The Human Genome Project (1990-2003) [1]


The Human Genome Project (HGP) was an international scientific effort to sequence the entire human genome [5], that is, to produce a map of the base pairs of DNA in the human chromosomes, most of which do not vary among individuals. The HGP started in the US in 1990 as a public effort and included scientists and laboratories located in France, Germany, Japan, China, and the United Kingdom. Scientists hypothesized that mapping and sequencing the human genome [6] would facilitate better theories of human development, the genetic causes and predispositions for a number of diseases, and individualized medicine. The HGP, alongside the private effort taken up by the company Celera Genomics, released a working draft of the human genome [7] in 2001 and a complete sequence in 2003. The history of the HGP ripples beyond biomedical science and technology into the social, economic, and political.

The conceptual foundations for the HGP emerged in the US in 1985, when the Office of Health and Environmental Research of the US Department of Energy (DOE) in Washington, DC, proposed the Human Genome Initiative. The DOE’s interest in the human genome [5] grew out of efforts to study DNA changes in atomic bomb survivors of Hiroshima and Nagasaki, Japan. In 1986 US federal advisors argued about whether the National Institutes of Health [6] (NIH), located in Bethesda, Maryland, or the DOE should undertake the project. Several agencies of the government became involved in the debate. Then in 1987 the director of the NIH, James Wyngaarden, testified before Congress that a new organization [7] within the NIH was needed to administrate the HGP. Congress agreed to fund the NIH for research on the human genome [8], and a lesser amount to the DOE. The DOE and NIH signed a memorandum of understanding in 1988 to work on the project together. James Watson [8], who had won a Nobel Prize for his work on discovering the structure of DNA and who at the time worked for the NIH, became the director of the new National Center for Human Genome Research (NCHGR), which became the National Human Genome Research Institute (NHGRI) in 1997, and was part of the NIH. The collective effort of the NIH, DOE, NCHGR, and their international partners constituted the Human Genome Project. These organizations set the agenda and distributed funding to HGP intramural researchers, including, for example, laboratories at the Whitehead Institute’s Center for Functional Genomics in Cambridge, Massachusetts.

In the late 1980s, the US National Research Council [9] and the US Office of Technology Assessment both located in Washington, DC, voiced concerns over the social implications of the HGP. In response, in 1988 Watson promised to allocate three to five percent of the HGP’s budget to research the ethical, legal and social implications (ELSI) of sequencing the human genome [5]. This became the ELSI research program. The goal of ELSI was to predict the social implications of the HGP, and to provide policy and regulatory options to prevent problems. In 1989, an ELSI working group, led by the geneticist Nancy Wexler, a professor at Columbia University [10], was formed to identify ethical, legal, and social issues that would be involved in sequencing the human genome [8].

The Human Genome Project officially began in 1990 as part of the International Human Genome Sequencing Consortium, a collection of labs and organizations funded by the NCHGR, NIH, DOE, and the Wellcome Trust in London, UK. The NIH and the DOE drafted an initial five-year plan, spanning from 1991 to 1995 establishing the goals for the HGP as first to improve and to develop the technology necessary to sequence the human genome [5]. The sequencing project was predicted to last fifteen years, with more detailed five-year plans spelled out for each five-year increment. The projected cost of the human genome [5] sequence was estimated at 200 million US dollars per year, totaling three billion dollars by 2005. Additionally, the HGP would sequence the genomes of a number of other organisms for comparative studies with the human genome [8]. In October of 1990, the NCHGR awarded its first grants, and the HGP began. The NCHGR’s first program announcements solicited research on sequencing the genomes of model organisms, such as a project on the genome [5] of the bacterium Mycoplasma capricolum by Walter Gilbert [11], at Harvard University [12], and colleagues. The HGP progressed more rapidly than planned after sequencing technologies improved. In the first few years of the project, besides mapping the human genome [8], researchers identified genes [13] associated with genetic conditions such as Menkes syndrome, the X-linked immune disorder agammaglobulinemia, Huntington’s disease, myotonic dystrophy, fragile X syndrome, and others.

