OPINION

Has the revolution arrived?

Looking back over the past decade of human genomics, **Francis Collins** finds five key lessons for the future of personalized medicine — for technology, policy, partnerships and pharmacogenomics.

On 26 June 2000, Craig Venter and I stood next to the President of the United States, in the same room of the White House where the explorers Meriwether Lewis and William Clark had unfurled their map of the Northwest Territories for Thomas Jefferson. "Today," Bill Clinton said, "the world is joining us here in the East Room to behold a map of even greater significance. We are here to celebrate the completion of the first survey of the entire human genome … With this profound new knowledge, humankind is on the verge of gaining immense, new power to heal. Genome science will have a real impact on all our lives — and even more, on the lives of our children. It will revolutionize the diagnosis, prevention, and treatment of most, if not all, human diseases."

I was honoured to be standing there, but also somewhat embarrassed: the milestone being reported was not yet attached to a publication — there was a lot of analysis still to do, and the paper would not appear in *Nature* until eight months later. Still, it was a heady moment.

Wisely, the president did not attach timetables to his bold predictions, even though in the early days of the millennium, everyone wanted to hear where this genome revolution was going. I even made my own predictions for 2010. Never having discarded a PowerPoint file, I can reproduce my list verbatim:

- Predictive genetic tests will be available for a dozen conditions
- Interventions to reduce risk will be available for several of these
- Many primary-care providers will begin to practise genetic medicine
- Preimplantation genetic diagnosis will be widely available, and its limits will be fiercely debated
- A ban on genetic discrimination will be in place in the United States
- Access to genetic medicine will remain inequitable, especially in the developing world

It is fair to say that all of these predictions have come true, with some caveats that offer important lessons about the best path forward for genomics and personalized medicine. The promise of a revolution in human health remains quite real. Those who somehow expected dramatic results overnight may be disappointed, but should remember that genomics obeys the First Law of Technology: we invariably overestimate the short-term impacts of new technologies and underestimate their longer-term effects.

Breathtaking acceleration

The decade from 2000 to 2010 was characterized by breathtaking acceleration in genome science. Thanks to advances in DNA sequencing technology that dropped the cost approximately 14,000-fold between 1999 and 2009, finished sequences are now available for 14 mammals, and draft or complete sequences have been done for many other vertebrates, invertebrates, fungi, plants and microorganisms. Comparative genomics has emerged as a powerful approach for understanding evolution and genome function at a level of detail barely imagined a few years ago.

For humans, the HapMap project produced a remarkable catalogue of common variation in the genome in just three years, from 2002 to 2005. As full sequencing has become more practical, researchers have been releasing complete genomes of individuals — a total of 13 at the time of this writing, including my personal hero, Archbishop Desmond Tutu of South Africa. In 2011, an international team is set to complete the data-production phase of the 1000 Genomes Project, designed to produce highly accurate assembled sequences from more than 1,000 individuals whose ancestors came from Europe, Asia and Africa.

The same determination to study the entire genome, not just isolated segments, has now been applied to understanding its function — although this quest is, of course, much more complicated and open-ended. The Encyclopedia of DNA Elements (ENCODE) project (started in pilot form in 2003 and slated to run at least until 2011) and the US National Institutes of Health (NIH) Roadmap Epigenomics Program (started in 2008 and funded until 2013) continue to define the 'parts list' of the human genome. These projects identify the locations of genes (protein coding and non-coding) and the patterns that determine whether genes are switched on or off in a given tissue — patterns of chromatin modification, transcription factors and DNA methylation.

With regard to medical applications, genome-wide association studies (GWAS) have now revealed an astounding number of common DNA variations that play a part in the risk of developing common diseases such as heart disease, diabetes, cancer or autoimmunity. To identify less common variations, methods to target DNA sequencing to subsets of the human genome have been developed. These methods can now sequence 80–90% of the protein-coding regions — the exons or 'exome' — of a human DNA sample for just a few thousand dollars.

