

eGrants:

[About](#)

[Login](#)

Gibbs Building, 215 Euston Road, London NW1 2BE, UK T:+44 (0)20 7611 8888

[Home](#) > [About us](#) > [75th anniversary](#) > Sir John Sulston and the Human Genome Project



# Sir John Sulston and the Human Genome Project



Mark Henderson (Science Editor of 'The Times') looks at Sir John Sulston, a Nobel Prize winner and the first Director of the Wellcome Trust Sanger Institute.

At the Laboratory of Molecular Biology (LMB) in Cambridge in the 1980s, scientists liked to repeat a saying that encapsulated the approach to research that has brought its alumni 14 Nobel Prizes. Sir John Sulston, one of those Nobel Laureates, attributes it to another: the great Francis Crick. "There is no point," it runs, "wasting good thoughts on bad data."

The dictum served as an admonition against hubris, a reminder to the LMB's supremely bright theorists never to get too carried away with an unsupported hypothesis. Yet it could also be interpreted another way. To Sulston, it was an exhortation. If a lack of data was preventing good thoughts from flourishing, somebody would have to go out and remedy that.

It was to be an inspiration for his pivotal role in the sequencing of the human genome.

Ten years now after the first draft of the 3 billion DNA letters that make up our species's genetic code was published, the LMB motto still captures something fundamentally important, both about that remarkable achievement and about many of the 'big science' genomic initiatives that have followed it. The worth of these projects - many of them funded by the Wellcome Trust - lie not so much in the triumph of clever thinking, but in the clever thinking they would facilitate.

The Human Genome Project is often described as a transformative moment in medical science, which is ushering in a new era of healthcare. That isn't wrong - it has already brought new approaches to diagnosis and drug design that are changing the way many diseases, particularly cancer, are managed - but its real significance is subtler. The reference sequence of *Homo sapiens* has not, in and of itself, revealed very many medical insights. Rather, the scientists who generated it - people like John Sulston and Bob Waterston, Eric Lander and Francis Collins - created a profoundly valuable resource, which others could use to perform science that would otherwise have been inconceivable. They began to end the era of poor genomic data.

When the idea of sequencing the human genome first began to be floated in the mid-1980s, the stated ambition was indeed to change medical science. "If we wish to learn more about cancer, we must now concentrate on the cellular genome," said Renato Dulbecco, who is usually credited along with Robert Sinsheimer as the first to think big about human genomics. Their vision gradually persuaded the US Department of Energy (which had a historic interest in DNA because of its responsibility for the health impacts of radiation), then the US National Institutes of Health (NIH), then Britain's Medical Research Council (MRC), to start committing funds. The Human Genome Organisation (HUGO) was formally founded in 1988, and the international sequencing project it would manage began in 1990.

As Sulston recalls, however, the nascent field of human genomics was seriously afflicted by the problem summed up by the LMB aphorism: lots of good ideas, but little good data. "It was absolutely the problem with genetics back then," he says. "There was no shortage of exceptionally clever people trying to find genes in the human, but they were wasting their time theorising about the bits of the sequence that might be necessary. There was tremendous thought potential, but it needed data. We had the potential to produce good data, that could enable good thinking, but there was a lack of willingness to go out and get it."



Sir John Sulston at the private viewing of exhibits from the DNA 50 Imagine Photographic Competition displayed at the Two Ten Gallery. Credit: Wellcome Library

Going out to get data that could underpin the efforts of other molecular biologists was an activity of which Sulston and his collaborator Bob Waterston, of Washington University in St Louis, had unique experience. In the 1980s, they and their colleagues had mapped the genome of the nematode worm *Caenorhabditis elegans*, a commonly used laboratory model organism, and then moved on to sequencing it. This resource served as a shortcut for gene discovery, and Sulston got a thrill from watching others use it to make discovery after discovery. "It got rid of a huge bottleneck in biology," he says. "People were suddenly isolating genes in weeks rather than years, because of the map we'd given them."

As the sequencing of the worm progressed, the pair's skills were in considerable demand, and not only from HUGO. Frederick Bourke, an entrepreneur who had

made millions from leather goods, saw the commercial potential of creating a DNA library that every human geneticist would want to use, and approached Sulston and Waterston through Leroy Hood, the first scientist to automate the sequencing process. He wanted them to move to a new centre he was establishing in Seattle, with generous funding.

