



# SCIENCE AFTER THE SEQUENCE

The completion of the draft human genome sequence was announced ten years ago. *Nature's* survey of life scientists reveals that biology will never be the same again. **Declan Butler** reports.

“With this profound new knowledge, humankind is on the verge of gaining immense, new power to heal. It will revolutionize the diagnosis, prevention and treatment of most, if not all, human diseases.” So declared then US President Bill Clinton in the East Room of the White House on 26 June 2000, at an event held to hail the completion of the first draft assemblies of the human genome sequence by two fierce rivals, the publicly funded international Human Genome Project and its private-sector competitor Celera Genomics of Rockville, Maryland (see *Nature* **405**, 983–984; 2000).

Ten years on, the hoped-for revolution against human disease has not arrived — and *Nature's* poll of more than 1,000 life scientists shows that most don't anticipate that it will for decades to come ([go.nature.com/3Aywn](http://go.nature.com/3Aywn)). What the sequence has brought about, however, is a revolution in biology. It has transformed the professional lives of scientists, inspiring them to tackle new biological problems and throwing up some acute new challenges along the way.

Almost all biologists surveyed have been influenced in some way by the availability of the human genome sequence. A whopping 69% of those who responded to *Nature's* poll say that the human genome projects inspired them either to become a scientist or to change the direction of their research. Some 90% say that their own research has benefited from the sequencing of human genomes — with 46% saying that it has done so “significantly”. And almost one-third use the sequence “almost daily” in their research. “For young researchers like me it's hard to imagine how biologists

managed without it,” wrote one scientist.

The survey, which drew most participants through *Nature's* print edition and website and was intended as a rough measure of opinion, also revealed how researchers are confronting the increasing availability of information about their own genomes. Some 15% of respondents say that they have taken a genetic test in a medical setting, and almost one in ten has used a direct-to-consumer genetic testing service. When asked what they would sequence if they

could sequence anything, many respondents listed their own genomes, their children's or those of other members of their family (the list also included a few pet dogs and cats).

Some are clearly impatient for this opportunity: about 13% say that they have already sequenced and analysed part of their own DNA. One in five

said they would have their entire genome sequenced if it cost US\$1,000, and about 60% would do it for \$100 or if the service were offered free. Others are far more circumspect about sequencing their genome — about 17% ticked the box saying “I wouldn't do it even if someone paid me”.

*Nature's* poll also gauged where the sequence has had the greatest effect on the science itself. Although nearly 60% of those polled said

they thought that basic biological science had benefited significantly from human genome sequences, only about 20% felt the same was true for clinical medicine. And our respondents acknowledged that interpreting the sequence is proving to be a far greater challenge than deciphering it. About one-third of respondents listed the field's lack of basic understanding of genome biology as one of the main obstacles to making use of sequence data today.

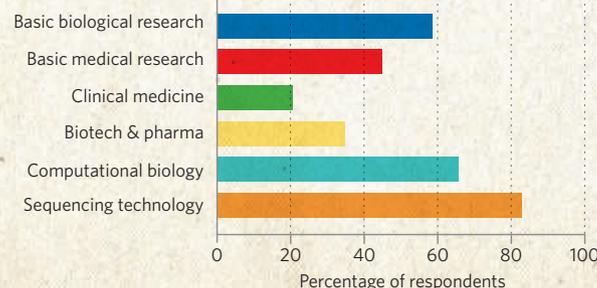
## Sequence is just the start

Studies over the past decade have revealed that the complexity of the genome, and indeed almost every aspect of human biology, is far greater than was previously thought (see *Nature* **464**, 664–667; 2010). It has been relatively straightforward, for example, to identify the 20,000 or so protein-coding genes, which make up around 1.5% of the genome. But knowing this, researchers note, does not necessarily explain what those genes do, given that many genes code for multiple forms of a protein, each of which could have a different role in a variety of biological processes. “The total sequence was needed, I think, to allow us to see that our one gene–one protein model of genetics was much too simplistic,” wrote one respondent.

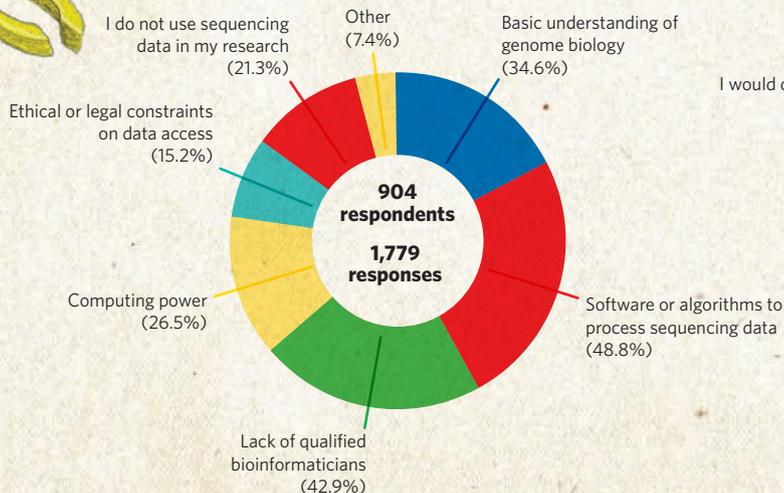
A decade of post-genomic biology has also focused new attention on the regions outside protein-coding genes, many of which are likely to have key functions, through regulating the expression of protein-coding genes and by making a slew of non-coding RNA molecules. “Now we understand,” wrote another survey respondent, “that, without looking at the dynamics of a genome, determining its sequence is of limited use.” Some big projects are under way to fill in the gaps, including the Encyclopedia of DNA Elements (ENCODE) and the Human Epigenome Project, an effort to understand the chemical modifications of the genome that are now thought to be a major means of controlling gene expression.

