ke MS, Parry EJ, Mark J, Cooper KD: Quantifying skin disease burden in mycosis fungoides-type cutaneous t-cell lymphomas. Arch Dermatol 138(1): 42-48, 2002.

Identification of Signature Profiles of Altered Gene Expression from Exposure to Specific Chemical Classes using Microarray Technology

Background: Numerous approaches are used to investigate the relationship between chemical exposure, toxicity, and disease states. One approach is to study the changes in gene expression in a biological model in response to a chemical exposure. These changes represent a "signature" of the cellular response to the effect of the compound. Unfortunately, the most well-understood chemical-responsive genes, such as genes coding for enzymes that regulate metabolism, are affected by many different compounds and can't be used individually to distinguish the properties of the response to one chemical versus another. However, it is now possible with DNA microarray technology to examine hundreds of genes at the same time, where the collective state (or expression) of these genes would define the profile or "signature" associated with exposure to a chemical. The concerted expression pattern across the full set of genes would thus constitute the expression profile of that compound at a given dose and time. Ultimately, it should be possible to detect similarities between profiles of chemicals that share mechanisms of action and/or toxicological responses, giving rise to a new genomics approach to predictive toxicology.

Advance: The application of these seminal studies to predict exposures to specific chemical classes was reported in three publications. First, researchers used microarray techniques to successfully generate gene expression profiles for mRNA obtained from animals exposed to four chemicals (three from a similar class and one distinct.) They were able to show, through the use of pattern recognition algorithms and computational analyses, that the patters were distinct for each chemical but that the profiles of chemicals from the same class were more similar to each other than to the one from a different class. In the second report, they tested the hypothesis that it would be possible to use the profile characteristics of a blind-tested compound to predict the nature of its toxicity. The results demonstrated that even with a minimum learning set with which to compare the profile of the unknown sample, it was possible to correctly determine patterns of altered gene expression which were concordant with previously determined patterns for 22 of 23 samples analyzed. These two reports were followed by a subsequent study to relate altered patterns of gene expression to specific toxic effects, a concept referred to as phenotypic anchoring of gene expression patterns. This study demonstrated a methodology for analysis of the liver toxicity of methapyrilene revealed by cDNA microarray analysis of samples of livers from exposed animals. This study demonstrated that it is possible to "phenotypically anchor" the results of microarray analysis of the tissues exposed animals to specific parameters of toxicity.

Implications: Development of novel approaches for high-throughput screening for potential adverse effects of chemicals is a major goal for both drug development and environmental health science. Many different endpoints have been used as measures for prediction of potential adverse effects of chemical exposure, but with limited success. The results reported here provide strong evidence in support of a toxicogenomic approach to classification of host effects from exposure to environmental agents.

Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Sieber S, Bennett L, Li L, Tennant RW, Stoll R, Barrett JC, Blanchard K, Paules RS, Afshari CA: Gene expression analysis reveals chemical-specific profiles. Tox Sci 67: 219-231, 2002.

Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Bennett L, Li L, Tennant RW, Stoll R, Barrett JC, Paules RS, Blanchard K, Afshari CA: Prediction of compound signature using high density gene expression profiling. Tox Sci 67: 232-240, 2002.

Hamadeh HK, Knight BL, Haugen AC, Sieber S, Amin RP, Bushel PR, Stoll R, Blanchard K, Jayadev S, Tennant RW, Cunningham ML, Afshari CA, Paules RS: Methapyrilene toxicity: Anchorage of pathologic observations to gene expression alterations. Tox Path 30(4): 470-482, 2002.

System Developed for the Rapid Modification of Genes Within Cells

Background: The completion of the sequence of the human genome, as well as those of other organisms, opens the way for new studies of the functions of genes in health and in disease. In order to capitalize on the rich possibilities of this new knowledge, better systems are needed for functional evaluation of genes within large DNA molecules. A disadvantage of present systems typically used for the manipulation and modification of such large molecules is the requirement for in vitro mutagenesis, which depends on laborious subcloning of the region to be mutagenized. If the mutation is created by transformation with (insertion of) a polymerase chain reaction product into the DNA, extensive sequencing is required.

Advance: A new system demonstrated in yeast, referred to as *delitto perfetto* provides a simple, highly efficient system for changing genes and chromosomes. It relies on the placement of a "cassette" in a gene and the subsequent removal with very small seqeuences (oligonucleotides) that provide the desired changes in the gene. Using this yeast-based system, it has been possible to study changes in the most prominent tumor suppressor gene p53 to understand better the consequences of the tumor associated mutations. The system can easily be applied to many genes that are associated with human disease.

Implications: The rapid modification of genes provides opportunities to study gene function and evaluate drug responsiveness.

Storici F, Lewis KL, Resnick MA: *In vivo* site-directed mutagenesis using oligonucleotides: a versatile approach based on recombination in yeast. Nat Biotech 19: 773-776, 2001.

Heart Assist Device Found to Extend and Improve Live of Patients with End-Stage Heart Failure

Background: Heart failure affects an estimated 4.7 million Americans, with 550,000 new cases diagnosed each year. It is a progressively debilitating disease in which the heart gradually loses its pumping ability and works less efficiently. The 5-year survival rate for all heart failure patients is approximately 50 percent. Patients with mild to moderate heart failure can be helped by drug treatments; but for patients with severe heart failure, heart transplantation is the only treatment that provides substantial benefit, and the small number of heart donors worldwide severely limits its availability.

Advance: Investigators recently demonstrated that despite several serious side effects, a type of implantable heart pump, called a left ventricular assist device (LVAD), can significantly extend and improve the lives of terminally ill patients with end-stage heart failure who are not eligible for cardiac transplantation. The LVAD tested is the size of a portable compact disc player and has a textured interior surface that helps prevent blood clots and reduces the need for blood-thinning medication. Participants in a clinical trial conducted over 3 years were randomly assigned to receive either the LVAD or existing optimal medical care. As a group, and regardless of age, 52 percent of the patients with the LVAD survived one year, compared with 25 percent of the patients in the comparison group. The 2-year survival rate was 23 percent for the LVAD group versus 8 percent for the comparison group.

Implications: The results of this study offer hope that LVADs can significantly extend and improve the quality of the lives of the many terminally ill patients with end-stage heart failure. Prior to this study, circulatory mechanical assistance had been used mainly as a short-term means to extending survival until a donor heart became available. Now it has been shown that, even with a high rate of complications, an LVAD can be of great benefit to those patients for whom transplantation is not an option.

Rose EA, Gelijns AC, Moskowitz AJ, et al.: Long-term use of a left ventricular assist device for end-stage heart failure. N Engl J Med 345(20): 1435-1443, 2001.

New Noninvasive Imaging Technique Can Pinpoint Tissue Inflammation in Mice

Background: Researchers are trying to develop more reliable, noninvasive tools to assess blood vessel inflammation and tissue injury, two events commonly associated with coronary heart disease. Previously, a group of investigators developed a technique that used ultrasound imaging to assess tissue inflammation in mice. The technique involved preparing microbubbles that attach themselves to white blood cells in inflamed tissue, and then using ultrasound to image the microbubbles. Because only a small proportion of the microbubbles were actually retained in the tissue, the resolution available from the technique was limited.

Advance: In further work with mice, the same group of investigators developed an improved imaging technique using microbubbles that results in a stronger ultrasound signal when imaging inflamed tissue. They developed microbubbles that carry antibodies against P-selectin, a molecule activated during inflammation and during periods of inadequate blood supply. The presence of the antibodies caused the microbubbles to be retained in inflamed tissue in sufficient numbers to provide an improved imaging capability.

Implications: These findings suggest a potential new imaging technique – the use of microbubbles combined with ultrasound – for determining the degree and extent of the inflammation and tissue injury commonly associated with coronary heart disease. Further research is needed to refine the technique and adapt it for use in humans. If successful, the technique could result in earlier, more reliable diagnoses for heart disease patients and enable physicians to tailor treatments to individual patients.

Lindner JR, Song J, Christiansen J, et al.: Ultrasound assessment of inflammation and renal tissue injury with microbubbles targeted to P-selectin. Circulation 104: 2107-2112, 2001.

Eyes are the Window to Organ Damage from Sickle Cell Disease

Background: Sickle cell disease is characterized by defective red blood cells. When those cells become lodged in small blood vessels, impaired blood circulation results, which causes pain and leads to organ damage. Heretofore, no good, noninvasive ways were available for studying blockages in small blood vessels to predict the extent of organ damage or to assess a patient's response to treatment.

Advance: Researchers recently demonstrated that a noninvasive method for monitoring blood flow through vessels on the surface of the eye can be useful for evaluating complications of sickle cell disease. Using a technique called intravital microscopy, they observed that, in general, blood flow was slower and blood vessels were fewer in number and smaller in diameter in people who had sickle cell disease compared with those who did not. Moreover, when patients were experiencing a sickle cell crisis, blood vessel diameter decreased even further and blood flow slowed to the extent that it was barely detectable in some cases. The changes were transient, however, and blood flow and vessel diameter reverted to their previous values after a crisis was resolved.

