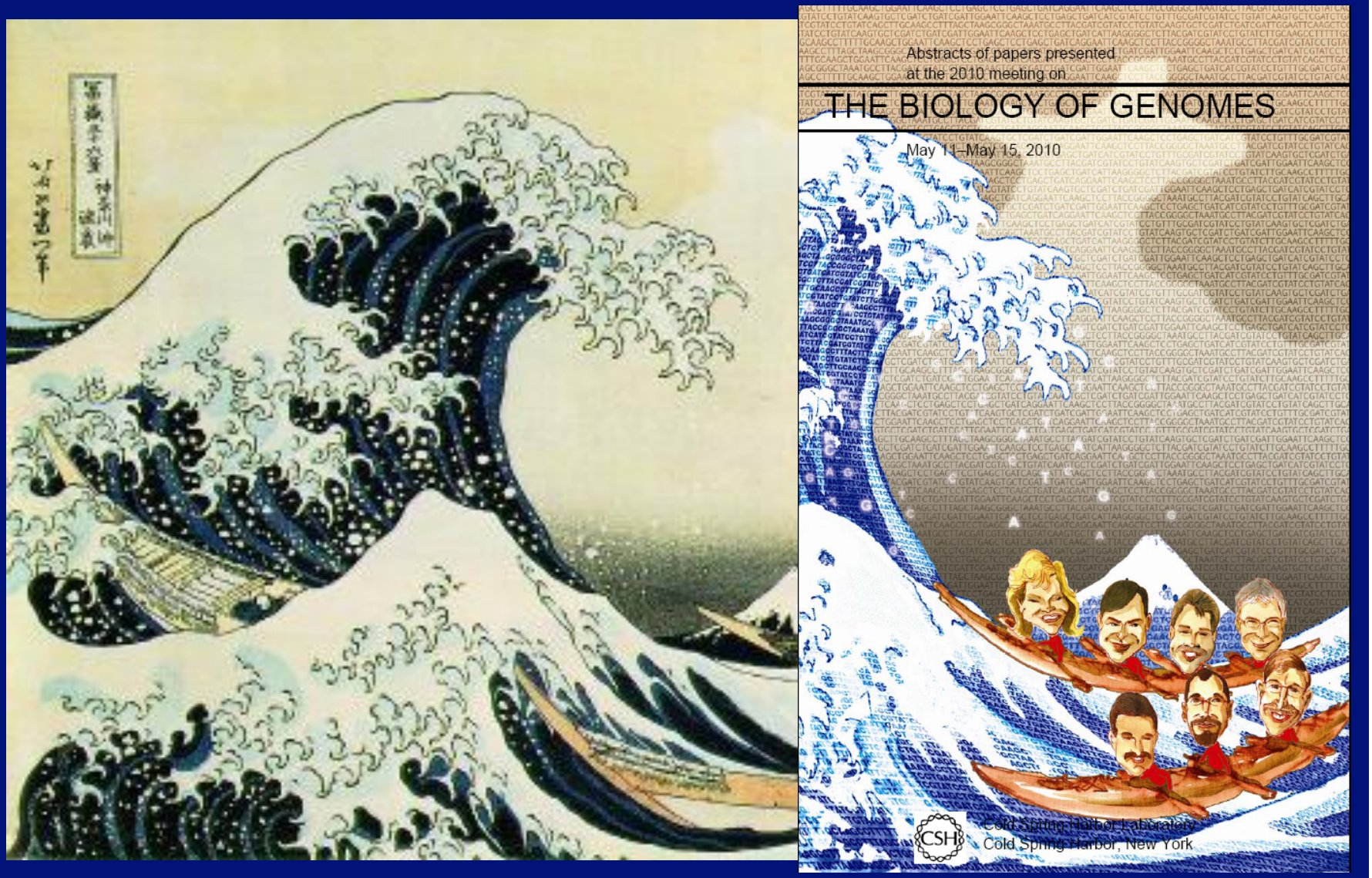
National Human Genome Research Institute

The Future of Genomics Research

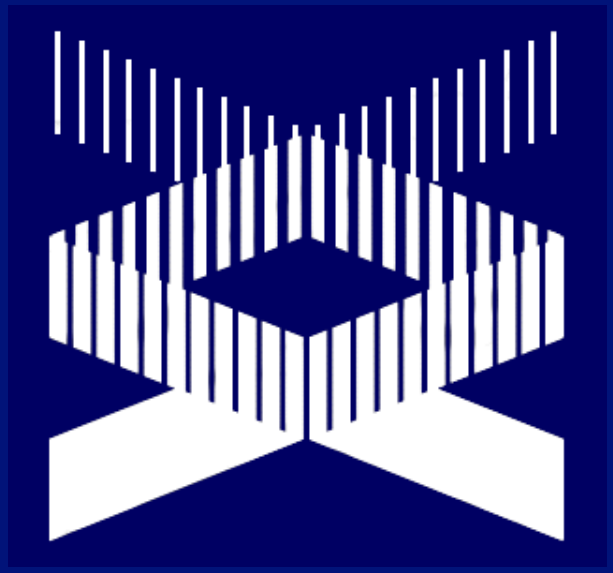
Eric Green, M.D., Ph.D.
Director, NHGRI

The Genomic Era



***The Great Wave* (K. Hokusai)**

NHGRI Strategic Planning Process



Understanding Our Genetic Inheritance

The U.S. Human Genome Project:

