

# National Human Genome Research Institute

National Institutes of Health



*Advancing human health through genomics research*

# National Human Genome Research Institute

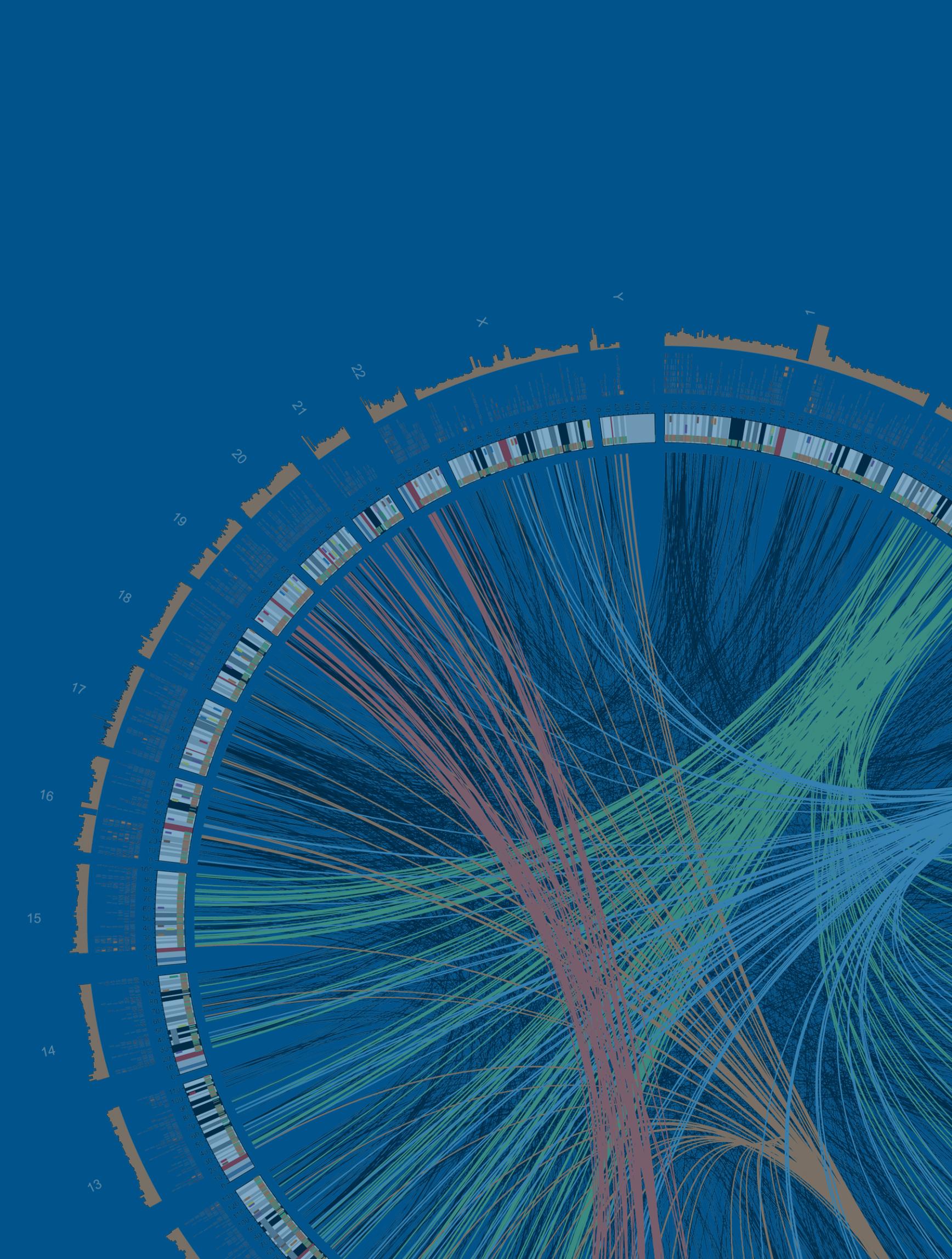


The National Human Genome Research Institute (NHGRI) is one of the 27 Institutes and Centers at the U.S. National Institutes of Health (NIH). NHGRI, an international leader in genomics research, develops resources, technologies, and policies for advancing genomics and its application to improving human health. The Institute also supports the training of investigators and the dissemination of genomic knowledge to the public and to health professionals. Additional information about NHGRI can be found at [genome.gov](http://genome.gov).

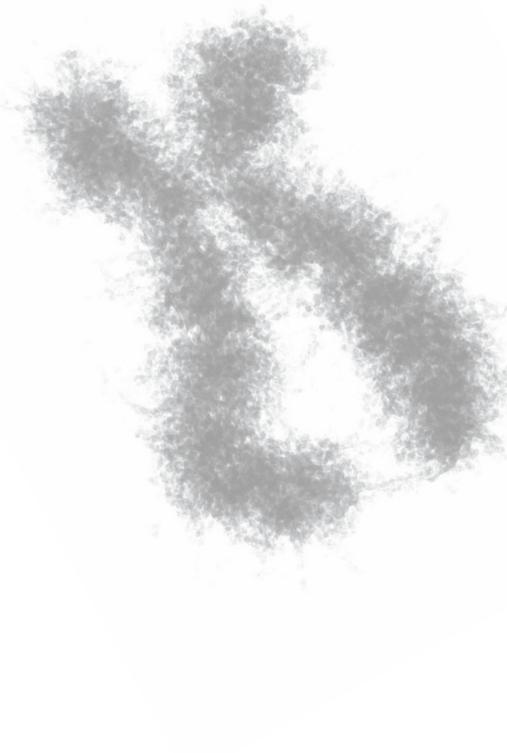
NIH, the nation's medical research agency, is a component of the U.S. Department of Health and Human Services. NIH is the primary federal agency conducting and supporting basic, translational, and clinical research, and investigates the causes, treatments, and cures for both common and rare diseases. For more information about NIH and its programs, visit [nih.gov](http://nih.gov).



1	Director's Message
2	Genomics Primer
3	Benefits of Genomics Research
4	History of Genomics
5	Overview of NHGRI's Organization
6	NHGRI Core Values
7	Priority Research Areas for NHGRI
8	Genome Structure and Function
12	Genomics and Human Disease
16	Genomic Medicine
20	Genomics and Society
24	Genomics and Data Science
25	Training in Genomics
26	Future Horizons



# Director's Message



Genomics is one of the most vibrant, compelling, and relevant scientific disciplines of the 21<sup>st</sup> century. I am proud to be leading the National Human Genome Research Institute (NHGRI), which has been a pioneer in genomics for more than a quarter century.

NHGRI was established in the late 1980s to lead NIH's efforts in the Human Genome Project, the audacious international endeavor that deciphered the order of the ~3 billion 'letters' that make up the human 'blueprint' (i.e., the human genome sequence). Completed in 2003, the Human Genome Project laid the foundation for the burgeoning field of genomics. Fast forward to today, and NHGRI is the largest organization in the world dedicated to genomics research.

Genomics is central to understanding human biology and human disease—everything starts with the genome's long strands of deoxyribonucleic acid (or DNA). The information encoded in DNA provides the basic instructions for our lives, and subtle variations in our genomes greatly influence our health, our risk for disease, and many of our features. While NHGRI receives less than two percent of NIH's total annual budget, its genomics research programs are important for many of the studies supported by other NIH Institutes and Centers whose missions are mostly focused on specific disease areas. In addition, NHGRI leads multiple research programs supported by the NIH Common Fund, a trans-NIH pool of funds used for short-term, exceptionally high-impact projects that aim to eliminate key scientific roadblocks ([commonfund.nih.gov](http://commonfund.nih.gov)).

The majority of NHGRI's funds are used to support genomics research at leading academic and commercial institutions across the United States and around the world. NHGRI is best known for funding and leading large, consortium-based programs, but it also supports the work of many individual investigators. Among the most impressive recent successes resulting from this support are spectacular advances in genomic technologies, particularly those aiming to reduce the cost of DNA sequencing.

In addition, roughly twenty percent of NHGRI's funds are used to conduct research in the Institute's laboratories in and around Bethesda, Maryland. These efforts are led by NHGRI's intramural investigators, who capitalize on the unique strengths of the broader NIH Intramural Research Program to pursue a wide range of genomics studies, from addressing very basic questions about genome structure and function to developing approaches for using genomic information in clinical research and medical care.