Implications: Because the microvascular organization on the eye surface mimics that found in internal organs, observations made using intravital microscopy to examine these vessels can provide some indication of what is likely to be occurring in other organs that are susceptible to damage from sickle cell crises. In addition to providing clinicians with a method for assessing patients' prognoses and responses to existing treatments, the approach also should be useful to investigators when testing new treatments for sickle cell disease that affect blood flow.

Cheung AT, Chen PC, Larkin EC, et al.: Microvascular abnormalities in sickle cell disease: a computer-assisted intravital microscopy study. Blood 99(11): 3999-4005, 2002.

Wanted: A Bone Marrow Transplant Protocol that Does Not Eliminate a Recipient's Marrow, Does Not Require a Matching Donor, and Cures Sickle Cell Anemia

Background: Sickle cell disease can be cured by bone marrow transplantation, but this procedure is unavailable except for patients who have a matched donor. It is also potentially lethal because a patient's own bone marrow must first be destroyed by toxic drugs and/or radiation.

Advance: Scientists have reported a new approach to bone marrow transplantation that cured 10 of 13 mice developed as "models" for sickle cell disease. Only a portion of the bone marrow of each recipient mouse was removed to make room for the new marrow, and along with the new bone marrow, the mice were given antibodies that stopped their immune systems from recognizing the donor marrow as foreign. This was particularly important because the bone marrow of the donors was not considered to be "a match" for the recipients.

Implications: Although many additional studies need to be done before clinical trials of the new transplantation procedure can be seriously considered, researchers have found a method of bone marrow transplantation that theoretically may provide a benefit with fewer risks than the one currently used. Larger animal models and humans are much more sensitive to mismatching than mice are, but the ability to circumvent the requirement for a matching donor in a small animal model in a laboratory indicates that the requirement may also be overcome someday in the clinic.

Kean LS, Durham MM, Adams AB, et al.: A cure for murine sickle cell disease through stable mixed chimerism and tolerance induction after nonmyeloablative conditioning and major histocompatibility complex-mismatched bone marrow transplantation. Blood 99(5): 1840-1849, 2002.

A New Test to Help Measure the Effectiveness of Novel HIV Vaccines

Background: Intense research is under way to design novel and effective vaccines against human immunodeficiency virus type 1 (HIV-1), the virus that causes Acquired Immunodeficiency Syndrome. One important goal of such efforts is to develop vaccines that can stimulate the production of so-called neutralizing antibodies, which bind to the virus and prevent it from infecting cells. To compare and optimize potential vaccines, researchers need a way to precisely measure neutralizing antibody responses in the laboratory. A number of approaches have been developed that involve exposing a sample of HIV to antibody and measuring a reduction in the virus's ability to infect white blood cells grown in the laboratory and the ability of antibodies to reduce infection. Typically, these approaches have been labor intensive and time consuming, provided only indirect measures of the number of HIV-infected cells, and yielded inconsistent results. More recent approaches are based on flow cytometry, a procedure that detects light absorbing or fluorescing properties of individual cells. Cells are labeled with a fluorescent dye and are counted as they pass through a laser-based detector system. Flow cytometry approaches have been used to detect directly HIV-infected white blood cells by fluorescent labeling of an HIV protein known as p24 inside the cells.

Advance: NIH scientists developed a strategy that allows the direct detection of HIV-1-infected white blood cells 2 days after the cells are exposed to the virus. The researchers infected the cells with HIV-1 in the laboratory and grew them in the presence of a protease inhibitor, a drug that prevents the virus from replicating. Scientists used a flow cytometry-based HIV-1 neutralization test that precisely and reproducibly measures the number of white blood cells infected with the virus and the extent of inactivation of infectious virus due to reaction with neutralizing antibodies. The addition of the protease inhibitor prevents secondary rounds of virus replication and allows clear discrimination of infected cells. The detection of fluorescently labeled p24 inside the cells by flow cytometry identifies over 90 percent of HIV-1-infected cells. Overall, this approach is faster, more accurate, and less expensive than other methods.

Implications: The accurate and precise measurement of neutralization by antibodies against HIV-1 is important for understanding how these antibodies neutralize the virus and for evaluating antibody responses to potential vaccines. The precision and reproducibility of this new method and its ability to quantify directly the number of white blood cells infected by HIV should be useful for comparing the neutralizing antibody response generated by various potential HIV-1 vaccines.

Mascola JR, Louder MK, Winter C, Prabhakara R, De Rosa SC, Douek DC, Hill BJ, Gabuzda D, Roederer M: Human immunodeficiency virus type 1 neutralization measured by flow cytometric quantitation of single-round infection of primary human T cells. J Virol 76(10): 4810-4821, 2002.

Sequencing Genomes: Progress in Model Organisms and Selection of New Target Genomes

Background: From the Human Genome Project's outset, the NIH and its partners have included among their research goals the mapping and sequencing of the genomes of several other organisms. Complete genomic sequences of many bacteria (e.g., *E. coli)*, several fungi (e.g., *S. cerevisiae)*, and two invertebrates (the roundworm *C. elegans*, and the fruit fly *D. melanogaster)* were generated in the past few years. Those sequencing efforts provided an enormous amount of information, including cost-effective ways to sequence entire genomes and greater understanding of the biological data encoded in genomic DNA sequence. Following the draft sequence of the human genome, interest in sequencing the genomes of many other organisms rose dramatically.

Advance: Genomic sequences for a number of important organisms, beyond those initially identified by the Human Genome Project, have been determined.

- Primary among these is the laboratory mouse. In May 2002, a draft sequence with 96 percent of the mouse genome in long, continuous stretches of DNA was placed in public databases. This major milestone provides a key tool for interpreting the human sequence.

- Since February 2001, the Rat Genome Sequencing Project has achieved about 90 percent of its overall goal and will complete a draft sequence by the end of 2002. The rat is important in physiological studies and the availability of its genomic sequence will greatly facilitate studies using the rat and accelerate the availability of new therapies.

- The zebrafish (*D. rerio*) has become an important model organism and is currently used widely in developmental biology. The Sanger Institute began sequencing its genome in 2001. In July 2002, they released an initial assembly of the zebrafish genome.

- The NIH, and other sequencing efforts produced, in the past year, genome sequence for rice, two sea squirts, two puffer fish, and a second round worm. The NIH also recently funded sequencing of a second species of fruit fly. Availability of the genomic sequences from closely related organisms, such as two round worms, two fruit flies, and human and other mammals, is a great advantage in interpreting sequences' meanings. Proliferation of new sequences has stimulated development of a new field, comparative genomics.

A peer review process for selecting which genomes to sequence has been established. Scientists who want to champion an organism for genomic sequence must write a "white paper" that clearly presents the arguments for putting their favorite at the head of the line. After two rounds of white paper submissions, this process has determined the highest priority as: chicken, chimpanzee, cow, dog, a set of fifteen fungi, honey bee, sea urchin and two protozoans (*Tetrahymena* and *Oxytricha*). Sequencing of the chicken and the chimpanzee has already begun. New cycles of this review process occur every four months, and it is likely that other genomes will be added to the high priority list soon.

Implications: Availability of genomic sequence and assemblies for multiple organisms provides crucial data for identifying functionally important human DNA sequences and, ultimately, understanding human disease. By comparing the sequences of different organisms, researchers gain valuable clues to understanding the human genome's structure and function.

Thomas JW, Touchman JW: Vertebrate genome sequencing: building a backbone for comparative genomics. Trends Genet 18(2): 104-108, 2002.

Model Organism Databases

Background: With the magnitude and diversity of new scientific data, research communities face challenges sharing data across disciplines, especially in genomics. Need exists to develop and maintain databases of genomic data to support the production of robust, exportable software that can be widely shared among different databases. To address these issues, Model Organism Databases (MODs), funded by the NIH, capture and organize key information, integrate data from disparate sources, and facilitate the formulation of new hypotheses and new perspectives.

Advance: The MODs have recently faced these data challenges by developing new database programs to help the research community. Each MOD has solved problems that are unique to its database. In addition, through the Generic Model Organism Database project, the NIH has encouraged development of shared database software to increase functionality of the databases.

An example of a MOD that is developing new ways to manage and share data is the Saccharomyces genome database (SGD). The SGD faces a challenge that all MODs will soon confront, the availability of genome sequences of closely related species. Several close relatives of *Saccharomyces cerevisiae* have been sequenced. The SGD, in collaboration with the Whitehead Institute, has developed a genome display that aligns the different yeast chromosomes and allows for easy identification of related genes, along with links to protein alignments. Availability of this central resource of comparative genomic information will enable better annotation of genes, the identification of potential regulatory elements, and the mapping of highly diverged amino acid residues in protein domains.

The MODs have dealt with the problem of database interoperability by forming the Gene Ontology (GO) Consortium to promote a shared vocabulary, or ontology, to describe the properties of gene products. Bioinformaticians have widely adopted the GO vocabularies as a standard method to annotate gene products, avoiding a potential Tower of Babel that might otherwise ensue. The GO consortium is leading an effort to establish additional ontologies to describe other information in gene databases. This effort includes the development of phenotype and anatomy ontologies by the mouse genome informatics group and sequence ontology to describe DNA sequence features by the fruit fly database called FlyBase.

