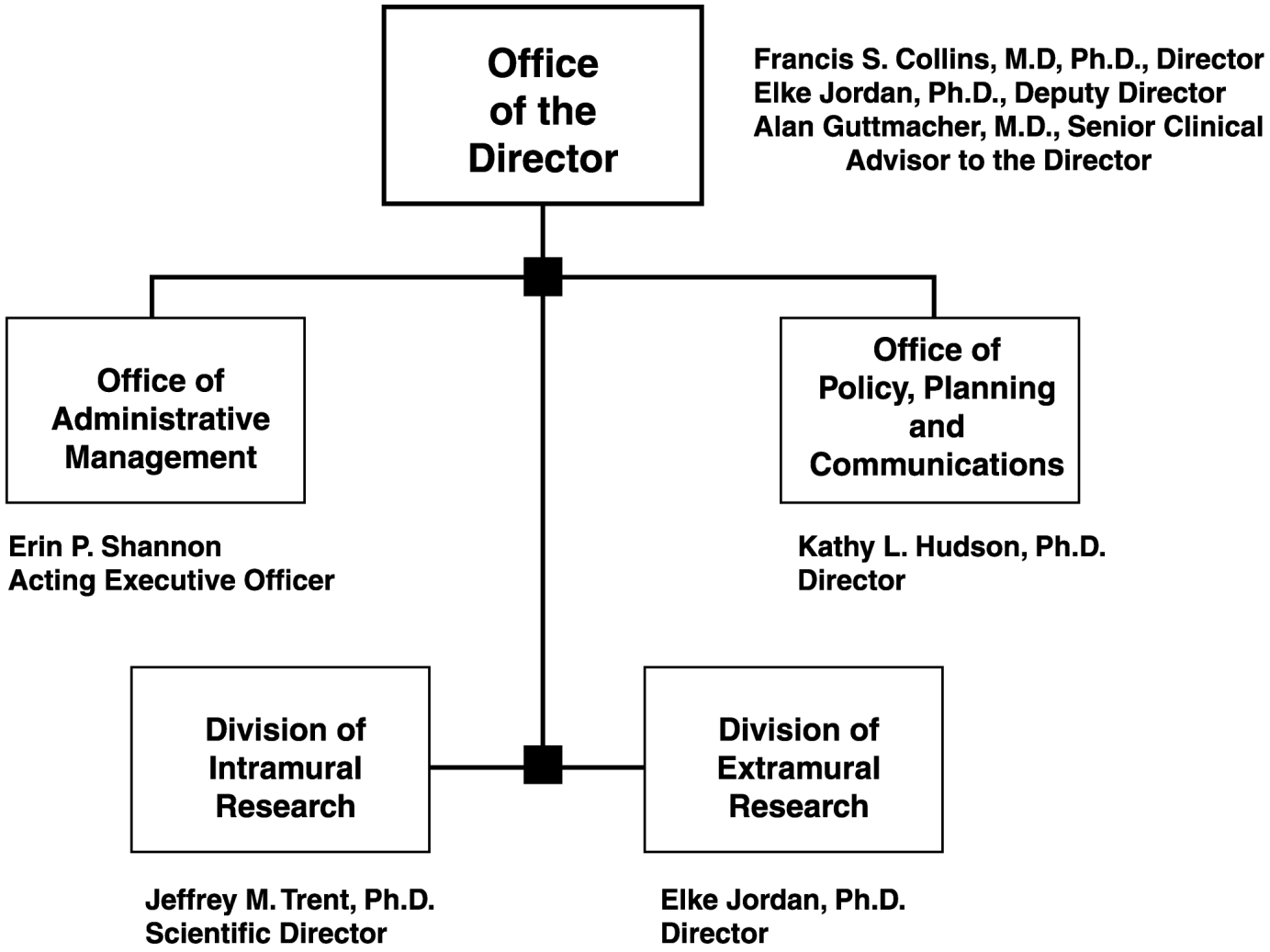
DEPARTMENT OF HEALTH AND HUMAN SERVICES

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute

NATIONAL HUMAN GENOME RESEARCH INSTITUTE
Organizational Structure



NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute

For carrying out section 301 and title IV of the Public Health Service Act with respect to human genome research, [$429,515,000] *$457,032,000*.

[Departments of Labor, Health and Human Services, Education, and Related Agencies Appropriations Act, 2002. (P.L. 107-116)]

National Institutes of Health

National Human Genome Research Institute
Amounts Available for Obligation 1/

Source of Funding	FY 2001 Actual	FY 2002 Estimate	FY 2003 Estimate
Appropriation	$382,384,000	$429,515,000	$455,474,000
Enacted Rescission	(192,000)	(203,000)	---
Subtotal, Adjusted Appropriation	382,192,000	429,312,000	455,474,000
Comparable adjustment for legislative proposal for accrued retirement costs	1,299,000	1,406,000	1,558,000
Real transfer to:			
Other HHS Agencies through Secretary's one-percent transfer authority	(72,000)	---	---
Real transfer to HHS for the Office of Human Research Protection	(80,000)	---	---
Comparative transfer from:			
National Cancer Institute for research activities	---	---	9,663,000
Comparative transfer to:			
The National Institute of Biomedical Imaging and Bioengineering (NIBIB) for grants transferred.	(939,000)	---	---
Subtotal, adjusted budget authority	382,400,000	430,718,000	466,695,000
Unobligated balance, lapsing	(69,000)	---	---
Total obligations	382,331,000	430,718,000	466,695,000

1/ Excludes the following amounts for reimbursable activities carried out by this account:
FY 2001 - $42,824,000; FY 2002 - $42,824,000; FY 2003 - $42,878,000
Excludes $44,448 in FY 2001 and $125,987 in FY 2002 for royalties.

Justification

National Human Genome Research Institute

Authorizing Legislation: Sections 301, 485B, and 487(d) of the Public Health Service Act, as amended. Reauthorizing legislation will be submitted.

Budget Authority:

	2001 Actual	2002 Appropriation	2002 Current Estimate	2003 Estimate	Increase or Decrease
Current Law BA	$381,101,000	$429,515,000	$429,312,000	$465,137,000	$35,825,000
Accrued Costs	1,299,000	1,406,000	1,406,000	1,558,000	+152,000
Proposed Law BA	$382,400,000	$430,931,000	$430,718,000	$466,695,000	$35,977,000
FTE	275	275	275	278	+3

This document provides justification for the Fiscal Year 2003 activities of the National Human Genome Research Institute, including HIV/AIDS activities. A more detailed description of NIH-wide Fiscal Year 2003 HIV/AIDS activities can be found in the NIH section entitled "Office of AIDS Research (OAR).

The President's appropriations request of $466,695,000 for this account includes current law adjusted by assuming Congressional action on the proposed Managerial Flexibility Act of 2001.

INTRODUCTION

The National Human Genome Research Institute (NHGRI) of the National Institutes of Health (NIH) leads the international Human Genome Project and conducts research to understand and treat genetic disorders.

During fiscal year 2003, the field of genetics will observe a major anniversary, and NHGRI will reach a landmark accomplishment with the Human Genome Project. Fifty years ago, in the spring of 1953, Drs. James D. Watson and Francis Crick reported the discovery of the double helix structure of DNA, a landmark paper in the annals of scientific research. Dr. Watson would later serve as the first director of what would become NHGRI. Also in 2003, the Human Genome Project expects to complete the final sequence of the human genome. NHGRI announced the first draft of the human genome sequence at the White House in June 2000, and said then that it would take an additional three years to remove all the spelling errors and fill in all the gaps. The Human Genome Project is on target to meet that deadline and expects to finish the analysis in time for the 50th anniversary of the Watson-Crick paper.

For the last decade, NHGRI has been the driving force behind the Human Genome Project–a massive undertaking that has revolutionized the way biomedical researchers study human illnesses. The project, begun in October 1990, seeks to identify the precise order of the letters used to write the genetic instructions for making a healthy human. In the United States, the work has been funded by NHGRI and the Department of Energy. The project's international partners include the United Kingdom, France, Germany, Japan, and China.

While sequencing the human genome is NHGRI's most visible goal, the Institute also conducts important genetic and genomic research in a variety of areas including working to understand the way individuals differ from each other at the genetic level and the impact these variations may have on health. In addition, the Institute leads in the development of new technologies, such as microarrays or DNA chips, and creates novel research strategies to study the function of genes and genomes. Because of its unique organization, the Human Genome Project has moved forward in a series of 5-year plans. NHGRI has now embarked on a new planning process that seeks to utilize the knowledge gathered through the Human Genome Project to rapidly expand the scientific understanding of genomics and to improve the health of all people.

Sequencing the Human Genome: Our Genetic Instruction Book

The Human Genome Project has, from its beginning, enjoyed remarkable success. In addition to creating the first draft of the human sequence, many of the project's initial goals have already been achieved, including building maps to localize genes in both the human and mouse genomes, and sequencing the genomes of model organisms including the bacterium *E. coli*, baker's yeast, the fruit fly, and the roundworm *C. elegans*. The work with these model organisms has helped interpret the human sequence and will provide important research tools for understanding, and ultimately treating, human illnesses.

The human sequence has always been the ultimate goal. Between March 1999 and June 2000, production of human genome sequence skyrocketed. During this time, scientists sequenced 1,000 DNA letters a second--24 hours a day, 7 days a week. The resulting working draft sequence covered over 94 percent of the human genome with 33 percent in highly accurate finished form by February 2001. By January 2002 the amount in highly accurate finished sequence was 65 percent. The average accuracy of all of the DNA sequence in this assembly is 99.9 percent.

The International Human Genome Sequencing Consortium published the sequence and initial analysis of the human genome on February 15, 2001. The sequence and analysis was a remarkable achievement, representing the work of thousands of scientists laboring in 20 associated genome centers around the globe.

The Book of Life, as some have termed the human genome, is actually three books: It's a history book that tells the narrative of the human species' journey through time. It's a shop manual that provides the parts list and an incredibly detailed blueprint for building every human cell. And finally, it's a transformative textbook of medicine that provides insights giving health-care providers immense new power to treat, prevent, and cure disease.

In the months following the completion of the working draft, Human Genome Project scientists–including many of the most brilliant genetics and computational experts in the world–began to scour the sequence to find answers to age-old questions in biology and, in the process, revealed new mysteries that will occupy researchers for years to come.

Here are some of the treasure trove of discoveries from this initial analysis:

1) Humans have only 30,000-35,000 genes, far fewer genes than the 80,000 -150,000 that had been predicted.
2) Individual human genes are able to produce multiple different proteins.
3) The architecture of human proteins is much more complex than those of simpler organisms.
4) The mutation rate in males is twice that found in females.
5) Much of the "junk" in the genome appears to have important functions.

The February 2001 publication, including these and other surprising facts, represented the end of the beginning of genomics as they described the first look at the entire human genome sequence. The human genome contains a vast amount of information that the scientific community can only now begin to understand. The effort to fully translate the mountains of raw genetic data into advances that will improve human health will continue to challenge the brightest minds for decades to come.

Because of the enormous value of sequence information to researchers around the world, Human Genome Project scientists have placed all DNA sequence data in public databases where it is immediately and freely available with no restrictions on its use or redistribution. The information is scanned daily by tens of thousands of scientists in academia and industry, as well as by commercial database companies providing information services to biotechnologists. More information about the sequencing and analysis of the human genome is available at: http://www.nhgri.nih.gov/genome_sequence.html.