In 1992, HGP researcher Craig Venter left the NIH after he disagreed with leadership about the use of gene sequencing techniques that he was developing. Years later, Venter’s own private scientific enterprise, Celera Genomics, would also complete the sequence of the human genome [8]. Around the same time, on 10 April 1992, James Watson [8] resigned as the director of the NCHGR over disagreements about the patenting of genetic sequences and the physician Michael Gottesman became acting director after working at the NIH’s National Cancer Institute [14]. Francis Collins became director of the NCHGR on 4 April 1993, leaving a job at the University of Michigan [15] in Ann Arbor, Michigan. Under the new leadership, a revised five-year plan of the HGP was announced on 1 October 1993. The new five-year plan, ranging from 1993 to 1998, focused on incorporating improved technology and insights from previous genetic studies. The plan aimed to increase the resolution of the...
Venter founded The Institute for Genomic Research in Rockville, Maryland, in 1992, after he left the NIH. In 1995, Venter’s group used a method commonly known as shotgun sequencing to sequence the \textit{Haemophilus influenzae genome} \cite{5}, which was 1.8 million base pairs of DNA in size. Shotgun sequencing breaks the \textit{genome} \cite{5} into many fragments and scientists sequence each fragment from both ends. For this method, scientists use fluorescent chemical labels that attach to the DNA to determine the sequence, which is called the chain termination method. Scientists sequence the several random fragmentations of the \textit{genome} \cite{5}, producing some overlap between the ends of fragments. Then, a computer assembles the sequenced pieces into one \textit{genome} \cite{5} by matching up overlapping fragment sequences. Shortly after the initial development of this sequencing method, the whole-genome \cite{5} shotgun sequencing method became applicable to larger genomes, even those as large as the human \textit{genome} \cite{5}. In 1996, Venter and his colleagues published a paper arguing that the advances made in shotgun sequencing made it feasible to apply the method to the human \textit{genome} \cite{5}. They predicted they could complete the project before 2005 under a cost of three billion dollars.

In January 1997, the NCHGR was elevated to the status of full institute of the NIH, becoming the National Human Genome Research Institute (NHGRI). This followed a number of advancements in basic science and clinical research of the human \textit{genome} \cite{5}. The NHGRI now had equal standing with the other NIH institutes. This status gave the NHGRI autonomy on budget development and priorities.

In 1997 scientists published the first complete sequence of the eukaryotic organism, \textit{Saccharomyces cerevisiae}, or common yeast, funded in part by NHGRI. HGP-funded projects led to several bacterial \textit{genome} \cite{5} sequences by the end of 1997, such as the sequence of \textit{Escherichia coli}. At this time the NHGRI and other researchers also published a map that pinpointed the locations of greater than 16,000 \textit{genes} \cite{13} in the human chromosomes, the first location of a gene associated with Parkinson’s disease, the first known gene for predisposition to prostate cancer, mutations on BRCA1 and BRCA2 \textit{genes} \cite{13} related to predisposition to breast cancer, the identification of a genetic mutation causing Pendred Syndrome, and a complete map of the \textbf{human chromosome 7} \cite{16}.

The NHGRI, NIH, and DOE’s final five-year plan for the HGP began on 1 October 1998 and planned to sequence the human \textit{genome} \cite{5} by 2003. The project had advanced faster than expected, with a projected completion date two years sooner than originally planned. This five-year plan expected to sequence the first one-third of the human \textit{genome} \cite{5} by the end of 2001, as well as to complete the sequence of other organisms such as \textit{C. elegans} and \textit{Drosophila} \cite{17}.

By 1998 Venter’s group had sequenced the genomes of several organisms, including the bacteria responsible for Lyme disease, syphilis, tuberculosis, and meningitis. That year Venter founded the company Celera Genomics, headquartered in Alameda, California, to finish the sequence of the human \textit{genome} \cite{5} by 2000. At that point, US Congress questioned the humangenome \cite{5} sequence mandate of the NIH, and it considered diminishing and eventually cutting funding for the HGP altogether to allow private industry to complete the HGP. Francis Collins, alongside the NIH director Harold Varmus, testified in 1998 before Congress on whether or not the NHGRI’s efforts were necessary in light of Celera’s project. Collins and Varmus cited worries over Celera’s approach and the public availability of the \textit{genome} \cite{5} sequence. They convinced Congressional subcommittee members that the HGP had room for both the NIH and Celera.