Implications: Integration of the diverse data generated as part of the HGP is critical for understanding human disease and developing treatments. The MODs are developing new methods to keep pace with the increasing amount of data and to assist further in their analysis. Many of these methods will be useful in future human genome databases that will convert findings in model organisms to new approaches to improving human health.

Blake JA, et al: The mouse genome database (MGD): the model organism database for the laboratory mouse. Nucleic Acids Res 30(1): 113-115, 2002.

The FlyBase Consortium. The FlyBase database of the Drosophila genome projects and community literature. Nucleic Acids Res 30(1): 106-108, 2002.

Dwight SS, et al: Saccharomyces genome database (SGD) provides secondary gene annotation using the gene ontology (GO). Nucleic Acids Res 30(1): 69-72, 2002.

Bacterial Artificial Chromosome Resource Network

Background: In the past several years, one system for recovering and purifying large DNA segments emerged as the method of choice for the construction of chromosomal DNA libraries – the bacterial artificial chromosome (BAC). BAC "libraries" that contain complete collections of DNA fragments from a given genome propagated in bacteria are important tools in the study of DNA sequence at a whole genome level. With increasing interest in genomic approaches to biological research, demand for new BAC libraries will increase rapidly in the next few years.

To date, the complete sequencing of all large eukaryotic genomes (including *C. elegans*, *D. melanogaster*, *A. thaliana*, M. musculus, and *H. sapiens*) has involved BAC clones. Recent experience in sequencing the human genome shows that determining the sequence of a large, repeat-rich genome requires an ordered map of clones, such as BAC clones, to assign stretches of DNA sequence to their location in the genome and to assemble regions containing repeated DNA sequences. While an alternative strategy, the "whole genome shotgun" approach, will likely be part of any project to sequence such genomes, it is not capable of producing a finished sequence, and so does not obviate the need for well-mapped BAC clones and, therefore, for BAC libraries.

Advance: To increase the number of available BAC libraries and to generate a large number of BAC libraries from many organisms, the NIH awarded a set of cooperative agreements to form the NIH BAC Resource Network. The network will enable researchers to sequence specific genomic regions and to find and characterize genes. It will produce at least 15 BAC libraries at ten-fold coverage of "mammalian-size" genomes or the equivalent. A panel of six scientists oversees the network, regularly evaluates the program's progress, and makes recommendations to the participating institutes about the program. A peer review committee assesses requests for organism candidates for BAC libraries on the basis of scientific interest and strategic feasibility. For each organism proposed, the committee recommends whether the NIH BAC Resource Network should accept the request and, if so, whether it should be assigned to a high or to a standard priority pool.

Through this program, new BAC libraries have already been constructed and made available to the research community for eight organisms, including four mouse strains, the vervet monkey, the Sumatran orangutan, the gorilla and the duck-billed platypus. Over 20 other organisms have been approved for BAC library construction and have entered the production pipeline.

Implications: Because BAC clones are relatively large and appear to represent an organism's genome faithfully, the BAC system will also be the choice for isolating targeted regions of genomic DNA from additional organisms used in specific biological studies, including a variety of mouse strains, and even from individual humans. Although worldwide genomic sequencing capacity has increased considerably in the last few years, certain research organisms will not be completely sequenced for some time. Availability of BAC libraries for many organisms will enable genomic studies to advance in the absence of complete genomic sequence.

http://www.genome.gov/page.cfm?pageID=10001844
http://www.genome.gov/page.cfm?pageID=10001852

Yeast Gene Deletion Mutant Set – A Powerful Tool for Understanding Genes and Proteins

Background: Sequencing its genome and identifying all its genes is just the beginning of understanding the biology of an organism. Scientists want to know the function of all the proteins the genes encode. They also want to know how the genes interact to form genetic networks and how the proteins work together in the cell to carry out biological processes. Deciphering this information is technically challenging, and much of the work to develop large-scale approaches to do so has occurred in model organisms, such as the yeast, *Saccharomyces cerevisiae*. Yeast has served as an important model organism for both developing new genomic technologies and understanding the genomes of higher organisms, including humans.

Advance: One approach used to study gene function and genetic regulatory networks is to make and analyze genetic mutations that alter protein expression. While the principle behind this approach is as old as classical genetics, the availability of a full genome sequence affords a new opportunity to assess systematically and rapidly the function of every gene in a cell. Following the completion of the sequence of *Saccharomyces cerevisiae* in 1996, the NIH supported the U.S. component of an international consortium formed to construct systematically a complete collection of gene-deletion mutants. To allow researchers to keep track of each deletion strain and to assess the effects of the mutation, short DNA sequences, "molecular bar codes," were inserted into each gene deletion. The bar codes can track mutant strains even within populations. Use of these molecular bar codes avoids the need to study the fitness of each deletion strain individually and provides an efficient, high-throughput approach. Construction of the set of deletion strains was recently completed, with 96 percent of genes in the yeast genome disrupted. These strains are available through distributors in both the U.S. and Europe. A number of investigators now use this set of deletions in a variety of studies aimed at understanding gene function.

While looking at single mutants is a powerful route to understanding function, it is clear that proteins do not act on their own; rather, numerous proteins act together to carry out a particular biological function. In addition, one difficulty in uncovering a gene's function is that a high degree of functional redundancy exists; in many cases another gene is similar enough to compensate if a gene is deleted. To address these questions, NIH recently funded a group at Johns Hopkins University to look systematically at how yeast deletions act in combinations by generating a second gene deletion in strains with existing deletions. Combinations that lead to a different phenotype from that produced by either individual deletion indicate that the proteins somehow act together to carry out a biological process. This work should aid in the development of a global protein interaction map.

Implications: The use of gene deletion strains provides an opportunity to employ a large-scale approach to determine the function of genes and their protein products and to start to determine how genes and their expressed proteins interact with each other.

Giaever G, et al: Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418: 387-391, 2002.

Winzeler EA, et al: Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science 285: 901-906, 1999.

Hemochromatosis and Iron Overload Study Update (HEIRS)

Background: Iron overload is a common disorder estimated to affect between 1 in 200 and 1 in 400 individuals. One of the major causes of iron overload is the genetic disorder, hereditary hemochromatosis (HH). It is estimated that one in every 400 Caucasians has HH due to variations in the gene, *HFE*. The genetic contributions to HH in non-Caucasian populations are not at all clear. When two specific *HFE* mutations were discovered in 1996, some observers suggested carrying out widespread screening for *HFE* mutations. In 1997, a panel of experts convened by the NIH and the Centers for Disease Control and Prevention concluded that it would be premature to consider widespread genetic screening for HH prior to answering important questions about prevalence, penetrance, genotype/phenotype correlations, and the psychosocial impact of such screening. Questions about prevalence, penetrance, and genotype/phenotype correlations were particularly problematic in diverse populations.

Advance: To answer these questions, the NIH is funding a large multi-center study designed to study iron overload and HH in diverse populations. Recruitment of participants began in February 2001, and continues at five field centers. As of September 2002, more than 80,000 participants have been recruited into the initial screening phase, almost 50 percent of whom are from minority populations. Comprehensive clinical exams of those screening positive are underway and the family study is beginning. Analysis of data will begin in the coming year.

Implications: Early findings from this study are expected within the next one to two years. The findings will inform health policy debates about whether more widespread screening for iron overload and hereditary hemochromatosis should be implemented.

User's Guide to the Human Genome

Background: Genomic information, and the methods that allow useful access to it, have revolutionized biology. The NIH's commitment to rapid and open data release and to providing usable tools that ensure access by both advanced genomic researchers and lay biologists has made genomic information available throughout the scientific community and to other interested parties. The availability of the sequences of many genomes through the Internet has made an extraordinary amount of information freely accessible to anyone with a desktop computer and a link to the World Wide Web. However, the information itself is not enough to allow efficient use. Interested people from outside the centers of genomic study need to know where best to view the information in a form suitable for their purposes and how to take advantage of the software provided for retrieval and analysis.

Advance: In September 2002, researchers from the NIH published a User's Guide to the Human Genome. The electronic version of the guide appears on *Nature* magazine's website (www.nature.com/cgi-taf/DynaPage.taf?file=/ng/journal/v32/n1s/index.html) and a hard copy version have been distributed to the subscribers of *Nature Genetics*. The guide offers help to those who might otherwise have trouble in using the products of genomics. The guide is an additional resource for scientists who wish to make use (or better use) of both sequence data and the major tools that can be employed to view these data. The guide provides information in a highly inviting and understandable format. It is written in a practical, question-and-answer format, with step-by-step instructions on how to approach a representative set of problems using publicly available resources.

Implications: The User's Guide to the Human Genome, which is freely available on the World Wide Web, will help investigators, scientists, and students better understand and capitalize on the enormous potential of the data from the HGP. By making the HGP data more accessible, it will speed the pace of discovery, while encouraging young scientists to enter the field of genomics.

http://www.nature.com/ng/

Direct Brain Control of Neuroprosthetic Devices

Background: Over the last three decades the NIH Neural Prosthesis program has pioneered the development of electronic and mechanical devices that help restore functions lost through damage to the brain, spinal cord, or nerves of the body. Cochlear implants improve hearing for thousands of people, illustrating the potential of neural prostheses in the sensory realm. For movement, the Freehand ® system allows some people with serious spinal cord injuries to grasp an object such as a comb or a cup, improving independence in daily living. However, devices such as this grasping prosthesis require some residual movement capability for control.