NHGRI Embarks On A New Research Plan

The Human Genome Project has, since its inception, been guided by a series of overlapping 5-year plans. These plans have laid out ambitious goals to advance our understanding of the human genome and the associated ethical, legal and social implications. The plans have been instrumental to the success of the Project by clearly enumerating our program objectives to the scientific community and the public, and by providing measurable objectives to guide our work and gauge our progress and success.

The major goal of the 1998-2003 plan, the completion of the working draft sequence of the human genome, has been achieved and many other goals in the plan have been completed or are on a clear path towards completion. Therefore, the time has come to take stock of where we are, to critically evaluate the challenges and opportunities that lie before us and to create a bold new vision for the future of genomics. NHGRI has already begun laying the foundation for a new 5-year plan and, over the coming year and a half, we will engage the scientific community to help us develop an ambitious set of new goals for the NHGRI.

In December 2001, the NHGRI convened about 200 outside experts, including scientists, researchers in the ethical, legal, and social implications (ELSI) of the Human Genome Project, consumers, and policy experts to think very broadly and creatively about the future of genomics. Over the course of the following months, we will host several workshops to explore specific topics in detail and enumerate specific goals appropriate for NHGRI.

The Institute had already begun this planning process through the convening of several workshops. In July 2001, NHGRI hosted a workshop to develop guidelines for prioritizing new organisms for large-scale genomic sequencing. That same month, the Institute also held a meeting on developing a new kind of genetic map called a "haplotype map" of the human genome, which may have profound consequences for identifying genetic contributors to common disease. Additional workshops planned for the next several months include those on large-scale protein analysis, regulation of gene expression, and the ethical, legal, social, and policy implications of genomics. In the fall of 2002, we will again convene outside experts to review the year's findings and provide input on the draft research plan. We anticipate publishing the new NHGRI plan in April 2003 coinciding with two historic events in genomics: the completion of the human genome DNA sequence and the 50th anniversary of the discovery of the double helical structure of the DNA molecule.

STORIES OF DISCOVERY

Mouse and Rat Genome Sequences: Tools for Understanding the Human Genome

The initial analysis of the human genome sequence revealed a wealth of information about the contents of the human genetic instruction book, but much more remains encrypted in the 3.1 billion letter code. A major tool for interpreting and understanding the human genome sequence is to compare it to the genome sequences of other organisms. This approach is based on the fact that functionally important regions of DNA are conserved over long periods of evolutionary time. By comparing the human genome sequence with those of rat and mouse, similar regions are readily apparent. Many of these identify protein and RNA coding regions and regulatory sequences that control the timing, level and location of gene expression, control chromosome structure, or specify other biological properties. Thus, identifying the regions of similarity in the sequences of different organisms is now a primary method of finding candidates for functionally important DNA sequences. In addition to their utility in interpreting the human genome, the genome sequences of rat and mouse will make it easier to use results from experiments with those model organisms to make inferences about human medicine and biology.

NHGRI has established two model organism research networks; the mouse research network was funded with contributions from three private companies and several other NIH Institutes, and the rat research network is an NHGRI/NHLBI collaboration. Both efforts are now fully operational and have made rapid progress. The mouse project has already produced a "shotgun coverage" of 95 percent of the genome and will generate a draft quality sequence early in 2002, a more detailed version of the sequence by 2003, and a finished sequence by 2005. The rat project currently has a more limited objective: to produce a draft quality assembled sequence by 2003. Both efforts are on target to meet these objectives. As with the human sequence, the data from both the mouse and rat projects are rapidly released to the public databases and the World Wide

Web with no restrictions. As they are produced, both data sets are being used extensively by the scientific community.

The mouse data is already finding multiple uses in research. For example, Merck & Co., Inc., of Whitehouse Station, NJ, has used the newly available data to find the mouse equivalent of a human gene that may be related to schizophrenia. Previous work by the company had identified a human gene, located at a chromosomal break point, but Merck scientists had been unable to find the mouse equivalent. As the mouse sequence became available, Merck researchers found a match that helped them locate the mouse gene. In turn, the discovery will help the company develop a mouse model to test new treatments for this mental disorder.

Mammalian Gene Collection

The initial sequence and analysis of the human genome sequence suggested that there are about 30,000-35,000 human genes. However, identifying the genes is only the first step in the pathway of using genomics to improve human health. The next step is to understand the function of each of the genes and their products. Many approaches are being pursued, which collectively have come to be known as "functional genomics."

Genes do their work by being transcribed into messenger RNA (mRNA), which in turn is translated into protein. One of the most important challenges in functional genomics is to understand how, where, and when each of the ~30,000 genes is active. To do this, scientists need a collection of mRNAs representing all human genes. However, because mRNAs are unstable, scientists use complementary DNAs (or cDNA), which are simply DNA copies of the mRNAs made in a test tube. Full-length cDNAs can be easily replicated and used in experiments to detect their cognate mRNA and monitor patterns of gene activity during normal development, in disease states, and in response to drugs. Full-length cDNAs can also be used to confirm the identity of genes predicted from the sequence and identify multiple mRNAs produced from a given gene. Even more importantly, full-length cDNAs can be used to synthesize pure protein products in the laboratory. These kinds of experiments will provide insight into how genes work together, which gene activities could be used as diagnostic markers, and which genes are appropriate targets for developing drugs.

The NIH Mammalian Gene Collection (MGC), managed by NHGRI and the National Cancer Institute with contributions from many other NIH institutes, is a project to isolate and sequence a representative full-length cDNA for each human and mouse gene. Initial collections of unsorted cDNAs, called libraries, undergo stringent quality control measures to identify those that are the richest sources of full length cDNAs, and only those that meet rigorous standards are used by the MGC in its gene discovery efforts. Because most of the cDNAs in even a good library are not of high enough quality to be part of the MGC, much effort has gone into making the selection process more efficient.

As of January 2002, over 13,000 full length cDNAs from human and mouse were identified and fully sequenced. All of the MGC products, including the full gene sequences and the clones themselves are available to the entire research community without restriction. The output of the MGC is already being used by a large of number scientists in studies such as those mentioned above. More information can be found on the project's website: http://mgc.nci.nih.gov.

SCIENCE ADVANCES

Advances in the Search for Prostate Cancer Genes

Prostate cancer is one of the most common cancers in American men, with over 175,000 new cases diagnosed in the United States each year.[1] The most significant of all known risk factors for prostate cancer is having two or more close relatives–a father or a brother–with the disease. Men whose fathers or brothers have prostate cancer are five times more likely to develop the disease themselves, as compared to men with no family history of prostate cancer. This observation strongly suggests that susceptibility to prostate cancer involves inherited factors.

In recent years, researchers have found several genes that may play a role in the development of hereditary prostate cancer. NHGRI scientists and their colleagues mapped the first such gene, called HPC-1 ("hereditary prostate cancer 1") to human chromosome 1 in 1996. Recently, NHGRI researchers and collaborators have made a number of significant advances in the search for prostate cancer susceptibility genes. Mutations have been found in a specific gene that may represent HPC-1. Regions of chromosomes 8 and 20 have also been implicated. A study of affected Finnish families has identified a region responsible for a distinct sub-group of late-onset prostate cancer on the X chromosome (HPC-X). Understanding the molecular basis of cancer is crucial to the development of a new generation of targeted treatments and preventive strategies. New genomic resources are accelerating these advances and hastening progress in cancer research.

New Tumor Suppressor Gene Involved in Breast, Prostate, and Other Cancers

One of the characteristics of cancer cells is that they have defects in the regulatory circuits that control normal growth. Scientists have made great strides in recent years in determining the key proteins within cells that control and comprise these regulatory circuits.

Scientists at NHGRI and the M. D. Anderson Cancer Center at the University of Texas found a novel tumor suppressor gene on human chromosome 7q31 that appears to be involved in a wide range of cancers. The gene, ST7, is widely expressed in normal tissues throughout the body and is often disrupted by mutation or deletion in tumors arising from epithelial cells, such as cancers of the breast, prostate, colon and ovary.

The discovery of ST7 demonstrates a new paradigm in molecular genetics research. Now that the working draft sequence of the human genome is available in public databases, a variety of computational tools can identify the sequence of the gene immediately after the gene's chromosomal location is determined. The ST7 gene was discovered by a single post-doctoral scientist using the new tools produced by the Human Genome Project. Such a project would previously have taken many man-years of work.

---

[1] Zheng SL, Xu J, Isaacs SD, Wiley K, Chiang B, Bleecker ER, Walsh PC, Trent JM, Meyers DA, Isaacs WB. Evidence for a prostate cancer linkage to chromosome 20 in 159 hereditary prostate cancer families. Hum Genet, 2001; 430-435.

Distinguishing Hereditary from Non-Hereditary Tumors

Differentiating between hereditary and non-hereditary breast cancer tumor types is not easy using traditional techniques. When these tumors are viewed under a microscope, it is very difficult to tell which tumors are caused by BRCA1, BRCA2, and/or non-inherited mutations. A newly developed DNA microarray technology offered the opportunity to study the expression of many genes within a tumor cell's metabolic pathways.

NHGRI scientists used DNA microarray technology to simultaneously assess the activity of 6,000 genes within breast cancer cells and generate a unique gene-expression profile of breast tumors. The research team examined samples of tumors from 22 breast cancer patients, some with family history and some with no family history of the disease. Using the microarry technology, the team was able to quickly and accurately differentiate the tumors arising in individuals with specific heritable mutations from the sporadic cases. The clear differences in the patterns of gene activity in breast tumors are as unique as a fingerprint, pinpointing into which group a woman's cancer belongs. These fingerprints also revealed key genes involved in tumor development and progression.

This new approach should make it possible for physicians to quickly and accurately diagnose the cause of an individual woman's disease. Ultimately, these DNA microarrays will allow physicians to know with much greater certainty the type and stage of the patient's tumor, making the selection of an individualized treatment regimen possible. This approach is not limited solely to cancer; the same techniques can be applied to any disease having a genetic or genomic underpinning.

Gene Chips Accurately Diagnose Four Complex Childhood Cancers

Scientists at the NHGRI and Lund University in Sweden developed a method of genetic fingerprinting that may help accurately discern between the four closely related and difficult to differentiate types of childhood cancer, neuroblastoma, rhabdomyosarcoma, non-Hodgkin lymphoma, and the Ewing Sarcoma. The method combines cutting-edge technology involving DNA microarrays (discussed above) with a form of artificial intelligence called an artificial neural network. The neural network, machine-learning approach automatically analyzes the enormous amounts of data produced by the microarray to make a highly accurate diagnosis. The research team surveyed over 6,000 genes and narrowed down the number of genes that differentiate these tumor types from one another to 93. Of these, 41 were previously uncharacterized and might provide new, important insights into the biology underlying these cancers. While it may take some time for this approach to move from the lab into the clinic, this newly-developed approach could be invaluable during diagnosis, particularly since the treatment of each of these cancer types is significantly different.

Using Adult Bone Marrow Stem Cells in Mice to Repair Damage From Heart Attacks

Scientists at the NIH and the New York Medical College, Valhalla, NY, have demonstrated for the first time that adult stem cells isolated from mouse bone marrow can become functioning

heart muscle cells when injected into a damaged mouse heart. More important for future clinical application in humans, the new cells at least partially restore the heart's ability to pump blood. Bone marrow stem cells ordinarily produce red and white blood cells, but the apparent ability of stem cells in the bone marrow of adult animals to rebuild the heart reveals the remarkable flexibility of adult stem cells.