Celera began to sequence the humangenome \cite{5} in 1999 using two strategies of the shotgun sequencing method: the whole-genome \cite{5} assembly (WGA) and the compartmentalized shotgun assembly (CSA). The WGA follows the same shotgun sequencing strategies described earlier, while the CSA subdivided the human \textit{genome} \cite{5} into segments, and then applied the shotgun sequencing method to each of the segments separately. Comparison of the two processes can reveal inconsistencies between the methods, and fills in the data gaps left by just one method. Overall, Celera found that CSA provided more consistent data and coverage of the entire \textit{genome} \cite{5} than did the WGA.

In 1997 an ELSI task force published a report on genetic testing, anticipating concerns about people’s genetic privacy. The NIH leadership agreed that regulations should protect people’s genetic information from undue access and abuse. Some people argued that workplaces and health insurance companies could discriminate based on genetic information. In 2000, US President Bill Clinton signed an executive order that prohibits federal departments and agencies from using genetic information to discriminate during hiring or promotion of employees. At the NHGRI, Collins focused the ELSI program efforts on a genetic privacy policy to propose to Congress. In 2008 US Congress passed GINA, the Genetic Information Non-discrimination Act. GINA drew from ELSI’s recommendations, and expanded Clinton’s executive order on genetic privacy.

Venter and Collins appeared alongside Clinton and British Prime Minister Tony Blair to announce the preliminary draft of the human \textit{genome} \cite{5} in 2000. A first draft of the humangenome \cite{5} was published ahead of schedule on 16 February 2001, simultaneously by the NIH and Celera in the journals \textit{Nature} and \textit{Science} respectively. Later that year, Venter and Collins also discussed their findings at the American Association for the Advancement of Science’s annual meeting in San Francisco, California. Scientists continued to sequence of all human base pairs by 2003 when the complete human \textit{genome} \cite{5} was published.
While the HGP completed in 2003, the NIH retained the NHGRI to administrate genome-related scientific projects, such as relating genetic variation to phenotypic outcomes. In 2003 Collins testified to Congress that the NHGRI would support research to translate genomic science into clinical practice, research on the interaction of the human genome with other determinants of health, and to continue to study the social implications of genetic and genomic science. Collins announced his resignation from the NHGRI on 28 May 2008, shortly after the passing of GINA. On 17 August 2009 Collins became the director of the NIH. On 17 November 2009 the NIH promoted Eric Green to become the new director of the NHGRI.

Since 2003, the NHGRI has focused on supporting genetic and genomic research related to cancer through The Cancer Genome Atlas project, which has supported studies on a number of cancers including, lung, brain (glioblastoma), and ovarian. Researchers at the NHGRI have sequenced the genomes of several eukaryotic organisms including drosophila, rat, mouse, dog, honeybee, and chimpanzee. In 2010, the NHGRI joined the British Wellcome Trust on a project to support population-based genetic studies in Africa, the Human Heredity and Health in Africa Project. Also in 2010, the NHGRI developed the Genotype-Tissue Expression Project, to elucidate gene activity and its relationship to development of disease phenotypes. As of 2014, the NHGRI continues to fund basic to clinical research, as well as social research and medical humanities.

Sources

The Human Genome Project (HGP) was an international scientific effort to sequence the entire human genome, that is, to produce a map of the base pairs of DNA in the human chromosomes, most of which do not vary among individuals. The HGP started in the US in 1990 as a public effort and included scientists and laboratories located in France, Germany, Japan, China, and the United Kingdom. Scientists hypothesized that mapping and sequencing the human genome would facilitate better theories of human development, the genetic causes and predispositions for a number of diseases, and individualized medicine. The HGP, alongside the private effort taken up by the company Celera Genomics, released a working draft of the human genome in 2001 and a complete sequence in 2003. The history of the HGP ripples beyond biomedical science and technology into the social, economic, and political.