The First Five Years FY 1991-1995

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Institutes of Health

U.S. DEPARTMENT OF ENERGY
Office of Energy Research
Office of Health and Environmental Research

1991-1995

POLICY FORUM

A New Five-Year Plan for the U.S. Human Genome Project

Francis Collins and David Galas*

The U.S. Human Genome Project is
of an international effort to develop
netic and physical maps and determine
DNA sequence of the human genome
the genomes of several model organ
Thanks to advances in technology a
tightly focused effort, the project i
track with respect to its initial 5-year g
Because 3 years have elapsed since t
goals were set, and because a much
sophisticated and detailed understandi
what needs to be done and how to do
now available, the goals have been re
and extended to cover the first 8
(through September 1998) of the 15
genome initiative.

In 1990, the Human Genome prog
of the National Institutes of Health (N
and the Department of Energy (DOE
veloped a joint research plan with sp
goals for the first 5 years [fiscal year
1991–95] of the U.S. Human Gen
Project (1). It has served as a valu
guide for both the research communit
the agencies' administrative staff in d
oping and executing the genome pr
and assessing its progress for the p
years. Great strides have been made to
the achievement of the initial set of
particularly with respect to constructin
tailed human genetic maps, impri
physical maps of the human genome
the genomes of certain model organ
developing improved technology for
sequencing and information handling
defining the most urgent set of ethica
gal, and social issues associated with th
quisition and use of large amounts o
netic information.

Progress toward achieving the first
goals for the genome project appears
on schedule or, in some instances,
ahead of schedule. Furthermore, tec
logical improvements that could not
been anticipated in 1990 have in som
eas changed the scope of the project a
lowed more ambitious approaches. E
this year, it was therefore decided to u
and extend the initial goals to addres
scope of genome research beyond

physical maps; (iii) the definition of the sequence tagged site (STS) (5) as a common unit of physical mapping; and (iv) improved technology and automation for DNA sequencing. Further substantial improvements in technology are needed in all areas of genome research, especially in

F. Collins is the director of the National Cen
Human Genome Research, National Institu
Health, Bethesda, MD 20892.
D. Galas was associate director, Office of Heal
Environmental Research, Department of E
Washington, DC 20585.

* Present address: Darwin Molecular, 2405 C
Point, Kirkland, WA 98033.

1993-1998

SPECIAL SECTION — GENOME

5. G. An, B. D. Watson, C. C. Chiang, *Plant Physiol.* 81, 301 (1986); A. M. Lloyd *et al.*, *Science* 234, 464 (1986); K. A. Feldmann and M. D. Marks, *Mol. Gen. Genet.* 208, 1 (1987).
6. E. M. Meyerowitz and R. E. Pruitt, *Science* 229, 1214 (1985).
7. G. R. Fink, *Genetics* 149, 473 (1998).
8. C. Koncz, N.-H. Chua, J. Schell, Eds., *Methods in Arabidopsis Research* (World Scientific, River Edge, NJ, 1992); J. M. Martinez-Zapater and J. Salinas, Eds., *Arabidopsis Protocols*, vol. 82 of *Methods in Molecular Biology* (Humana, Totowa, NJ, 1998).
9. For example, see S. A. Kempin *et al.*, *Nature* 38
10. N. Bechtold, J. Ellis, G. Pelletier, *C. R. Acad. Sci. P*
et al., *Plant J.* 5, 551 (1994).
11. V. Sundaresan, *Trends Plant Sci.* 1, 184 (1996).
12. D. Meinke and M. Koornneef, *Plant J.* 12, 247 (1
13. P. Franz *et al.*, *Bhil.* 13, 867 (1998).
14. C. Lister and C. Dean, *Bhil.* 4, 745 (1993); C. A
(1998).
15. E. J. Finnegan, R. K. Genger, W. J. Peacock, E. S. Dea
Mol. Biol. 49, 223 (1998).
16. D. Preuss, S. Y. Rhee, R. W. Davis, *Science* 264, 14
Browne, D. Preuss, *Proc. Natl. Acad. Sci. U.S.A.* 9
17. For recent examples of identifying knockouts in c
al., *Plant J.* 8, 613 (1995); P. J. Krysan, J. C. Young
Acad. Sci. U.S.A. 93, 8145 (1996).
18. For information, contact http://www.bio.net/hy
19. D. Meinke *et al.*, Eds., "Multinational coordina
research project, progress report, year six" (Natio
97-131, Arlington, VA, 1997).
20. M. Bevan *et al.*, *Plant Cell* 9, 476 (1997).
21. H. Hofte *et al.*, *Plant J.* 4, 1051 (1993); T. Newm
(1994).
22. S. Choi, R. A. Creelman, J. E. Mullet, R. A. Wing, *Plant Mol. Biol. Rep.* 13, 124 (1995); F. Creusot *et al.*, *Plant J.* 8, 763 (1995).
23. R. Schmidt *et al.*, *Science* 270, 480 (1995); E. A. Zachgo *et al.*, *Genome Res.* 6, 19 (1996); R. Schmidt, K. Love, J. West, Z. Lenehan, C. Dean, *Plant J.* 11, 563 (1997).
24. M. Bevan *et al.*, *Nature* 391, 485 (1998).
25. S. Sato *et al.*, *DNA Res.* 4, 215 (1997).
26. C. Chang, S. F. Kwok, A. B. Bleecker, E. M. Meyerowitz, *Science* 262, 539 (1993); G. E. Schaller and A. B. Bleecker, *Bhil.* 270, 1809 (1995).
27. J. U. P. Nagpal, V. Vibart, T. C. McMorris, J. Chory, *Bhil.* 272, 398 (1996); M. Szekeres