While they listened to Bourke, the offer on the table "was never a real prospect as it turned out," Sulston said. "We explored the possibility, but Rick Bourke had different goals to Bob and I. He was strictly about commercial sequencing of the human, and the worm was nowhere." When Sulston turned Bourke down in a hotel room in Berkeley, California, he remembers the entrepreneur saying: "I hope this isn't going to damage you, John." Nothing could have been further from the truth. "That someone else thought I was an attractive investment did me no harm, though I hasten to add that wasn't deliberate," Sulston says.

Bourke's offer proved catalytic when word of it reached Jim Watson, who had deciphered the double helix of DNA with Crick and who now headed the NIH's sequencing operations. "I worried that the NIH might lose its most successful genome-sequencing effort, and the UK government might abandon large-scale genome research," Watson said. The Genome Project would then lose the great intellectual resources nurtured by the MRC at the LMB. "I knew that John Sulston would prefer to stay in Cambridge, but he was dependent on procuring committed funding from a UK source."<sup>[1]</sup>

Watson couldn't give NIH money to a British-based scientist, but he could work his contacts, and he found receptive ears at the Wellcome Trust. The charity had recently sold 288 million shares in the pharmaceutical company Wellcome plc, raising £2.3 billion - at the time, the largest cheque ever written. It had the money to support Sulston's work and, thanks to the prompting of Watson and others such as the LMB's Aaron Klug, the inclination. Bridget Ogilvie, the Trust's new Director, agreed to support a new sequencing centre led by Sulston that would begin work on the human code. The centre would also play host to the completion of the worm sequencing, funded by the MRC.

The next big question, according to Michael Morgan, who was appointed to oversee the Trust's investment in sequencing, was where to put it. "We essentially toured the countryside," he said. "One site that sticks in my mind is a chicken farm, where John and I discussed putting sequencers in the coops! Eventually John found an abandoned scientific site at Hinxton Hall. The idea was to build a temporary facility, as no one at that time thought this was going to be a big deal. John submitted a grant application to set up a sequencing facility on the site, and subsequently, in 1992, the Trust made its biggest grant up to that point, of £46.5m. The Sanger Centre was born."<sup>[2]</sup>

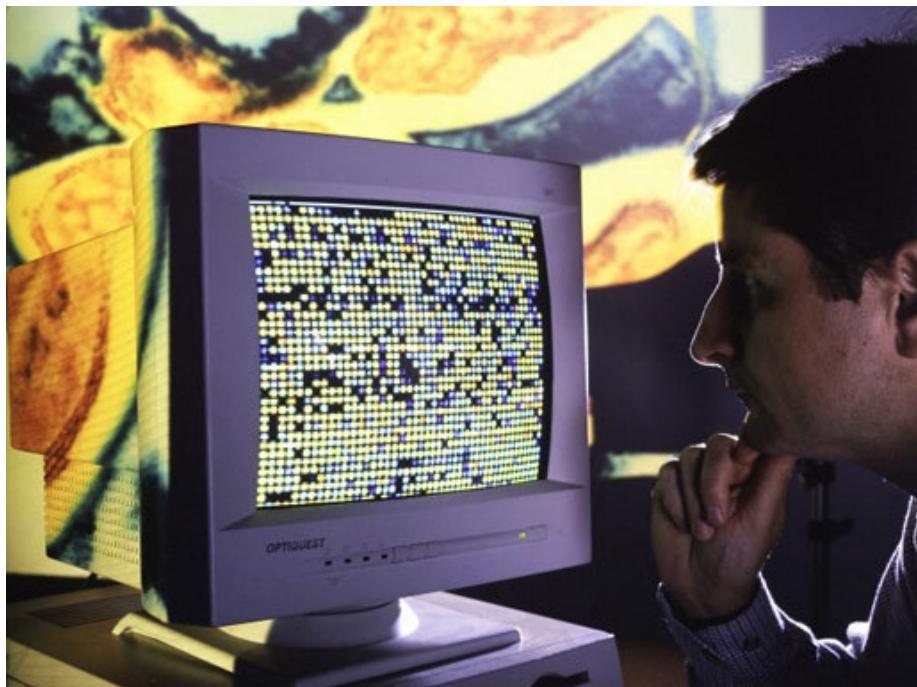
The Centre, now the Wellcome Trust Sanger Institute, took its name from Sir Fred Sanger, who had developed DNA sequencing in the 1970s and remains the only Briton to have won two Nobel Prizes. He was the obvious choice to be honoured, yet he is apt to shun publicity, and Sulston remembers asking nervously for his blessing. "It had better be good," the great man replied.

With the Sanger Centre up and running, Britain had a world-class genome institution to supplement the efforts of Waterston's lab in St Louis and other big US players, particularly those at Baylor and MIT. As human sequencing began, the parallel success of the worm project began to hint at just how much it might achieve. "People started to see that we were getting great sequence out, that was being put to good use," Sulston says. "The worm was a critical milestone.