At NHGRI, we believe that studying the broader societal implications of genomics and genomic advances is a critical component of our research program. From the beginning, the Institute has dedicated a fixed portion of its budget to study the ethical, legal, and social implications of genomics research. In addition, NHGRI staff regularly engage in active dialogue with many societal audiences, including educational institutions, community organizations, healthcare professionals, and the general public. To facilitate these interactions, we use many different communication tools—including our widely respected website ([genome.gov](http://genome.gov)), our GenomeTV channel on YouTube, and other social media sites (such as Facebook and Twitter).

The impact of genomics over the past quarter century has been remarkable. Going forward, NHGRI remains 'laser focused' on helping to fulfill the promise of the Human Genome Project, which is largely synonymous with our fundamental mission—advancing human health through genomics research. I invite you to learn more about our research portfolio and our many associated programs by reading the following pages and visiting our website at [genome.gov](http://genome.gov).

Eric D. Green, M.D., Ph.D.  
Director, National Human Genome Research Institute

# Genomics Primer

Some basics about the human genome:

- To build a house or a car, you need a blueprint—a detailed parts list with assembly instructions. For all living creatures, the genome functions as that blueprint.
- Our blueprint—the human genome—resides within tiny double-stranded fibers of deoxyribonucleic acid (DNA) that are packed into chromosomes within the nucleus of cells.

## **ge·nome**

noun [jē' nōm']: all of the DNA in a cell

## **ge·no·mics**

noun [jē nō' miks]: the study of genomes

- The typical human cell contains 23 pairs of chromosomes, with one set of 23 chromosomes coming from each parent.
- Males and females have two copies of 22 of these chromosomes (called autosomes); for the 23<sup>rd</sup> pair, known as the sex chromosomes, the two sexes differ—females have two X chromosomes and males have one X chromosome and one Y chromosome.
- DNA is made of four different building blocks, each containing one of the following chemicals: adenine (**A**), thymine (**T**), cytosine (**C**), and guanine (**G**).
- The order of the **A**, **T**, **C**, and **G** building blocks ('letters') in the human genome encodes the biological instructions that tell each cell what to do and when to do it.
- Each set of 23 chromosomes (essentially one copy of the human genome) contains ~3 billion 'letters' (so the typical human cell contains ~6 billion genomic 'letters' in total).
- The genome of each person is slightly different from the genome of every other person (i.e., it is unique). For example, when the sequences (the order of 'letters') of any two people's genomes are compared, they have a different 'letter' roughly once every thousand positions.
- The great majority of genomic differences (or variants) are inconsequential, but some influence our physical traits and health.



# Benefits of Genomics Research

*Genomics is an engine for both scientific and economic growth*

The scientific leaders who proposed the Human Genome Project in the 1980s envisioned numerous benefits from the multi-billion dollar investment: (1) mapping and sequencing the human genome would increase basic understanding about how the genome works and, therefore, how cells work; (2) genomic knowledge would accelerate medical research, yielding fundamental insights about inherited diseases and disorders like cancer; and (3) technological advances coupled with new genomics-enabled scientific opportunities would stimulate the biotechnology industry (then less than a decade old), helping to propel the world economy into the next century. The predictions of these scientific visionaries were remarkably on target. Since that time, NHGRI has played a pivotal role in realizing the scientific and economic benefits brought about by genomics research, mostly in the biomedical research arena. At the same time, genomic advances are also benefiting many other important areas (e.g., agriculture, livestock, energy, and forensics).

## Increasing basic understanding of the genome

Genomics research has expanded our grasp of the interplay between genome structure and function and has clarified many misconceptions. As the Human Genome Project got under way, many scientists estimated that humans—being the complex creatures that we are—carry 100,000 or more genes in our genomes. Today, we know that number is much lower, more like ~20,000 genes. We have also learned that those vast stretches of DNA that do not code for proteins, previously called “junk DNA” by some, are far from useless. In fact, studies over the last decade have revealed the presence of hundreds of thousands of functional elements within the large portions of the human genome that do not directly code for proteins. Many of these latter elements play central roles in controlling the activity of our genes.



## Accelerating medical research

Genomics has become a central discipline of biomedical research, quickly spreading across the entire research landscape. Virtually all other NIH Institutes and Centers, as well as many other private and public institutions in the United States and around the world, have made major investments in genomics. NHGRI and other components of NIH have launched important partnerships to use genomics to study areas of longstanding interest. These have included The Cancer Genome Atlas (TCGA), a joint endeavor between NHGRI and the National Cancer Institute to investigate the genomics of cancer, and a program exploring the implications, opportunities, and challenges of using genome sequence information in the newborn period, which is being pursued collaboratively between NHGRI and the Eunice Kennedy Shriver National Institute of Child Health and Human Development. NHGRI also supports multiple efforts to enhance the use of genomic medicine to improve patient care.

## Stimulating the biotechnology industry

Since the completion of the Human Genome Project, the largest driver of genomic advances has been the stunning progress in developing more powerful technologies for sequencing DNA. Catalyzed by an NHGRI program in technology development coupled with significant investments by the private sector, the costs of sequencing DNA have plummeted at a pace far exceeding Moore's Law (the well-known observation that computing power doubles roughly every 24 months). Along with cost reductions, the speed of sequencing genomes has increased substantially.

These technological developments, in conjunction with other genomic advances, have been a boon for the economy. A 2013 analysis by the Battelle Technology Partnership Practice determined that from 1988 to 2012, the U.S. government invested \$11.3 billion (\$14.5 billion in 2012 dollars) in the Human Genome Project and related areas of genomics research. This investment generated directly or indirectly nearly \$1 trillion in economic activity, more than 4.3 million job-years of supported employment, and \$54.8 billion in tax revenues from genomics research, development, and commercial activities. The 2013 report also noted that for every \$1 invested in genomics by the U.S. government, there was a return on investment of ~\$65 for the U.S. economy. While there are different ways to assess the specific economic output of the federal investment in genomics, it is clear that the cumulative yield has been substantial.

# History of Genomics

*Genomics is a young discipline built on the work of early geneticists and, later, molecular biologists*

In the 1860s, Gregor Mendel, an Austrian monk, studied pea plants to decipher patterns of inheritance, but he had no idea what carried traits from one generation to the next. The young Swiss physician Friedrich Miescher discovered DNA in the 1870s; he named it “nuclein” since it existed in the cell’s nucleus. Decades later, in the 1940s, Oswald Avery and colleagues proved that DNA is the molecule containing life’s inherited information.

In 1953, James Watson and Francis Crick described the double-helical structure of DNA, a finding deduced from the X-ray diffraction images of DNA generated by Rosalind Franklin. This key insight provided the definitive piece of the puzzle about how DNA serves as the molecule of heredity, carrying genetic information from one cell to the next and one generation to the next. The 1960s then brought key insights about DNA structure and function. This included elucidation of the genetic code—the fundamental rules about how DNA’s nucleotides (represented by As, Ts, Cs, and Gs) encode the instructions for making proteins—and development of a more refined view of the units of DNA responsible for encoding proteins (i.e., genes).

A detailed understanding of the workings of genes remained beyond reach, however, until the molecular biology revolution of the 1970s and 1980s, an era that brought powerful new tools for studying and manipulating DNA. Among the critical advances of that time were the contributions of Fred Sanger and Walter Gilbert, who independently developed the first techniques for determining the sequence of nucleotides in DNA.

Improvements to the early methods for isolating, analyzing, and sequencing DNA led to the notion of comprehensively studying an organism’s DNA (i.e., its genome), leading to the launching of the Human Genome Project in 1990. This remarkable international endeavor galvanized interest in genomics among scientists and the public alike.

Within 10 years, the first draft of the human genome sequence was generated, and in April 2003, leaders of the Human Genome Project revealed a finished sequence of the ~3-billion-letter human genome. The Human Genome Project was complete. In many ways, that historic milestone marked the ‘starting line’ for what has since transpired, as knowledge about genome function and the genome’s role in health and disease has soared.