New Goal

Francis S. Collins,* Ari Patrin
and

REVIEW

**The Human Genome Project has suc
the major goals in its current 5-year p
1993–98. A new plan, for 1998–2003
human DNA sequencing will be the n
bitious schedule has been set to com
by the end of 2003, 2 years ahead of
the course of completing the sequenc
the human sequence will be produced
plan also includes goals for sequenci
ment; for studying human genome
developing technology for functional
ing the sequence of *Caenorhabditis*
melanogaster and starting the mouse
the ethical, legal, and social implicatio
for bioinformatics and computational
of genome scientists.**

The Human Genome Project (HGP) is f
single most important project in biology
es—one that will permanently change biol

F. S. Collins and E. Jordan are with the National H
National Institutes of Health, Bethesda, MD 20892,
of Biological and Environmental Research, Departm
20585, USA. A. Chakravarti is with the Department o
Genetics, Case Western Reserve University and U
Cleveland, OH 44106, USA. R. Gesteland is at the
University of Utah, Salt Lake City, UT 84112, USA
Institute of Ethics, Georgetown University, Washing

*To whom correspondence should be addressed:

682 23

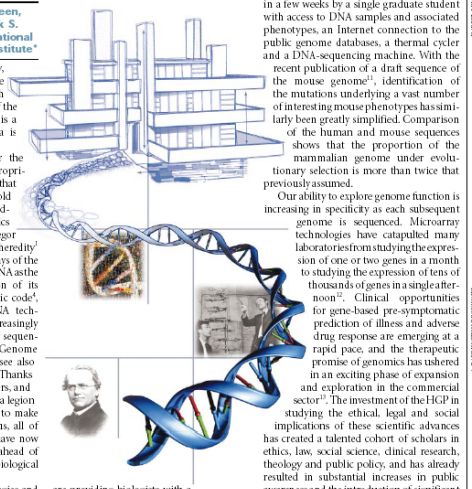
1998-2003

feature

A vision for the future of genomics research

A blueprint for the genomic era.

Francis S. Collins, Eric D. Green, Alan E. Guttmacher and Mark S. Guyer on behalf of the US National Human Genome Research Institute*



DARIUS LEBA

A. BARRINGTON BROWNSELL

The completion of a high-quality, comprehensive sequence of the human genome, in this fiftieth anniversary year of the discovery of the double-helical structure of DNA, is a landmark event. The genomic era is now a reality.

In contemplating a vision for the future of genomics research, it is appropriate to consider the remarkable path that has brought us here. The rollfold (Figure 1) shows a timeline of landmark accomplishments in genetics and genomics, beginning with Gregor Mendel's discovery of the laws of heredity[1] and their rediscovery in the early days of the twentieth century. Recognition of DNA as the hereditary material[2], determination of its structure[3], elucidation of the genetic code[4], development of recombinant DNA technologies[5,6], and establishment of increasingly automatable methods for DNA sequencing[7–10] set the stage for the Human Genome Project (HGP) to begin in 1990 (see also www.nature.com/nature/DNA50). Thanks to the vision of the original planners, and the creativity and determination of a legion of talented scientists who decided to make this project their overarching focus, all of the initial objectives of the HGP have now been achieved at least two years ahead of expectation, and a revolution in biological research has begun.

The project's new research strategies and experimental technologies have generated a steady stream of ever-larger and more complex genomic data sets that have poured into public databases and have transformed the study of virtually all life processes. The genomic approach of technology development and large-scale generation of community resource data sets has introduced an important new dimension into biological and biomedical research. Interwoven advances in genetics, comparative genomics, high-throughput biochemistry and bioinformatics are providing biologists with a markedly improved repertoire of research tools that will allow the functioning of organisms in health and disease to be analysed and comprehended at an unprecedented level of molecular detail. Genome sequences, the bounded sets of information that guide biological development and function, lie at the heart of this revolution. In short, genomics has become a central and cohesive discipline of biomedical research.

The practical consequences of the emergence of this new field are widely apparent. Identification of the genes responsible for human mendelian diseases, once a herculean task requiring large research teams, many years of hard work, and an uncertain outcome, can now be routinely accomplished in a few weeks by a single graduate student with access to DNA samples and associated phenotypes, an Internet connection to the public genome databases, a thermal cycler and a DNA-sequencing machine. With the recent publication of a draft sequence of the mouse genome[11], identification of the mutations underlying a vast number of interesting mouse phenotypes has similarly been greatly simplified. Comparison of the human and mouse sequences shows that the proportion of the mammalian genome under evolutionary selection is more than twice that previously assumed.