Advance: Decades ago, when scientists first recorded the activity of nerve cells in the brain that control movement, researchers speculated that these brain signals might someday be harnessed to directly control neuroprosthetic devices. Scientists have now devised a system using electrodes implanted in the brain and a sophisticated computer program to interpret the recorded signals, which brings this goal much closer. With this device, monkeys can quickly and accurately move a cursor on a computer screen by thought alone. The investigators first studied brain signals while the monkeys moved the cursor with a hand control, then wrote a computer program to translate those brain signals directly into cursor movements. While direct brain control devices have been previously tested in animals and in paralyzed humans, this new system worked essentially instantaneously, without requiring any special training, and allowed control in a virtual 3-dimensional world. Signals from far fewer nerve cells were required for accurate movement control, bringing the number within the range obtainable with devices that might be implanted in the human brain. Success relied upon learning by the brain itself as well as "learning" by the computer controllers, which adapted to the changing brain signals.

Implications: These results suggest that it may be possible to develop neural prosthetic devices that are controlled directly by the brain. Substantial obstacles remain, but the combination of mathematical algorithms, computer and electronic technology, biomaterials, and understanding of how the nervous system controls movement may someday be translated into effective systems to help restore function to neurologically impaired people.

Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP. Instant neural control of a movement signal. Nature 416: 141-142, 2002.

Taylor DM, Helms Tillery SI, Schwartz AB: Direct cortical control of 3D neuroprosthetic devices. Science 296: 1829-1832 2002.

Tracing Brain Activation Using Manganese-enhanced Magnetic Resonance Imaging

Background: Brain imaging techniques such as magnetic resonance imaging (MRI) have revolutionized clinical practice, as well as opening new approaches to understand how the normal brain works. MRI can non-invasively reveal brain anatomy with impressive detail in living people. Functional MRI shows not just brain structure, but also which areas of the brain are active during various behaviors. To understand how brain circuits work, it would be useful to further develop brain imaging techniques to trace the pathways of activation in the brain during various tasks, in order to see not only the anatomical connections among functional brain areas, but also the specific routes in the brain that are activated during a specific stimulus or behavior.

Advance: Scientists have developed a method to enhance MRI using manganese. With this method, researchers can trace the nerve pathways activated by specific odors, beginning with nerve cells in the nasal cavity that detect odor molecules in the air, and passing along their nerve fibers to distinct areas in the olfactory bulb of the brain that process the information. Previous studies had shown that manganese ions can enter excitable cells through calcium channels, which normally allow calcium ions to pass through activated cell membranes. In addition, neuroanatomists have used manganese ions to trace neuronal connection in the mouse brain because these ions are transported along nerve fibers. Because manganese is paramagnetic (can be induced to be weakly magnetic) it is detectable by MRI imaging. Manganese-enhanced MRI relies upon these fortuitous properties of manganese by introducing trace amounts of this ion to the nasal cavity along with aerosolized solutions containing specific odors and monitoring the brain activation with MRI imaging.

Implications: Using this approach, it may someday be possible to map the activation of the olfactory system exposed to a wide variety of odors and thus help explain how the system processes odor information. It is worth noting that processing of smell is often disrupted early in the course of neurodegenerative disease, so better understanding of olfaction may have many implications. There is a natural trade-off between the exposure time of an MRI image and the detail obtained, so the fact that the manganese enhancement lasts for a long time compared to other methods that measure brain activation should allow very detailed images to be constructed. More generally, this approach illustrates how MRI imaging tools are continuing to evolve, often first as a tool to understand brain function, as in this case by allowing scientists to trace activated neural pathways in the brain, and subsequently as a diagnostic tool.

Paulter, RG, Koretsky AP. Tracing odor-induced activation in the olfactory bulbs of mice using manganese-enhanced magnetic resonance imaging. NeuroImage 16: 441-448, 2002.

Chemical from Sea Anemone Ameliorates Animal Model of Multiple Sclerosis

Background: Multiple sclerosis (MS) is a chronic, inflammatory disease of the brain and spinal cord that commonly begins in young adults. The disease destroys myelin, which electrically insulates nerve fibers and gives "white matter" its appearance. MS also damages the nerve fibers themselves. Symptoms are usually recurrent, ranging from mild to very severe, but growing progressively worse through the course of the disease. These problems may include visual disturbances, weakness and coordination problems, abnormal sensations such as numbness, "pins and needles" pain ,or loss of feeling, and mild cognitive impairments. The discovery that the patients' own immune system is responsible for much of the damage led to treatments that can reduce symptoms and somewhat slow the course of the disease, but these therapies are still far from adequate.

Advance: Scientists have identified a component from the venom of the sea anemone that can ameliorate both moderate and severe experimental autoimmune encephalomyelitis (EAE), a rat model of multiple sclerosis. In EAE, the rats' T-cells, a type of immune cell, are provoked to attack myelin, as in the human disease. By studying T-cells in cell culture, the researchers identified ion channels whose activity is crucial to the activation of the cells' damaging activity. Ion channels allow specific charged ions, in this case potassium, to enter or leave the cell, thereby regulating various functions. Venoms of snakes, insects, and sea creatures such as snails and sea anemones often contain highly specific ion channel toxins, apparently because the crucial regulatory roles of ion channels make them a good target for disabling their prey. Once the team identified a specific toxin (ShK) that acted on the channel (Kv1.3), they found that this chemical, and another chemical (TRAM-34) with similar effects on the same channel, greatly reduced the clinical course of both moderate and severe EAE in rats.

Implications: Although more experiments are needed in other types of animal models, these findings suggest that ShK, or drugs with similar effects, might help treat multiple sclerosis in people. More generally, these results illustrate how progress in the fundamental scientific studies of ion channels is leading to progress against specific disorders. Advances in genetics have allowed scientists to identify the many types of ion channels, and techniques such as patch clamp recording, used in the T-cell studies, allow scientists to monitor their activity directly, enabling researchers to unravel the crucial roles channels play in controlling a wide range of cellular functions, in virtually all organ systems.

Beeton C, et al.: Selective blockade of T lymphocyte K channels ameliorates experimental autoimmune encephalomyelitis, a model of multiple sclerosis. Proc Nat Acad Sci 98(24): 13942-13947, 2002.

Microarray "Fingerprinting" of Disease

Background: Microarray technology, or "gene chips," allow scientists to simultaneously monitor the activity of thousands of genes in a single tissue sample. The patterns of gene expression (activity), for example, not only differ in tumors compared to normal tissue, but also between tumors that do and do not respond to particular drugs. Because blood is the tissue most easily obtained from patients, it is noteworthy that even white blood cells demonstrate different gene expression patterns under different conditions. This led researchers to wonder whether the gene expression patterns in white blood cells could serve as a "fingerprint" to help identify various medical states that affect the brain and are currently difficult to assess.

Advance: Using microarrays, scientists have demonstrated the potential of gene expression profiles of white blood cells to identify a variety of neurological conditions. Rats were subjected to ischemic strokes (reduced blood flow to brain cells), hemorrhagic strokes (bleeding into the brain), sham surgeries, drug induced seizures, hypoxia (oxygen deprivation), or insulin induced hypoglycemia (low blood sugar), and compared with normal rats. One day later scientists analyzed gene expression profiles in the rats' white blood cells using microarrays that measured the activity levels of more than 8,000 genes. Each condition produced reliable changes in gene expression characteristic of the particular condition, including substantial increases or decreases in the activity of particular genes.

Implications: With confirmation in humans, it may be possible to use microarray analysis of blood to diagnose the occurrence of many medical conditions that affect the brain. Physicians might, for example, detect recent hypoglycemia in diabetic persons, seizures in patients with epilepsy, or brain hypoxia or ischemia in people with cardiovascular problems. Although only a few conditions were examined in this study, these results reinforce an emerging picture that microarray analysis of gene expression may become an increasingly useful tool for disease analysis, not only for research but also in the clinic.

Tang Y, Lu A, Aronow BJ, and Sharp FR. Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease. Ann Neurol 50(6): 699-707, 2001.

Spatial Language and the Development of Wayfinding Systems

Background: A major goal of blindness rehabilitation programs is the development of wayfinding systems. This requires a better understanding of the cognitive requirements that underlie the effective use of navigation systems by the visually impaired. Recent work has shown that designers of wayfinding systems can use the mind's ability to use spatial language to update its internal mental image or map of a location or object relative to the body's location. Spatial language describes the location of an object relative to the location of a person's position (e.g. "a chair at 3 o'clock"). Up to now it has not been clear whether mental images based on spatial language cues could be used to guide one's navigation or action through space with the same degree of precision as achieved by perceptually-based images using auditory cues.

Advance: Vision researchers studied the role of spatial updating by experimental subjects as they moved using spatial language cues. Spatial updating refers to a person's ability, while moving, to update the location of a target previously perceived while stationary. The researchers were interested in knowing if a subject could navigate to a target using spatial language and updating as effectively as they could using auditory cues. Both blindfolded normally-sighted subjects and blind subjects were used for these studies. All participants were tested on their ability to walk toward a target location. They were cued on the target location one of two ways: an auditory 3-D stimulus from a loudspeaker at the target location, or an experimenter providing directions using spatial language with "o'clock" for direction and linear feet for distance (e.g. 3 o'clock, 20 feet). In some trials, the subjects were asked to walk directly towards the target location; in other trials they walked forward from the starting point, then had to turn and walk toward the target after being tapped on the shoulder.