# Overview of NHGRI's Organization

Seven divisions serve to align the Institute's structure with its mission

NHGRI leads the field of genomics by strategically investing in highly innovative studies, technologies, and data resources needed to understand how the genome works and to use that knowledge for advancing human health. The Institute was established as a Center at the National Institutes of Health (NIH) in 1989 to carry out NIH's role in the Human Genome Project, then was elevated to an Institute and renamed the National Human Genome Research Institute, or NHGRI, in 1997. The Institute funds a wide range of research on the structure and function of the human genome and the genomes of other animals, on the genome's role in health and disease, and on the ethical, legal, and social implications of genomics research. NHGRI also supports the training of the next generation of genomics investigators and the dissemination of genomic information to the public and to healthcare professionals.

## National Human Genome Research Institute



	Strategic Plan: <i>New Goals for the U.S. Human Genome Project</i>	Strategic Plan: <i>A Vision for the Future of Genomics Research</i>	Eric Green appointed NHGRI Director	NHGRI reorganizes to accommodate expanding research mission	President Obama announces the U.S. Precision Medicine Initiative
	<b>1998</b>	<b>2003</b>	<b>2009</b>	<b>2012</b>	<b>2015</b>
<b>1997</b>	<b>2003</b>	<b>2008</b>	<b>2011</b>	<b>2013</b>	
NCHGR becomes National Human Genome Research Institute (NHGRI)	Human Genome Project completed	Genetic Information Nondiscrimination Act (GINA) becomes U.S. law	Strategic Plan: <i>Charting a Course for Genomic Medicine from Base Pairs to Bedside</i>	NHGRI-Smithsonian exhibition <i>Genome: Unlocking Life's Code</i> opens	

# NHGRI Core Values

Successful organizations are guided by core values that shape how they pursue their missions and influence their cultures and people. Some of NHGRI's core values are common to research organizations, while others are less typical.

## **Improving human health**

*Genomics offers tremendous potential for deciphering the molecular bases of human disease and, eventually, improving the health of people around the world and reducing health disparities.*

## **Elucidating a basic understanding of biology**

*Genomic approaches are proving to be powerful for gaining fundamental new insights about biological systems.*

## **Translating basic knowledge into medical advances**

*The translation of basic genomic knowledge into tools and approaches for improving the diagnosis, treatment, and prevention of human disease is a high priority.*

## **Fostering interdisciplinary and collaborative research**

*Genomic studies prosper when they are pursued in a highly collaborative fashion by researchers from multiple disciplines.*

## **Embracing consortia-based science**

*Highly coordinated and carefully managed research consortia have proven to be extremely effective and productive for pursuing certain large genomics projects.*

## **Maximizing data sharing**

*Widespread sharing of research data is a basic tenet for all genomics (and, increasingly, all biomedical) research, with attention to appropriate protections for the human subjects participating in the research.*

## **Investing in technology development**

*All major successes in genomics have been driven by the development of new technologies; while high-risk in nature, investment in innovative genomic technology development is critical.*

## **Addressing the societal implications of genomics research**

*The nature of genomics demands attentiveness to the larger implications of research advances for individuals, communities, and societies; relevant issues relate to providing equitable access to genomic technologies and the outcomes of genomics research, informing and protecting research participants, and educating healthcare professionals and the public about this rapidly emerging area of science.*

# Priority Research Areas for NHGRI

Realizing the health benefits of genomics will take decades of dedicated work and a global effort that builds on past findings and expands as new productive avenues emerge. NHGRI has a long-term commitment to this promising enterprise through the following four priority research areas.

## Genome Structure and Function



### Understanding how the human genome works

The Human Genome Project deciphered the order of the 'letters' in the human genome sequence, but it will take decades of research to fully reveal all the complexities of how the genome functions. These efforts involve using laboratory and computational approaches to assemble inventories of functional elements in the human genome, to establish the choreography by which these elements confer biological function, and to catalog the differences among people's genomes. Key to these advances is the ongoing development of new technologies and approaches for studying genome structure and for elucidating genome function.

## Genomics and Human Disease



### Establishing the role of genomic variants in health and disease

Diseases are a consequence of a complex choreography of influences from our genomes and our environmental and social exposures. Rare diseases typically result from the presence of genomic variants (mutations) in a single gene, with environmental and social influences playing a lesser role. Common diseases typically result from the presence of multiple risk-conferring genomic variants in conjunction with environmental and social influences. Large-scale genomic studies can establish the role that genomic variants play in rare and common diseases, in the response to medications, and in the preservation of health.

## Genomic Medicine



### Using genomic information to advance medical care and human health

Genomic medicine is an emerging medical discipline that involves using an individual's genomic information as part of his or her clinical care (e.g., for diagnostic or therapeutic decision-making) and the other implications of that clinical use. A foundation for the systematic implementation of genomic medicine is being built by research programs that are establishing tools, resources, and a knowledge base to empower healthcare professionals to capitalize on genomic information in the delivery of clinical care. Helping patients, their families, and their friends understand the role that genomics will play in making healthcare decisions is also vital.

## Genomics and Society



### Addressing the societal impact of genomic advances

Studying the ethical, legal, and social implications of genomics research has been a cornerstone of the field since its inception. As new technologies increase our capability to generate genomic information and research increases our understanding of what that information might mean, society needs to determine how to use the technologies and information responsibly. Genomics is the study of commonality and of differences, and ensuring that research moves forward in ways that remain mindful of the implications of the generated knowledge is fundamental to the ultimate success of the enterprise.

# Genome Structure and Function



# Understanding How the Human Genome Works

## *Establishing how genomes confer biological information*

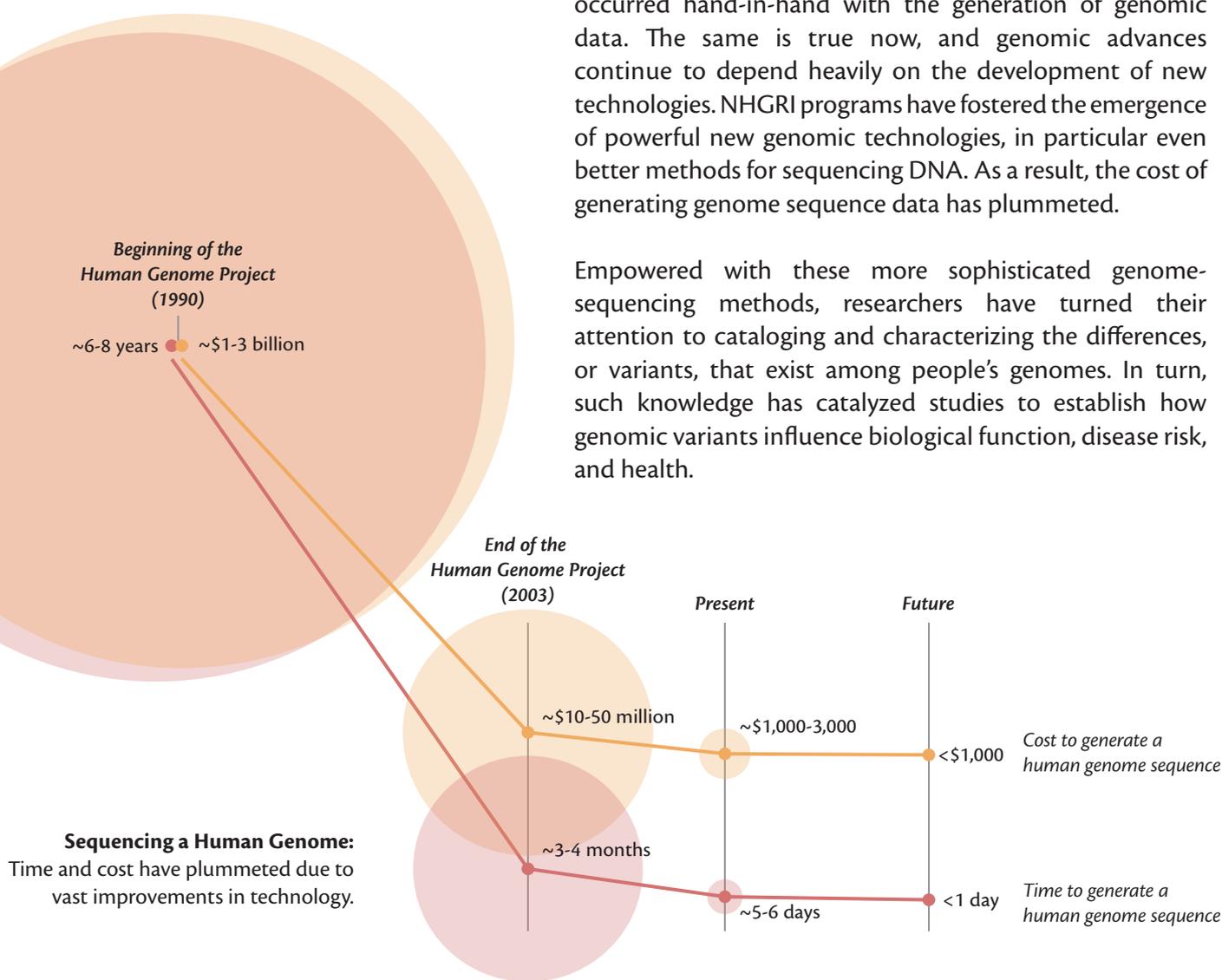
After completing the Human Genome Project in 2003 and with a generated human genome sequence in hand, genomic researchers immediately turned their attention to understanding what that sequence means in order to acquire a deeper understanding of genome function and to aid studies to uncover the genomic bases of human health and disease.