Genome research has already had a profound impact on scientific progress. The combination of new technologies and freely accessible databases of high-quality genomic information has enabled the average investigator to make discoveries much more quickly than would otherwise have been possible. For example, the search for the cystic fibrosis gene finally succeeded in 1989 after years of effort by my lab and several others, at an estimated cost of US$50 million.

> **"This profusion of therapeutic opportunites is a challenge to prioritize."**

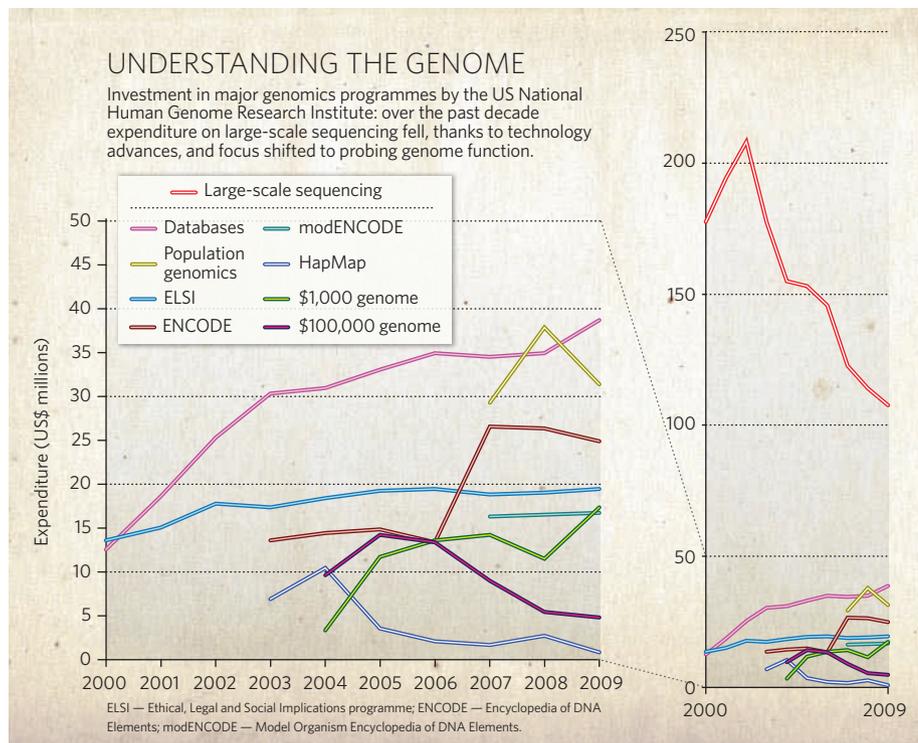Such a project could now be accomplished in a few days by a good graduate student with access to the Internet, appropriate DNA samples, some inexpensive reagents, a thermal cycler and a DNA sequencer (see graphic).

The consequences for clinical medicine, however, have thus far been modest. Some major advances have indeed been made: powerful new drugs have been developed for some cancers; genetic tests can predict whether people with breast cancer need chemotherapy; the major risk factors for macular degeneration have been identified; and drug response can be predicted accurately for more than a dozen drugs. But it is fair to say that the Human Genome Project has not yet directly affected the health care of most individuals.

GWAS have so far identified only a small fraction of the heritability of common diseases, so the ability to make meaningful predictions is still quite limited, even using chips that test for a million or more common variants. Nonetheless, direct-to-consumer marketing of genetic risk prediction, based on the rapidly growing

M. GUYER, NHGRI

UNDERSTANDING THE GENOME

Investment in major genomics programmes by the US National Human Genome Research Institute: over the past decade expenditure on large-scale sequencing fell, thanks to technology advances, and focus shifted to probing genome function.

ELSI — Ethical, Legal and Social Implications programme; ENCODE — Encyclopedia of DNA Elements; modENCODE — Model Organism Encyclopedia of DNA Elements.

database of GWAS results, is attracting early adopters. Having gone through that process myself, I can report that I found the opportunity to view my own personal genotype results rather riveting, despite the limited clinical validity and utility of many of these predictions.

