

# Human Genome Project: Implications for Healthcare

Eric Green, MD

## Summary

The human genome project successfully identified the composition of human and some other species' DNA. The next phase of genomics research is to begin identifying which DNA sequences are functional and what precisely the sequences code. The future of genomics is personalized medicine with improved diagnostics and treatment based on an individual's genetic makeup.

## Key Points

- The Human Genome Project succeeded in decoding the human genome in April 2003.
- A complete working knowledge of our genetic instruction book provides unprecedented information and opportunities for knowing the intricacies of human physiology, thereby advancing our understanding of human health and disease.
- Now that the human genome has been sequenced, personalized diagnostics, preventive strategies, and medications are on the horizon.

THE 1990s WAS THE DECADE of genomic revolution, with the centerpiece being a large international effort aimed at mapping and sequencing the human genome.<sup>1</sup> Known as the Human Genome Project, this effort began in the fall of 1990 and was regarded by many as the Lewis and Clark expedition of the 21st century. At its core, the Human Genome Project sought to characterize, in essence, to de-code the human body's genome. The rationale for the project was that a complete working knowledge of our own genetic instruction book would provide unprecedented information and opportunities to know the intricacies of human physiology and also advance our understanding of human health and disease.

The project's name was a misnomer, however, because it was not limited to the study of human DNA. It included unicellular organisms, such as common yeast, a fly, a nematode worm, and a mouse. A key goal of the Human Genome Project was to characterize the genomes of these organisms in order to provide detailed views of their genetic blueprints, and fundamental insights into unicellular, multi-cellular, mammalian, and human biology.

In April 2003, two years ahead of schedule, the human genome was completely deduced. That date will forever be regarded as a monumental, perhaps

pivotal, time in the history of bio-medical research, with the completion of the human genome sequence and the official end of the Human Genome Project.

## The Next Step in Genomics

The next great frontier in fundamental genomics is actually interpreting the human genome sequence. The challenge is being able to determine the subset of bases that are active because not all base sequences are functionally important. This work is being done by comparative sequence analysis, which is based on knowledge that sequences known to be common among species that are separated over very large evolutionary distances are more likely to be functionally important; otherwise evolution would have changed them. In essence, scientists are reviewing evolutionary changes for clues as to how the human genome actually functions. This process involves extracting sequencing from human, mouse, and dog genomes and making comparisons via computer modeling.

Genome comparisons have shown that humans and chimpanzees are 98 percent identical. The baboon and human genome are less related. In comparing the human genome to mouse, rat, cat, and dog genomes, some parts are very similar and other parts are not similar at all. The opossum, duck-billed platypus, chicken, and fish

genomes are far removed from the human genome. The similarities that do exist between the human genome and the genomes of birds and fish, are found in exons, which are the part of the gene that code for protein.

By having all of these different species sequences together for comparison, scientists can start to develop experimental and computational models that allow identification of functionally important sections of DNA (e.g., the section that turns on the cystic fibrosis gene). Sequences that match among different species are called multi-species conserved sequences, which are preserved during evolution. Researchers are looking for the 5 percent most conserved sequences because only about 5 percent of our genome is thought to be functionally important (see Exhibit 1). These conserved sequences will then be characterized in the laboratory to determine their function.

The National Institutes of Health has invested in sequencing various animal genomes in order to gather additional clues as to how the human genome functions. NIH anticipates that a broader understanding of how the human genome functions will ultimately shed light on human health and how gene alterations can actually lead to disease.

The Human Genome Project was originally conceived as a means of improving human health. A key part of this vision is being able to make the connections that are most relevant to humans. Buried in human genomes are outright alterations that lead to disease or at least predispose the body to disease. Virtually all dis-

eases have a genetic component (see Exhibit 2). Cystic fibrosis, for example, is a genetic disease with a small environmental component. Even in diseases such as AIDS, in which environment is the predominant cause, there are relevant genetic underpinnings, especially with respect to morbidity and mortality.

The purpose of the genome project was to give tools, technologies, and information so that geneticists, clinicians, and biologists can begin to untangle the genetic components of a range of diseases. The genome project has delivered on that promise when it comes to rare genetic diseases. Individual single genes of rare diseases often are much easier to find than the multiple small genetic defects that individually may not cause a disease but are a contributing factor. Obviously, it is more difficult to determine which subtle genetic defects in which combinations cause disease. The major contributing genes for many of the most common diseases will be identified in the next five to 15 years (see Exhibit 3). Whole genome association studies are a strategy that is being used to tease out the genetic contributions of complex diseases. In the last couple of years, a number of significant studies have begun identifying the genetic lesions that are associated with complex diseases such as Crohn's.