Embedded in the billions of As, Ts, Cs, and Gs across the human genome is a fundamental, yet complicated, code for human biology. Elucidating that code requires novel approaches for performing laboratory and computational studies. Part of that work involves understanding the

genomic differences and similarities among a vast array of organisms—an area known as comparative genomics. Overall, researchers have made substantial progress in identifying and characterizing the thousands and thousands of functional elements in the human genome. While a complete understanding of human genome function will take decades to achieve, the information collected to date is already providing critical insights for scientists and clinicians studying genomic contributions to human health and disease.

One key to the phenomenal success of the Human Genome Project was the dedicated technological innovations that occurred hand-in-hand with the generation of genomic data. The same is true now, and genomic advances continue to depend heavily on the development of new technologies. NHGRI programs have fostered the emergence of powerful new genomic technologies, in particular even better methods for sequencing DNA. As a result, the cost of generating genome sequence data has plummeted.

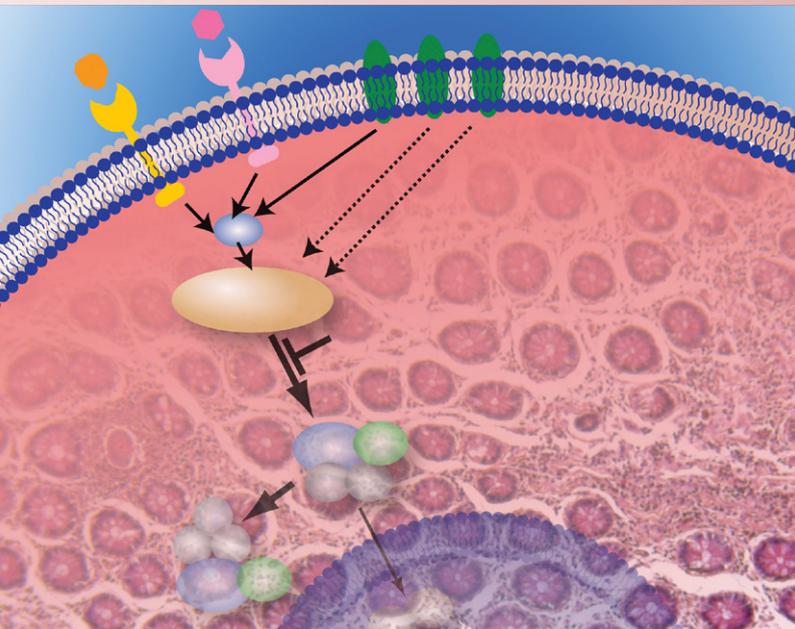
Empowered with these more sophisticated genome-sequencing methods, researchers have turned their attention to cataloging and characterizing the differences, or variants, that exist among people's genomes. In turn, such knowledge has catalyzed studies to establish how genomic variants influence biological function, disease risk, and health.





## Cataloging Genomic Variation

Near the end of the Human Genome Project, several systematic efforts began to catalog genomic variants (i.e., genome sequence differences) in the human population. Two early efforts were the Single-Nucleotide Polymorphism (SNP) Consortium and the International HapMap Project. More recently, the 1,000 Genomes Project sequenced the genomes of ~2,500 individuals from 26 geographically dispersed populations, yielding information about close to 100 million genomic variants in the human population. This effort also provided more complete views of the content and pattern of variants within a typical person's genome, yielding early insights into how each person's genome uniquely operates. In a broader context, knowledge of genomic variation is also used for searching for genes that play a role in human disease, for performing forensic DNA analysis, for establishing more refined views of our near and distant ancestors, and for revealing important aspects of human population history.

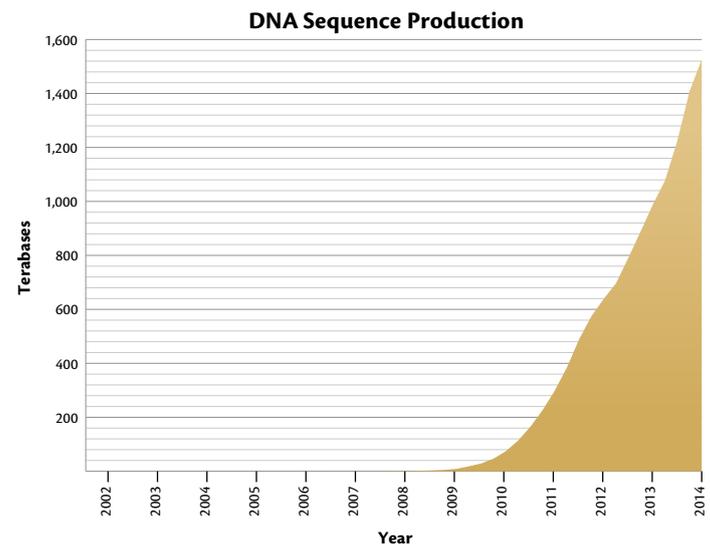


## Understanding How Variation Affects Function

Compiling catalogs of genomic variants represents an initial step in understanding how variation affects biological function. Some variants are inconsequential, some confer risk for disease, and some are beneficial, offering protection against disease or making someone a good candidate for a particular medication. The NIH Common Fund's Genotype-Tissue Expression (GTEx) project, led by NHGRI, aims to increase understanding about how genomic variants contribute to human disease. Specifically, GTEx is establishing how some genomic variants influence gene expression. A complementary NHGRI program is developing highly innovative computational approaches for characterizing variants in non-protein-coding regions of the human genome that play a role in diseases or other traits.

## Performing Genomics Research at Scale

NHGRI research programs provide ever-growing capacities for genome sequencing and genome analyses—at increasing speeds and decreasing costs. Such efforts have made seminal contributions to the understanding of genome structure and evolution, for example by sequencing the genomes of hundreds of organisms and thousands of humans. Such efforts dovetailed with large programs to identify functional genome sequences and to catalog human genomic variation. The key characteristic of all of these programs has been scale—each being capable of generating prodigious amounts of genomic data and doing so in a way that is much more complete, rapid, and cost-effective than cumulative small-scale efforts. An inherent (but exciting) consequence of such scale is the challenge of then handling and analyzing the resulting massive genomic datasets.



