

The  
Economist

A special report on the human genome

## Biology 2.0

A decade after the human-genome project, writes Geoffrey Carr (interviewed here), biological science is poised on the edge of something wonderful

Jun 17th 2010 | from PRINT EDITION

0 Like



TEN years ago, on June 26th 2000, a race ended. The result was declared a dead heat and both runners won the prize of shaking the hand of America's then president, Bill Clinton, at the White House. The runners were J. Craig Venter for the private sector and Francis Collins for the public. The race was to sequence the human genome, all 3 billion genetic letters of it, and thus—as headline writers put it—read the book of life.

It quite caught the public imagination at the time. There was the drama of a maverick upstart, in the form of Dr Venter and his newly created firm, Celera, taking on the medical establishment, in the form of Dr Collins's International Human Genome Sequencing Consortium. There was the promise of a cornucopia of new drugs as genetic targets previously unknown to biologists succumbed to pharmacological investigation. There was talk of an era of "personalised medicine" in which treatments would be tailored to an individual's genetic make-up. There was the frisson of fear that a genetic helotry would be created, doomed by its DNA to second-class health care, education and employment. And there was, in some quarters, a hope that a biotech boom based on genomics might pick up the baton that the internet boom had just dropped, and that lots and lots of money would be made.

And then it all went terribly quiet. The drugs did not appear. Nor did personalised medicine. Neither did the genetic underclass. And the money certainly did not materialise. Biotech firms proved to be just as good at consuming cash as dotcom start-ups, and with as little return. The casual observer, then, might be forgiven for thinking the whole thing a damp squib, and the \$3 billion spent on the project to be so much wasted money. But the casual observer would be wrong. As *The Economist* observed at the time, the race Dr Venter and Dr Collins had been engaged in was a race not to the finish but to the starting line. Moreover, compared with the sprint they had been running in the closing years of the 1990s, the new race marked by that starting line was a marathon.

The new race has been dogged by difficulties from the beginning. There was a false start (the announcement at the White House that the sequence was complete relied on a generous definition of that word: a truly complete sequence was not published until 2003). The competitors then ran into numerous obstacles that nature had strewn on the course. They found at first that there were far fewer genes than they had expected, only to discover later

that there were far more. These discoveries changed the meaning of the word "gene". They found the way genes are switched on and off is at least as important, both biologically and medically, as the composition of those genes. They found that their methods for linking genetic variation to disease were inadequate. And they found, above all, that they did not have enough genomes to work on. Each human genome is different, and that matters.

### **All is revealed**

One by one, however, these obstacles are falling away. As they do so, the science of biology is being transformed. It seems quite likely that future historians of science will divide biology into the pre- and post-genomic eras.

In one way, post-genomic biology—biology 2.0, if you like—has finally killed the idea of vitalism, the persistent belief that to explain how living things work, something more is needed than just an understanding of their physics and chemistry. True, no biologist has really believed in vitalism for more than a century. Nevertheless, the promise of genomics, that the parts list of a cell and, by extension, of a living organism, is finite and cataloguable, leaves no room for ghosts in the machine.

Viewed another way, though, biology 2.0 is actually neo-vitalistic. No one thinks that a computer is anything more than the sum of its continually changing physical states, yet those states can be abstracted into concepts and processed by a branch of learning that has come to be known as information science, independently of the shifting pattern of electrical charges inside the computer's processor.

So it is with the new biology. The chemicals in a cell are the hardware. The information encoded in the DNA is the preloaded software. The interactions between the cellular chemicals are like the constantly changing states of processing and memory chips. Though understanding the genome has proved more complicated than expected, no discovery made so far suggests anything other than that all the information needed to make a cell is squirreled away in the DNA. Yet the whole is somehow greater than the sum of its parts.

Whether the new biology is viewed as rigorously mechanistic or neo-vitalistic, what has become apparent over the past decade is that the process by which the genome regulates itself, both directly by one gene telling another what to do and indirectly by manipulating the other molecules in a cell, is vastly more complicated and sophisticated than anybody expected. Yet it now looks tractable in a way that 20 years ago it did not. Just as a team of engineers, given a rival's computer, could strip it down and understand it perfectly, so biologists now believe that, in the fullness of time, they will be able to understand perfectly how a cell works.

And if cells can be understood completely in this way, then ultimately it should be possible to understand assemblages of cells such as animals and plants with equal completeness. That is a much more complicated problem, but it is different only in degree, not kind. Moreover, understanding—complete or partial—brings the possibility of manipulation. The past few weeks have seen an announcement that may, in retrospect, turn out to have been as portentous as the sequencing of the human genome: Dr Venter's construction of an organism with a completely synthetic genome. The ability to write new genomes in this way brings true biological engineering—as opposed to the tinkering that passes for biotechnology at the moment—a step closer.