Implications: Two results emerged from these studies. First, participants reached approximately the same stopping point whether they took a direct or an indirect route. This result provides evidence of spatial updating, the capacity to keep track of the target location without receiving perceptual input. This suggests that once a subject formed a spatial image it stayed fixed in the environment. The second finding showed that both groups, the blindfolded sighted and the blind subjects, reached the target locations equally well whether they used spatial language or 3-D auditory input. This suggests that once a mental image of a location is formed one can update the representation using either mode of input. Blind subjects were equally adept at spatial updating using spatial language, leading the researchers to conclude that visual experience may not be required for the brain to develop spatial updating skills. The findings also suggest that spatial language can convey information about important off-route landmarks (e.g. phone at 3 o'clock). The results of this study will impact significantly on the development of navigation systems for the blind and visually impaired. Wayfinding systems that use spatial language may be the best approach for development of relatively simple wayfinding systems. This is important because many of the growing number of visually impaired are elderly and require system designs that are easy to use.

Loomis J, Lippa Y, Klatzky R, Golledge R: Spatial updating of locations specified by 3-D sounds and spatial language. Journal of Experimental Psychology 28(2): 335-345, 2002.

New Approach to Identifying Disease Genes in the Visual System

Background: An important barrier to therapeutic intervention in human retinal disease is the identification of the gene or genes that cause vision loss. Gene cloning efforts to identify disease genes largely rely on the availability of large, informative human pedigrees; expressed sequence tag databases; or complete genomic sequences to reveal candidate genes. Small sample sizes and the possible incomplete penetrance of the phenotype, however, make gene identification using these strategies an imperfect science. An additional confounding variable is the involvement of multiple organ systems in the disease that hampers candidate gene hunting. For the visual system however, visual loss and the degenerative and other changes in the retina are largely linked to rod and cone photoreceptors.

Advance: Scientists have recently undertaken a comprehensive genetic analysis of rod photoreceptors, the most abundant sensory neuron in the retina, in order to identify all the possible genes expressed in these cells. Rod cells play an essential role in the visual pathway as phototransducers, and thus may be especially vulnerable to any genetic defect involving the retina or other visual centers. Moreover, for many identified retinal disease genes, either a photoreceptor gene is mutated or photoreceptor expression profiles are altered due to the mutation. A complete database of rod-specific genes would simplify the identification of candidate retinal disease genes. Using the technique of serial analysis of gene expression, or SAGE, the expression of rod-specific genes is enriched above their normal levels in tissue, increasing their sensitivity for screening. Using this technique in both mature and developing retinal tissues, 264 new genes that are enriched in rod photoreceptors have been identified.

Implications: This work represents an important new strategy for the cloning of retinal disease genes. The success of this endeavor is distinguished from similar efforts because it focuses on cell-type specific gene expression. Moreover, the use of such a complete cell-specific gene expression profile will eliminate reliance on large human pedigrees or incomplete genetic maps in the search for disease genes that affect limited cells or tissues and provide new insights into photoreceptor biology.

Blackshaw S, Fraioli RE, Furukawa T, Cepko CL: Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. Cell 107: 579-589, 2001.

Crucial Signaling Events are Visualized in Single Living Cells

Background: Cells constantly receive messages from their environment that determine their growth and function. External molecules attach to highly specific sites (receptors) on the cell's surface. The activated receptors then generate signals that send messages into the cell. The cells respond to these signals by initiating a chain of events. These events are called signaling cascades. For example, when insulin binds to its receptor, it activates signaling pathways that instruct the cell to take in glucose and to grow. The organization and sequencing of the signaling events determine the accuracy and specificity of the cell's numerous signaling processes. A small group of molecules, called inositol phospholipids, are responsible for directing the activities of the cell.

Advance: Researchers developed new methods to study signaling events within a single living cell, without disrupting the cell. They used a green fluorescent protein to create "reporter tag" proteins. These "reporter tags" interact with inositol phospholipids and make it possible to follow the dynamic changes that occur when the cell is activated. The tags permit researchers to see and follow highly localized changes in inositol phospholipid levels in many cell responses, ranging from ovum fertilization to bacteria ingestion. The technique has been used for other cellular processes as well. Using fluorescent protein tags, researchers have visualized the activity of a tumor-producing protein called Ras, a key regulator of cell growth and tumor formation. A fluorescent-tagged angiotensin receptor (a major factor in cardiovascular disease) has enabled researchers to identify the compartments within the cell through which the angiotensin receptor is recycled to the cell membrane.

Implications: This new method of tagging biological signal molecules allows scientists to identify small molecules that affect specific signaling steps and to study how these molecules function both in normal and diseased cells. Furthermore, this technique will help researchers to understand how cellular invaders, such as bacteria (or their toxins), use the cellular machinery to enter the cell. This technique can also be used to identify molecular mechanisms as potential targets for drugs to control human disease. Furthermore, visualizing activated Ras in live cells will help shed light on tumor formation.

Balla T, Varnai P: Visualizing cellular phosphoinositide pools with GFP-fused protein-modules. Science's Stke 125: PL3, 2002.

Bondeva T, Balla A, Várnai P, Balla T: Structural determinants of Ras-Raf interaction analyzed in live cells. Molecular Biology of the Cell 13: 2323-2333, 2002.

Hunyady L, Baukal AJ, Gáborik Z, Olivares-Reyes JA, Bor M, Szaszák M, Lodge R, Catt KJ, Balla T: Differential PI 3-kinase dependence of early and late phases of recycling of the internalized AT_1 angiotensin receptor. J Cell Biol 157(7): 1211-1222, 2002.

Technique to Identify Shortened Cervix Predicts Premature Birth

Background: Premature birth (before 38 weeks of gestation) complicates from 8 to 10 percent of all births. Most deaths of premature infants occur in infants born before the 32nd week of pregnancy. Infants born before the 37th week of pregnancy are at greater risk for life-threatening infections, respiratory complications, central nervous system immaturity, and serious damage to the digestive system. The cause of prematurity is unknown in most cases. In an attempt to address this, researchers recently used endovaginal sonography to measure the length of the cervix. The cervix, or "neck" of the uterus, normally shortens late in pregnancy in preparation for birth. The researchers discovered that a shortened cervix in women early in the second trimester of pregnancy appeared to be a warning sign of impending premature birth among women with a history of premature birth. The researchers wondered if cervical shortening earlier in pregnancy could also be predictive of premature birth.

Advance: Researchers examined 183 pregnant women who had previously experienced a premature birth before the 35th week of pregnancy. They wanted to know if shortened cervical length beginning as early as the 16th week of pregnancy predicts spontaneous premature birth. The scientists took the initial cervical measurements between 16 and 18 weeks, 6 days of pregnancy. They took three additional ultrasound measurements before the 23rd week of pregnancy. The researchers found that women whose cervixes had shortened to less than 25 millimeters in length by 16 to 18 weeks of pregnancy were 3.3 times more likely to give birth prematurely than were other women in the study. Furthermore, taking several measurements significantly improved the prediction of spontaneous premature birth. The study confirmed that shortened cervical length early in pregnancy can predict spontaneous premature birth in high-risk women.

Implications: This knowledge lays a strong foundation for future research to identify and to prevent preterm birth in high-risk women. Furthermore, endovaginal sonography proved to be a useful tool to identify women whose cervical anatomy would make them ideal candidates for this research. The next step may be to investigate whether preterm birth can be prevented early in pregnancy in women with a shortened cervix. For example, certain interventions, such as cerclage (encircling the cervix with a ring or loop), might be more effective if applied early in gestation.

Owen J, Yost N, Berghella V, Thom E, Swain M, Dildy III GA, Miodovnik M, Langer O, Sibai B, McNellis D: Mid-trimester endovaginal sonography in women at high risk for spontaneous preterm birth. JAMA 286(11): 1340-1348, 2001.

Building Comfortable Sockets for Amputees

Background: Lower-limb amputees must carry the full load of their body weight, and sometimes up to three times their body weight on the soft tissues of the residual limb during daily activity. Unfortunately, the soft tissue is not designed to be weight bearing. Precise fitting of the socket and maintaining this fit are extremely important to the amputee's level of comfort, skin integrity, stability, sense of security and overall performance. Yet, in most amputees, it has not been possible to create a consistent, precision fit of a socket because of natural fluctuations in the size of an amputee's stump. For example, when an amputee first gets up in the morning, the stump is at its greatest volume. As the day progresses, however, pressure on the stump from the artificial limb causes the stump to lose fluids and to shrink. As a result, it doesn't fit as well into the socket of the artificial limb. The socket rubs against the stump, eventually causing blisters and painful sores. In fact, 75 percent of amputees report that their sockets do not fit well. Many amputees try different prosthetic limbs in a costly and unsuccessful attempt to solve the problem. As a result, the amputees often cut back on their activities. This sedentary lifestyle often leads to weight gain and associated diabetes, heart disease, and other health problems.

