Human Genome Project completed: an extraordinary scientific achievement

Frank Gaglioti
7 May 2003

The publication of the detailed structure of 99 percent of the human genome on April 14 is the culmination of one of the largest scientific undertakings in history. Initiated in 1990, the Human Genome Project (HGP) involved the cooperative work of hundreds of scientists in 20 sequencing centres in countries including China, France, Germany, Great Britain, Japan and the United States.

The achievement is significant. Not only has it given a glimpse into the extraordinary complexity of the structure of human DNA but it has also demonstrated the huge potential of cooperative scientific endeavour organised on an international basis. The HGP was finished two and a half years ahead of schedule with all goals completed and for considerably less than the estimated budget.

The whole project stands as a refutation of the prevailing nostrums that private profit, personal advantage and competition are the only driving forces for scientific research. The plethora of private biotechnology companies set up to patent genes and other therapeutic products have only acted as a barrier to scientific advance by preventing the sharing of information and collaboration between research teams.

US National Human Genome Research Institute (NHGRI) director Francis Collins, who led the project, summed up its significance on April 14, by declaring: "The HGP has been an amazing adventure into ourselves, to understand our own DNA instruction book, the shared inheritance of all humankind."

The project's completion is testimony to the enormous scientific advances in the 50 years since James Watson and Francis Crick revealed the basic double-helix structure of DNA—the molecule that controls all cellular functions including reproduction.

Although the DNA sequence is not entirely complete, scientists have mapped 2.9 billion base pairs, the chemical units that make up DNA, to an accuracy of 99.99 percent. It is estimated that the human genome has 30,000 genes coding for various cellular functions. The whole HGP database is so large that if it were printed, it would require 750,000 A4 pages.

The project revolutionised DNA-sequencing technology. In 1990, 10,000 base pairs could be scanned in a day, but today robotic sequencers can process 10,000 base pairs a second.

The announcement marks a huge advance on the "working draft" published in 2000, which covered 85 percent of the genome. Scientists who used this relatively inaccurate map were often forced to re-sequence the specific segment of interest to obtain sufficient precision for their research. Now researchers can access a highly accurate sequence from the HGP database that is freely available on the Internet.

The HGP web site receives 600,000 visits a week from scientists in 120 different countries. Robert Waterson, a medical scientist from the University of Washington, pointed out: "Not only does the rapid release of data promote the best interests of science, it also maximises the benefits that the public receives from such research."

The missing sections, representing less than 1 percent of the genome, could not be sequenced because the necessary technology is not yet available. The genetic material in these segments is toxic to the bacteria used to replicate the DNA in sufficient quantities for analysis. Work is continuing to develop the new techniques required. Dr Huntington Willard of Duke University, an expert on the X chromosome, commented: "We shouldn't declare the job 'complete' until it is." He speculated that it could take another 10 to 20 years to provide a map of the whole human genome.

The HGP has already produced important medical breakthroughs. On April 24, scientists at the University of Michigan in Ann Arbor and the Clinical Research Institute of Montreal in Canada simultaneously published results that could assist research into cancers such as

leukemia. They identified the Bmi-1 gene that plays a key role in the growth of stem cells involved in forming the blood and immune system.

On April 16, a NHGRI research team announced the discovery of the genetic basis of progeria, a disease that causes children to age at 5 to 10 times the normal rate. NHGRI director Francis Collins commented: "The implications of our work may extend far beyond progeria to each and every human being. What we learn about the molecular basis of this model of premature aging may provide us with a better understanding of what occurs in the body as we all grow older."

To date, 1,400 disease genes have been identified; enabling the production of specifically targeted drugs and treatments. More than 350 new drugs derived from the HGP research are currently undergoing tests.

Future research projects

Even the staggering scale of the HGP research is dwarfed by plans for future investigations, which have the potential to cure genetic diseases and offer deeper insights into human development and evolution. HGP scientists published a discussion paper in the April 24 issue of *Nature* entitled *A Vision for the future of genomics research*, which outlined new research directions.

The great achievement of the HGP research has been to reveal the complex structure of the human genome. But scientists still do not understand the function of its vast tracts. The current stage of research could be compared to the tasks confronting linguists who have identified all the letters of an unknown language but only have a limited knowledge of its words, grammar and sentence structure.

Long lengths of the human genome consist of continuous repetitions, while others have a regulatory function—turning genes on or off and regulating various cellular functions. It is currently estimated that only one to two percent of DNA base pairs has the information required to manufacture different proteins. A full account of protein-coding genes does not exist.

Future research is aimed at elucidating the structure and function of the human genome. NHGRI has launched an Encyclopedia of DNA Elements (ENCODE) project to identify all the functional elements in the human genome. Scientists want to unravel how genes control the complex operations of living cells. The project will start by examining a relatively simple organism such as yeast or bacteria that will be used as a model for the human system.

When the ENCODE project turns to humans, it will initially investigate just one percent of the vastly more complex human genome to determine its interaction with the rest of its genetic and cellular environment. The project will look at specific genetic regions of interest, but will ultimately be expanded to eventually include the whole genome—a process that could span decades.

A fuller appreciation of cell functions will emerge, enabling scientists to more closely appreciate how one section of the genome interacts with other parts and the cellular environment. The knowledge will have profound consequences for the treatment of genetic disorders, as scientists will be able to identify their source in the genome and specifically target treatments to turn off harmful genes.

A second research area, the International HapMap Project, is aimed at establishing a catalogue of all common variants in the human population. These are known as single-nucleotide polymorphisms and involve small variations in the DNA code, which are responsible for genetic diseases, racial differences and subtle differences in the structure of proteins. It is estimated the HapMap Project will be concluded in three years.

These projects will have a far-reaching impact on the study of evolution, as scientists will be able to determine where mutations occur in the genome and their impact on the organism. Genome maps of several organisms such as the roundworm, laboratory fruit flies, the mouse and nematode worms have been completed. Their comparison with the human genome will help to determine the genetic basis for the formation of new species.

The huge advances made by the HGP researchers confirm, if only in embryo, the ability of mankind to unravel the most complex scientific mysteries. This gain in human self-understanding stands in stark contrast to the irrationality of the present capitalist order, which is incapable of resolving the most basic social problems—poverty, disease, inequality and war.