In their search for a way to reverse heart muscle damage, the team began by isolating bone marrow stem cells from male mice. The isolated stem cells carried a newly inserted gene that produces green fluorescent protein, a marker that enabled the researchers to identify the transplanted cells. The researchers blocked blood flow through the coronary artery in mice to mimic a heart attack and then injected the labeled stem cells into the heart muscle next to the damaged tissue. Over the next several days, the stem cells began to multiply, transform themselves into heart muscle cells and migrate into the damaged area. After an average of nine days, the newly formed heart muscle cells occupied 68 percent of the damaged portion of the heart. In addition, the stem cells also began producing smooth muscle cells and endothelial cells that organized themselves into new blood vessels. These results reveal the great potential of adult bone marrow stem cells to differentiate into other cell types and repair a damaged heart.

Mouse Model Provides Insights about the Inner-Ear Defects in Pendred Syndrome

Pendred syndrome is a genetic disorder associated with deafness and goiter. Hearing loss is generally profound and occurs before age three, although occasionally it is later in onset and progressive. For over 100 years, the molecular basis of Pendred syndrome remained largely unknown until 1997, when NHGRI scientists identified the defective gene, PDS, on chromosome 7 (*PDS*).

The PDS gene was discovered to encode a protein named pendrin. To further investigate the functional role of pendrin and to provide a system for more detailed study of the inner-ear defects that occur in the absence of pendrin, NHGRI scientists developed a mouse model. Experiments using this mouse model showed the disruption of the PDS gene prevented any of the pendrin protein from being made. While the mice appear normal at birth, they develop early-onset, profound deafness. The inner-ear abnormalities seen in the mice generally mirror the defects seen in individuals with PDS mutation, and thus provide a valuable model for studying hearing loss associated with PDS mutations and a valuable experimental tool for investigating possible therapeutic options.

A Mouse Model for a Cancer Syndrome that Results in Multiple Endocrine Tumors

Multiple endocrine neoplasia type 1 (MEN1) is an autosomal dominant cancer syndrome characterized by multiple tumors of the parathyroid, pancreas and pituitary. It is caused by mutations in the MEN1 gene, which was identified by NHGRI, NIDDK, and NCI scientists in 1997. Menin, the protein encoded by MEN1, appears to function as a classic tumor suppressor that plays a role in preventing unregulated cell division. Using "knockout" technology, a mouse model of MEN1 was generated to examine the cellular transitions that lead to pancreatic tumors when menin is absent. The tumor incidence and distribution pattern in the mice correlate closely

with the human MEN1 phenotype. The availability of a mouse model for MEN1 will be an asset for testing possible therapeutic approaches for this inherited disorder and its sporadic counterparts.

Bacterial Artificial Chromosome Resources

Over the past several years, a particular cloning system, known as the bacterial artificial chromosome (BAC), has emerged as the vector system of choice for the construction of the large-insert chromosomal DNA libraries that are needed in genomic studies. BAC clones have been used for the complete sequencing of all large eukaryotic genomes to date and are also the vehicle of choice for isolating targeted regions of genomic DNA from humans and other species.

BAC libraries representing the genomes of a small number of organisms are publicly available. However, the current national capacity for making BAC libraries is too low to meet future demand, and BAC library construction could become a bottleneck in the application of genomic approaches to biomedical and biological research. Therefore, the NHGRI and a number of other institutes at NIH–including the National Center for Research Resources–have initiated programs to generate and characterize BAC libraries. These programs will ensure an adequate capacity for BAC library construction and characterization, and will generate the large number of BAC libraries needed for contemporary biological and biomedical research. The programs will support construction of an estimated ten libraries of human-sized genomes per year.

At present, the relatively high cost of making a BAC library is the major factor limiting the number of libraries that can be constructed each year. Thus, the new programs include a technology development component intended to improve methods for constructing and characterizing BAC libraries so that the number of useful libraries produced each year can be increased without an increase in overall costs. To ensure that the new capacity is used in the most effective way possible, the NHGRI has also implemented a new process to establish priorities for producing new BAC libraries. The increased capacity for generating high-quality BAC libraries will provide a critical resource for genomic studies.

Human Genome Project Multimedia Educational Kit

With the sequencing of the human genome and accelerated pace of genetic research, innovative new tools are needed to help the public, both old and young, understand how genomics will improve health and affect our lives. The NHGRI, with several cosponsors, created a free limited-edition educational multimedia kit to improve life sciences education in the nation's schools by ensuring that science teachers throughout the country have better access to the latest information about the Human Genome Project. Designed primarily for high school students, the kit has also proven useful for college students, voluntary health organizations, and the general public. Since its release in February 2001, nearly 60,000 free kits have been distributed to high school teachers and the public. The material from the kit is now available on the World Wide Web at www.nhgri.nih.gov/educationkit.

NHGRI has recruited members of the scientific community to become participants in a mentorship network program to help local high school classrooms effectively use the kit. Teachers will be able to contact scientists in their local community to clarify the underlying science in the kit, to increase his/her comfort level in teaching the materials, to invite them to speak to his/her classroom, or to inquire about bringing his/her classroom to their lab/clinic to learn, first hand, what they do. To date, over 600 volunteers have signed onto the mentorship network program.

As genetics is increasingly integrated into regular medical care, it will be important for the public to have a basic understanding of genomics. This tool kit is helping to bridge that gap in knowledge and inspire students to pursue careers in scientific research with a specific emphasis on genomics.

NHGRI/ORD Genetic and Rare Diseases Information Center

The NHGRI and the Office of Rare Diseases (ORD) have established the NHGRI/ORD Genetic and Rare Diseases Information Center to provide information on genetic and rare disorders to the public. A contract was awarded in FY 2001 and the center will be operational by the end of November 2001. The Information Center will meet the ever-increasing information needs of the general public, including patients and their families, health care professionals, and biomedical researchers by: 1) serving as a central, national repository of information materials and resources on genetic and rare diseases, conditions, and disorders; 2) collecting and disseminating information on the diagnosis, treatment, and prevention of genetic and rare disorders; and 3) coordinating with organizations and associations interested in genetic and rare disorders to explore networking capabilities, avoid duplication of effort, and identify information gaps.

With the high public profile of newly acquired genetics knowledge and ongoing genetic research, there is an increasing public demand for cutting edge, accurate, and timely information. Obtaining this information is essential for optimizing decisions about genetic testing, treatments, and participating in research. The NHGRI/ORD Genetic and Rare Diseases Information Center will provide inquiry response services and information resources that will help the public make informed choices based on new understandings about genetic and rare diseases.

NEW INITIATIVES

New Tools for Finding Genes that Contribute to Inherited Diseases

Any two people have the same DNA sequence for about 99.9 percent of their DNA. The 0.1 percent difference includes genetic variation that can lead to differences in the risk of getting various diseases and having adverse drug reactions. Some diseases, such as cystic fibrosis and Huntington's disease, result from differences in the DNA sequence in single genes. However, many common diseases such as diabetes, cancer, heart disease, psychiatric disorders, and asthma are influenced by complex interactions between multiple genes as well as by non-genetic factors such as diet, exercise, smoking, and exposure to toxins. With the tools of the Human Genome Project, finding the genes for diseases caused by alterations in single genes has become relatively

straightforward. Finding the genes that contribute to common diseases remains extremely difficult. To make this task easier, faster, and more efficient, the Human Genome Project is creating a new set of powerful genomic tools that, together with the reference genome sequence, will dramatically enhance scientists' ability to identify the genetic contributions to common diseases.

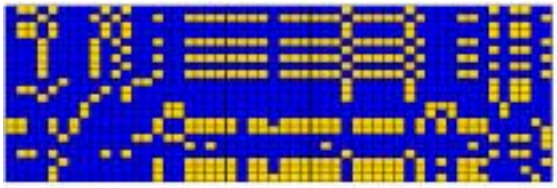
Single Nucleotide Polymorphisms (SNPs)
The Human Genome Project and its partners are creating a catalogue of the places in the genome where the DNA sequence differs among individuals. The most common variations are single nucleotide polymorphisms (SNPs), or places where the DNA sequence varies by a single base pair or DNA letter. Between two unrelated individuals, these occur approximately once every 1000 bases.

The NHGRI studies to identify SNPs have been complemented by the efforts of The SNP Consortium (TSC), a non-profit consortium whose members include the Wellcome Trust, 10 pharmaceutical companies, IBM, Motorola, and Amersham. Through these combined efforts, nearly 3 million SNPs have been identified and are available in public databases. A high-density map of SNPs is expected to be a valuable research tool that will help scientists pinpoint genetic differences that predispose some, but not others to disease, and underlie variability in individual response to treatment.

Scientists use the SNPs to scan the entire genome to find chromosomal regions that are statistically associated with disease. Scientists can then refer to the working draft of the human genome to find the genes in those regions and narrow their search for the disease-causing alterations. Whole-genome association studies such as this will be used to identify genes that are risk factors for disease and eventually to develop novel diagnostics and drugs that are tailored to patients' genetic profiles.

Haplotypes
While whole genome association studies are potentially extremely valuable, they remain dauntingly expensive, requiring additional technical improvements. The next key step of the Human Genome Project is the generation of a new map of the patterns of genetic variations across populations, a so-called "haplotype map." The SNP variants do not occur at random, but are correlated in important ways with their neighbors. By finding the pattern of variation along chromosomes, scientists can select a much smaller "gold standard" set of SNPs distributed along the chromosomes that will represent nearly all of the underlying pattern of variation. Thus, testing for one or two SNPs can reveal the information about variation for a large region surrounding the indicator SNPs.

Although it is still in its formative stage, the development of a haplotype map will provide a critical resource for researchers seeking to identify and understand the genetic basis of common human diseases.



Haplotypes (rows) for 22 chromosomes in the region of the ACE gene. Each column is for a SNP; the more common SNP allele is in black and the rarer SNP allele is in white. From Mark Rieder *et al.*, Nature Genetics, May 1999, vol. 22, page 61.

Proteomics: Understanding Protein Function

Success in genomics has stimulated the development of a new field of research known as "proteomics", the simultaneous study of a large numbers of proteins, or even the complete set of an organism's proteins–its "proteome". As complex as genomics is, proteomics presents a scientific problem of considerably greater scope and magnitude. While there appear to be between 30,000–50,000 human genes, the number of human proteins is much higher–in the range of 300,000–because each gene is capable of encoding several proteins, and any given protein can be modified in different ways by the cell. The number of proteins is not the only important issue; their distribution is critical to biology. Each cell type (e.g., skin cells, heart cells, etc) produces a characteristic subset of all of the possible proteins and specific proteins are located in different parts of the cell (nucleus, cytoplasm, cell membrane, and so on). The timing of protein expression is also critical; different subsets of proteins are made at different times during an organism's development and life span. To add a final level of complexity, many proteins carry out their work in interaction with other proteins, either serially (in time) or spatially (in protein complexes, or even protein "machines"). Together, these complexities present a challenge to our ability to study and understand how proteins work and how disease results when they don't work properly.

In the future, proteomics will have many important applications in the study and treatment of human disease. For the discovery of the causes of different diseases, proteomics will provide tools to investigate the precise molecular defect in diseased tissues. For diagnosing disease states, proteomics will be used in the development of reagents that will allow clinicians to precisely define the molecular characteristics of a particular disease. Finally, for drug discovery, proteomics will be a tool to identify new drug targets and to develop assays for new drugs.

Protein Sequence Databases

With the increasing number and variety of both protein sequences and functional information, the research community will become increasingly dependent on the availability of a stable database that collects and annotates protein sequences. At a basic level, scientists need a database of all possible proteins encoded by the genome of an organism to understand how these proteins function in making up a living cell, knowledge that is essential for the mission of NIH. The NHGRI and other institutes are planning to fund a centralized protein sequence database that will include information on protein sequence, nomenclature, and functional information. Importantly, scientists throughout the world will have unrestricted access to the information in this database.