### The Future of Genomics

Identifying the genetic cause of human disease is not the endpoint, but a new beginning. New tools and better diagnostics will emerge so that preventive strategies can be implemented as well as better therapeutic strategies, be they gene therapy or personalized medication (see Exhibit 4). New technologies in the clinical setting will allow determination of DNA sequence data from individual patients. In turn, clinicians will be better equipped to predict susceptibility to certain diseases and determine medications that will work best for the individual patient. In the future, genomic technologies will be applied to perplexing problems in clinical medicine, and human genome will become increasingly relevant to practicing physicians.

Pharmacogenomics is the idea of being able to stratify individuals based on their genetic background and thus being able to predict whether they'll be good or bad responders to a particular medication. A recent

#### Exhibit 1: The Human Genome...by the Numbers

##### ~5% of the Human Genome Is Functionally Important

- 5% of 3B bases = ~150M bases
- Do not yet know the position of these ~150M functional bases

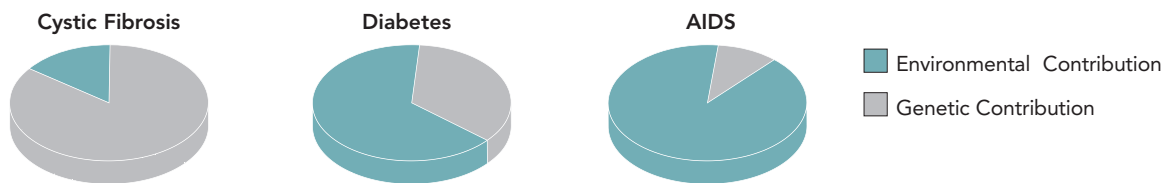
##### ~1.5% Encodes for Protein (Genes)

- Corresponds to ~20-25K genes
- Many more than 25K different proteins

##### ~3.5% Functional But Non-Coding

- Gene regulatory elements
- Chromosomal functional elements
- Undiscovered functional elements (not yet in textbooks)

#### Exhibit 2: Virtually All Diseases Have a Genetic Component



article in the *New England Journal of Medicine* clearly showed specific genetic alterations predicted exactly what the response was going to be to the drug warfarin.<sup>2</sup> In thinking broadly about this particular area of genomics and its relationship to health, science will go from the most basic DNA sequence to predicting what are the best medications for individuals based on their genetic profile, and also designing better medications based on the intricacies of genetics.

### Other Issues

From its outset, the genome project invested 5 percent of its total budget in studying the ethical, legal, and social implications of genetic research. There are many issues in genomics that have to be grappled with, including genetic discrimination. Although there is much excitement that could be capitalized on, the scientific community needs to be careful to properly utilize information gained from the project and to make sure that harm does not occur. For additional information on the

Human Genome Project and ongoing studies, consult [www.genome.gov](http://www.genome.gov).

### Conclusion

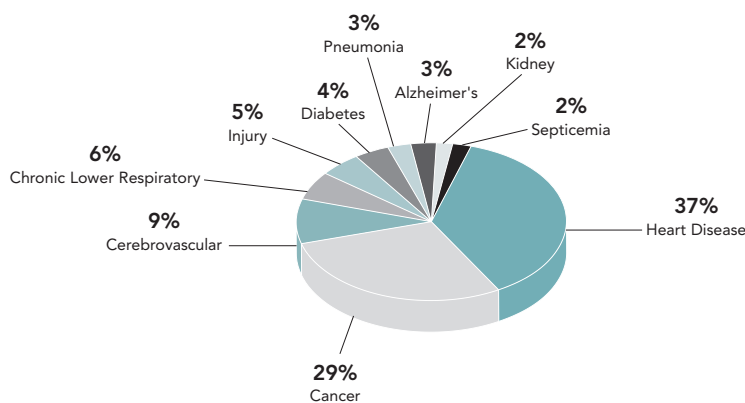
The year 2003 will stand in history as a pivotal time by scientists and all of healthcare. Many aspects of medicine will change over the next decade and beyond. Advances in genomics are opening new opportunities for applying genetic information to diagnostics, risk prevention, and pharmacotherapy, thus ensuring better health outcomes, drug efficacy, and safety in clinical practice. **JMCM**

**Eric Green, MD**, is scientific director of the National Human Genome Research Institute (NHGRI) at the National Institutes of Health (NIH) in Bethesda, Md. In addition, he serves as chief of the NHGRI's Genome Technology Branch and director of the NIH Intramural Sequencing Center.

### References

1. National Institutes of Health. National Human Genome Research Institute. Available at: [www.genome.gov](http://www.genome.gov).
2. Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, Blough DK, Thummel KE, Veenstra DL, Rettie AE. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Eng J Med*. 2005;352:2285-93.

**Exhibit 3: Top Ten Leading Causes of Mortality**



The major contributing genes for many common diseases will be identified within the next five to 10 years.

**Exhibit 4: The Start of Personalized Medicine**