# Genomics and Human Disease



# Establishing the Role of Genomic Variants in Health and Disease

## *Determining the genomic bases of rare and common diseases*

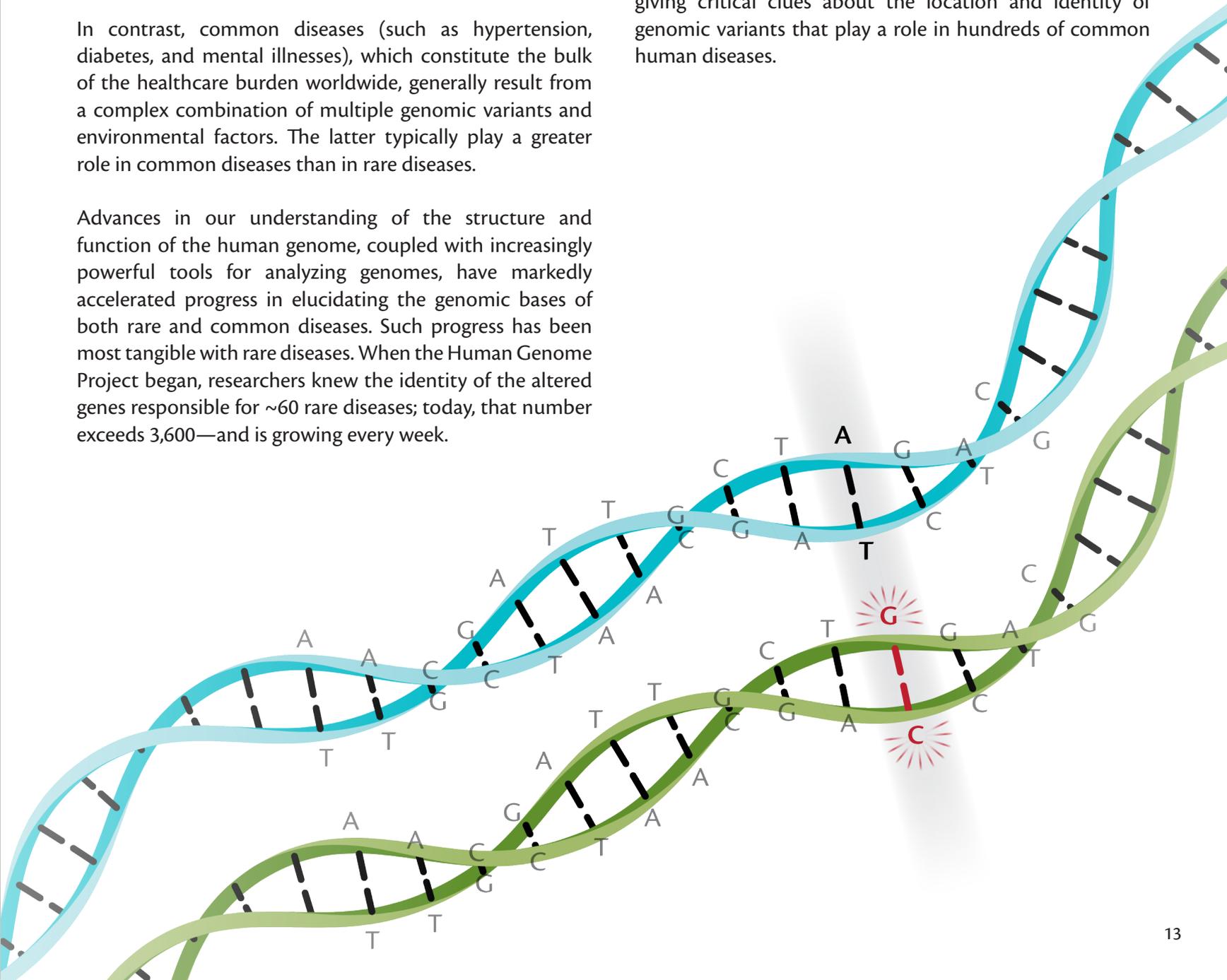
Genomic variants play a key role in essentially all human diseases. But the nature of that role is generally different for rare diseases versus common diseases.

Rare diseases (such as sickle cell anemia, cystic fibrosis, and Huntington's disease) are most often caused by genomic mutations that disrupt a single gene. While other genomic variants and environmental factors may influence the severity of the disease, the fundamental cause is an altered single gene.

In contrast, common diseases (such as hypertension, diabetes, and mental illnesses), which constitute the bulk of the healthcare burden worldwide, generally result from a complex combination of multiple genomic variants and environmental factors. The latter typically play a greater role in common diseases than in rare diseases.

Advances in our understanding of the structure and function of the human genome, coupled with increasingly powerful tools for analyzing genomes, have markedly accelerated progress in elucidating the genomic bases of both rare and common diseases. Such progress has been most tangible with rare diseases. When the Human Genome Project began, researchers knew the identity of the altered genes responsible for ~60 rare diseases; today, that number exceeds 3,600—and is growing every week.

Studying common diseases is complex, but has been enormously aided by the advent of a detective-like strategy called a genome-wide association study (or GWAS), which involves scanning the genomes of thousands of affected and unaffected individuals in search of particular regions that harbor variants conferring risk for a common disease. The first successful genome-wide association study, published in 2005, revealed the genomic cause of a type of blindness called age-related macular degeneration; since then, more than 2,000 such studies have been reported, collectively giving critical clues about the location and identity of genomic variants that play a role in hundreds of common human diseases.



# Genomics and Human Disease

## Rare Diseases

An estimated 25-30 million Americans suffer from some type of rare disease (generally considered to be a disorder that affects fewer than 200,000 people). Of the roughly 7,000-8,000 rare diseases (also referred to as Mendelian disorders) described to date, researchers have determined the genomic bases for more than 3,600. The NHGRI Centers for Mendelian Genomics program aims to discover the underlying genomic causes of the remaining rare diseases using powerful new methods for genome sequencing and data analysis. These centers are scaling approaches for disease-gene discovery and developing collaborative partnerships with investigators around the world involved in studying and caring for patients with rare diseases.

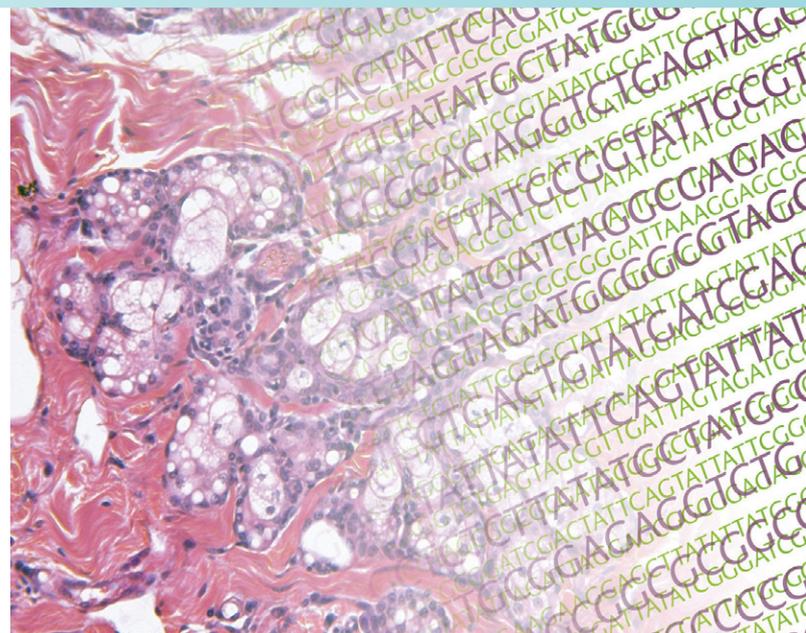


## Common Diseases

Tackling the monumental challenges associated with elucidating the genomic bases of common diseases requires cutting-edge approaches and sheer scale for genome analyses. The NHGRI Centers for Common Disease Genomics program is using these approaches to pursue very large studies of common diseases (including Alzheimer's disease, diabetes, autism, and cardiovascular disease). The centers provide particular expertise in the development of novel study designs, the refinement of new genome-sequencing methods for use at large scale, and the increasingly challenging bioinformatics and data-integration components of such studies.

## The Genomics of Cancer

Cancer is a disease of the genome, a direct consequence of genomic changes that cause cells to grow in an uncontrolled manner. The Cancer Genome Atlas (TCGA)—a joint venture of the National Cancer Institute and NHGRI—has cataloged the genomic alterations found in more than 30 cancer types, facilitating studies of how those changes may play a role in disease. TCGA has used the newest genomics methods (in particular, genome sequencing) and data-analysis approaches to rigorously interrogate the genomes of more than 10,000 cancer specimens. TCGA has changed the face of cancer research in many ways, defining new research directions and accelerating the development of new cancer diagnostics and therapeutics. The resulting information has allowed specific cancer types (e.g., breast cancer) to be more accurately diagnosed as a collection of distinct diseases, with each major subtype requiring different clinical management.





