Human Molecular Genetics and Genomics — Important Advances and Exciting Possibilities

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

JANUARY 7, 2021

A HALF-CENTURY OF PROGRESS IN HEALTH: THE NATIONAL ACADEMY OF MEDICINE AT 50

Human Molecular Genetics and Genomics — Important Advances and Exciting Possibilities

Francis S. Collins, M.D., Ph.D., Jennifer A. Doudna, Ph.D., Eric S. Lander, Ph.D., and Charles N. Rotimi, Ph.D.

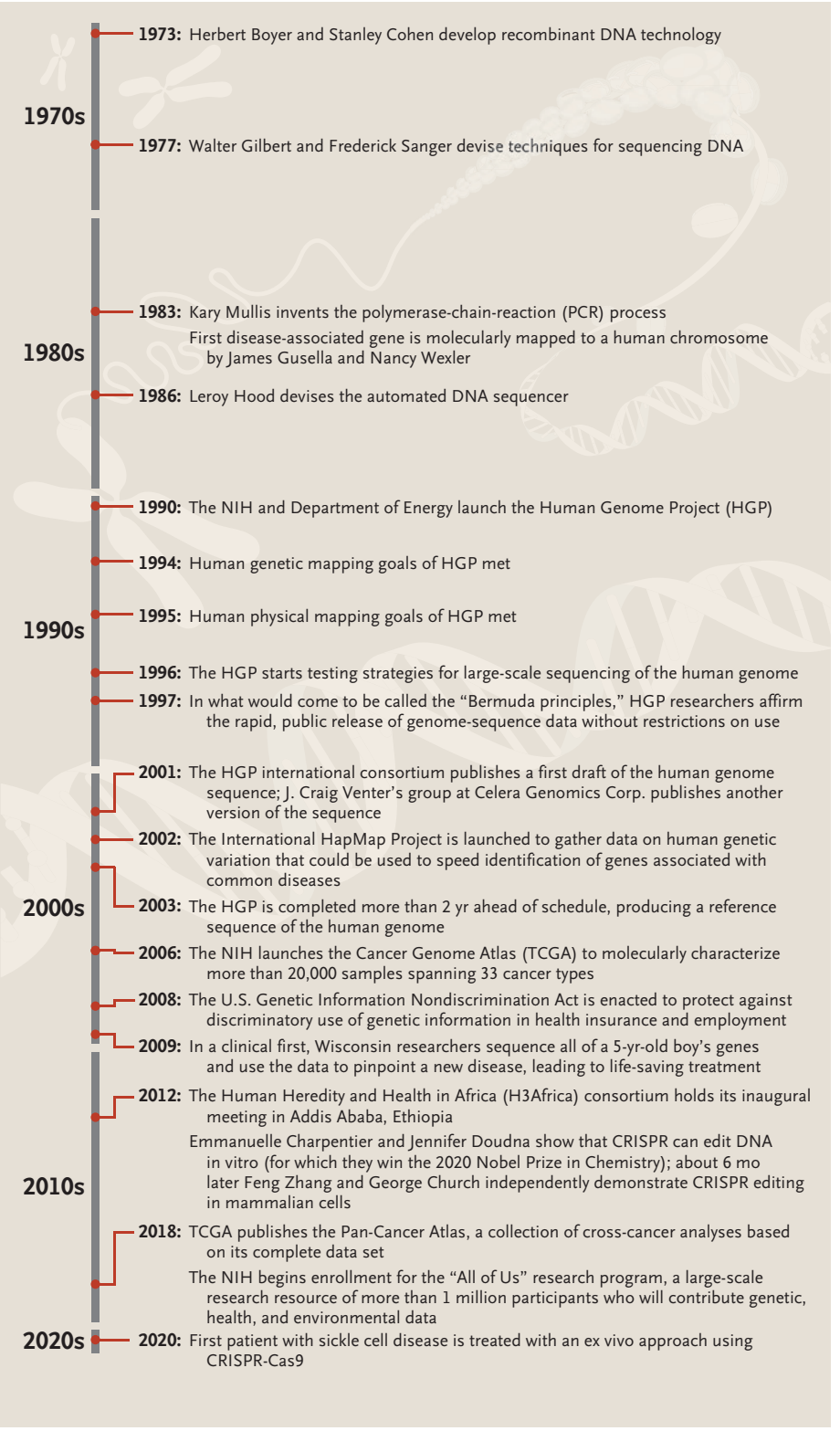
The breathtaking progress in molecular genetics that has occurred over the past five decades and the transition to genomic medicine would have been difficult to imagine in 1970, when the Institute of Medicine (IOM), now the National Academy of Medicine (NAM), was formed. The term "genomics" hadn't yet been coined, the tools and technologies that are the foundation of modern biotechnology were in their infancy, and methods for sequencing even a few nucleotides were barely workable.[1]

The IOM's early years coincided with paradigm-shifting discoveries related to DNA, as biologic research swiftly incorporated Boyer and Cohen's recombinant method, Sanger's DNA-sequencing work, and Mullis's introduction of polymerase chain reaction (PCR) technology (see timeline). Yet even against this backdrop, the notion of a "big science" endeavor to sequence the human genome seemed radical.