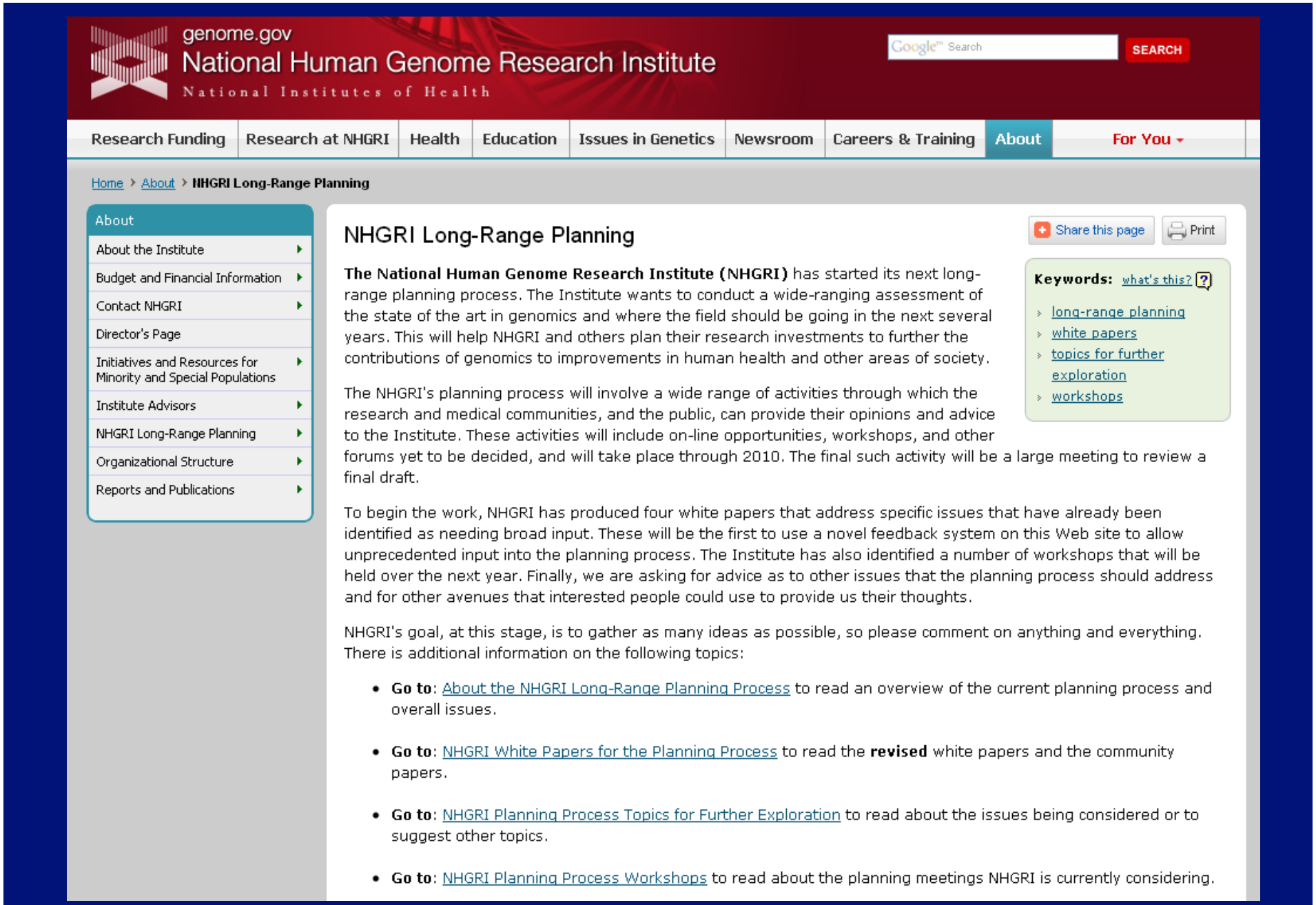
Our ability to explore genome function is increasing in specificity as each subsequent genome is sequenced. Microarray technologies have catapulted many laboratories from studying the expression of one or two genes in a month to studying the expression of tens of thousands of genes in a single afternoon[12]. Clinical opportunities for gene-based pre-symptomatic prediction of illness and adverse drug response are emerging at a rapid pace, and the therapeutic promise of genomics has ushered in an exciting phase of expansion and exploration in the commercial sector[13]. The investment of the HGP in studying the ethical, legal and social implications of these scientific advances has created a talented cohort of scholars in ethics, law, social science, clinical research, theology and public policy, and has already resulted in substantial increases in public awareness and the introduction of significant (but still incomplete) protections against misuses such as genetic discrimination (see www.genome.gov/PolicyEthics).

These accomplishments fulfil the expansive vision articulated in the 1988 report of the National Research Council, *Mapping and Sequencing the Human Genome*[14]. The successful completion of the HGP this year thus represents an opportunity to look forward and offer a blueprint for the future of genomics research over the next several years.

