

UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF NEW YORK

-----X  
ASSOCIATION FOR MOLECULAR PATHOLOGY;  
AMERICAN COLLEGE OF MEDICAL GENETICS;  
AMERICAN SOCIETY FOR CLINICAL PATHOLOGY;  
COLLEGE OF AMERICAN PATHOLOGISTS;  
HAIG KAZAZIAN, MD; ARUPA GANGULY, PhD;  
WENDY CHUNG, MD, PhD; HARRY OSTRER, MD;  
DAVID LEDBETTER, PhD; STEPHEN WARREN, PhD;  
ELLEN MATLOFF, M.S.; ELSA REICH, M.S.;  
BREAST CANCER ACTION; BOSTON WOMEN'S  
HEALTH BOOK COLLECTIVE; LISBETH CERIANI;  
RUNI LIMARY; GENAE GIRARD; PATRICE FORTUNE;  
VICKY THOMASON; KATHLEEN RAKER,

09 Civ. 4515 (RWS)

Plaintiffs,

ECF Case

v.

UNITED STATES PATENT AND TRADEMARK  
OFFICE; MYRIAD GENETICS; LORRIS BETZ,  
ROGER BOYER, JACK BRITAIN, ARNOLD B.  
COMBE, RAYMOND GESTELAND, JAMES U.  
JENSEN, JOHN KENDALL MORRIS, THOMAS PARKS,  
DAVID W. PERSHING, and MICHAEL K. YOUNG,  
in their official capacity as Directors of the University  
of Utah Research Foundation,

DECLARATION OF  
SIR JOHN E.  
SULSTON, Ph.D.

Defendants.

-----X  
I, Sir John Edward Sulston declare under penalty of perjury:

1. I am Chair of the Institute for Science, Ethics and Innovation (iSEI) at the University of Manchester.
2. I received a B.A. from the University of Cambridge, UK in 1963. I received a Ph.D. from the University of Cambridge, UK in Chemistry for research in nucleotide chemistry in 1966.

3. In 1969, I joined the Medical Research Council Laboratory of Molecular Biology, Cambridge, where I researched the cellular and genetic structure of the nematode worm, *C. elegans*.

4. I became involved in genomics starting in 1983, and played a central role in both the *C. elegans* worm and human genome projects. In 1998 I and my colleagues published the sequence of the nematode worm, *C. elegans*. This was the first animal to be sequenced.

5. From 1992-2000 I served as the first Director of the Wellcome Trust Sanger Institute in Cambridgeshire. During my tenure as Director, the Institute grew from 15 staff to more than 500. Ours was the largest genome sequencing center outside of the United States and the biggest producer of genome sequence in the world. Ultimately, we were responsible for the UK's entire contribution to the international Human Genome Project – the complete sequencing of one-third of the human genome – which was completed in 2003.

6. I was elected a Fellow of the Royal Society in 1986, and was knighted in 2001 for my contributions to genome research.

7. In 2002, I was awarded the Nobel Prize for Physiology or Medicine jointly with Sydney Brenner and Bob Horvitz, for our work in understanding the development of the nematode worm and in particular the role of “programmed cell death,” or the ways in which genes regulate tissue and organ development by causing certain cells to die during the normal differentiation process.

8. I am co-author, with Georgina Ferry, of *The Common Thread: A Story of Science, Politics, Ethics and the Human Genome*, published by Bantam Press in 2002.

The book tells the story of and the role of my research institute in the sequencing of the human genome, and discusses the importance of ensuring that information contained in our genome be freely available for the benefit of all.

9. A full copy of my current curriculum vitae is attached as an Exhibit.

### **Genes and Genetic Sequences**

10. Genes and human genetic sequences are not inventions. They are naturally occurring. They are the most fundamental information about humanity, information that is – or should be – common heritage.

11. Genes are the basic units of heredity in all living organisms. A gene is a segment of DNA, the molecule that makes life possible. DNA encodes the instructions for the development and functioning of each of our cells. As a result of the human genome project, we have come to estimate that humans have approximately 25,000 genes that make up our genome.

12. We all have essentially the same set of genes in our respective genomes, but each gene has small variations in its sequence when we compare one individual with another. The differences between us amount to only about 1 letter in 1000.

13. Scientists have long recognized the role of genes in heredity. But it wasn't until 1953 – the year of the discovery of the structure of DNA – that we came to understand how DNA played its role. This central discovery for modern biology made it immediately apparent that the structure embodies a *linear digital code*. This code – nucleic acid sequence – gets copied more or less faithfully from one generation to the next.

14. The genetic code is similar to the English alphabet, except that it consists of four letters (A, T, C, and G) rather than 26 (A through Z). A sequence of alphabetical letters can be read out as text, stored in the volatile memory of my computer, moved to a hard disc where it can survive power failure, printed out on paper, photographed, and spoken aloud, transmitted electronically, or sent to a publisher for binding in a book. Similarly, genetic sequences can also be read out into computers, and from there, we can move the resulting strings of four letters to any medium.