## Genome Sequencing at the NIH Clinical Center

The NIH Clinical Center is America's research hospital—and the largest hospital in the world dedicated to clinical research. It has become a valuable venue for major clinical genome-sequencing programs. For example, the ClinSeq® Project is investigating the use of genome sequencing as a tool for clinical research and clinical care. By establishing a ~1,500-adult cohort, evaluating their health and disease status, and conducting detailed genomic analyses, ClinSeq® is studying the genomic architecture of disease, the clinical implementation of genomic technologies, and the disclosure of genomic information to patients. Another initiative, called the Clinical Center Genomics Opportunity, aims to generate genome-sequence data for >1,000 patients being studied by investigators across the NIH, thereby helping to jump-start genomic medicine implementation more broadly throughout the NIH Clinical Center.

## Global Genomics

Understanding that the benefits of genomic advances must be realized throughout the world, NHGRI aims to stimulate genomics research on a global level. For example, Human Heredity and Health in Africa (H3Africa) is an NIH Common Fund program, led by NHGRI, that aims to enhance the use of genomic approaches to study the genetic and environmental determinants of disease in Africa. The program provides direct support to studies led by African scientists that use genomic, clinical, and epidemiologic methods to identify hereditary and environmental contributions to the risk of common, non-communicable disorders (such as heart and kidney disease) as well as communicable diseases (such as tuberculosis). H3Africa is enabling African scientists to use cutting-edge genomic and genetic approaches for disease studies, empowering African researchers to make important genomic discoveries, and establishing effective collaborations among African researchers on the African continent.



## Electronic Medical Records and Genomics Research

Research that involves the integrative analysis of detailed clinical data and genomic information offers great promise for the study of human disease and the realization of genomic medicine. The Electronic Medical Records and Genomics (eMERGE) Network is developing best practices for using the electronic medical record (EMR) as a tool for genomics research. eMERGE investigators are also pursuing the discovery of genomic variants associated with clinical conditions identified using EMRs, studying the ethical, legal, and social issues involved in the use of EMRs for genomics research, and working to incorporate actionable (i.e., relevant) genomic risk factors into EMRs for use in clinical care.

# Genomic Medicine



# Using Genomic Information to Advance Medical Care and Human Health

*Applying genomics to clinical care and improving the effectiveness of healthcare*

Early examples of the exciting potential of genomic medicine have already emerged. By using genomic technologies and findings, clinical care can be tailored to an individual's unique, genomically influenced predispositions to disease and disability. Already a number of genomic applications have been deployed into routine clinical practice, such as the use of specific tumor mutations for guiding cancer treatment and genomic-based drug hypersensitivity testing prior to treatment for HIV/AIDS. These examples provide added enthusiasm for the additional research needed to capitalize on numerous genomic discoveries and to enable the realization of genomic medicine's full potential.

NHGRI is investing in a number of research efforts to establish a robust foundation for genomic medicine. Studies are being performed to evaluate the impact of using genomic information for clinical care, operationalize genome analyses as a diagnostic tool, and assess the generalizability of genomic findings across populations and clinical settings. In fact, these nascent NHGRI genomic medicine programs are proving to be prototypes for broader efforts that aim to individualize medical care to advance human health, such as the recently launched Precision Medicine Initiative. In this

way, NHGRI is not alone; other NIH Institutes and Centers as well as other government agencies, funding organizations, not-for-profit groups, and the private sector are actively working to advance genomic medicine.

The journey to improved medical care and human health through effective application of genomics requires other important steps as well. Healthcare professionals need easy, accessible tools and systems, so that they can use genomic information for clinical decision-making; they also require training and education to learn how to use these new tools effectively. Patients need a basic understanding of genomics to communicate effectively with their healthcare providers and their own families, so that they can make key decisions about their medical care and other important life choices. Finally, the broader medical system needs to adapt to the ongoing genomic advances, so that clinical benefits are made rapidly and fairly available to all people.



# Genomic Medicine



## Rare and Undiagnosed Diseases

Undiagnosed diseases are conditions that even skilled physicians struggle to diagnose despite extensive clinical investigation. The disease may be one that is rarely seen, has never been described, or is a rare form of a more common disorder. Newly developed methods for genome sequencing now provide powerful approaches for deciphering the causes of such very rare, undiagnosed conditions. The Undiagnosed Diseases Network (UDN) is an NIH Common Fund program, led by NHGRI, involving a collaborative group of medical centers that systematically evaluates patients with serious conditions that have been undiagnosed despite intensive, sometimes years-long clinical examinations. With an emphasis on cutting-edge genomic studies combined with rigorous clinical evaluations, the program aims to identify best practices for the diagnosis and management of rare and new diseases and to improve the lives of patients living with undiagnosed conditions.

## Genome Sequencing of Newborns

Genomic technologies have advanced so dramatically over the past decade that the value of incorporating genomic information into medical care early in life (e.g., during the newborn period) can now be seriously considered, especially in light of the existing newborn genetic screening programs that have been in place for decades. The Newborn Sequencing in Genomic Medicine and Public Health (NSIGHT) program is a joint endeavor between NHGRI and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development that is conducting pilot research projects to investigate the implications, challenges, and opportunities associated with the clinical use of genomic information in the newborn period. This program is exploring, in a limited but deliberate manner, opportunities to use genomic information for broadening the identification and understanding of diseases that occur in the newborn period, and is also examining the legal and ethical issues surrounding such use.



## Test-Driving Genomic Medicine

Recognizing the opportunities for the clinical implementation of genomics in the coming decade, NHGRI is supporting studies that are essentially taking genomic medicine out for a 'test drive.' The Clinical Sequencing Exploratory Research (CSER) program supports a number of projects that collectively aim to develop technical specifications and standards for genome sequencing in the clinical setting, investigate methods for transmitting genomic data to healthcare professionals in a way that fits into the normal clinical workflow, explore regulatory requirements for the clinical applications of genomics, and study the ethical implications of returning unexpected genomic information to patients. Capitalizing on the lessons learned from these and other studies, the Implementing Genomics in Practice (IGNITE) network of researchers aims to examine and refine best practices for the effective implementation, diffusion, and sustainability of genomic medicine in diverse clinical settings; the information generated from the program will contribute to the growing knowledge base for the use of genomic information in patient care.

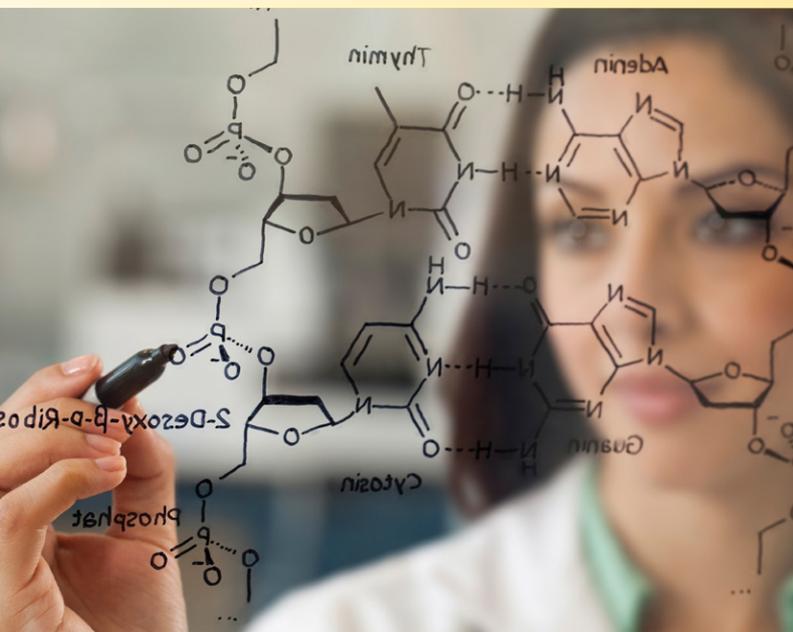
## Readying Healthcare Professionals for Genomics

To effectively move genomics into routine medical practice, clinicians need to be able to interpret genomic data and to make medical decisions based on reliable evidence. NHGRI is supporting several endeavors intended to ready front-line healthcare professionals for genomic medicine. The goal of the Clinical Genome Resource (ClinGen) is to establish a knowledge base of carefully curated information about genomic variants and their clinical significance. To enhance genomics education for healthcare professionals, NHGRI is developing online resources—such as the Genetics/Genomics Competency Center (G2C2)—that provide self-directed genetics and genomics learning materials for genetic counselors, nurses, physician assistants, pharmacists, and physicians. NHGRI also co-leads the Inter-Society Coordinating Committee on Practitioner Education in Genomics to promote the sharing and dissemination of genomic medicine educational materials among the numerous health professional societies that are the authoritative sources of information in their respective fields.