The vision presented here addresses a different world from that reflected in earlier plans published in 1990, 1993 and 1998 (refs 15–17). Those documents addressed the goals of the 1988 report, defining detailed paths towards the development of genome-

*Endorsed by the National Advisory Council for Human Genome Research, whose members are Vickie Yates Brown, David R. Bergsma, Wylie Burke, Ronald W. Davis, William M. Gelbart, Eric T. Juengst, Bronya J. Keats, Bartharshaputi, Richard F. Lifton, Kim L. Nickerson, Maynard V. Olson, Janet D. Rowley, Robert Tepper, Robert B. Waterston and Tadataka Yamada.

2003-2010



genome.gov/Planning

NHGRI Strategic Planning Process

- **Topic-Specific Workshops**
- **White Papers and Web-based Feedback**
  - *Applying Genomics to Clinical Problems: Diagnostics, Preventive Medicine, and Pharmacogenomics*
  - *Applying Genomics to Clinical Problems: Therapeutics*
  - *A Vision for the Future of Genomics: Education and Community Engagement*
  - *The Future of Genome Sequencing*
- **Engagement of Institute Advisors**
- **Internal Discussions and Synthesis**
- **Few External Town Halls**
- **Finale Meeting (July 2010)**
- **Publication of New Plan (December 2010)**

Five Themes of the NIH Research Agenda

POLICYFORUM

RESEARCH AGENDA

Opportunities for Research and NIH

The promise of fundamental advances in diagnosis, prevention, and treatment of disease has never been greater.

Francis S. Collins

The mission of the National Institutes of Health (NIH) is science in pursuit of fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to extend healthy life and to reduce the burdens of illness and disability. The power of the molecular approach to health and disease has steadily gained momentum over the past several decades and is now poised to catalyze a revolution in medicine. The foundation of success in biomedical research has always been, and no doubt will continue to be, the creative insights of individual investigators. But increasingly those investigators are working in teams, accelerated by interdisciplinary approaches and empowered by open access to tools, databases, and technologies, so a careful balance is needed between investigator-initiated projects and large-scale community resource programs. For both individual and large-scale efforts, it is appropriate to identify areas of particular promise. Here are five such areas that are ripe for major advances that could reap substantial downstream benefits.

High-Throughput Technologies

In the past, most biomedical basic science projects required investigators to limit their scope to a single aspect of cell biology or physiology. The revolution now sweeping the field is the ability to be comprehensive—for example, to define all of the genes of the human or a model organism, all of the human proteins and their structures, all of the common variations in the genome, all of the major pathways for signal transduction in the cell, all of the patterns of gene expression in the brain, all of the steps involved in early development, or all of the components of the immune system. Further development of technologies in areas such as DNA sequencing, imaging, nanotechnology, proteomics, metabolomics, small-molecule screening, and RNA interference are ripe for aggressive investment. Furthermore, these technologies will spur the production of massive and complex data sets and will require major investments in computational biology.

As one example, the Cancer Genome Atlas (*1*) is now poised to derive comprehensive information about the genetic underpinnings of 20 major tumor types. This information will likely force a complete revision of diagnostic categories in cancer and will usher in an era where abnormal pathways in specific tumors will be matched with the known targets of existing therapeutics. Another example is the opportunity to understand how interactions between ourselves and the microbes that live on us and in us (the "microbiome") can influence health and disease (*2*).

Translational Medicine

Critics have complained in the past that NIH is too slow to translate basic discoveries into new diagnostic and treatment advances in the clinic. Some of that criticism may have been deserved, but often the pathway from molecular insight to therapeutic benefit was just not discernible. For many disorders, that is now changing. Three major factors have contributed to this: (i) the discovery of the fundamental basis of hundreds of diseases has advanced dramatically; (ii) with support from the NIH Roadmap, academic investigators supported by NIH now have access to resources to enable them to convert fundamental observations into assays that can be used to screen hundreds of thousands of candidates for drug development; (iii) public-private partnerships are being more widely embraced in the drug-development pipeline to enable biotech and pharmaceutical companies to pick up promising compounds that have been effectively "de-risked" by academic investigators and to bring them to clinical trials and U.S. Food and Drug Administration (FDA) approval.

As one example, the NIH Therapeutics for Rare and Neglected Diseases (TRND) (*3*) program will allow certain promising compounds to be taken through the preclinical phase by NIH, in an open environment where the world's experts on the disease can be involved. Furthermore, as information about common diseases increases, many are being resolved into distinct molecular subsets, and so the TRND model will be even more widely applicable.

The first human protocol (for spinal cord injury) involving human embryonic stem cells (hESCs) was approved by the FDA in 2009, and the opening up of federal support for hESC research will bring many investigators into this field. The capability of transforming human skin fibroblasts and other cells into induced pluripotent stem cells (iPSCs) opens up a powerful strategy for therapeutic replacement of damaged or abnormal tissues without the risk of transplant rejection (*4*–*6*). Although much work remains to be done to investigate possible risks, the iPSC approach stands as one of the most breathtaking advances of the last several years, and every effort should be made to pursue the basic and therapeutic implications with maximum speed.