15. The letters of the genetic alphabet correspond to 4 chemical bases (adenine (A), thymine (T), cytosine (C) and guanine (G)). Each gene is typically thousands of bases long, and its sequence of As, Ts, Cs and Gs usually encodes a protein. The code is a set of three-letter words – for example TTT, CAG – each of which corresponds to one of the twenty amino acids that are the building blocks of proteins. Each human gene has its place on one of the twenty-four chromosomes (numbered 1-22, plus the X and Y sex chromosomes), which together constitute the whole genome.

16. A genetic sequence is the sequence of letters of a specified section of the human genome. But beyond this, it is the *biological information itself*. Like strings of alphabetical text, the genetic sequences are the same, regardless of the medium. The physical form in which they occur is unimportant; what matters is the *informational content*. Whether the data resides in the DNA of an organism, in a computer, or as letters on a printed page, the information is the same. The entire human genome sequence consists of approximately 3 billion letters.

17. What distinguishes human genetic sequences from the English alphabet is that the information contained in the genetic sequence is a product of nature. Unlike

alphabetical text, which can be arranged by my own inventive choice (and thus copyrighted), the informational content of a human gene sequence is fixed. While many inventive steps have been necessary to allow us to extract and read a genetic sequence, the ordering of the 4 letters is determined by nature.

18. The slight variations that occur among individual genomes are of great interest to some scientists, because they are thought to account for some of the differences that we see among us. These “typos” or mutations can be in the form of the insertion or deletion of a single letter, or rearrangements, deletions or repeated segments of groups of letters. Some of these mutations have been found to have clinical significance, such as some of the mutations that have been found along the BRCA1 and BRCA2 genes.

19. Mutations are products of nature just as much as genetic sequences are. They are dictated by nature. Similarly, correlations between mutations and disease are scientific facts, or laws of nature. There is no inventive step in determining the variations, or in correlating them with medical conditions.

### **Genetic Sequencing**

20. Genetic sequencing is the process by which one “reads”, or determines the ordering of the 4 letters (A, T, C, and G) within a specified part of the genome. It is a way of transferring the DNA data contained in each and every one of our cells into a format where it can be looked at. Sequencing of a gene can easily be done by several processes that are all well known and understood by scientists, and is performed routinely in both research and clinical laboratories all over the world.

21. Progress in sequencing was slow the first two decades following the discovery of DNA. The first practical method that allowed sequences of many thousands of bases to be read out was invented by Fred Sanger and his colleagues at Cambridge England, published in 1977. It involved smashing the target DNA sample, so that many overlapping pieces were produced, and cloning them randomly (the so-called shotgun approach). Each piece was sequenced by an ingenious method involving replication enzymes, radioactive labels and size separation of the resulting complex mixtures in an electric field. The beautiful patterns that emerge look rather like a bar code, and can be read out to yield the sequence of each piece. The individual sequences can be reassembled, using the overlaps, to yield the sequence of the entire target. In the following two decades Sanger's basic chemistry continued to be used, but many of the steps were automated, radioactive labeling was replaced with fluorescence, and the scale attainable was gradually increased. By the 1980s techniques for DNA manipulation were advanced enough so that geneticists could go fishing for genes among families of people with inherited diseases. Once they narrowed down the location of the gene, they could then try to clone the gene with a view to reading its sequence.

22. The Human Genome Project officially began in 1990, with a target of a complete human sequence by 2005. In 2000, the first human genome sequence was determined in draft form, with a more accurate and complete reference sequence achieved by 2003. Since then, progress has continued with new generations of sequencing machines following one another in rapid succession, and additional human sequences are read out routinely.

23. Our ability to sequence genes has been predicated on advances in numerous areas, including chemistry, biochemistry, instrumentation, and computing. Some of these highly inventive advances have been -- and deserved to be -- patented. These inventions apply to *processes*. They do not apply to the data flowing through them.

24. Gene sequencing is used in diagnostic testing. A gene sequence can be examined to determine if it contains any alterations or mutations that have been associated with a particular genetic condition.

25. In order to sequence, or read a gene, we have to remove it from the cell of an organism and place it in a form so that it can be replicated outside of the body. Most commonly, we use a technique called PCR to replicate many times over small segments of the gene. Amplifying these segments allows us to read out the genetic code.

26. Promoters of gene patents argue that these steps of "isolating and purifying" a gene (removing it from the body and placing it in a form so that they can be sequenced and possibly used in other ways) is sufficient for allowing a gene to be patented. But "isolating and purifying" a gene is simply copying it into another format. It's like taking a hardback book written by someone else, publishing it in paperback and then claiming authorship because the binding is different.

27. The process of sequencing a gene does not change the informational content of that gene. The resultant sequence is informationally and functionally identical to the sequence found inside the body. The alterations or mutations in the gene that we are able to see after sequencing the gene were made by nature, not by the process of sequencing or by me or other scientists, and the effect of those alterations or mutations is dictated by nature, not by any scientist. A patent on a gene sequence and any mutations of that gene

gives a monopoly over this information, regardless of the person from whom the gene is taken or the sequencing process that is used.