In 1987, the *New York Times Magazine* characterized the Human Genome Project as the "biggest, costliest, most provocative biomedical research project in history."[2] But in the years between the project's launch in 1990 and its completion in 2003, genomic technology advanced dramatically. DNA-sequencing throughput increased from 1000 base pairs per day to more than 1000 base pairs per second, which opened the door for low-cost sequencing techniques that are enabling genomic advances to be incorporated into routine medical care. Genomic research has evolved from seeking to understand the fundamentals of the human genetic code to examining the ways in which this code varies among people, and then applying this knowledge to interventions that are tailored to target, with precision, the underlying causes of disease.

The development of genomic tools and data sets has transformed the nature of medical discovery, enabling scientists to undertake comprehensive and powerful explorations rather than being confined to testing hypotheses focused on candidate pathways. With the completion of the first reference sequence of the human genome,[3] attention shifted from searching for genes to discovering their functions. Systematic genetic mapping in families and populations helped scientists pinpoint the genetic variants that contribute to human disease.

The effects have been profound.



Highlights in Human Molecular Genetics and Genomics.

Cas9 denotes CRISPR-associated protein 9, CRISPR clustered regularly interspaced short palindromic repeats, HGP Human Genome Project, and NIH National Institutes of Health.

The discovery of genes responsible for more than 5000 rare mendelian diseases has facilitated genetic diagnostics for many patients, pregnancy-related counseling, new drug treatments, and in some cases, gene therapies. The discovery of more than 100,000 robust associations between genomic regions and common diseases has pointed to new biologic mechanisms, such as the role of microglia in Alzheimer's disease, autophagy in inflammatory bowel disease, and synaptic pruning in schizophrenia. It has also enabled the development of polygenic risk scores to identify patients at increased risk for heart disease, breast cancer, and other conditions, although additional rigorous testing of such scores is needed, including evaluation of clinical outcomes. Studies of cancer genomes have revealed hundreds of genes in which somatic mutations propel tumor initiation and growth, information that has fueled the development of new drugs. Genomic analysis is also helping to explain why some people have responses to certain therapies or survive certain infections, whereas others do not.

The focus of genomics research has recently moved beyond analyzing DNA variation to studying patterns of gene expression in individual cells, a step that has been driven by new methods for single-cell RNA sequencing and chromatin analysis. Tens of millions of cells have been characterized thus far en route to a complete cell atlas of the human body. This effort is revealing hundreds of new cell types and characterizing the ways in which cell types differ between healthy people and people with various diseases.

With the cost of sequencing a complete genome having dropped from $3 billion during the Human Genome Project to $600 today, there are growing efforts to create large-scale biobanks of complete genome-sequencing and phenotype information from hundreds of thousands of people. Examples include the U.K. Biobank (https://www.ukbiobank.ac.uk/) and the U.S. "All of Us" research program (https://allofus.nih.gov/). The ultimate goal is for health care systems to couple genomic information with medical records.

Much work remains to be done to enhance the study of human genetic variation. Despite the promise of insights into biology and health disparities offered by studying people of diverse backgrounds, both the investigators and the participants involved in genomic research have largely been of European ancestry. This lack of diversity hinders our understanding of biology, exacerbates already unacceptable health disparities, and raises the question of whether polygenic risk scores, diagnostics, and therapeutics derived from genomic research will benefit all populations equally.

Studies of people of diverse ancestral backgrounds have revealed the ways in which genomic variation contributes to population-level differences in disease susceptibility, drug responses, and the diagnostic accuracy of clinical approaches guided by genomic research.[4] One insight from such studies has been the identification of African ancestry–specific *APOL1* variants that protect against African sleeping sickness but increase the risk of kidney failure; these variants account for about 70% of cases of nondiabetic kidney failure in people of African ancestry in the United States. In addition, a risk haplotype for type 2 diabetes in *SLC16A11* that is present in about half of Indigenous peoples of the Americas and rare in peoples of European or African ancestry explains about 20% of the increased type 2 diabetes prevalence among Mexican Americans as compared with European Americans. Genomic research has also shown that *PCSK9* loss-of-function mutations are more common in people of African ancestry than in other populations; such mutations reduce cholesterol levels and the risk of heart disease and are providing new insights for drug development.

Several initiatives designed to increase the involvement of investigators and study participants from previously underrepresented populations are under way. The Human Heredity and Health in Africa (H3Africa) initiative (https://h3africa.org/), for example, has developed a pan-African consortium of laboratories that has ensured access to genomic technologies for more than 500 African scientists, enrolled more than 60,000 research participants, and established a bioinformatics network and three regional biorepositories. Precision medicine's benefits are expected to be more equitably shared when long-overdue steps are taken to close gaps in genomic-research participation.