Benefiting Health Care Reform

U.S. expenditures on health care now represent 17% of our Gross Domestic Product, are continuing to grow, and are excessive as a percentage of per capita gross income com-

National Institutes of Health, Bethesda, MD 20892, USA. E-mail: collinsf@mail.nih.gov

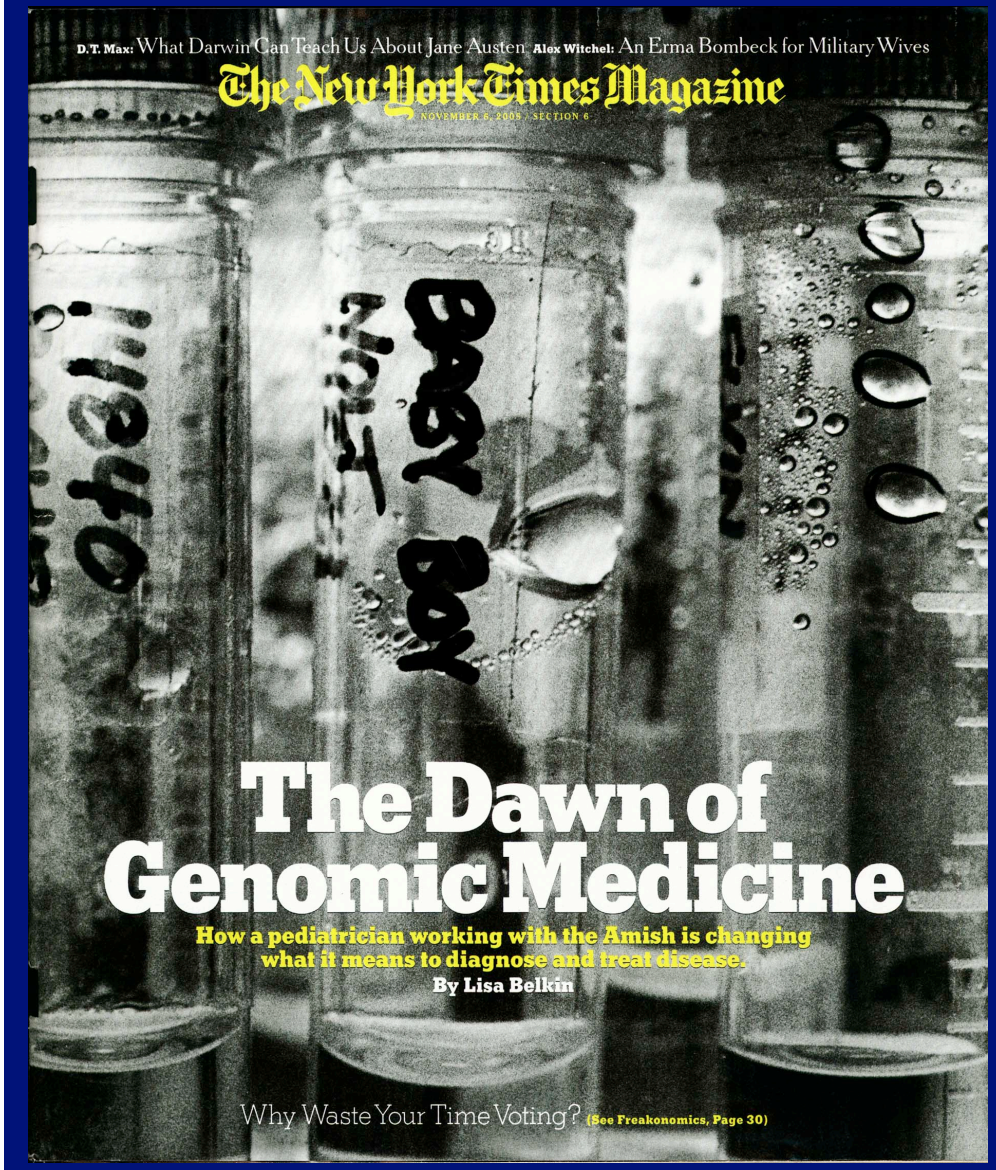
CREDITS (LEFT TO RIGHT): DOOM AND CHALLORAN/NORTHWESTERN UNIVERSITY; LINDA BARTLETT/NCI, NIH; BILL BRANSON/NCI/NIH; NASA; BILL BRANSON/NIH/NCRR, NIH

Science (2010)

- High-Throughput Technologies
- Translational Medicine
- Benefiting Health Care Reform
- Focusing More on Global Health
- Reinvigorating and Empowering the Biomedical Research Community

NIH…
Turning Discovery Into Health

U.S. Department of Health & Human Services
National Institutes of Health


D.T. Max: What Darwin Can Teach Us About Jane Austen. Alex Witchel: An Erma Bombeck for Military Wives
The New York Times Magazine
NOVEMBER 6, 2005 · SECTION 6
The Dawn of Genomic Medicine
How a pediatrician working with the Amish is changing what it means to diagnose and treat disease.
By Lisa Belkin
Why Waste Your Time Voting? (See Freakonomics, Page 30)