### **Information Sharing in Genomic Research**

28. From the point of view of scientific research, human genetic sequences are as basic as you can get in terms of biological information. They are as basic as the elements in the periodic table. Patenting a gene or genetic sequence impedes scientific progress much the same way that patenting a naturally occurring element such as oxygen or gold would impede science.

29. From the very beginning of the Human Genome Project, most scientists and even some private companies recognized the importance of keeping the genome freely available to all. In 1994, the pharmaceutical company Merck funded a massive drive to generate genetic sequences and place them into public databases. By doing this Merck not only gave the entire research community, public and private, free access to valuable genomic data; it also made those sequences much more difficult to patent.

30. In November 1995, a team of researchers at the United Kingdom-based Institute of Cancer Research (ICR) led by Michael Stratton (who went on to become the head of the Cancer Genome Project at the Sanger Centre) found a mutation in some of their breast cancer patients, which appeared to lie in BRCA2. Michael Stratton was collaborating with Mark Skolnick at the University of Utah. He ended the collaboration when he learned from Skolnick that the company he had started, Myriad Genetics, planned to patent the gene and own exclusive rights to exploit it both for diagnosis and therapy if the collaboration were to find and clone the BRCA2 gene. Shortly thereafter

BRCA2 was sequenced by the Sanger Centre. Over the next two weeks, the ICR team confirmed their results and identified five additional mutations. But the day before their findings were published, Mark Skolnick filed a patent application for BRCA2.

31. Myriad used its patent applications to claim rights over the entire BRCA2 gene, including the mutations that had been identified by ICR. Myriad has since claimed proprietary rights for all diagnostic testing for the BRCA genes. One of their tests focuses on one of the mutations discovered by the ICR team that is commonly found among Ashkenazi Jews from central and eastern Europe. Thus, by having a patent on the gene as a whole, Myriad was able to claim scientific findings made by others. Myriad has benefited directly from the work of the international scientific community, while their practices have driven up health care costs and impeded further research on these genes that might lead to future therapies.

32. Michael Stratton concluded from this experience that the only way he could protect his team's discovery from commercial exploitation by others was to patent it himself. As a result, the Institute took out one patent on the first mutation as soon as it was discovered, and another later covering more mutations.

33. This experience with Myriad in the mid-nineties prompted some of us to seek a commitment from the international sequencing community that genomic information would be made publicly available and not patented. In 1996, we organized an international meeting in Bermuda to overcome rivalries in the large international consortium and ensure that data was released in a timely fashion. The meeting was attended by 50 scientists from around the world, and established the so-called Bermuda principles: public sequencing labs would release data within 24 hours of its collection,

by deposit into a public database, and would not take out patents on it. We stated: “All human genomic sequence information should be freely available and in the public domain in order to encourage research and development and to maximise its benefit to society.”

34. Although the genome as a whole is in the public domain, patents are now issued or pending on some 25% of human genes. These patents stand in sharp contrast to the tenor of the Bermuda meeting and threaten to undermine scientific and medical progress.

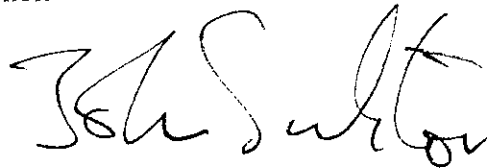
35. There is still much to learn about the products of our genes – what they look like, when or where they are produced, and how they interact with one another. In order to translate this information into medical advances, this basic data must be freely available to everyone to interpret, update and share. The situation is too complex for a piecemeal approach, in which a single entity holds the key to any given gene.

36. Data sharing is also key to the future of genetic discoveries and bioinformatics. This is because a lot of value comes from comparing one sequence with another, both within a genome and between genomes. Many genomes other than human have now been sequenced in whole or part, and much of this information is available on the web for all to use. A remarkable finding from all this data is the tremendous unity of life at the sequence level, incidentally confirming the fact of evolution in a way that is complementary to relationships deduced from the physical appearance of organisms. It turns out for example that half of our human genes have recognizable counterparts in insects.

37. Patents on human genes and genetic sequences are deleterious to the practice of science. Because gene patents tend to cover all uses of that sequence, they are a disincentive to further research on those genes. Patents on genes damage accessibility to this most basic information and discourage scientific communication and data sharing.

38. Patents on human genes will be deleterious to unraveling their role in medical conditions. As we move into an era where the sequencing of all the genes of an individual is more efficient than obtaining the sequence for a gene or two, scientists will discover individual variations in genes whether they are patented or not. Free sharing of this information will be vital in understanding the role of these variations in the human disease and other traits.

I declare, pursuant to 28 U.S.C. §1746, under penalty of perjury under the laws of the United States, that the foregoing is true and correct to the best of my knowledge and belief.



John E. Sulston, Ph.D.

Executed on 17 August, 2009