Technology Development
A major obstacle to proteomics is the absence of the tools and technologiesto explore protein expression, structure, and function in high-throughput, low cost manner. In FY 2003, NHGRI is particularly interested in funding technology development projects. Examples include:

- A technique for differentially labeling the proteins in cells under two different conditions (e.g., before and after administration of a drug, or a normal cell and a diseased cell). These techniques can be used, for example, to reveal which proteins are overproduced or under produced in disease; such proteins would become important targets for drug design.
- A technique known as the protein "chip" that will allow investigators to measure the activity of thousands of proteins at a time.
- A technique known as "two-hybrid genetics" that allows comprehensive studies of protein-protein interactions in cells, with the aim of understanding how proteins normally work together.

The improvements in these, and many other techniques, will open up significant new avenues of understanding of human health and disease and will be the basis of many new approaches to the control of disease.

Sequencing New Genomes

Much of our current understanding of the molecular mechanisms underlying many important biological processes is due to the study of non-human, or model, organisms. This approach has become even more important recently as comparison of genome sequences has revealed much more similarity than anticipated between human genes and the genes in other organisms. With the extraordinary progress made in genomic research in the past few years, there are now opportunities for sequencing the genomes of additional organisms. Possible organisms for which genomic sequence would be extremely valuable include chimpanzee, rhesus monkey, dog, frog, and chicken.

With the completion of the human DNA sequence in 2003 and the rapid progress being made toward completion of the sequence of the mouse, the NHGRI sequencing program will have the opportunity to initiate major new sequencing projects in FY2003. However, there is a need to choose sensibly which genomes to sequence among many candidates. Accordingly, the NHGRI has implemented a systematic, peer review-based process for choosing new sequencing targets. The process requires investigators to submit a request articulating specific, well-defined scientific goals which will then be evaluated by a rigorous peer review process. This new selection process during FY2002 will assure a prioritized set of candidates will be available as the necessary capacity becomes available.

The Gene Ontology Consortium

As our appreciation for the close relationship of biological systems increases, investigators must be able to compare information in different databases and the databases must be able to readily exchange information about related genes and their potential relevance to human disease.

Unfortunately, this is currently very difficult to do. Historically, each of the fields of research using a different model organism has developed independently and has used a different system for naming genes. To add to the confusion, multiple names have been given to what have turned out to be the same gene. Many genes have also been given names that do not appropriately denote their function. A veritable Tower of Babel is threatening to impede progress in medicine and biology.

One approach to this problem would be to develop a standardized nomenclature across all organisms. However, because these naming systems are embedded in the research literature, this approach is unlikely to succeed. With funding from NHGRI, the Gene Ontology (GO) Consortium has taken a different approach. This consortium, which was started by members of the model organism database community, is developing controlled vocabularies–or gene ontologies–that will allow investigators to use the same terms to describe comparable molecular functions, biological components, and cellular components from any organism. Thus, any gene can be defined by a set of specific GO terms that will uniquely describe its attributes and allows the linking of similar genes despite their muddled nomenclature. Thus, by using GO terms, scientists can apply information learned from model organisms to humans and the understanding of human disease, a central goal of the NIH.

Barbados Prostate and Breast Cancer Study

The population of the island of Barbados represents an ideal resource for genetic-based studies of diseases prevalent in populations of West African ancestry. Prostate and breast cancer are the most common cancers in Barbados. Prostate cancer is substantially more common in African-American (and particularly Afro-Caribbean men) than in Caucasian or Asian men. Breast cancer, while not more common in blacks than whites, leads to higher mortality in black populations.

In collaboration with the Health Ministry of Barbados, NHGRI proposes to complete a national population-based case-controlled epidemiological study of all prostate and breast cancer cases with age/gender matched population controls over a 4-year period. A familial study component is also planned to explore the possibility of common genetic susceptibility in these diseases. Cases will be identified from the hospital in Barbados from which over 90 percent of the residents receive their care, and from local physicians. This study will build on cohorts studied in earlier NIH-supported studies on glaucoma. The study of an isolated, high-risk population such as this is made more valuable by the meticulous family history records maintained on the island, allowing researchers to easily determine and follow familial relationships that may be important in understanding the mode of transmission of these genetically-based diseases.

This study will examine a broad range of risk factors plus known and novel gene loci with state-of-the-art analytical tools, resulting in a comprehensive approach to the study of prostate and breast cancer. The overall goals will be to identify specific factors involved in genetic susceptibility for these cancers in Barbados, which can then be used to increase overall understanding of disease in this and other populations. This study will be a cooperative effort by the State University of New York (SUNY) at Stony Brook, the National Human Genome Research Institute, and the University of the West Indies.

OTHER AREAS OF INTEREST

Expansion of NHGRI Centers of Excellence In Genomic Science (CEGS)

The purpose of the CEGS program is to encourage academic centers to pursue advanced genome research using the new technologies developed by and the large databases produced by the Human Genome Project. The CEGS program challenges the entire biomedical research community to form multi-investigator, interdisciplinary teams to develop novel and innovative genomic research projects. CEGS projects will develop new methods and technologies for collecting, interpreting, and/or using genomic data sets that can be employed by the broader biomedical research community. The CEGS program is also a centerpiece for NHGRI's genomics training programs in general, and more specifically, for training of scientists from communities of underrepresented minorities.

The goals of the first three CEGS grants awarded in FY 2001 are:

- Analysis of how one person's genome is different from another person's, and how these differences affect their health, as well as reducing the cost of analyzing the medically relevant variation in an individual's genome to make genetically personalized medicine practical;
- Development of modular micro-scale instrumentation systems for the detection and analysis of how, when, and why very small populations of living cells interact with each other and their environment, to be able to relate specific genes to specific cellular processes essential for health or disrupted in disease;
- Simultaneous analysis of thousands of small DNA segments from the human genome and development of methods for discovering where key regulatory proteins bind throughout the genome, to help understand how sets of genes are regulated, which is important for understanding human development and disease.

The response to this program from the research community has been very strong, and NHGRI anticipates that community interest will remain high. We therefore expect to receive many high quality applications for funding in FY2003. These projects are pushing the cutting edge of genomic science and engage the talents of highly skilled scientists and engineers.

Health Disparities Strategic Plan

NHGRI has developed a Strategic Plan for Reducing Health Disparities that lays out a multifaceted approach to address issues of health disparities. The plan encompasses research , training, and education/outreach activities. The NHGRI staff recognizes the inherent value of increasing diversity among the research workforce as well as engaging and empowering people from minority communities through joint research projects, information sharing, dialogue and the development of partnerships.

Training
In order to address the significant under-representation of minorities conducting NHGRI research, a workshop was held to brainstorm about creative ideas and models for increasing the number of underrepresented minorities pursuing research careers in genomics and related sciences. The goals established during that workshop were:

- to increase the diversity of investigators participating in genomic and ELSI research;
- to increase the diversity of students trained in genomic and ELSI research;
- to expose a greater diversity of students and faculty to genomic and ELSI approaches to research.

One initiative is a partnership NHGRI has launched with NIGMS and the Indian Health Service to support research into the ethical, legal and social implications of genetic research as part of the new Native American Research Centers for Health initiative. This initiative supports partnerships between American Indian or Alaska Native tribes or of tribal-based organizations such as the National Indian Health Board and Area Health Boards, and institutions that conduct intensive academic-level biomedical and behavioral research.

Research
During the past several years NHGRI and the National Center for Minority Health and Health Disparities (formerly the Office of Research on Minority Health) have supported innovative research collaborations between investigators from Howard University and scientists in the intramural research program of the NHGRI. The collaboration involves support for projects involving African Americans affected with diabetes and hereditary prostate cancer. In addition, Howard University and the NHGRI are serving jointly as research training sites for African Americans involved in these projects. Among the goals of these studies has been to collect family and population-based information in a way that maximizes the participation of minority physicians, research scientists, and the community and to establish a center at Howard University for collaborative research on genomic analyses of diseases that disproportionately affect African Americans. This latter goal was realized on May 1, 2001 when The National Human Genome Center at Howard University was formally dedicated.

Africa America Diabetes Mellitus Study (AADM)
Because of the high frequency of environmental risk factors (diet, obesity) for type 2 diabetes in the African American population, it may well be more productive to study genetic risk factors in West Africans, since they are thought by many anthropologists to be the founding population of modern African Americans and have fewer dietary and nutritional confounding variables. To establish recruitment sites for the AADM study, five sites were selected through a peer review process from a total of 24 applications, three of them in Nigeria and two in Ghana. A total of 400 affected sib pairs has now been collected, are fully studied in terms of clinical parameters, and DNA obtained. Genotyping at Howard University has revealed evidence of genes involved in diabetes and obesity on Chromosome 20, and a full genome scan is underway. The study has not only started to yield high quality data, but has assisted in the recruitment of several top-flight scientists to the National Human Genome Center at Howard University.

African American Hereditary Prostate Cancer Study Network (AAHPC)
The National Human Genome Center at Howard is also coordinating a linkage study of hereditary prostate cancer, the African American Hereditary Prostate Cancer Study Network (AAHPC). African American prostate cancer families are almost completely missing from other pedigree collections, despite the higher incidence and higher lethality of prostate cancer in black men. Through a competitive review, the AAHPC study network has funded seven centers around the country. For most of these, the Principal Investigator is an African American urologist. Community acceptance and participation has been good and over 50 families have been identified with at least four affected males. DNA from these families is being studied to see if linkage can be confirmed to a known hereditary prostate cancer location on chromosome 1, as well as whether other linkages exist.

As the first large-scale genetic study of African Americans conducted almost entirely by African American clinical investigators, the AAHPC study has provided a foundation and productive environment for the exploration of all aspects of the involvement of African Americans in genetic research. Both of these projects are having a significant impact on the training of minority researchers as well. As a direct result of the initiation of these two studies, five recently recruited doctoral level young scientists are being supported to pursue these investigations.

MANAGEMENT INNOVATIONS

Bioinformatics and Scientific Programming Core

With the explosion of sequence and structural information available to researchers, the field of bioinformatics is playing an increasingly large role in the study of fundamental biological problems. The challenge facing computational biologists, especially in light of the vast amount of data produced by the Human Genome Project and other systematic sequencing projects, is to aid in the design of experiments that can potentially reveal previously unknown relationships with respect to gene function and structure. To that end, the Bioinformatics and Scientific Programming Core of the NHGRI intramural program provides expertise and assistance in bioinformatics and computational analysis for genome research being performed by NHGRI investigators.

The Bioinformatics and Scientific Programming Core provides access to specialized software and hardware; develops new “generalized solutions” for sequence-based analysis; develops databases for efficient archiving and retrieval of genomic and genetic data; and disseminates new software and database solutions to the genome community at large. Information on all software developed by this Core can be found at *http://genome.nhgri.nih.gov*.

Center For Inherited Disease Research

The Center for Inherited Disease Research (CIDR), located on the Bayview campus of The Johns Hopkins University, provides high-throughput genotyping services, study design advice, sophisticated data warehousing technologies, and database assistance to numerous investigators attempting to identify genetic loci and allelic variants involved in human disease. Most of these

investigators lack access to this kind of high through-put laboratory assistance. CIDR is a joint effort by twelve NIH institutes, with the National Human Genome Research Institute (NHGRI) serving as the lead agency.