**69%**  
were inspired by the genome to become a scientist or change their research direction.

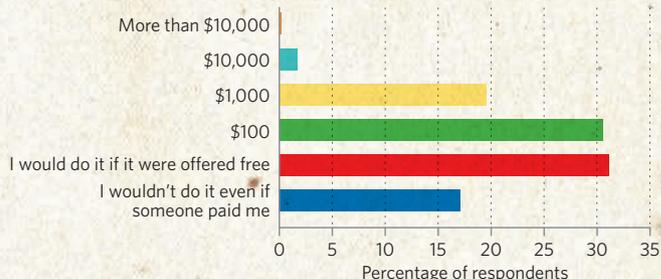
Which areas have benefited significantly from the sequencing of human genomes?



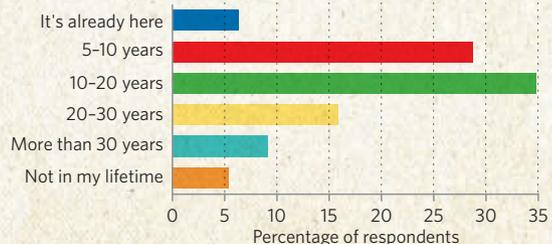
**What are the main obstacles to making use of the flood of sequencing data? More than one answer can be selected.**



**How much would you pay to have your genome sequenced?**



**How soon do you expect personalized medicine, based on human genetic information, to become commonplace?**



The biggest effects of the genome sequence, according to the poll, have been advances in the tools of the trade: sequencing technologies and computational biology. Technological innovation has sent the cost of sequencing tumbling, and the daily output of sequence has soared (see *Nature* 464, 670–671; 2010). “Deep sequencing technology is now becoming a staple of scientific research. Would this have occurred if it wasn’t for the technological push required to finish the human genome?” read one response.

**Data dreams, analysis nightmares**

Cheaper and faster sequencing has brought its own problems, however, and our survey revealed how ill-equipped many researchers feel to handle the exponentially increasing amounts of sequence data. The top concern — named by almost half of respondents — was the lack of adequate software or algorithms to analyse genomic data, followed closely by a shortage of qualified bioinformaticians and to a lesser extent raw computing power. Other concerns include data storage, the quality of sequencing data and the accuracy of genome assembly. Commenting on the survey results, David Lipman, director of the US National Center for Biotechnology Information in Bethesda, Maryland, says that the worries about data handling and analysis were an issue even in the earliest discussions of the genome project. Perhaps, he suggests, “there’s a sort of disappointment that despite having so much data, there is still so much we don’t understand”.

Eric Green, director of the National Human Genome Research Institute (NHGRI) in

Bethesda, says that the institute is well aware of the need for more bioinformatics experts, better software and a clearer understanding of how the differences between genomes influence human health. He says the institute is planning to publish in late 2010 its next strategic five-year plan for the genomics field. One possible solution to the computing challenge, which was discussed at an NHGRI workshop in late March, is cloud computing, in which laboratories buy computing power and storage in remote computing farms from companies such as Google, Amazon and Microsoft. The European Nucleotide Archive, launched on 10 May at the European Molecular Biology Laboratory’s European Bioinformatics Institute in Cambridge, UK, will also offer labs free remote storage of their genome data and use of bioinformatics tools.

Given ten years’ of hindsight and the current set of obstacles, it’s no surprise that researchers now state somewhat modest expectations for what human genomics can deliver and by when. The rationale for sequencing and exploring the human genome — to revolutionize the finding of new drugs, diagnostics and vaccines, and to tailor treatments to the genetic make-up of individuals — is the same today. But almost half of respondents now say that the benefits of the human genome were oversold in the lead up to 2000. “While I do feel that the gains made by the human genome project are extraordinary and affect my research significantly, I still feel that it was overhyped to the general

population,” read one typical response. More than one-third of respondents now predict that it will take 10–20 years for personalized medicine, based on genetic information, to become commonplace, and more than 25% even longer than that. Some 5% don’t expect it will happen in their lifetime. “Our understanding of the genome will not come in a single flash of insight. It will be an organized hierarchy of billions of smaller insights,” says David Haussler, head of the Genome Bioinformatics Group at the University of California, Santa Cruz.

Green says that when the Human Genome

Project was envisioned, scientific leaders of the day predicted that it would take 15 years to generate the first sequence, and a century for biologists to understand it. “I think they got that about right,” he says. “While we still don’t have all the answers — being a mere 10% of the way into the century

with a human genome sequence in hand — we have learned extraordinary things about how the human genome works and how alterations in it confer risk for disease.”

Haussler agrees. “All that happened in the first ten years is still just early rumblings of much more dramatic changes to come when we begin to truly understand the genome,” he says.

**Declan Butler is a senior correspondent at Nature based in France.**

**Survey work was aided by Sara Grimme.**

**See full survey results at [go.nature.com/3Ayuwn](http://go.nature.com/3Ayuwn).**

**See Nature’s human genome special at [www.nature.com/humangenome](http://www.nature.com/humangenome)**

**13%**  
have sequenced  
part of their  
own DNA.