Science
24 October 2003
Vol. 302 No. 5645
Pages 517–728 $10
Genomic Medicine
American Association for the Advancement of Science

Genomic Medicine

Healthcare tailored to the individual based on genomic information

The Path to Genomic Medicine

Human Genome Project

→ → ? → →

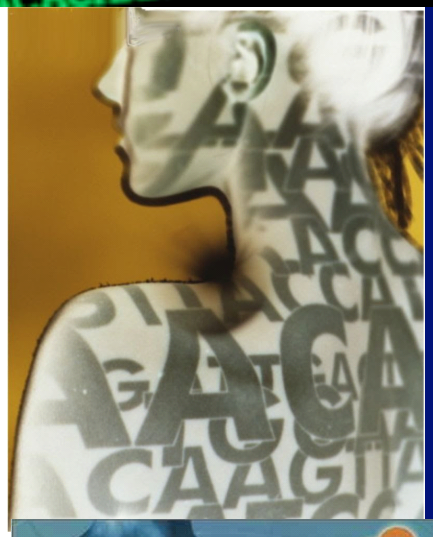
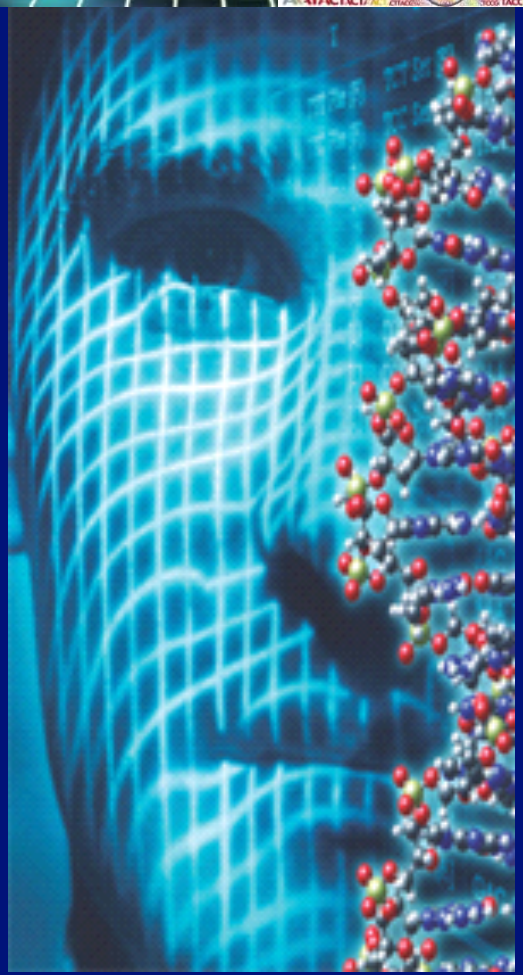
Realization of Genomic Medicine

✓ *"Fulfilling the Promise"* ?

A central component of NHGRI's future mission will be to foster the maturation and practice of genomic medicine






Many
Incoming
Technologies…

How many human genomes can you sequence for $10M?