To date, 72 projects (out of a total of 137 different projects requesting access) have been accepted for genotyping at CIDR, including studies of colon cancer, lung cancer, schizophrenia, Alzheimer's disease, Type II diabetes, bipolar disorder, non-syndromic deafness, obesity, HNPCC, osteoporosis, and dystonia. A rapidly growing list of exciting linkages is the result. The current genotyping capacity of CIDR is over 5.7 million genotypes per year and is projected to increase to 8.5 million genotypes per year in the fall of 2002, and 11 million in FY2003.

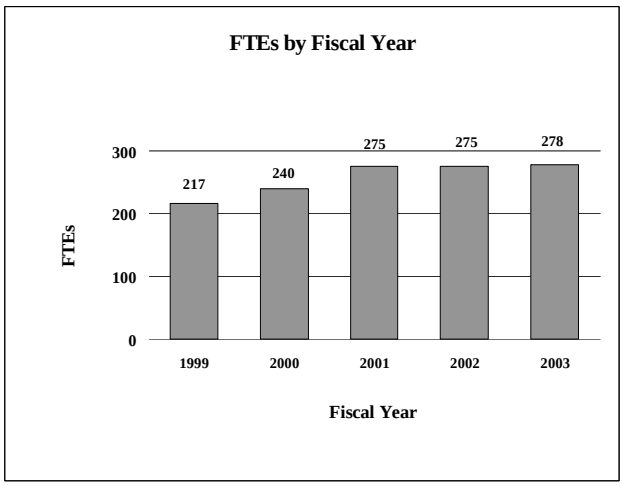
NIH Intramural Sequencing Center

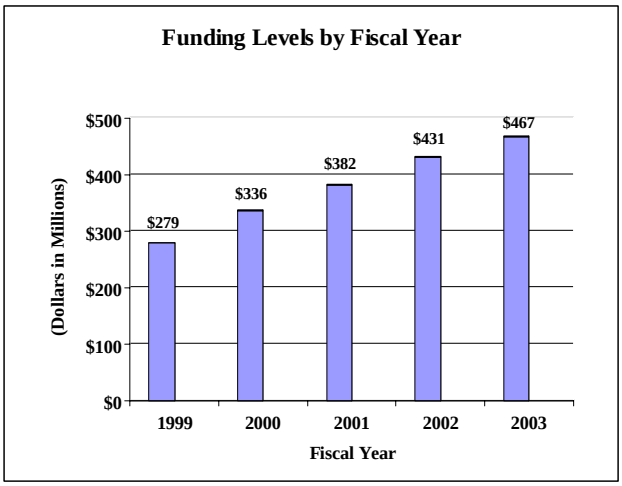
The NIH Intramural Sequencing Center (NISC) was established in 1997 with an initial mandate of providing NIH intramural investigators access to large-scale DNA sequencing and sequence analysis. This service component of NISC operates exclusively via a fee-for-service reimbursement program and has been remarkably stable since its inception. Over the past 24 months, NISC has undergone a significant metamorphosis with respect to both its size and scientific portfolio. Located in an off-campus facility in Gaithersburg, NISC now employs 35 people (including 10 Ph.D.-level scientists) and will obtain roughly $10 million in fees during the current fiscal year. NISC's total sequence production has increased 20-fold over the past 24 months, with the generation of 2,199,437 sequence reads last fiscal year and the projected generation of 3,000,000 sequence reads during this current fiscal year. NISC is now playing a central role in exploring the utility of comparing the sequence of numerous vertebrates to understand the function of the human genome. For more information see http://www.nisc.nih.gov.

Budget Policy

The Fiscal Year 2003 budget request for the NHGRI is $466,695,000, including AIDS, an increase of $35,977,000 and 8.4 percent over the FY 2002 level.

A five year history of FTEs and Funding Levels for NHGRI are shown in the graphs below. Note that Fiscal Years 2000 and 1999 are not comparable for the Managerial Flexibility Act of 2001 legislative proposal.



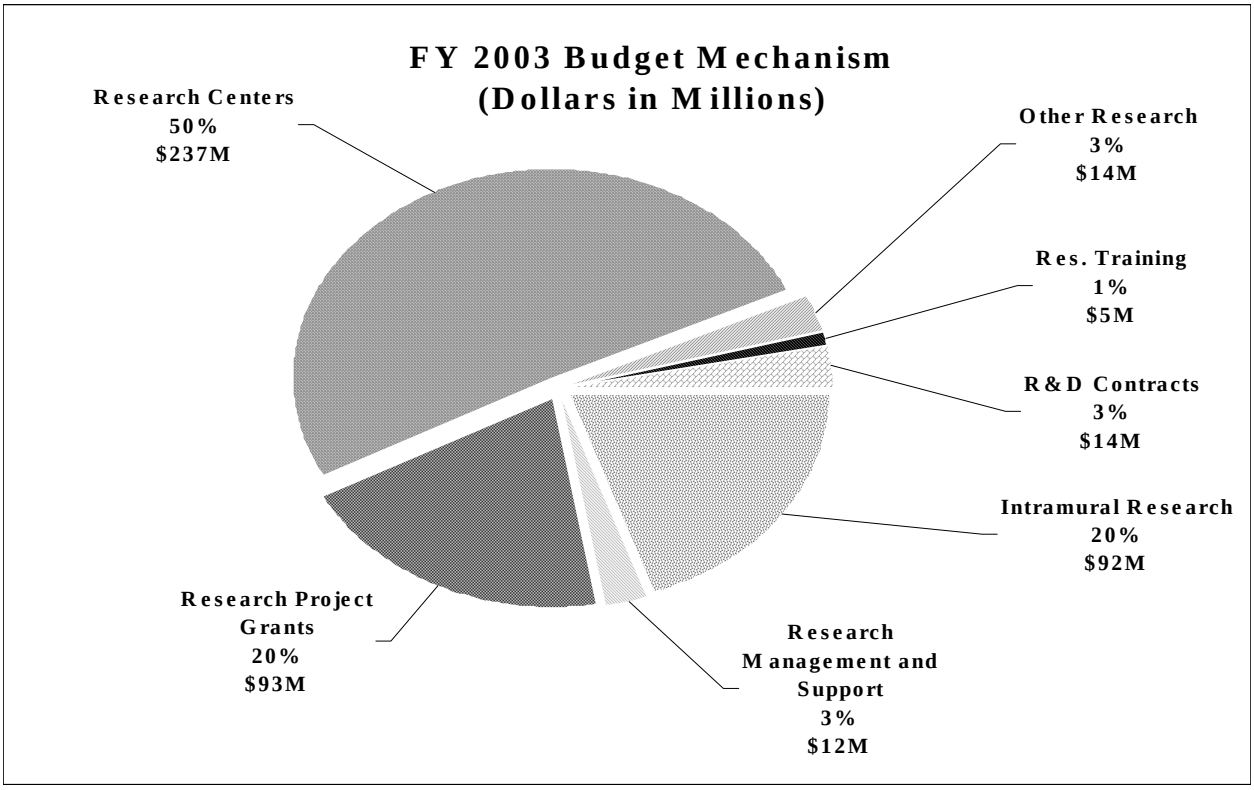


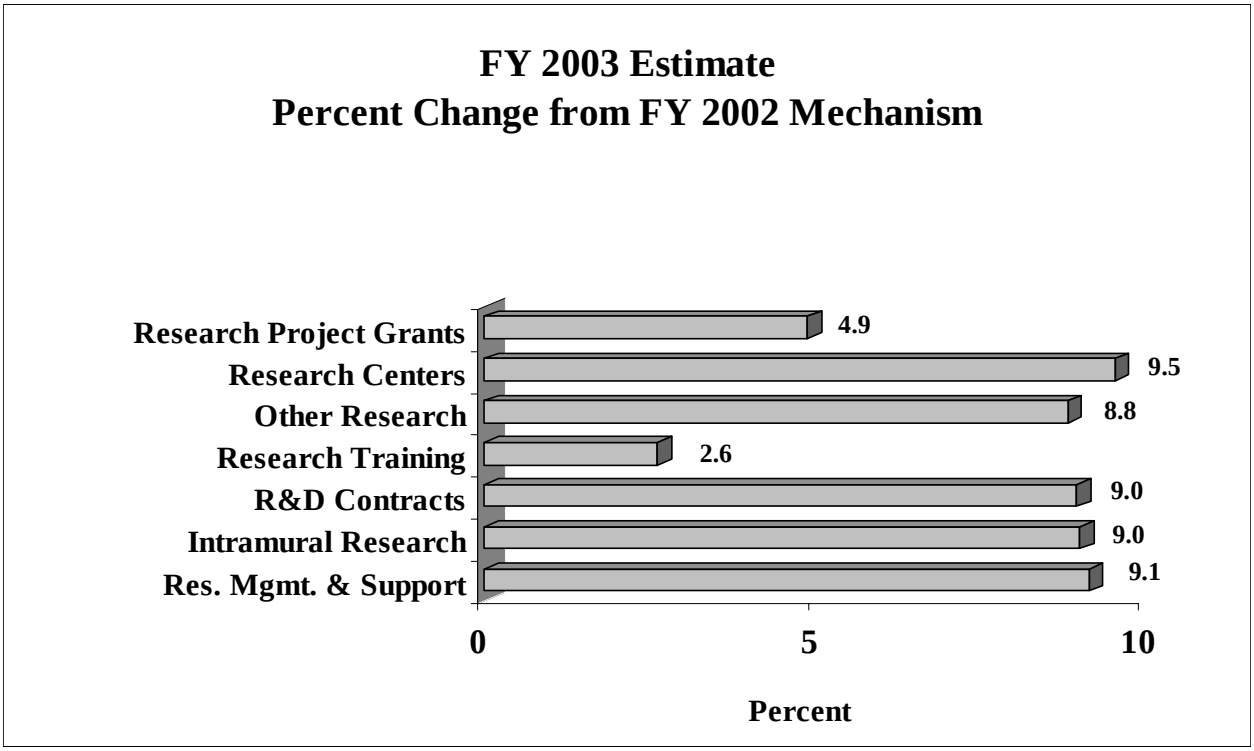
One of NIH's highest priorities is the funding of medical research through research project grants (RPGs). Support for RPGs allows NIH to sustain the scientific momentum of investigator-initiated research while providing new research opportunities. The Fiscal Year 2003 request provides average cost increases for competing RPGs equal to the Biomedical Research and Development Price Index (BRDPI), estimated at 4.0 percent. Noncompeting RPGs will be funded at committed levels which include increases of 3 percent on average for recurring direct costs.

Future promises for advancement in medical research rest in part with new investigators with new ideas. In the Fiscal Year 2003 request, NHGRI will support 110 pre- and postdoctoral trainees in full-time training positions, the same number as in FY 2002 . Stipend levels for NRSA trainees will increase by 4 percent over Fiscal Year 2002 levels.

The Fiscal Year 2003 request includes funding for 35 research centers, 53 other research grants, including 25 research career awards, and 16 R&D contracts. Intramural Research and Research Management and Support receive increases of 9 percent over FY 2002.