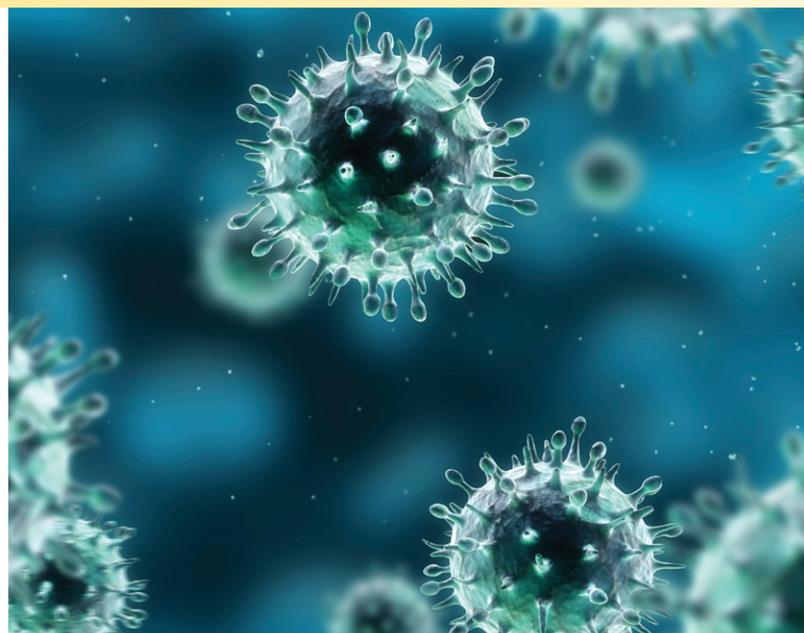
## Pharmacogenomics

Genomic variants play an important role in how our bodies metabolize drugs, which is why a drug may work for one individual but not for another or may have devastating side effects in only a small subgroup of patients. NHGRI is supporting several studies investigating how to use genomic information effectively to tailor decisions about the selection and dosing of drug therapies, an area known as pharmacogenomics. For example, the Electronic Medical Records and Genomics Pharmacogenomics (eMERGE-PGx) project is identifying gene variants involved in the response to commonly used drugs in nearly 9,000 individuals and testing how best to integrate that information into their electronic medical records for use in their clinical care. NHGRI is supporting this and similar programs (e.g., the IGNITE network) that share methods and tools so as to develop best practices for using genomic information to improve drug safety and effectiveness. Computerizing and streamlining such approaches, and making them work across diverse clinical settings and providers, will enable genomic medicine to be implemented broadly and simultaneously tailored to individual patients.



## Changing Face of Infectious Disease Diagnostics

The advent of faster and cheaper DNA sequencing methods has created exciting new opportunities in clinical microbiology, especially related to public health and hospital epidemiology. The ability to very rapidly sequence microbial DNA from clinical samples is transforming infectious disease diagnostics and bringing systematic genomic surveillance of infectious agents into the mainstream. Examples are plentiful: elucidating how infections spread among patients and healthcare workers, identifying new and unexpected infection-causing microbes, and discovering new genes and modes by which microbes develop antibiotic resistance. Today, epidemiologists can much more rapidly trace the origins of food-borne illnesses and compare the genomes of infection-causing microbes from different patients. Finally, genome sequencing is being used for biodefense and public health purposes, such as identifying the genomic features of bacterial and viral strains that make them more infectious and tracing their origins in a detective-like fashion. Overall, genomic technologies are changing many traditional practices in clinical microbiology, providing a more rigorous means for detecting and monitoring infectious agents.



# Genomics and Society





# Genomics and Society

## Transdisciplinary ELSI Studies

Many of the studies funded by the NHGRI ELSI Research Program involve investigators from multiple disciplines. The transdisciplinary nature of ELSI research allows highly innovative concepts, methods, and analyses to be focused on problems of broad relevance to genomic researchers and to society. The NHGRI Centers of Excellence in ELSI Research (CEER) program is a prototype effort to foster transdisciplinary research and training to facilitate the translation of ELSI findings to research, healthcare, and public policies. These centers promote intensive and sustainable interactions among basic genomic and genetic researchers, clinical and social scientists, and scholars from law, bioethics, and the humanities. CEER-based projects are investigating a range of multifaceted issues, including those related to clinical applications of genomics. CEER groups are also involved in training the next generation of genomics and society researchers.



## Informing Public Policy

NHGRI has helped inform policies surrounding human subjects research, the incorporation of genomics into healthcare, and the sharing of genomic data in a manner that enables scientific discovery while protecting research participant privacy. NHGRI played a key role in the development and eventual passage of the Genetic Information Nondiscrimination Act (GINA) of 2008. The Institute also helped to shape policies on the patenting and licensing of intellectual property based on genomic discoveries. NHGRI collaborates with other federal partners, including the Federal Trade Commission (to examine advertising for products claiming a genomic basis for their use) and the Food and Drug Administration and the Centers for Medicare and Medicaid Services (to inform the oversight of genomic tests and targeted genomic-based therapies and how they are reimbursed by health insurers). The Institute also engages with private-sector and other non-governmental stakeholders (e.g., biotechnology companies and health advocacy organizations) to pursue common interests, promote discussions about important genomics issues, and identify potential partnership opportunities.

## Engaging Students in Genomics

National DNA Day, held each April, commemorates the reporting of DNA's double-helical structure by James Watson and Francis Crick in 1953 and the successful completion of the Human Genome Project in 2003. This annual celebration offers students, teachers, and the public exciting opportunities to learn about the latest advances in genomics research and explore what these advances may mean for their lives and, possibly, future careers. Each year, NHGRI partners with professional societies and advocacy groups to promote DNA Day, and supports DNA Day activities at schools, museums, libraries, and community centers around the country. The Institute also offers educational resources, sends NHGRI researchers as DNA Day ambassadors to speak at various events, and communicates about DNA Day via social media.



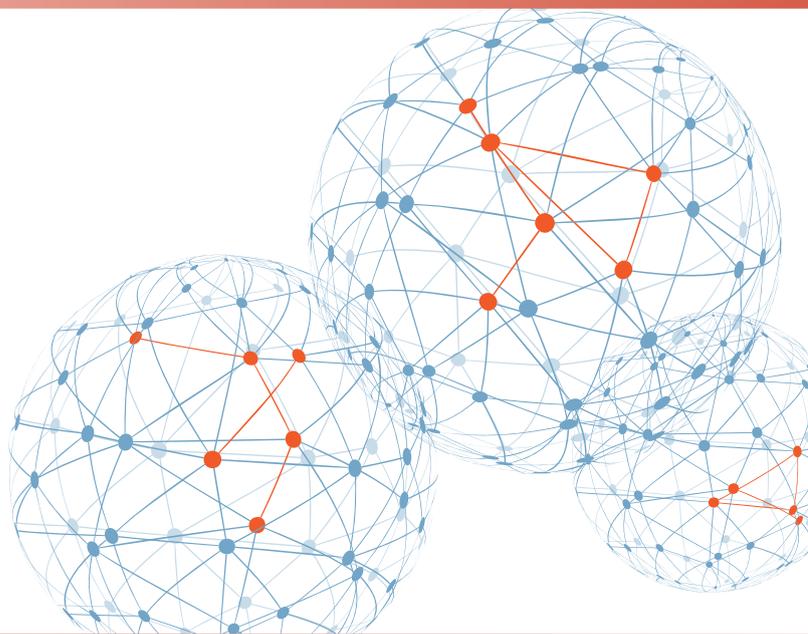


## Improving Public Genomic Literacy

NHGRI is helping to prepare the public for a future in which the use of genomic information is a routine part of medical care and other aspects of life. Through community engagement activities, workshops, and forums, Institute staff work with diverse communities to identify opportunities to enhance genomic literacy. For example, in 2011, NHGRI launched an exciting partnership with the Smithsonian Institution to engage the public and promote genomic literacy. The centerpiece of this effort is a high-tech, high-intensity exhibition *Genome: Unlocking Life's Code*, which was visited by more than 3.6 million people at the Smithsonian's National Museum of Natural History in 2013-2014, and is now traveling to several venues around North America. The exhibition shows the revolutionary nature of genomics and the impact on fields as varied as biodiversity and medicine; it also raises important issues relevant to individuals, communities, and society. Affiliated with the exhibition have been extensive public engagement programs, many of which have been recorded and made freely available on the Internet.

