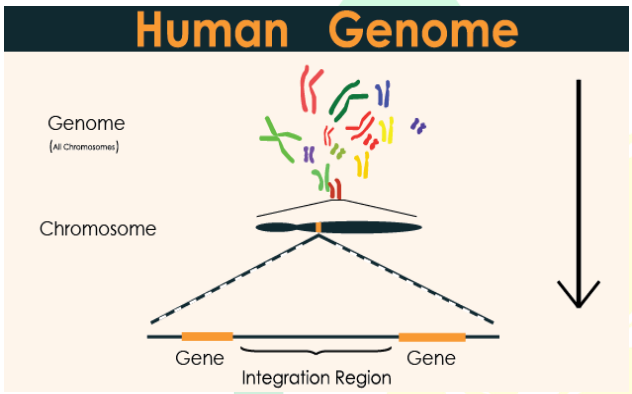
JUST AGRICULTURE
multidisciplinary e-Newsletter

Human Genome Project: Cracking Down the Life's Code

Jogender* and Vishal Saini
Dept. of Genetics & Plant Breeding
College of Agriculture, CCS Haryana Agricultural University, Hisar (125004)

ARTICLE ID: 57

Introduction:



All organisms, from bacteria to elephants have genomes. A genome contains the genetic information needed to make a living organism, written in the DNA four-letter code of bases, or nucleotides. Sequencing an organism's genome gives us a comprehensive view of this information, with which we can better understand their evolution, development and biological functions. The human genome contains more than three billion DNA base pairs and all of the genetic information needed to make us. The human genome was first mapped and sequenced over a period of 13 years from 1990 to 2003. The Human Genome Project is an ambitious research effort aimed atdeciphering the chemical makeup of the entire human genetic code (*i.e.*, the genome). Sequencing the human genome has helped researchers to identify important genes and genetic sequences, to better understand their role in disease, and to investigate our origins using variations in the DNA sequence.

Definition:

The **Human Genome Project (HGP)** is an international, interdisciplinary, scientific research project aimed at determining the sequence of chemical base pairs which make up human DNA, mapping the entire human genome, and identifying its complex structures and functions. The Human Genome Project (HGP) was a ground-breaking international

JUST AGRICULTURE
multidisciplinary e-Newsletter

initiative, considered to be one of the most ambitious scientific projects undertaken in the twentieth century. It was the first major global collaboration of its kind and the largest biological research project ever undertaken, involving thousands of staff in institutes across the globe.

The idea was picked up in 1984 by the US Government when planning started. In total the HGP took 13 years; it was expected to take more than 15 years. The project officially started in October 1990, a first "draft" was announced in June 2000 and a "finished" sequence was completed in April 2003 which was published in 2004.

Goals of HSP:

- To identify and map all the 20000-25000 genes in the human DNA from a physical and functional standpoint.
- To determine the sequence of the 3 billion chemical bases pairs that makes up the human DNA.
- To store this information's in databases.
- To discover more efficient technologies for data analysis.
- Allow the private sector access to the information and technologies that arise from this project.
- Also, to sequence the genome of other organisms that are important in medical research such as mouse, drosophila etc.
- To address ethical, legal and social issues.

Participants of HGP:

Twenty institutes from six different countries (China, France, Germany, Japan, UK and USA) were involved in the HGP making it a truly international collaboration.

The five biggest contributors were:

- Welcome Trust Sanger Institute
- Washington University School of Medicine
- Whitehead Institute/MIT centre for Genome research
- The DOE's Joint Genome Institute
- Baylor College of Medicine

Pioneers in HGP:

1 Robert Sinsheimer proposed the idea of sequencing the human genome in year 1985.

JUST AGRICULTURE
multidisciplinary e-Newsletter

2 Charles DeLisi and David Smith proposed the budget for HGP.

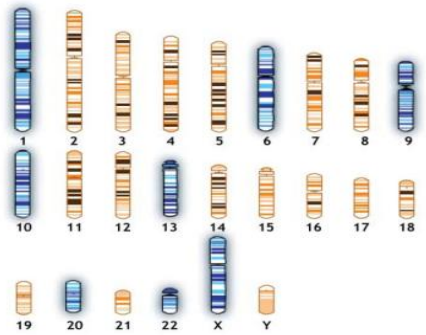
3 HGP act was passed in the US congress under President Regan in 1988.