Advance: A scientist, who is also a lower limb amputee, developed an innovative, inexpensive and reliable socket that creates a truly integrated relationship between the stump and the socket. The Active Contact System (ACS) utilizes liquid filled bladders that automatically regulate their volume as the amputee walks. This provides a socket that continuously fits securely and comfortably. The ACS eradicates the discomfort that often leads to skin breakdown, curtailed activity, and other health complications.

Implications: Each year, more than 60,000 Americans lose all or some portion of a lower limb. In all, there are about 220,000 lower limb amputees living in the United States. Amputees now using the ACS report that they enjoy greater mobility and more confidence with the new socket than they had with other such devices. This should increase their willingness to use the prosthesis to perform activities of daily living, reduce their health risks associated with immobility and improve their overall quality of life.

U.S. Patent Application #20020099450, July 25, 2002, Robert C. Dean, Jr.

New Brain Imaging Techniques Allow Insight into Neuronal Functioning

Background: Functional magnetic resonance imaging (fMRI) has been used in recent years to examine whole brain function by measuring changes in blood flow and volume. While this technique is relatively common and provides good localization of brain activity, fMRI devices are bulky, expensive to use, and cannot resolve either very fast neuronal activity or very slow changes in blood flow dynamics. Yet, the ability to measure the full range of brain function is crucial for enabling clinical neurologists and cognitive-neuroscientists to better understand various disease states and to detect early abnormal brain functioning. Two disease examples in which this type of technology would be relevant are Parkinson's and Alzheimer's diseases. However, near infrared spectroscopy (NIRS) was recently introduced as a non-invasive and inexpensive, imaging modality capable of measuring changes in the optical properties in the surface of the brain. This technology holds great potential to resolve both fast neuronal signals and slow hemodynamic (blood flow) fluctuations.

Advance: It is well established that brain activity changes optical properties near the surface of the brain, due both to changes in blood flow and scattering from brain cells. It has also been shown that near-infrared light can pass easily through the skull and reach the surface of the brain. Therefore, measurements of slow changes in blood flow can be readily measured by NIRS methods. However, measurements of fast neuronal activity have yet to be demonstrated. The recent development of a new sensor technology by NIH investigators may make this type of observation not only possible, but widely affordable and accessible. In addition, researchers have proven that this new NIRS methodology correlates well with traditional fMRI techniques. The development of such multimodality devices (combining fMRI with NIRS) significantly increases the researchers' and clinicians' ability to resolve high-resolution images of brain function.

Implications: The introduction of this new, complementary imaging modality opens the door to a new approach for observing brain function. It also has the potential to provide a wealth of new knowledge about the progression, detection and treatment of a wide range of neurological disorders. In addition, this technique could be expanded beyond the realm of neurology to applications in cardiology and vascular biology. The development of non-invasive imaging techniques affordable to a wide range of medical centers, clinics, and health facilities is vital to the wide translation and adoption of new advancements for treating and diagnosing disease. Furthermore, as high resolution imaging methods provide greater insight into a wider range of physiological functions, researchers will be better poised to further decipher the mysteries of illnesses such as Parkinson's and Alzheimer's diseases.

Toronov V, Webb A, Choi JH, Wolf M, Safonova L, Wolf E, Gratton E: Study of local cerebral hemodynamics by frequency-domain near-infrared spectroscopy and correlation with simultaneously acquired functional magnetic resonance imaging. Optics Express 9(8): 417-427, 2001.

Toronov V, Webb A, Choi JH, Wolf M, Michalos A, Gratton E: Investigation of human brain hemodynamics by simultaneous near-infrared spectroscopy and functional magnetic resonance imaging. Med Phys 28(4), 521-527, 2001.

Mouse Model Implicates Enzyme in Human Development and Disease

Background: Recently scientists have identified a family of proteins called type II transmembrane serine proteases (TTSPs). Studies show that the seven currently known members of this family encode uniquely shaped enzymes that are adept at cleaving and thus activating other proteins present on the cell surface. What happens after a TTSP enzyme cleaves one of its target proteins remains largely unknown, though the initial data indicate that these proteins might activate cellular processes that are involved in development and maintaining cell homeostasis. Establishing the precise interactions of these enzymes on the cell surface could point the way to important and previously unknown signaling pathways within the cell. Such information is vital not only in defining the biocircuitry of the cell, now a major area of interest in biology, but in learning which circuits to exploit in the future to design more targeted molecular therapies.

Advance: NIH scientists have generated a mouse model in which both copies of the Matriptase/MT-SP1 gene are inactivated. A member of the TTSP family, Matriptase/MT-SP1 was identified in 1998 as a novel, tumor-associated protein. Subsequent studies have shown that it may be associated specifically with tumor invasion and metastasis, though little else is known about the protein in health or disease. In evaluating the newborn mice deficient in Matriptase/ MT-SP1, NIH scientists and collaborators noted saw the pups had unusually dry, red, wrinkly skin. This was caused by malformation of the stratum corneum, the layer of cells that prevents the skin from losing fluid and dehydrating. The scientists also detected that the newborns had severely malformed hair follicles and a depletion of precursor immune cells in the thymus. Though the lack of Matriptase/MT-SP1 was not fatal during gestation, the scientists found that the skin abnormality caused such severe dehydration that the mice did not survive beyond two days.

Implications: This finding demonstrates that Matriptase/MT-SP1 is required for survival after birth. It also shows that the protein has a complex, multi-faceted effect during normal mammalian development. The scientists noted, because the expression of the protein is so high in the thymus and skin-producing cells, future work to delineate the specific molecular interactions of the protein should focus there, an important finding to accelerate research.

List K, Haudenschild CC, Szabo R, Chen WJ, Wahl SM, Swaim W, Engelhorn LH, Behrendt N, Bugge TH: Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. Oncogene 21: 3765-3779, 2002.

Important New Mouse Model for Ovarian Cancer

Background: Ever since Alfred Knudsen's deduction in the 1970s that cancer arises from a series of "hits," or gene inactivations, scientists have tried to define the possible patterns of hits that might create a tumor cell. While fairly detailed progression models have been developed for some tumors, including colon and oral squamous cell cancer, that has not been the case for ovarian cancer. Progress in the field has been limited due to the lack of a suitable model that would allow scientists to systematically study the genetic changes that causes of this often deadly disease.

Advance: NIH scientists and their collaborators recently developed a unique mouse model for ovarian cancer. It contains mutations in or overexpression of genes that previous studies suggest trigger tumor development in the ovary, such as p53, c-myc, K-ras, and AKT. The model is also particularly interesting because the tumors arise in adult mice, not newborns, which better mimics ovarian cancer in people. In their initial studies with the model, the scientists found that tumors arise in cells of the ovarian surface epithelium, a point of controversy in the field, and resemble the development and spread of human ovarian carcinoma.

Implications: With its unique combination of well-defined genetic alterations, the mouse model will allow scientists to better explore the basic biology of ovarian cancer. It should also be helpful in developing a progression model for the cancer that defines the specific genetic hits that are required for tumor initiation and development. It is also possible that the model will help scientists define the many potential molecular pathways that might be involved in causing ovarian cancers. This information will allow them to begin to better gauge the behavior of an individual tumor and its ability to grow, spread, or respond to specific types of cancer treatments.

Orsulic S, Li Y, Soslow RA, Vitale-Cross, Gutkind JS, Varmus HE: Induction of ovarian cancer by defined multiple genetic changes in a mouse model. Cancer Cell 1: 53-62, 2002.

Delivering Dual Growth Factors to Engineer New Tissues

Background: Tissue engineering is a relatively new field that will provide the scientific basis for future attempts to grow replacement tissues and organs. Scientists have published over the last few years several proof-of-principle findings, and work in the field continues to advance at a brisk pace. However, with early success comes new fundamental challenges. An example is the initial success in implanting a porous polymer scaffold within an injury site. As the tiny scaffold biodegrades, it releases a single growth factor, a natural protein that prompts the development of new tissue directly into the wound. Despite the promise of these initial results, the first-generation scaffolds have had limited therapeutic utility in subsequent studies. This maybe due to shortcomings in how the protein is delivered and/or it may be due to the fact that multiple growth factors are required to properly regenerate tissue, an indication that further refinements will be needed to enhance the scaffold.

Advance: A team of NIH grantees has taken an important next step in further developing polymer scaffolds. The group successfully implanted a biodegradable scaffold that provided sustained, localized release of two growth factors in animals. Importantly, the proteins were energetically distinct, meaning the device will not be limited to delivering kinetically compatible proteins, thus increasing its potential utility. In the study, the scientists evaluated the dual-release device in prompting the regeneration of blood vessels, a much-studied process in biology. The group found that, once implanted, the device reliably delivered vascular endothelial growth factor (VEGF), which stimulates the outgrowth of endothelial channels, followed by platelet-derived growth factor (PDGF)-BB, which stabilizes the nascent vessels by recruiting smooth muscle cells. Once delivered, the growth factors triggered the rapid formation of a mature vascular network.

Implications: This is the first report in the medical literature of a device that delivers multiple growth factors with distinct kinetics. It is also the first report of a dual-protein delivery device that can direct the formation of a new tissue. The researchers envision that the system may also have applicability to regenerating other tissue types, including bone, muscle, and nerve.