The mechanism distribution by dollars and percent change are displayed below:





NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
TOTAL - Current Law
Budget Mechanism

MECHANISM	FY 2001 Actual		FY 2002 Appropriation		FY 2002 Current Estimate		FY 2003 Estimate	
Research Grants:	No.	Amount	No.	Amount	No.	Amount	No.	Amount
Research Projects:								
Noncompeting	73	$38,948,000	76	$50,820,000	76	$50,820,000	70	$53,493,000
Administrative supplements	(30)	6,540,000	(25)	5,151,000	(25)	5,151,000	(26)	5,372,000
Competing:								
Renewal	13	13,452,000	12	13,474,000	12	13,474,000	13	14,910,000
New	25	9,412,000	26	10,373,000	26	10,373,000	25	9,891,000
Supplements	0	0	0	0	0	0	0	0
Subtotal, competing	38	22,864,000	38	23,847,000	38	23,847,000	38	24,801,000
Subtotal, RPGs	111	68,352,000	114	79,818,000	114	79,818,000	108	83,666,000
SBIR/STTR	25	8,030,000	28	8,755,000	28	8,755,000	30	9,231,000
Subtotal, RPGs	136	76,382,000	142	88,573,000	142	88,573,000	138	92,897,000
Research Centers:								
Specialized/comprehensive	16	179,647,000	22	195,739,000	22	195,739,000	24	213,489,000
Clinical research	0	0	0	0	0	0	0	0
Biotechnology	8	15,643,000	10	20,669,000	10	20,669,000	11	23,558,000
Comparative medicine	0	0	0	0	0	0	0	0
Research Centers in Minority Institutions	0	0	0	0	0	0	0	0
Subtotal, Centers	24	195,290,000	32	216,408,000	32	216,408,000	35	237,047,000
Other Research:								
Research careers	22	3,558,000	24	4,162,000	24	4,162,000	25	4,631,000
Cancer education	0	0	0	0	0	0	0	0
Cooperative clinical research	0	0	0	0	0	0	0	0
Biomedical research support	0	0	0	0	0	0	0	0
Minority biomedical research support	0	0	0	0	0	0	0	0
Other	23	7,841,000	26	8,579,000	26	8,579,000	28	9,234,000
Subtotal, Other Research	45	11,399,000	50	12,741,000	50	12,741,000	53	13,865,000
Total Research Grants	205	283,071,000	224	317,722,000	224	317,722,000	226	343,809,000
Training:	FTTPs		FTTPs		FTTPs		FTTPs	
Individual awards	19	662,000	20	766,000	20	766,000	20	762,000
Institutional awards	66	2,691,000	90	3,905,000	90	3,905,000	90	4,031,000
Total, Training	85	3,353,000	110	4,671,000	110	4,671,000	110	4,793,000
Research & development contracts	14	11,454,000	15	12,660,000	15	12,634,000	16	13,765,000
(SBIR/STTR)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	FTEs		FTEs		FTEs		FTEs	
Intramural research	210	73,698,000	210	83,647,000	210	83,491,000	212	91,000,000
Research management and support	65	9,525,000	65	10,815,000	65	10,794,000	66	11,770,000
Cancer prevention & control	0	0	0	0	0	0	0	0
Construction		0		0		0		0
Total, NHGRI	275	381,101,000	275	429,515,000	275	429,312,000	278	465,137,000
(Clinical Trials)		(6,342,000)		(7,122,000)		(7,122,000)		(7,724,000)

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
TOTAL - Accrued Costs for Retirement and Health Benefits
Budget Mechanism

MECHANISM	FY 2001 Actual		FY 2002 Appropriation		FY 2002 Current Estimate		FY 2003 Estimate	
Research Grants:	No.	Amount	No.	Amount	No.	Amount	No.	Amount
Research Projects:								
Noncompeting								
Administrative supplements								
Competing:								
Renewal								
New								
Supplements								
Subtotal, competing								
Subtotal, RPGs								
SBIR/STTR								
Subtotal, RPGs								
Research Centers:								
Specialized/comprehensive								
Clinical research								
Biotechnology								
Comparative medicine								
Research Centers in Minority Institutions								
Subtotal, Centers								
Other Research:								
Research careers								
Cancer education								
Cooperative clinical research								
Biomedical research support								
Minority biomedical research support								
Other								
Subtotal, Other Research								
Total Research Grants								
Training:	FTTPs		FTTPs		FTTPs		FTTPs	
Individual awards								
Institutional awards								
Total, Training								
Research & development contracts (SBIR/STTR)								
	FTEs		FTEs		FTEs		FTEs	
Intramural research	0	1,042,000	0	1,130,000	0	1,130,000	0	1,246,000
Research management and support	0	257,000	0	276,000	0	276,000	0	312,000
Cancer prevention & control	0	0	0	0	0	0	0	0
Construction								
Total, NHGRI	0	1,299,000	0	1,406,000	0	1,406,000	0	1,558,000
(Clinical Trials)		(0)		(0)		(0)		(0)

National Human Genome Research Institute
TOTAL - Proposed Law
Budget Mechanism

MECHANISM	FY 2001 Actual		FY 2002 Appropriation		FY 2002 Current Estimate		FY 2003 Estimate	
Research Grants:	No.	Amount	No.	Amount	No.	Amount	No.	Amount
Research Projects:								
Noncompeting	73	$38,948,000	76	$50,820,000	76	$50,820,000	70	$53,493,000
Administrative supplements	(30)	6,540,000	(25)	5,151,000	(25)	5,151,000	(26)	5,372,000
Competing:								
Renewal	13	13,452,000	12	13,474,000	12	13,474,000	13	14,910,000
New	25	9,412,000	26	10,373,000	26	10,373,000	25	9,891,000
Supplements	0	0	0	0	0	0	0	0
Subtotal, competing	38	22,864,000	38	23,847,000	38	23,847,000	38	24,801,000
Subtotal, RPGs	111	68,352,000	114	79,818,000	114	79,818,000	108	83,666,000
SBIR/STTR	25	8,030,000	28	8,755,000	28	8,755,000	30	9,231,000
Subtotal, RPGs	136	76,382,000	142	88,573,000	142	88,573,000	138	92,897,000
Research Centers:								
Specialized/comprehensive	16	179,647,000	22	195,739,000	22	195,739,000	24	213,489,000
Clinical research	0	0	0	0	0	0	0	0
Biotechnology	8	15,643,000	10	20,669,000	10	20,669,000	11	23,558,000
Comparative medicine	0	0	0	0	0	0	0	0
Research Centers in Minority Institutions	0	0	0	0	0	0	0	0
Subtotal, Centers	24	195,290,000	32	216,408,000	32	216,408,000	35	237,047,000
Other Research:								
Research careers	22	3,558,000	24	4,162,000	24	4,162,000	25	4,631,000
Cancer education	0	0	0	0	0	0	0	0
Cooperative clinical research	0	0	0	0	0	0	0	0
Biomedical research support	0	0	0	0	0	0	0	0
Minority biomedical research support	0	0	0	0	0	0	0	0
Other	23	7,841,000	26	8,579,000	26	8,579,000	28	9,234,000
Subtotal, Other Research	45	11,399,000	50	12,741,000	50	12,741,000	53	13,865,000
Total Research Grants	205	283,071,000	224	317,722,000	224	317,722,000	226	343,809,000
Training:	FTTPs		FTTPs		FTTPs		FTTPs	
Individual awards	19	662,000	20	766,000	20	766,000	20	762,000
Institutional awards	66	2,691,000	90	3,905,000	90	3,905,000	90	4,031,000
Total, Training	85	3,353,000	110	4,671,000	110	4,671,000	110	4,793,000
Research & development contracts	14	11,454,000	15	12,660,000	15	12,634,000	16	13,765,000
(SBIR/STTR)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	FTEs		FTEs		FTEs		FTEs	
Intramural research	210	74,740,000	210	84,777,000	210	84,621,000	212	92,246,000
Research management and support	65	9,782,000	65	11,091,000	65	11,070,000	66	12,082,000
Cancer prevention & control	0	0	0	0	0	0	0	0
Construction		0		0		0		0
Total, NHGRI	275	382,400,000	275	430,921,000	275	430,718,000	278	466,695,000
(Clinical Trials)		(6,342,000)		(7,122,000)		(7,122,000)		(7,724,000)

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Budget Authority by Activity 1/
(dollars in thousands)

ACTIVITY	FY 2001 Actual		FY 2002 Estimate		FY 2003 Estimate		Change	
	FTEs	Amount	FTEs	Amount	FTEs	Amount	FTEs	Amount
Extramural Research:								
Human Genome Research		$297,878		$335,027		$362,367		$27,340
Subtotal, Extramural research		297,878		335,027		362,367		27,340
Intramural research	210	74,740	210	84,621	212	92,246	2	7,625
Research managament and support	65	9,782	65	11,070	66	12,082	1	1,012
Total	275	382,400	275	430,718	278	466,695	3	35,977

1/ Please see the following tables for the crosswalk from current law to proposed law to reflect the administration's proposal for full accrued retirement and health benefits.

National Institutes of Health

National Human Genome Research Institute

2001 Crosswalk for Accrued Retirement and Health Benefit Costs
(Dollars in thousands)

	2001 Actual Current Law	2001 Additional Accrual Costs	2001 Actual Proposed Law
Extramural Research:			
Human Genome Research	$297,878	$0	$297,878
Subtotal, extramural research	297,878	0	297,878
Intramural Research	73,698	1,042	74,740
Research management and support	9,525	257	9,782
Total	381,101	1,299	382,400

National Institutes of Health

National Human Genome Research Institute

2002 Crosswalk for Accrued Retirement and Health Benefit Costs
(Dollars in thousands)

	2002 Current Estimate Current Law	2002 Additional Accrual Costs	2002 Appropriation Proposed Law
Extramural Research:			
Human Genome Research	$335,027	$0	$335,027
Subtotal, extramural research	335,027	0	335,027
Intramural Research	83,491	1,130	84,621
Research management and support	10,794	276	11,070
Total	429,312	1,406	430,718

National Institutes of Health

National Human Genome Research Institute

2003 Crosswalk for Accrued Retirement and Health Benefit Costs
(Dollars in thousands)

	2003 Estimate Current Law	2003 Additional Accrual Costs	2003 Estimate Proposed Law
Extramural Research:			
Human Genome Research	$362,367	$0	$362,367
Subtotal, extramural research	362,367	0	362,367
Intramural Research	91,000	1,246	92,246
Research management and support	11,770	312	12,082
Total	465,137	1,558	466,695

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Summary of Changes

2002 Estimated budget authority				$430,718,000
2003 Estimated budget authority				466,695,000
Net change				35,977,000
	2002 Current Estimate Base		Change from Base	
CHANGES	FTEs	Budget Authority	FTEs	Budget Authority
A. Built-in:				
1. Intramural research:				
a. Within grade increase		$21,402,000		$304,000
b. Annualization of January 2002 pay increase		21,402,000		257,000
c. January 2003 pay increase		21,402,000		430,000
d. Payment for centrally furnished services		14,755,000		1,328,000
e. Increased cost of laboratory supplies, materials, and other expenses		47,334,000		1,016,000
f. Accrued costs for retirement and health benefits		1,130,000		116,000
Subtotal				3,451,000
2. Research Management and Support:				
a. Within grade increase		6,082,000		103,000
b. Annualization of January 2002 pay increase		6,082,000		72,000
c. January 2003 pay increase		6,082,000		121,000
d. Payment for centrally furnished services		739,000		67,000
e. Increased cost of laboratory supplies, materials, and other expenses		3,973,000		119,000
f. Accrued costs for retirement and health benefits		276,000		36,000
Subtotal				518,000
3. Cancer Prevention and Control:				
a. Within grade increase		0		
b. Annualization of January 2002 pay increase		0		
c. January 2003 pay increase		0		
d. Payment for centrally furnished services		0		
e. Increased cost of laboratory supplies, materials, and other expenses		0		
f. Accrued costs for retirement and health benefits		0		
Subtotal				0
Subtotal, Built-in				3,969,000

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Summary of Changes--continued