Genetics was coming of age."

Another lesson from the worm was also to prove vital. As Sulston and Waterston had produced the worm's gene map and DNA sequence, they had released new data publicly along the way, allowing the research community to use the information as soon as it was ready. At the 1996 meeting of HUGO in Bermuda, the same policy was adopted for the human sequence. "The project wouldn't have been the same without that," Morgan says.

For Sulston, who drove through the data-release policy with Waterston and Morgan, it had both principled and practical merit. It was right, he thought, that DNA sequences ought not to belong to anyone. And regular release into the public domain would bring scientific discoveries sooner. "It was very apparent to me that it had to be released as we went along. If we were doing this to support biomedical applications, the data had to be shared."



A scientist examining the hybridisation pattern to a DNA microarray. Credit: Wellcome Library

This sharing philosophy, however, was not universal: in 1998, a new commercial player emerged. Craig Venter was a brilliant but somewhat maverick geneticist who had left the NIH and established a private sequencing company called Celera. His business plan - backed by PerkinElmer, the instruments giant that made the sequencers that would be needed to complete the human genome - was to race the public consortium to publish first. The Human Genome Project had a target date of 2005; Celera promised a "substantially complete" sequence by 2001.

Celera deployed a different technical approach to the public project, skipping a genome-mapping stage that ensured thoroughness at the expense of speed. A still greater difference lay in its philosophy. It aimed to patent 300 clinically important genes, and to charge subscribers to interpret the genomic data the company would hold and own. Venter often spoke of wanting to emulate Bloomberg, the financial information provider. This paid-for model posed a profound challenge to the public project: if Venter succeeded, the goal of

universal access agreed in Bermuda would fail. To prevent this, the initiative would need to scale up its sequencing effort quickly, to place maximum data in the public domain before Celera could stake a claim.

Once again, it was the Wellcome Trust that stepped up to the plate. The Sanger Centre had already submitted an application to accelerate its sequencing effort, and to take responsibility for one-third of the genome instead of one-sixth; within days of Celera's launch, Morgan found the funds.

"The Trust's intervention was absolutely critical," says Sulston: its aggressive counter-play bounced the NIH and DoE into increasing their support for the project, when some figures in the US were arguing for a deal with Celera. "Michael Morgan said Wellcome would fund half of it, or even all of it, if needs be, to keep it in the public domain," Sulston said. "I think it was the Trust's finest hour in many ways."

Under the revised plan, the public consortium would seek to publish a 'working draft' of the genome in 2001, and then finish the sequence later. They reached this interim finish line in virtually a dead heat with Celera: on June 26, 2000, Bill Clinton and Tony Blair announced that both groups had finished a working draft, and they published together in February 2001. Celera sold subscription access to interpretive software for a while, but eventually pulled out of genomics, and its sequence was released openly in 2004. The Wellcome Trust's move to force the pace of the public project had worked magnificently: the reference human genome sequence was available to all.

Celera is often credited with giving much-needed impetus to the public initiative, but Sulston disagrees. Celera's version was not only inferior, it diverted its rivals' resources towards speed rather than accuracy. "It was a distraction," he says. "We were already funded to provide a finished sequence by 2003, and that's exactly when we finished. The draft was a sideshow, but we had no choice. The reward for having a public genome out first was so huge we would do anything to keep it."

"I sat next to Max Perutz and Fred Sanger at Downing Street when they made the announcement, and one of them, I forget which, said to me: 'Why are we publishing something that's incomplete?' I said: 'This isn't science, it's politics.'"

Through his ability, and the Trust's, to play politics, Sulston and his colleagues had delivered the genomic data the biomedical research community craved. What was complete was an anatomical resource: as Norton Zinder, a founding member of HUGO, put it, it would do for genomics what Vesalius had done for anatomy. "This is the beginning of the beginning," Zinder told the 'New Yorker' in 2000.[3] "Before Vesalius, people didn't even know they had hearts and lungs. With the human genome, we finally know what's there, but we still have to figure out how it all works. Having the human genome is like having a copy of the Talmud but not knowing how to read Aramaic." The good thinking could begin.

One of the first examples came from a team led by Mike Stratton, now Director of the Sanger Institute. In 2002, they used the genome sequence to discover that a gene called BRAF is mutated in about half of all cases of malignant melanoma - the deadliest form of skin cancer. By 2008, a biotech company called Plexxicon had developed an agent, PLX4032, that can treat melanoma by inhibiting this BRAF mutation. It was licensed by Roche, which at the time of writing is preparing to publish final trial results that it claims show significant

survival benefits. The genome hasn't cured this form of cancer - most patients on PLX4032 still relapse. But it has made some types of advanced melanoma treatable for the first time.