Richardson TP, Peters MC, Ennett AB, Mooney DJ: Polymeric system for dual growth factor delivery. Nat Biotechnol 19: 1029-1034, 2001.

Critical Technical Advance in Extending Bone Forming Capacity of Adult Stem Cells

Background: Scientists have dreamed of healing bone fractures by inserting directly into the wound adult bone marrow stromal stem cells, the progenitors that create the body's bone-forming cells. The adult stem cells would replicate in the wound, create millions of new bone cells, and heal the fracture rapidly and efficiently. As appealing as this approach is, however, technical challenges have emerged to slow the research. One of the most formidable is the discovery that adult bone marrow stromal stem cells stop growing soon after they are introduced into cell culture and quickly lose their ability to form new bone. Because hundreds of thousands of these stem cells are required to heal even a minor bone fracture, scientists have been hard pressed to generate an adequate supply of these precursors.

Advance: For the first time, NIH scientists and grantees reported they have more than doubled the life span of adult bone marrow stromal stem cells, under laboratory conditions. At the same time, they also enhanced the natural ability of these cells to form new bone and possibly cartilage. In followup animal studies, the scientists found the newly formed bone, generated from the stem cells, to have all of the hallmarks of normal bone, including organized collagen fibers and various mineral components. Specifically, the NIH scientists and grantees incorporated the hTERT gene into adult bone marrow stromal stem cells, an extremely challenging technical feat, to extend their life spans.

Two previous research findings have contributed to a critical advance in extending re-forming capacity of adult stem cells. The first is the popular theory in biology that each time a cell divides, the ends of its chromosomes – called telomeres – shorten by several base pairs, or units of DNA. After multiple cell divisions the shortening telomeres reach a critical length that signals the cells to stop growing or die. The second line of research is the relatively recent finding that scientists had succeeded in transplanting a gene called hTERT into a bone-forming cell, called an osteoblast, and extend its life span. This was particularly interesting because hTERT is the catalytic, or active, component of a much-studied enzyme called telomerase. Telomerase has been shown to counter telomere shortening by triggering a chemical reaction that adds base pairs to the end of telomeres.

Implications: This finding marks a critical technical advance for scientists who hope one day to use these adult stem cells to treat people with bone fractures, age-related bone loss, or other skeletal conditions. The advance also clears the way for scientists to extend these findings to studies of larger animals and, if all goes well, initial testing in people.

Shi S, Gronthos S, Chen S, Reddi A, Counter CM, Robey PG, Wang C-Y: Bone formation by human postnatal bone marrow stromal stem cells is enhanced by telomerase expression. Nat Biotechnol 20: 587-591, 2002.

MEDLINE and MEDLINEplus: A Continuing Story of Discovery

Medical practitioners and medical researchers both depend on the accumulated wisdom of those who have gone before. It was not so long ago that one could assemble this wisdom only by poring over printed bibliographies, usually the *Index Medicus*, which the NIH's National Library of Medicine (NLM) began publishing in 1879. Today, virtually all biomedical scientists, many health practitioners, and an increasing number of consumers use a variety of methods to search the MEDLINE database to learn about published research findings.

The pioneering MEDLINE project, begun in the early seventies, evolved from the computerized system used to produce the *Index Medicus*, which the NLM had installed in 1964. MEDLINE was the first successful marriage of a large reference database with a national telecommunications network. The eighties saw the introduction of "Grateful Med," a software program created by the NLM that one could load onto a PC and, equipped with a modem and a password, search MEDLINE right from one's home, office, or laboratory. Grateful Med was eagerly snapped up not only by librarians but also by health professionals, scientists, students, lawyers, medical journalists and others, who saw the average charge of $2 per MEDLINE search as a bargain.

This, of course, is the age of the Internet and the World Wide Web. The NIH Web site is the second most heavily trafficked site in the Federal Government, and the NLM databases account for the majority share of that use. MEDLINE searching via the "Internet Grateful Med" was introduced in 1996. The following year, free MEDLINE searching via the Web, began using a new system called PubMed. For the first time, anyone with access to the Web could search through an immense database of references and abstracts to now 12 million medical journal articles. The response was immediate and startling: from 7 million MEDLINE searches (1996) the usage climbed dramatically and now stands at more than 400 million searches annually.

Improvements to PubMed continue to be introduced, and today it offers a high degree of flexibility to users. For example, there are now Web links to about the two-thirds of the journals represented in MEDLINE, allowing users to have access to the full text of articles referenced in the database. A user's view of PubMed can be customized to highlight electronic and print journals available through the local hospital or academic health sciences library with which he or she is affiliated. Where such links are not available, the user may use PubMed to place an online order for an article directly from a library in the National Network of Libraries of Medicine. A popular and expanding feature of PubMed is the ability to limit a MEDLINE search to articles on selected broad topics, including complementary and alternative medicine, AIDS, and bioethics.

The introduction of PubMed simplified MEDLINE searching to the point where the general public could search for relevant references on any biomedical subject from the literature. Since, as it turned out, about 30 percent of all MEDLINE searches were being done by consumers, this presented the NLM with a wonderful opportunity. Why not create a service that not only will provide selective MEDLINE results that are useful to the consumer, but also link the Web user to

authoritative, full-text health information. Such a service, called MEDLINE*plus*, was introduced in October 1998.

Where did such "authoritative" information come from? Since the NIH publishes a wealth of consumer health information based on the medical research it sponsors, it was natural to start there. Also "recruited" were other HHS agencies, professional societies, and voluntary health agencies, many of which issue authoritative information that the public can trust. With help from members of the National Network of Libraries of Medicine across the country, NLM information specialists select and organized this information and extensively cross-link it. In the four years since its introduction, MEDLINEplus has grown tremendously, both in terms of its coverage of health and its usage by the public. As of August 2002 it was being consulted some 11 million times each month by 1.3 million unique visitors. The original two dozen "health topics," containing detailed consumer information on various diseases and health conditions, have been increased to 560. Other information available through MEDLINEplus: medical dictionaries and an encyclopedia, detailed information about prescription drugs, interactive patient tutorials, health-related news items from the daily print media, directories of health professionals and hospitals, links to organizations and libraries that provide health information for the public, and preformulated searches of MEDLINE on various aspects of the health topics to find scientific articles. MEDLINEplus also contains preformulated searches of the NIH ClinicalTrials.gov database for each health topic. Developed by NLM for NIH, ClinicalTrials.gov is a registry of some 7,000 trials for both federally and privately funded trials of experimental treatments for serious or life-threatening diseases, designed to assist the public in locating trials that may be appropriate for themselves, family members, or friends.

In September 2002, in response to many user requests, NLM launched a Spanish version of MEDLINEplus, including a comprehensive encyclopedia with images, nearly 500 health topics, and 150 interactive slideshows with animation and sound. Also new this year is the NIHSeniorHealth website, developed in collaboration with the National Institute on Aging, which applies research on how seniors learn to the organization and presentation of electronic information. NLM has also worked with the University of North Carolina to develop a model for integrating access to the authoritative information in MEDLINEplus with access to local information about neighborhood clinics, nurses, doctors, counseling programs, support groups and classes of interest to patients, their families, and the general public. NCHealth is the first implementation of this model.

The task now is to ensure that these health information services are known and used even more widely. To that end, NLM is working through the eight Regional Medical Libraries, the thousands of member institutions of the National Network of Libraries of Medicine, even larger numbers of public library systems, and an increasing array of community and faith-based organizations to reach into communities around the Nation. There are special outreach programs directed to African Americans, Latinos, Native Americans, and rural populations. The taxpayer's investment in medical research will be fully repaid only when the most authoritative and understandable health information is readily accessible to every American.

First Analysis of the Mouse Genome Sequence

A concerted effort to sequence the entire mouse genome began in October 2000. Three private companies, six NIH institutes, and the Wellcome Trust supported the initial phase of the project, which achieved its first goal, three-fold coverage of the genome, in May 2001. Members of the Mouse Genome Sequencing Consortium (MGSC), with support from the NHGRI and the Wellcome Trust, have continued working towards achieving a fully finished sequence by 2005.

In April 2002, the MGSC completed the mouse draft sequence and made available the first public assembly of the entire mouse genome. This marked the end of the whole genome shotgun phase of mouse sequencing, which generated enough sequence to represent the genome seven times over. The MGSC will soon publish a paper reporting production of a high quality, highly ordered, representative assembly of the genome and an initial analysis of its content.

The sequence data and assemblies have been freely available throughout the course of the project. The next step of the project, already underway, is to convert the draft sequence into a finished sequence to provide a permanent foundation for biomedical research in the 21st century.

Previous comparisons of mice and humans showed that they share gene sequences to a high degree and share virtually the same overall set of genes. The availability of the assembly of the mouse genome allows for a closer analysis of these similarities, and of differences, on a genome-wide scale. Already, this comparison has produced a number of interesting observations. Perhaps the most striking observation is that fully two-thirds of the fraction of the human genome that is most similar to the mouse, implying that it serves some critical function, falls in parts of the genome not previously known to be important.