CHANGES	2002 Current Estimate Base		Change from Base	
	No.	Amount	No.	Amount
B. Program:				
1. Research project grants:				
a. Noncompeting	76	55,971,000	(6)	2,894,000
b. Competing	38	23,847,000	0	954,000
c. SBIR/STTR	28	8,755,000	2	476,000
Total	142	88,573,000	(4)	4,324,000
2. Centers	32	216,408,000	3	20,639,000
3. Other research	50	12,741,000	3	1,124,000
4. Research training	110	4,671,000	0	122,000
5. Research and development contracts	15	12,634,000	1	1,131,000
Subtotal, extramural		335,027,000		27,340,000
	FTEs		FTEs	
6. Intramural research	210	84,621,000	2	4,174,000
7. Research management and support	65	11,070,000	1	494,000
8. Construction	0	0	0	0
Subtotal, program		430,718,000		32,008,000
Total changes	275		3	35,977,000

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Budget Authority by Object

		FY 2002 Appropriation	FY 2002 Current Estimate	FY 2003 Estimate	Increase or Decrease
Total compensable workyears:					
Full-time employment		275	275	278	3
Full-time equivalent of overtime and holiday hours		2	2	2	0
Average ES salary		$135,035	$135,035	$138,546	$3,511
Average GM/GS grade		10.7	10.7	10.7	0.0
Average GM/GS salary		$58,971	$58,971	$60,504	$1,533
Average salary, grades established by act of July 1, 1944 (42 U.S.C. 207)		$66,864	$66,864	$68,602	$1,738
Average salary of ungraded positions		$92,439	$92,439	$94,842	$2,403
	OBJECT CLASSES	FY 2002 Appropriation	FY 2002 Estimate	FY 2003 Estimate	Increase or Decrease
	Personnel Compensation:				
11.1	Full-Time Permanent	$10,107,000	$10,107,000	$10,655,000	$548,000
11.3	Other than Full-Time Permanent	8,847,000	8,847,000	9,338,000	491,000
11.5	Other Personnel Compensation	659,000	659,000	689,000	30,000
11.8	Special Personnel Services Payments	2,618,000	2,618,000	2,626,000	8,000
11.9	**Total Personnel Compensation**	**22,231,000**	**22,231,000**	**23,308,000**	**1,077,000**
12.1	Personnel Benefits	5,253,000	5,253,000	5,471,000	218,000
12.1	Personnel Benefits, Accrued Retirement Costs	951,000	951,000	1,090,000	139,000
13.0	Benefits for Former Personnel	0	0	0	0
	Subtotal, Pay Cost, Current Law	**27,484,000**	**27,484,000**	**28,779,000**	**1,295,000**
	Subtotal, Pay Cost, Proposed Law	**28,435,000**	**28,435,000**	**29,869,000**	**1,434,000**
21.0	Travel and Transportation of Persons	1,207,000	1,207,000	1,253,000	46,000
22.0	Transportation of Things	152,000	152,000	157,000	5,000
23.1	Rental Payments to GSA	3,000	3,000	3,000	0
23.2	Rental Payments to Others	262,000	262,000	268,000	6,000
23.3	Communications, Utilities and Miscellaneous Charges	530,000	530,000	546,000	16,000
24.0	Printing and Reproduction	102,000	102,000	107,000	5,000
25.1	Consulting Services	494,000	494,000	506,000	12,000
25.2	Other Services	7,658,000	7,658,000	8,377,000	719,000
25.3	Purchase of Goods and Services from Government Accounts	42,265,000	42,088,000	47,305,000	5,217,000
25.3	Accrued Retirement Costs	455,000	455,000	468,000	13,000
25.4	Operation and Maintenance of Facilities	3,347,000	3,347,000	3,423,000	76,000
25.5	Research and Development Contracts	3,321,000	3,295,000	6,565,000	3,270,000
25.6	Medical Care	774,000	774,000	788,000	14,000
25.7	Operation and Maintenance of Equipment	1,308,000	1,308,000	1,339,000	31,000
25.8	Subsistence and Support of Persons	0	0	0	0
25.0	**Subtotal, Other Contractual Services, Current Law**	**59,167,000**	**58,964,000**	**68,303,000**	**9,339,000**
25.0	**Subtotal, Other Contractual Services, Proposed Law**	**59,622,000**	**59,419,000**	**68,771,000**	**9,352,000**
26.0	Supplies and Materials	9,829,000	9,829,000	10,117,000	288,000
31.0	Equipment	8,637,000	8,637,000	8,908,000	271,000
32.0	Land and Structures	0	0	0	0
33.0	Investments and Loans	0	0	0	0
41.0	Grants, Subsidies and Contributions	322,138,000	322,138,000	346,692,000	24,554,000
42.0	Insurance Claims and Indemnities	0	0	0	0
43.0	Interest and Dividends	4,000	4,000	4,000	0
44.0	Refunds	0	0	0	0
	Subtotal, Non-Pay Costs, Current Law	**402,031,000**	**401,828,000**	**436,358,000**	**34,530,000**
	Subtotal, Non-Pay Costs, Proposed Law	**402,486,000**	**402,283,000**	**436,826,000**	**34,543,000**
	Total Budget Authority by Object, Current	**429,515,000**	**429,312,000**	**465,137,000**	**35,825,000**
	Total Budget Authority by Object, Proposed	**430,921,000**	**430,718,000**	**466,695,000**	**35,977,000**
	Total Accrued Retirement Costs	**1,406,000**	**1,406,000**	**1,558,000**	**152,000**

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Salaries and Expenses

OBJECT CLASSES	FY 2002 Appropriation	FY 2002 Current Estimate	FY 2003 Estimate	Increase or Decrease
Personnel Compensation:				
Full-Time Permanent (11.1)	$10,107,000	$10,107,000	$10,655,000	$548,000
Other Than Full-Time Permanent (11.3)	8,847,000	8,847,000	9,338,000	491,000
Other Personnel Compensation (11.5)	659,000	659,000	689,000	30,000
Special Personnel Services Payments (11.8)	2,618,000	2,618,000	2,626,000	8,000
Total Personnel Compensation (11.9)	**22,231,000**	**22,231,000**	**23,308,000**	**1,077,000**
Civilian Personnel Benefits (12.1)	5,253,000	5,253,000	5,471,000	218,000
Accrued Costs of Retirement Benefits (12.1)	951,000	951,000	1,090,000	139,000
Benefits to Former Personnel (13.0)	0	0	0	0
Subtotal, Pay Costs, Current Law	**27,484,000**	**27,484,000**	**28,779,000**	**1,295,000**
Subtotal, Pay Costs, Proposed Law	**28,435,000**	**28,435,000**	**29,869,000**	**1,434,000**
Travel (21.0)	1,207,000	1,207,000	1,253,000	46,000
Transportation of Things (22.0)	152,000	152,000	157,000	5,000
Rental Payments to Others (23.2)	262,000	262,000	268,000	6,000
Communications, Utilities and Miscellaneous Charges (23.3)	530,000	530,000	546,000	16,000
Printing and Reproduction (24.0)	102,000	102,000	107,000	5,000
Other Contractual Services:				
Advisory and Assistance Services (25.1)	440,000	440,000	450,000	10,000
Other Services (25.2)	7,658,000	7,658,000	8,377,000	719,000
Purchases from Govt. Accounts (25.3)	32,019,000	31,842,000	36,137,000	4,295,000
Accrued Retirement Costs (25.3)	455,000	455,000	468,000	13,000
Operation & Maintenance of Facilities (25.4)	3,347,000	3,347,000	3,423,000	76,000
Operation & Maintenance of Equipment (25.7)	1,308,000	1,308,000	1,339,000	31,000
Subsistence & Support of Persons (25.8)	0	0	0	0
Subtotal, Other Contractual Services, Current Law	**44,772,000**	**44,595,000**	**49,726,000**	**5,131,000**
Subtotal, Other Contractual Services, Proposed Law	**45,227,000**	**45,050,000**	**50,194,000**	**5,144,000**
Supplies and Materials (26.0)	9,828,000	9,828,000	10,116,000	288,000
Subtotal, Non-Pay Costs, Current Law	**54,600,000**	**54,423,000**	**62,095,000**	**7,672,000**
Subtotal, Non-Pay Costs, Proposed Law	**55,055,000**	**54,878,000**	**62,563,000**	**7,685,000**
Total, Administrative Costs, Current Law	**82,084,000**	**81,907,000**	**90,874,000**	**8,967,000**
Total, Accrued Costs	**1,406,000**	**1,406,000**	**1,558,000**	**152,000**
Total, Administrative Costs, Proposed Law	**83,490,000**	**83,313,000**	**92,432,000**	**9,119,000**

NATIONAL INSTITUTES OF HEALTH
National Human Genome Research Institute
Significant Items in House and Senate Appropriations Committee Reports

FY 2002 House Appropriations Committee Report Language (H. Rpt.107-229)

Item

Ethical, Legal and Social Implications S The Committee commends the Institute for its leadership in research and policy development related to the ethical, legal and social implications (ELSI) of the human genome project. The early commitment to devote resources to study ELSI issues has generated important information and recommendations concerning both research and public policy. The Committee encourages NHGRI to move forward with ELSI research and policy initiatives outlined in the budget justifications and in testimony before the Committee. (p. 87)

Action taken or to be taken

The ELSI Research Program has continued to support significant and innovative research on the ethical, legal, and social implications of human genome research. Research projects supported in FY 2001 included projects in the areas of the privacy and fairness in the use and interpretation of genetic information; clinical integration of new genetic technologies; issues surrounding genetics research; and public and professional education. A three-day conference attended by over 400 people was held in FY 2001 to reflect on the past, present, and future of ELSI research; the conference featured approximately 90 panel and plenary presentations and over 50 posters. The ELSI Program continues to co-fund a contract for a large multi-center epidemiological study of genotypic and phenotypic screening for iron overload and hereditary hemochromatosis, focusing on the ethical and social implications of such screening. At the end of FY 2001, approximately 25,000 participants had been recruited, with about 50 percent of the participants coming from minority populations.

The ELSI program has also continued to support the Genetic Variation Consortium, a consortium of investigators conducting studies on ELSI issues arising from research on human genetic sequence variation. In FY 2001 the ELSI program released a new Request for Applications (RFA) on the ELSI implications of human genetic variation research for individuals and diverse racial and ethnic groups. ELSI investigators conducting research in this area have now begun to work closely with genomics researchers in an initiative to develop additional resources for the study of human variation in an ethically appropriate way.

The ELSI program has also become a partner on two new minority training initiatives: 1) an initiative to recruit minority college students into ELSI research careers; and 2) an initiative to support the development of partnerships between American Indian/Alaska Native tribes or tribal-based organizations, and institutions that conduct ELSI research. In February 2002, the ELSI program will launch a new initiative for ELSI dissertation research grants for underrepresented minorities.

Item

Human Genome Project S . . . The Committee encourages the Institute to pursue the development of the next generation of genomic tools and technologies needed to study the human genome and understand its role in human health and disease. (p. 87)

Action taken or to be taken

The NHGRI considers the next phase of genomic research to be profoundly exciting for the future of science and medicine. A wide-ranging and ambitious effort is underway to capitalize on these new opportunities. As one means of promoting the development of the next generation of genomic tools and technologies, the NHGRI has created the Centers of Excellence In Genomic Science (CEGS) program. This program encourages academic centers to pursue advanced and innovative genome research using large databases produced by the Human Genome Project (HGP) and the new technologies developed by the HGP. The CEGS program challenges the entire biomedical research community to form multi-investigator, interdisciplinary teams to develop novel genomic research projects. CEGS projects will develop new methods and technologies for collecting, interpreting, and/or using genomic data sets that can be employed by the broader biomedical research community. The CEGS program is also a centerpiece for NHGRI's genomics training programs in general, and more specifically, for training of scientists from communities of underrepresented minorities.