## Embedding ELSI Research

NHGRI supports an increasing number of ELSI research projects that are 'embedded' within larger programs. For these projects, behavioral and social scientists, legal scholars, and bioethicists work side-by-side with genomics researchers to identify and address the ethical, legal, and societal issues that arise and to anticipate downstream issues that may emerge. These ELSI research projects have become a vital component of several large NHGRI initiatives, including sequencing the genomes of newborns, incorporating genomic information into electronic medical records, and cataloging human genomic variation. Directly coupling ELSI and genomics research has proven to be mutually beneficial to all researchers and helps to ensure that the resulting work addresses the pressing genomics and societal issues of the day.



## Genomics and Health Disparities

NHGRI is committed to supporting research that increases our understanding of health disparities. This work includes studying the diseases and conditions that contribute to health disparities and examining the barriers to ensuring that genomic medicine is available to all patient populations. The Institute seeks to advance research to define relevant gene-environment interactions, including the identification of inherited susceptibility factors for disease and the development of technologies for the reliable and reproducible measurement of relevant environmental contributions. NHGRI is also committed to increasing the inclusion of research participants from diverse ancestral backgrounds in large genomic studies, so as to improve our knowledge about the role of genomic variation in health and disease across all populations.

# Genomics and Data Science

*Facilitating the assimilation, management, and analysis of large amounts of genomic data*



Modern-day genomic technologies generate massive amounts of data at lower and lower costs. In fact, genomic researchers now readily produce far more data than they can analyze in real time. For example, while a human genome can now be sequenced in a day, analyzing the resulting data to obtain a complete and meaningful interpretation remains a lengthy and laborious task. Similar situations are now regularly encountered with other research technologies.

Data science in the context of genomics (and biomedical research more broadly) involves the study of increasingly large, complex, and diverse datasets. This overlaps other related areas, such as bioinformatics, computational biology, biomedical informatics, information science, biostatistics, and quantitative biology—all of which, in turn, relate to the now-popular term ‘Big Data.’

Data science is a key component of projects across the spectrum of NHGRI-supported research. These efforts range from developing computational tools for analyzing genome-sequence data, to establishing and maintaining genomic databases that are accessible to scientists worldwide, to annotating reference genome sequences, to developing tools that facilitate clinical decision-making. In these studies, critical attention is being paid to genomic-specific issues related to data security, access, privacy, and management. With a growing focus on understanding the genomic bases of human diseases, researchers are developing computational approaches to predict the biological consequences of genomic variants; eventually, this work will lead to the design of computational tools that facilitate the implementation of genomic medicine.

NHGRI’s data science efforts are also becoming increasingly aligned with broader NIH programs, including the Big Data to Knowledge (BD2K) effort and the Precision Medicine Initiative. Such synergies reflect the new data-intensive nature of all areas of biomedical research.

# Training in Genomics

*Developing a strong workforce for genomics research and genomic medicine*

NHGRI's mission includes cultivating a large, strong, and diverse pool of future researchers, healthcare providers, and educators with appropriate expertise in genomics. For many years, NHGRI has supported training and workforce-development programs in the fundamental aspects of genomics—areas that were essential for completing the Human Genome Project and for the genomics efforts that immediately followed. To keep pace with the widening scope of the field, today's programs have expanded to include training in genomic medicine.

NHGRI has a multitude of mechanisms to support genomics training: grants to teaching institutions, fellowships to individual trainees, loan repayment, internships, courses, and online resources. Such training opportunities are pervasive within both the Institute's Extramural Research Program and Intramural Research Program.

The NHGRI Extramural Research Program offers a set of grant options that provide support to academic institutions for both pre- and postdoctoral trainees and for researchers already established in their careers. The various training programs are designed to develop trainees' expertise in broad areas, such as genome structure and function, computational biology, and/or genomic medicine.

The NHGRI Intramural Research Program offers a diverse array of training programs. Pre- and postdoctoral trainees can conduct research in the laboratories of NHGRI investigators. In addition, the Institute offers a number of specific fellowship opportunities in areas such as medical genetics, medical biochemical genetics, genetic counseling, health disparities, education, and policy. NHGRI also conducts short-term summer programs for students, teachers, and faculty.

All NHGRI training programs include a core commitment to maximize the diversity of the genomics workforce. The Institute develops research partnerships with minority-serving institutions to enhance their technical capabilities for conducting genomics research, builds relationships with individuals from underserved communities and minority organizations so as to be aware of their issues and concerns as they relate to genomics research, and provides opportunities for underserved students and researchers to pursue careers in genomics.



# Future Horizons

*The growing relevance of genomics*

In less than three decades, genomics has progressed from a just-named, emerging discipline led by a small community of trailblazing researchers to a vital area of biomedical research. Genomics has quickly become profoundly important for most aspects of biological inquiry, and is now poised to become a fundamental component of medical care.

With this transition, the field has gained relevance to a much broader audience. Once only relevant to biomedical researchers and, later, healthcare professionals, genomics is increasingly becoming important to patients (and friends and relatives of patients)—which means all of us. Indeed, the words “genome” and “genomics” are now regularly seen in the popular press, heard on the radio, and mentioned on television. In short, genomics is now highly relevant to society more broadly.

Since its inception, NHGRI has been a world leader in genomics research, training, education, and outreach. While the Institute’s origins were deeply rooted in the highly focused goals of the Human Genome Project, its mission and corresponding programs have grown in parallel with the expanding relevance of genomics; these now touch the entire landscape of biomedical research—from very basic science to clinical implementation.

Looking onward to future horizons, the scope and relevance of genomics will no doubt continue to grow. Genomics will catalyze important scientific and medical advances on a regular basis, bringing new opportunities along with new challenges. NHGRI stands ready to pursue those opportunities and to address the associated challenges through its research programs, engagement efforts, and outreach initiatives.

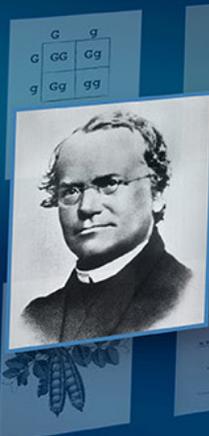
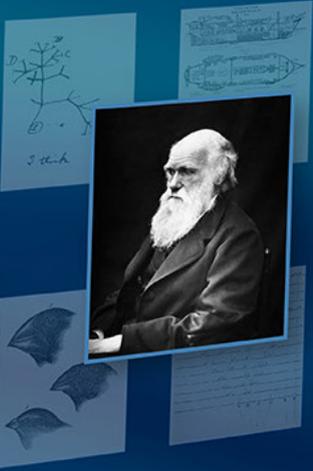
At the launch of the Human Genome Project, the full potential of genomics was hard to fathom. Roughly 25 years later, progress in genomics in terms of research achievements, technological developments, and societal integration has been nothing short of spectacular, bolstering NHGRI’s fundamental and core belief that genomics offers great promise and hope for improving the human condition.





National Human Genome  
Research Institute

[genome.gov](http://genome.gov)



	U	C	A	G
U	phe	leu	ser	tyr
C	pro	his	pro	arg
A	asn	thr	lys	asp
G	gln	glu	asp	glu



**NIH** National Human Genome Research Institute

NIH Publication Number: 15-8011

2015