This dynamic is likely to change in the next five years. Much of the missing heritability (the 'dark matter' of the genome) will probably turn up as the technology advances. Whole-genome sequencing, coming into its own as the cost per genome falls below $1,000 in the next three to five years, will identify rare variants of larger effect and the copy number variants that GWAS may have missed. With an increasing inventory of these discoveries, prediction of disease risk and drug response will continue to improve.

As the cost falls and evidence grows, there will be increasing merit in obtaining complete-genome sequences for each of us, and storing that information, with appropriate privacy protections, in our medical records, where it will be quickly available to guide prevention strategies or medication choice.

Perhaps the most profound consequence of the genome revolution in the long run will be the development of targeted therapeutics based on a detailed molecular understanding of pathogenesis. However, this is also the goal most challenged by long timelines, high failure rates and exorbitant costs. Despite those obstacles, inspiring examples of success are in hand, many of them (trastuzumab, imatinib, gefitinib and erlotinib) for the treatment of cancer. Furthermore,

the identification of new cancer drug targets is accelerating rapidly as a consequence of the ability to do deep genome sequencing of many tumours to identify recurrent mutations. Projects such as the Cancer Genome Atlas, which is carrying out the equivalent of 20,000 genome projects on matched tumour and blood DNA samples from 20 common types of cancer, have begun to reveal numerous opportunities for therapeutic development. And GWAS have pointed to hundreds of previously unrecognized drug targets for dozens of other diseases.

This profusion of therapeutic opportunities is a challenge to prioritize. Efforts are now under way to forge innovative partnerships between the traditional strengths of the private sector and academic labs. The NIH has provided new resources to catalyse such partnerships, including access by academic investigators to high-throughput screening through the Molecular Libraries Roadmap project, and to preclinical testing of promising lead compounds through the Therapeutics for Rare and Neglected Disease initiative.

Enabling the future

I propose five major lessons that could be gleaned from this first decade of the genome era. First, free and open access to genome data has had a profoundly positive effect on progress. The radical ethic of immediate data deposit, adopted by the Human Genome Project in 1996 and now the norm for other community resource projects, empowers the best brains

on the planet to begin work immediately in analysing the massive amounts of genomic data now being produced. It is a very good thing that the 'race for the genome' in 1998–2000 resulted in the human genome sequence being immediately and freely available to all, rather than becoming a commercial commodity.

Second, technology development for sequencing and functional genomics — key to the success achieved thus far — must continue to be a major focus of investment by both the public and private sectors. Although huge leaps have been made in increasing the speed and reducing the costs of DNA sequencing, expression analysis and methods to assess the epigenome, the limits are still nowhere near being reached.

Third, the success of personalized medicine will depend on continued accurate identification of genetic and environmental risk factors, and the ability to utilize this information in the real world to influence health behaviours and achieve better outcomes. This will require well designed, large-scale research projects, for discovering risk factors and for testing the implementation of prevention and pharmacogenomic programmes.

Fourth, achieving the enormous promise of the myriad new drug targets emerging from genomic analysis of common and rare diseases requires new paradigms of public–private partnership. Academic investigators will have a much more important role in the early stages, effectively 'de-risking' projects for downstream commercial investment. Closer relationships between the US Food and Drug Administration and the NIH, announced this February, will assist this process.

Finally, good policy decisions will be crucial to reaping the benefits that should flow from the coming revelations about the genome. These will include protection of individual privacy, effective education of health-care providers and the public about genomic medicine, and appropriate health-care system reimbursement for the cost of validated preventive measures.

In *The Wisdom of the Sands*, author Antoine de Saint-Exupéry wrote: "As for the future, your task is not to foresee, but to enable it." Genomics has had an exceptionally powerful enabling role in biomedical advances over the past decade. Only time will tell how deep and how far that power will take us. I am willing to bet that the best is yet to come.   ∎

Francis Collins is director of the National Institutes of Health, Bethesda, Maryland 20892, USA. Between 1993 and 2008 he was director of the National Human Genome Research Institute. e-mail: francis.collins@nih.gov