In 2001, NHGRI awarded the first three CEGS grants. Two separate awards went to the University of Washington; the third was to Yale University. The response to this program from the research community has been very positive, and NHGRI anticipates that community interest will remain high. We therefore expect to receive many high quality applications for funding in FY 2003. These projects are pushing the cutting edge of genomic science and will engage the talents of many highly skilled scientists and engineers.

Item

Human Genome Project S The Committee is pleased that NHGRI has upheld the policy that all sequence information is provided freely to all via the World Wide Web, with no restrictions on its use or redistribution. The Committee understands that ready access to this fundamental sequence information has already led to the identification of tens of thousands of genes, including in excess of 30 over the past two years that play a direct role in human disease. The Committee encourages the Institute to continue this policy of free and unfettered access to genomic data. (p. 87)

Action taken or to be taken

The NHGRI continues its data release policies to ensure that all sequence information produced by the HGP is made available in an unrestricted manner to all users through public databases and the World Wide Web. In the past year, in addition to the results of the human sequencing component of the HGP, large amounts of sequence data for the mouse, the rat, and several other

important experimental organisms have become available. In the latter cases, the data release practices have recently been expanded to include the actual raw data (the "traces" from the automated sequencing instruments) in addition to the processed sequences. The free access policy has also been applied to additional types of genomic data, such as SNPs (single nucleotide polymorphisms, or individual differences in the sequence between individuals) and the full-length cDNA products of the Mammalian Gene Collection program. The Institute is also actively developing appropriate policies for unrestricted public access to yet other data types such as gene expression data and the data that will be produced by the Centers of Excellence in Genome Science (CEGS), which are the next generation of genome centers.

FY 2001 Senate Appropriations Committee Report Language (S. Rept. 107-84)

Item

Epilepsy S . . . The Committee urges the Institute to make research in epilepsy a priority and to coordinate research efforts with other Institutes through the Interagency Epilepsy Coordinating Committee comprised of agency scientists and industry and patient representatives. (p. 171)

Action taken or to be taken

Hereditary factors play a significant role in epilepsy, and the study of the human genome should have direct benefit to elucidating the causes of the epilepsy syndromes. At its core, the Human Genome Project is generating genomic information and developing tools for understanding the instructions encoded in human DNA. Having the sequence, or spelling, of the human genome and the tools to interpret its meaning, will enable biomedical researchers in all areas to answer questions about disease processes and help develop new strategies for their prevention, diagnosis and treatment. Last year, Human Genome Project participants announced that they had completed a “working draft” of the human genome and that the draft was freely available in public databases. While this marked a significant milestone of the Human Genome Project, much work remains to be done in order for disease investigators to make sense of the information. The working draft, while not as refined as the final version, includes sequence covering most of the genome and represents the raw data needed to find most of the human genes. The final human genome sequence will be highly refined (99.99 percent accurate), and is expected to be completed by the spring of 2003. Furthermore, a catalogue of information on variations in this sequence is being rapidly developed. It is likely that such variants play a role in susceptibility to common illnesses such as epilepsy.

NHGRI is committed to both understanding the causes and developing effective therapies for all forms of diseases, including epilepsy. NHGRI will coordinate with NINDS, the lead NIH institute working on epilepsy research, as well as other institutes as needed.

Item

Human Genome Project S The Committee commends the international Human Genome Project, led by the NHGRI, for sequencing 95 percent of the 3.1 billion bases of the human genome years ahead of schedule. This was a remarkable achievement that could not have occurred without Federal support. At the same time, the Committee recognizes that substantial effort will be required to produce the final, highly accurate sequence by 2003. The Committee encourages NHGRI to pursue the development of the next generation of genomic tools and technologies needed to study the human genome and understand its role in human health and disease. (p. 171)

Action taken or to be taken

The NHGRI would like to recognize the profoundly important role of the U.S. Congress in initiating and sustaining the Human Genome Project, making it possible to meet and exceed these ambitious milestones. The process of finishing the sequence to high accuracy standards is on track for completion in 2003.

The NHGRI considers the next phase of genomic research to be profoundly exciting for the future of science and medicine. A wide-ranging and ambitious effort is underway to capitalize on these new opportunities. As one means of promoting the development of the next generation of genomic tools and technologies, the NHGRI has created the Centers of Excellence In Genomic Science (CEGS) program. This program encourages academic centers to pursue advanced and innovative genome research using large databases produced by the Human Genome Project (HGP) and the new technologies developed by the HGP. The CEGS program challenges the entire biomedical research community to form multi-investigator, interdisciplinary teams to develop novel genomic research projects. CEGS projects will develop new methods and technologies for collecting, interpreting, and/or using genomic data sets that can be employed by the broader biomedical research community. The CEGS program is also a centerpiece for NHGRI's genomics training programs in general, and more specifically, for training of scientists from communities of underrepresented minorities.

In 2001, NHGRI awarded the first three CEGS grants. Two separate awards went to the University of Washington; the third was to Yale University. The response to this program from the research community has been very positive, and NHGRI anticipates that community interest will remain high. We therefore expect to receive many high quality applications for funding in FY 2003. These projects are pushing the cutting edge of genomic science and will engage the talents of many highly skilled scientists and engineers.

Item

Clearinghouse for rare and genetic disorders S Approximately 6,000 of the 7,000 genetic disorders are rare disorders. Unfortunately, there is no centralized clearinghouse that can provide information to health professionals and the public about such disorders. In previous years, the Committee has encouraged the NHGRI, in collaboration with the Office of Rare Diseases (ORD), to create an information center to disseminate information, knowledge, and understanding of rare and genetic disorders. The Committee is pleased to learn that the NHGRI and the ORD plan to establish such a center this fiscal year. (p. 170)

Action taken or to be taken

The National Human Genome Research Institute (NHGRI) and the Office of Rare Diseases (ORD) have established the NHGRI/ORD Genetic and Rare Diseases Information Center in response to a Congressional directive in this area. A contract was awarded in FY 2001 and the center became operational in January 2002.

The purposes of the Information Center are to: 1) serve as a central, national repository of information materials and resources on genetic and rare diseases, conditions, and disorders; 2) collect, produce, update, and disseminate information on the diagnosis, treatment, and prevention of genetic and rare disorders; and 3) coordinate with organizations and associations interested in genetic and rare disorders to explore networking capabilities, avoid duplication of effort, and identify information gaps.

The Information Center focuses on addressing questions from the general public, including patients and their families, health care professionals, and biomedical researchers by carrying out the following main activities:

- Compile and store materials on genetic and rare diseases and disorders;
- Develop and maintain a computerized bibliographic database;
- Operate an information service with toll-free telephone service;
- Respond to information inquiries via phone, mail, and e-mail;
- Distribute materials; and
- Identify potential users of the Information Center and to provide information about the services of the Information Center to the intended audience(s).

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Authorizing Legislation

	PHS Act/ Other Citation	U.S. Code Citation	2001 Amount Authorized	2002 Estimate	2003 Amount Authorized	2003 Budget Estimate 1/
Research and Investigation	Section 301	42§241	Indefinite	$426,047,000	Indefinite	$461,902,000
National Human Genome Research Institute	Section 417B	42§285	Indefinite		Indefinite	
National Research Service Awards	Section 487(d)	42§288	2/	4,671,000	3/	4,793,000
Total, Budget Authority				430,718,000		466,695,000

1/ Reflects proposed transfer from the National Cancer Institute.
2/ Funding provided under the Departments of Labor, Health and Human Services, Education, and Related Agencies Appropriations Act, 2002 (P.L. 107-116).
3/ Reauthorizing legislation will be submitted.

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Appropriation History

Fiscal Year	Budget Estimate to Congress	House Allowance	Senate Allowance	Appropriation 1/
1994	$134,594,000	$119,030,000	$131,925,000	$127,112,000
1995 2/	152,010,000	151,878,000	151,878,000	151,518,000 3/
Rescission				(331,000)
1996	166,678,000 2/	170,041,000	163,943,000 2/	169,041,000
Rescission				(266,000)
1997	177,788,000 2/	189,267,000	180,807,000 2/	189,657,000 4/
1998	202,197,000 2/	211,772,000	218,851,000	217,704,000
1999	236,275,000 2/5/	246,111,000	249,891,000	264,892,000
Rescission				(185,000)
2000	271,536,000 2/	308,012,000	337,322,000	337,322,000
Rescission				(1,795,000)
2001	353,427,000 2/	386,410,000	385,888,000	382,384,000
Rescission				(192,000)
2002	426,739,000 2/	423,454,000	440,448,000	429,515,000
Rescission				(203,000)
2003	466,695,000			

1/ Reflects enacted supplementals, rescissions and reappropriations.
2/ Excludes funds for HIV/AIDS research activities consolidated in the NIH Office of AIDS Research.
3/ Excludes enacted administrative reductions of $ 161,000
4/ Excludes enacted administrative reductions of $ 128,000
5/ Excludes reductions of $721,000 for the budget amendment for bioterrorism.

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Detail of Full-Time Equivalent Employment (FTEs)

OFFICE/DIVISION	FY 2001 Actual	FY 2002 Estimate	FY 2003 Estimate
Office of the Director	7	7	7
Office of Administrative Management	23	23	23
Office of Policy, Planning and Communications	9	9	9
Division of Intramural Research	210	210	212
Division of Extramural Research	26	26	27
Total, NHGRI	275	275	278
Statutorily-ceiling exempt FTEs not included above	(3)	(1)	(1)

Funds to support these FTEs are provided by Cooperative Research and Development

FISCAL YEAR	Average GM/GS Grade
1999	10.9
2000	10.8
2001	10.7
2002	10.7
2003	10.7

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
Detail of Positions

GRADE	FY 2001 Actual	FY 2002 Estimate	FY 2003 Estimate
ES-6	0	0	0
ES-5	1	1	1
ES-4	0	0	0
ES-3	0	0	0
ES-2	0	0	0
ES-1	0	0	0
Subtotal	**1**	**1**	**1**
Total - ES Salary	$133,700	$138,200	$141,793
GM/GS-15	19	19	19
GM/GS-14	12	12	12
GM/GS-13	25	25	25
GS-12	38	38	38
GS-11	24	24	24
GS-10	1	1	1
GS-9	28	30	33
GS-8	14	14	14
GS-7	14	12	12
GS-6	7	7	7
GS-5	8	8	8
GS-4	2	2	2
GS-3	1	1	1
GS-2	0	0	0
GS-1	0	0	0
Subtotal	**193**	**193**	**196**
Grades established by Act of July 1, 1944 (42 U.S.C. 207):			
Assistant Surgeon General			
Director Grade	1	1	1
Senior Grade	2	2	2
Full Grade			
Senior Assistant Grade			
Subtotal	3	3	3
Ungraded	94	97	98
Total permanent positions	162	171	174
Total positions, end of year	304	304	308
Total full-time equivalent (FTE) employment,end of year	275	275	278

NATIONAL INSTITUTES OF HEALTH

National Human Genome Research Institute
New Positions Requested

	FY 2003		
	Grade	Number	Annual Salary
Biologist	GS-9	1	39,405
Biological Lab Technician	GS-9	1	39,405
Program Analyst	GS-9	1	39,405
Total Requested		3	