The increase in the amount of genomic sequence data for many organisms, including multiple mammals, has spawned a new generation of gene finding computer programs that use genome comparisons. These programs were tested using the mouse and human data and appear promising, making approximately 1,500 gene predictions additional to those from previous methods. Experimental validation suggests that these programs are more than 85% accurate.

Some striking examples of gene families (sets of genes whose members are structurally and functionally related) illuminate differences between the two species. In several instances, the number of genes in a gene family in the mouse is greater than in the human. It has been shown that the mouse genome has many more genes for olfactory receptors, genes that affect taste and smell, than the human genome. The new analysis shows that several different gene families controlling various aspects of reproduction in mouse have also expanded, including genes controlling pregnancy and selection of mates.

An annotated, highly continuous mouse sequence provides a powerful tool for finding genes involved in simple and complex diseases and for identifying other sequence elements that

regulate those genes. One can move directly from genetic mapping to identification of candidate genes, using only a few deft clicks of a computer mouse applied to the public databases.

The availability of an advanced draft of the mouse genome sequence will accelerate progress in biomedical research by improving the utility of the mouse as a model research system to study and understand human disease, and to develop and test new treatments in ways not possible with humans.

The Creation of a Haplotype Map

The NIH and researchers in the United States, Canada, China, Japan, and the United Kingdom are preparing to develop a "haplotype map" (HapMap) of the human genome by identifying how variation in the genome is organized along all of our chromosomes. The HapMap should be a key resource for finding genes affecting health, disease, and response to drugs and environmental factors.

Biomedical researchers have previously developed highly successful positional cloning methods to find the genetic basis of rare diseases that single genes strongly affect. However, multiple genetic and environmental factors influence many common diseases, such as diabetes, cancer, stroke, psychiatric disorders, heart disease, arthritis, and asthma. Linkage strategies that work well for single-gene "Mendelian" disorders lack power to map such multi-gene disorders, and, far too often, such studies have yielded only weak linkages that follow-up studies fail to confirm. Thus, relatively little is known about the genetic basis of these common diseases, or of the factors that determine individual risk of disease, clinical course, or response to treatment. Discovering the particular DNA sequence variants that contribute to common disease risk and response offers a unique opportunity for illuminating pathways of disease causation in humans.

Sites in the genome where individuals differ in their DNA spelling by a single letter are called single nucleotide polymorphisms (SNPs). Recent work has shown that about 10 million SNPs are common in human populations. SNPs are not inherited independently; rather, sets of adjacent SNPs are inherited in blocks. The specific pattern of particular SNP spellings in a block is called a haplotype. Recent studies have also shown that although a region of DNA may contain many SNPs, it takes only a few SNPs to uniquely identify or "tag" each of the haplotypes in the region. This presents the possibility of a major shortcut in identifying hereditary factors in disease – instead of testing 10 million SNPs, a carefully chosen subset of about 400,000 SNPs would essentially provide all of the critical information. This should make approaches to finding regions with genes that affect diseases much more efficient and comprehensive, since effort will not be wasted typing more SNPs than necessary and all regions of the genome can be covered.

Most common haplotypes occur in all human populations, although their frequencies may vary considerably. Initial studies also indicate that the boundaries between the blocks are remarkably similar among populations in Europe, Asia, and Africa. These data provide strong support for the idea that a human haplotype map built with samples from these three geographic areas would apply to most populations in the world, although further testing of this conclusion is needed.

This new initiative will support the development of the HapMap, a catalog of the haplotype blocks and the SNPs that tag them. In addition to its use in studying genetic associations with disease, the HapMap will be a powerful resource for studying the genetic factors contributing to variation in response to environmental factors, in susceptibility to infection, in host immune responses, and in the efficacy of, and adverse responses to, drugs and vaccines.

One should note that a set of haplotype tag SNPs will generally not include the actual causative variant that influences disease risk or drug response. But, the responsible variant will travel on a particular haplotype, whose overrepresentation in individuals with the particular phenotype will point to that small segment of the genome.

The international consortium of countries committed to conduct and fund the HapMap Project, will work in partnership with the SNP Consortium, consisting of the Wellcome Trust and several pharmaceutical companies. A steering committee will include members from the funded groups, the funding agencies, and appropriate others. The NIH began making grant awards for the HapMap project on September 30, 2002.

The NIH plans to require rapid data release of the SNPs, genotypes, quality scores, haplotypes, haplotype blocks, and associated information that the HapMap project generates. A protocol for the quickest release of data that allows adequate checking of their quality will be developed.

An interdisciplinary advisory group has been formed to provide advice on the scientific and ethical issues the HapMap Project raises, especially with regard to sampling populations. Ultimately, the development of this powerful tool will allow the biomedical research community to understand complex genetic diseases much more fully and will lead to improved treatments and, ultimately, cures for many of these disorders.

Early Cochlear Implantation Yields Best Results

Since the beginning of the 19th century, scientists have attempted to stimulate hearing by delivering electrical currents to the ears of deaf individuals. The first published report of an experiment to excite the auditory nerve by directly applying an electrode was documented by two French scientists in 1957. This research eventually led the way to the cochlear implant, the only sensory neural prosthesis in widespread clinical use today. This device converts sound into electrical impulses on an array of electrodes surgically inserted into the inner ear, bypassing the damaged hair cells that normally converts sound into neural impulses, stimulating the auditory nerve directly and restoring the perception of sound to individuals who are deaf. Many adults who lose their hearing suddenly can use the telephone again after receiving a cochlear implant.

The NIH has played a significant and important role in sponsoring the development of cochlear implant technology. Two NIH Institutes have been particularly important in this development: the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Deafness and Other Communication Disorders (NIDCD). Cochlear implant research in the U.S. significantly increased after the NIH Neural Prosthesis Program (NPP) was established in 1970 and again when the NIDCD was established in 1988. Under the leadership of the NIH over the past three decades, support and funding of research grants to universities have built a community of scientists and led to the establishment of cochlear implant programs in the U.S. In addition, NIH support provided a forum for scientists to share information through NIH-sponsored conferences and workshops. The support of research on cochlear implants provided critical information that allowed for the acceptance of auditory prostheses. This eventually resulted in the issuance of regulations by the Food and Drug Administration that ensured pre-market approval for safety and effectiveness of a cochlear implant in 1984. To date, over 70,000 individuals have received cochlear implants worldwide, including 21,000 Americans, with about one-half of the recipients being children.

One of the anticipated benefits of implantation in children is the improved acquisition of spoken language. Recent data indicate that the improvements in speech perception and language production in deaf children with cochlear implants have also resulted in better language and reading performance. With many states screening newborn infants for hearing impairment before discharge from the hospital, many more infants with hearing impairment will be identified at an early age when appropriate intervention, such as a cochlear implant, can be started that will optimize their long-term speech and language skills.

When is the best time to implant a deaf child so that he or she will acquire a normal level of speech and language skills? The prevailing wisdom is before the effects of deafness substantially alter the brain's ability to adapt to the stimulus of sound introduced through a cochlear implant. However, the time line for adaptiveness (plasticity) in the central auditory nervous system of humans is not known. NIDCD-supported scientists are measuring electrical activity of the brain in response to sound in children who received an implant in early childhood and in late childhood and have found that the brain can survive up to 3.5 years of sound

deprivation and still remain plastic, that is, can change quickly once sound is introduced through a cochlear implant. Immediately after young children are fitted with a cochlear implant, the brain's responses to sound resemble those of normal hearing newborns. Once sound stimulation is started with a cochlear implant, young children's brains develop at a very rapid rate. As a result of this rapid development or plasticity, young implanted children usually show age-appropriate brain responses within six to nine months after their cochlear implant is turned on. These results also show that the older the child gets before getting a cochlear implant, the less plastic the brain responds to sound. Children who are implanted between the ages of four- to seven-years old show varying amounts of plasticity. In children who are implanted after eight-years old, the scientists found an extremely reduced level of plasticity in response to sound stimulation. Overall, these results suggest that the best time to implant a deaf child is by three-and-a-half years old. These findings are consistent with recent research demonstrating that implanting a child early maximizes the probability of a successful outcome in the child's speech and language development. The time line for brain plasticity that is being defined should help parents and clinicians make informed decisions about the best time to surgically insert a cochlear implant in a deaf child.

In addition, two new exciting advancements in cochlear implants are currently undergoing investigation by NIH-supported scientists: binaural implants and a short electrode implant allowing for the combination of electric and acoustic hearing. Binaural (both ears) implants provide a potential avenue for better speech perception in noise, while the short electrode is being designed to be used in experienced, yet unsuccessful, adult hearing aid users with severe-to-profound hearing impairment. Scientists are currently collecting data and the preliminary results are promising. As technology continues to provide advances in cochlear implant design, additional people will have the potential to benefit from these remarkable devices.

With over 30 years of NIH research investment, the cochlear implant has evolved from an experimental device to a miniature hearing prosthesis that is commercially available to assist those who are profoundly deaf or severely hearing impaired. NIH-supported research has been instrumental in showing the appropriateness of cochlear implants in infants with profound hearing impairment, the age when children could safely be implanted and the importance of early intervention for language development. Without such research, cochlear implants might have remained experimental devices available only to a few, and not become commercially viable.