Diagnosis of genetic diseases has advanced rapidly because of genomic-sequencing technology. But developing and validating treatments has been more challenging. For some mendelian disorders, molecularly targeted drugs have been developed using detailed understanding of pathophysiology. Thanks to work building on the 1989 discovery of the *CFTR* gene, for example, safe and

effective molecularly targeted drugs can be offered to 90% of people with cystic fibrosis. But this approach can take decades and doesn't scale well to the thousands of genetic disorders for which the precise molecular cause is known. A strategy enabling effective treatment targeted directly at the gene would have important advantages. After years of ups and downs, some dramatic successes of gene therapy are emerging, such as for spinal muscular atrophy and hemophilia. The pace of this research could increase dramatically in the future; precisely targeted genome-editing technologies now provide new avenues to therapeutics.

Over the past 8 years, CRISPR (clustered regularly interspaced short palindromic repeats)–Cas9 (CRISPR-associated protein 9) technologies have emerged as accessible and adaptable tools for studying and altering genomes.[5] CRISPR-Cas9 can be used to induce genome edits by creating targeted DNA breaks that trigger site-specific DNA repair. In next-generation formats, it can also control the transcriptional output of genes or alter genome sequences using a process of nucleotide base editing that does not require repair of DNA breaks. As these technologies continue to mature, it will become increasingly possible to alter cellular genomes efficiently and accurately.

An audio interview with Dr. Rotimi is available at NEJM.org

Coming on the heels of engineered nucleases, CRISPR-Cas9 tools have accelerated the pace of genomic research by permitting highly efficient knockouts or edits of virtually any gene in cells or model organisms. Multiple CRISPR-Cas9–based clinical trials are in progress or are expected to begin soon. Although Cas9-engineered cells haven't yet demonstrated efficacy at scale, early trial results suggest that such cells are stable and don't cause acute adverse reactions in humans. Long-term safety is yet to be determined. Current applications largely focus on single-gene disorders for which gene editing can be carried out ex vivo on appropriate cells, such as bone marrow hematopoietic stem cells in the case of sickle cell anemia. Exploration is under way to develop delivery systems that can target the gene-editing apparatus to the appropriate tissue in vivo.

Genomic-technology advances will continue to move basic science forward in powerful ways, not all of which can be anticipated. Already, current trainees cannot imagine how research in human biology was ever done without immediate and free access to vast quantities of data on genomes, transcriptomes, and chromatin marks — and increasingly these data are available for single cells. Such data sets will provide opportunities for biologic insights that require sophisticated computational analysis, for which all biologists will need to be prepared.

Understanding of human genetic variation and its biologic consequences will also advance, which will provide the groundwork for the International Common Disease Alliance's Maps to Mechanisms to Medicine vision (https://www.icda.bio/). Increasing abilities to target genetic mutations in vivo with oligonucleotides or gene editing should put many mendelian disorders within reach for therapies — and maybe cures.

Many uncertainties remain, however, and not all the big questions can be answered by science alone. For example, how do we ensure equitable and inclusive access to the research opportunities and the benefits of therapies generated by the genomic revolution? How do we balance scientific progress with emerging ethical issues, such as questions regarding the use of genome-editing technologies for heritable genetic changes? And how do we sustain the data-sharing ethos that has fueled genomic science while also protecting participants' privacy and respecting cultural norms? Such questions underscore why the NAM will be needed more than ever over the next 50 years.

The series editors are Victor J. Dzau, M.D., Harvey V. Fineberg, M.D., Ph.D., Kenneth I. Shine, M.D., Samuel O. Thier, M.D., Debra Malina, Ph.D., and Stephen Morrissey, Ph.D.

Disclosure forms provided by the authors are available at NEJM.org.

From the National Institutes of Health, Bethesda, MD (F.S.C., C.N.R.); the Department of Molecular and Cell Biology and the Department of Chemistry, University of California, Berkeley, Berkeley (J.A.D.); and the Broad Institute of the Massachusetts Institute of Technology and Harvard, Cambridge, MA (E.S.L.).

This article was published on January 2, 2021, at NEJM.org.

1. Cook-Deegan RM. The gene wars: science, politics, and the human genome. New York: W.W. Norton & Company, 1994.
2. Kanigel R. The genome project. New York Times Magazine. December 13, 1987 (https://www.nytimes.com/1987/12/13/magazine/the-genome-project.html).
3. Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. Science 2003;300:286-90.
4. Hindorff LA, Bonham VL, Brody LC, et al. Prioritizing diversity in human genomics research. Nat Rev Genet 2018;19:175-85.
5. Broeders M, Herrero-Hernandez P, Ernst MPT, van der Ploeg AT, Pijnappel WWMP. Sharpening the molecular scissors: advances in gene-editing technology. iScience 2020;23:100789.

DOI: 10.1056/NEJMp2030694