Other direct medical benefits have been slower to materialise. Biological insights revealed by the genome, however, have proceeded at an exciting pace, as have initiatives to make the reference sequence still more useful as a resource. The Wellcome Trust has been in the vanguard, funding project after project to provide more and more good data.

The SNP Consortium, a £30m partnership between the Trust and the pharmaceutical industry, was one of the first. It involved comparing individual genomes to reveal 1.5m single nucleotide polymorphisms (SNPs) - single-letter spelling changes that commonly differ between people. This was followed in 2005 by the International Haplotype Map, or HapMap, which catalogues how these SNPs are generally inherited together. "The human genome sequence provided us with the list of many of the parts to make a human," explained Peter Donnelly, of the University of Oxford, a leader of the project. "The HapMap provides us with indicators - like Post-It notes - which we can focus on in looking for genes involved in common disease."<sup>[4]</sup>

The HapMap enabled a new approach to hunting for genes, called the genome-wide association study (GWAS). This involves comparing the genomes of thousands of people with and without a certain disease and looking for SNPs that are more common in cases or controls. One of the first large GWAS projects, the Wellcome Trust Case Control Consortium (which Donnelly led), published its first results in 2007, revealing dozens of new genetic variants that contribute to disease risk. These included FTO, a gene that in some versions makes people slightly more prone to obesity, and others linked to both types of diabetes, heart disease and rheumatoid arthritis. Hundreds more have since been identified using a similar approach, made possible by investment in resources like the Human Genome Project and the HapMap.

Another Wellcome Trust-funded initiative is now providing a still more detailed map of genetic variation from person to person. The 1000 Genomes Project, which published its first results in October 2010, has identified 15m SNPs, including about 95 per cent of the variants present in any given person. This catalogue of difference creates a valuable short-cut for research. At present, when scientists link a part of the genome to a disease, they must conduct detailed studies to pinpoint the precise DNA variant that causes the effect. They can now use the 1000 Genomes Project as an index, to look up SNPs that might be responsible.

Another exciting initiative has examined what was once labelled 'junk DNA' - the 98 per cent of the genome that does not code for proteins. The Encyclopedia of DNA Elements (ENCODE) consortium has revealed that much of this is in fact transcribed into RNA, suggesting that it plays an important biological role. Some of the junk isn't junk at all: it is intimately involved in controlling gene activity and expression. RNA, indeed, is turning out to be much more significant than was once thought, performing signalling functions well beyond the translation of DNA into protein set out in Crick's central dogma.

The International Cancer Genome Consortium is extending the type of work that led to Stratton's BRAF discovery, sequencing the entire genetic codes of thousands of tumour samples in search of the mutations that drive their unchecked growth. Most recently, the Wellcome Trust has supported a move to

bring genomics to low-income countries, through the Human Heredity and Health in Africa project.

Despite all this activity, most of the challenges of understanding the genome remain to be resolved. We do not yet know what explains the 'missing heritability' - the way the variants unlocked by GWAS research account for only a fraction of the known inherited components of risk. Genetic insights into the pathways behind disease have yet to yield many new drugs, with a few exceptions such as PLX4032. Genomic variation is now known to affect the way patients respond to many medicines - including Plavix, the world's second-biggest-selling drug - but the clinical implications of this are still unclear. Falling sequencing costs will probably bring us the \$1000 genome within a year or two, but interpreting this information to improve healthcare will be difficult and expensive.

The genomic research supported by the Wellcome Trust can be expected to allow many of these questions to be answered. It is providing ever-better data, so that fewer and fewer good thoughts about genetics need be wasted.

Mark Henderson is Science Editor of 'The Times' and author of ['50 Genetics Ideas You Really Need to Know'](#) (Quercus, 2009).

Top image: Sir John Sulston working in the laboratory to purify DNA. Credit: Wellcome Library, London.

## References

- 1Ferry G. [Human Genome Project history 4: Enter the Wellcome Trust.](#) Wellcome Trust Human Genome website 28 February 2001.
- 2Giles C. ["From then on, it was daggers drawn..."](#) Wellcome Trust blog 24 June 2010.
- 3Preston R. [The genome warrior.](#) New Yorker 12 June 2000.
- 4Henderson M. [Genetic map shows route to new world of medicine.](#) Times 27 October 2005.

All Sulston quotes are from an interview with the author, 11 January 2011.

Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, UK T:+44 (0)20 7611 8888