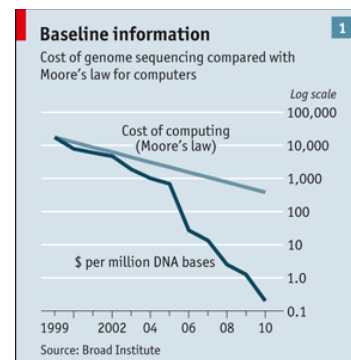
A second portentous announcement, of the genome of mankind's closest—albeit extinct—relative, Neanderthal man, shows the power of biology 2.0 in a different way. Putting together some 1.3 billion fragments of 40,000-year-old DNA, contaminated as they were with the fungi and bacteria of millennia of decay and the personal genetic imprints of the dozens of archaeologists who had handled the bones, demonstrates how far the technology of genomics has advanced over the course of the past decade. It also shows that biology 2.0 can solve the other great question besides how life works: how it has evolved and diversified over the course of time.

As is often the way with scientific discovery, technological breakthroughs of the sort that have given science the Neanderthal genome have been as important to the development of genomics as intellectual insights have been. The telescope revolutionised astronomy; the microscope, biology; and the spectroscope, chemistry. The genomic revolution depends on two technological changes. One, in computing power, is generic—though computer-makers are slaving at the amount of data that biology 2.0 will need to process, and the amount of kit that will be needed to do the processing. This torrent of data, however, is the result of the second technological change that is driving genomics, in the power of DNA sequencing.

### The new law

Computing has, famously, increased in potency according to Moore's law. This says that computers double in power roughly every two years—an increase of more than 30 times over the course of a decade, with concomitant reductions in cost.

There is, as yet, no sobriquet for its genomic equivalent, but there should be. Eric Lander, the head of the Broad Institute, in Cambridge, Massachusetts, which is America's largest DNA-sequencing centre, calculates that the cost of DNA sequencing at the institute has fallen to a hundred-thousandth of what it was a decade ago (see chart 1). The genome sequenced by the International Human Genome Sequencing Consortium (actually a composite from several individuals) took 13 years and cost \$3 billion. Now, using the latest sequencers from Illumina, of San Diego, California, a human genome can be read in eight days at a cost of about \$10,000. Nor is that the end of the story. Another Californian firm, Pacific Biosciences, of Menlo Park, has a technology that can read genomes from single DNA molecules. It thinks that in three years' time this will be able to map a human genome in 15 minutes for less than \$1,000. And a rival technology being developed in Britain by Oxford Nanopore Technologies aspires to similar speeds and cost.



This increase in speed and reduction in cost is turning the business of biology upside down. Up until now, firms that claim to read individual genomes (see [article](#)) have been using a shortcut. They have employed arrays of DNA probes, known as gene chips, to look for pre-identified variations in their clients' DNA. Those variations have been discovered by scientific collaborations such as the International HapMap Project, which search for mutations of the genetic code called single-nucleotide polymorphisms, or SNPs, in blocks of DNA called haplotypes. A SNP (pronounced "snip") is a place where a lone genetic letter varies from person to person. Some 10m SNPs are now known, but in the forest of 3 billion genetic letters there is reason to believe they are but a smattering of the total variation. Proper sequencing will reveal the lot.

Finding the sequence—even the full range of sequences—is, though, just the beginning. You then have to do something useful with the result. This is where the computing comes in. Computers allow individual genomes—all 3 billion base pairs of them—to be compared. And not only human genomes. Cross-species comparisons are enormously valuable. Laboratory experiments on creatures ranging from yeast to mice can reveal the functions of genes in these species. Computer comparison then shows which human genes correspond in DNA sequence and thus, presumably, in function, to the genes in these "model" organisms.

Cross-species comparison also shows how species differ, and thus how they have diverged. Comparing DNA from populations within a species can show how that species is evolving. Comparing DNA from individuals within a population can explain why those individuals differ from one another. And comparing the DNA from cells within an individual can show how tissues develop and become differentiated from one another, and what goes wrong in diseases like cancer.

Even before cheap sequencing became available, huge databases were being built up. In alliance with pathology samples, doctors' notes and—most valuable of all—long-term studies of particular groups of individuals, genetic information can be linked to what biologists refer to as the phenotype. This is an organism's outward expression: its anatomy, physiology and behaviour, whether healthy or pathological. The goal of the new biology is to tie these things together reliably and to understand how the phenotype emerges from the genotype.

That will lead to better medical diagnosis and treatment. It will result in the ability to manipulate animals, plants, fungi and bacteria to human ends. It will explain the history of life. And it will reveal, in pitiless detail, exactly what it is to be human.

Special reports

0

Like

[About \*The Economist\*](#) [Media directory](#) [Staff books](#) [Career opportunities](#) [Contact us](#) [Subscribe](#)

[\[+\] Site feedback](#)

Copyright © The Economist Newspaper Limited 2010. All rights reserved. [Advertising info](#) [Legal disclaimer](#) [Accessibility](#) [Privacy policy](#) [Terms of use](#)

[Help](#)