2000

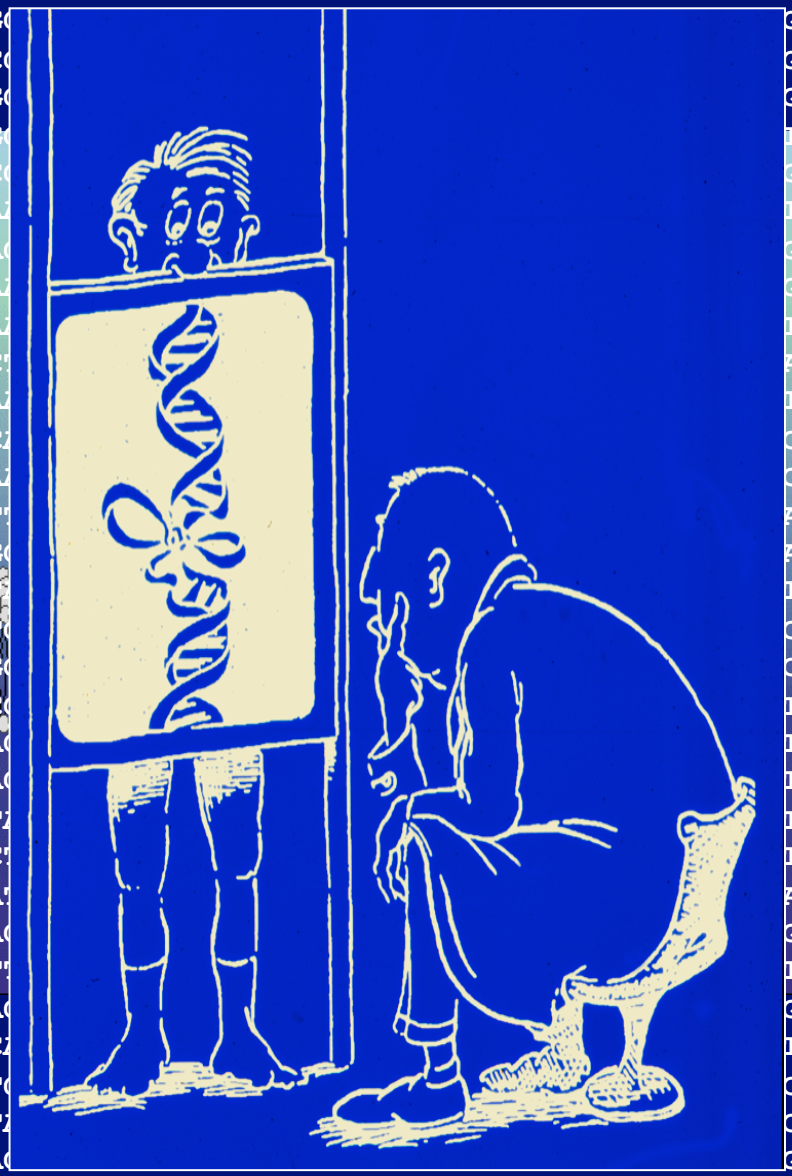
<1 Human Genome

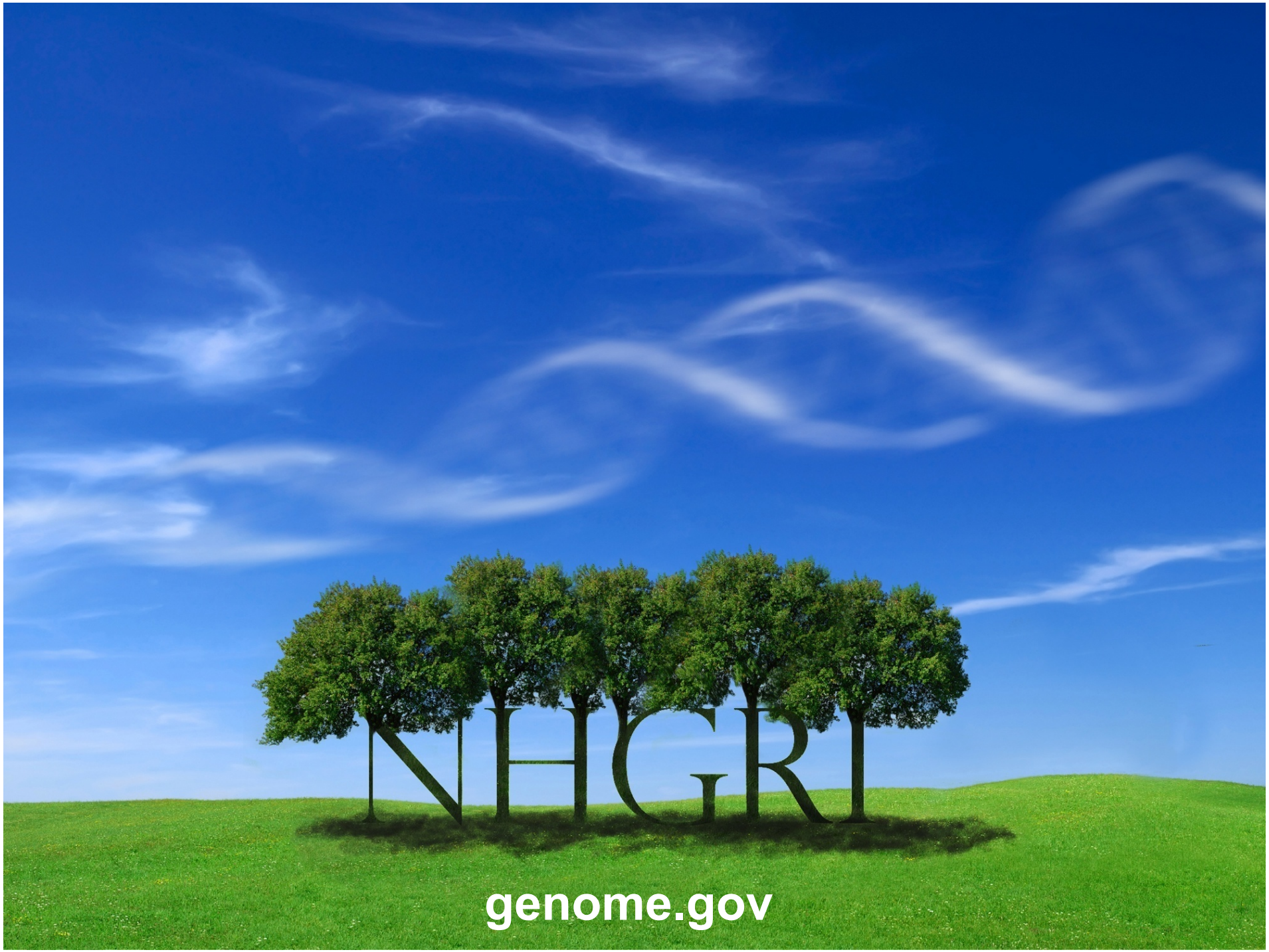
2010

200-400 Human Genomes

The Computational Bottleneck

The Informational Bottleneck




NHGRI
genome.gov