4 James Watson headed NIH Genome Program

5 Francis Collins succeeded James Watson in 1993 as the overall Project Head and the Director of the NIH (which later become the National Human Genome Research Institute NHGRI) and was in power until the completion of HGP in 2003.

6 Jim Kent, a PhD scholar in the University of California Santa Cruz, in May 2000, developed a software, GigAssembler, that allowed the publicly funded HGP to assemble and publish the Human genome sequence.

Sanger Institute's contribution:



The Wellcome Trust Sanger Institute (WTSI) was the single largest contributor to the HGP. Initially funded to sequence 1/6th of the genome it acquired additional funding to sequence just under a third of the entire human genome (29%). The diagram on the left shows the chromosomes that were sequenced, or partially sequenced, by the Sanger Institute coloured in blue.

Technical aspects in HGP:

- ❖ The process of determining the human genome first involves genome mapping or characterizing the chromosomes. This is called a genetic map.
- ❖ The next step is DNA sequencing or determining the order of DNA bases on a chromosome. These are physical maps.

How was the human genome sequenced?

Sequencing technology can only sequence a few hundred base pairs of DNA at a time. This meant that the three billion base pair human genome had to be broken up into small pieces for sequencing, which were then reassembled like a giant jigsaw puzzle.

The genome was first broken into 200,000 base pair sections (clones) and inserted into bacterial DNA, creating living libraries of the DNA clones. These libraries could be copied and shipped between collaborating institutes. The clones could then be broken into

JUST AGRICULTURE
multidisciplinary e-Newsletter

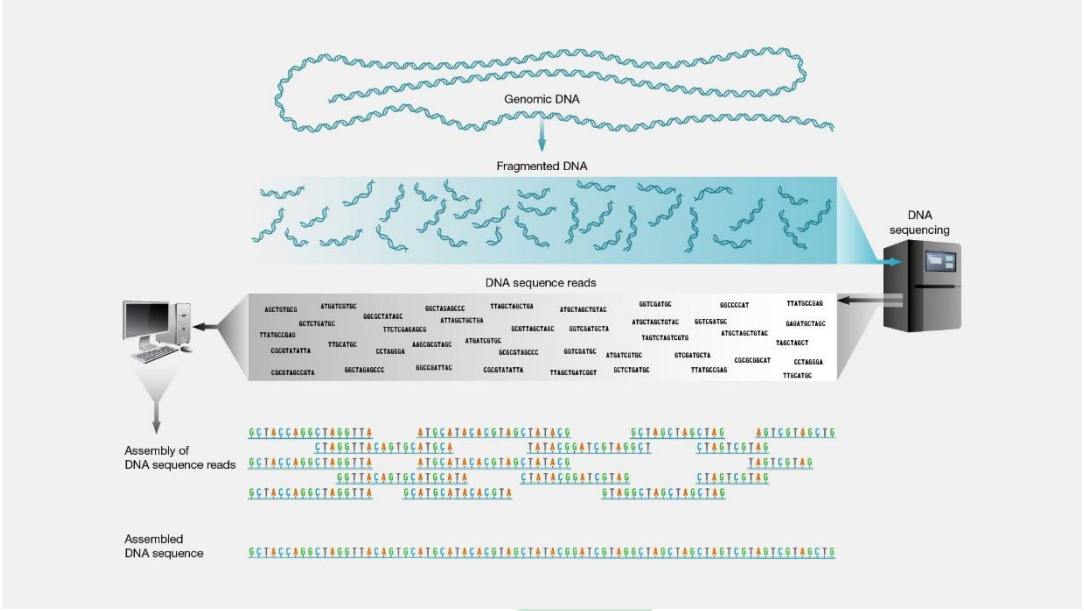
smaller pieces (4000-6000 base pairs), re-inserted into bacteria and cultured to make enough DNA for sequencing.

To sequence the DNA, bacterial colonies were transferred to tubes where the cells were lysed (split open) and the DNA extracted. These sections of DNA were then sequenced by machine using the Sanger sequencing method. The resulting data was pieced together by computers and researchers, to form the whole genome sequence.

Sequencing techniques used in HGP:

- Shotgun sequencing method
- Sanger sequencing method

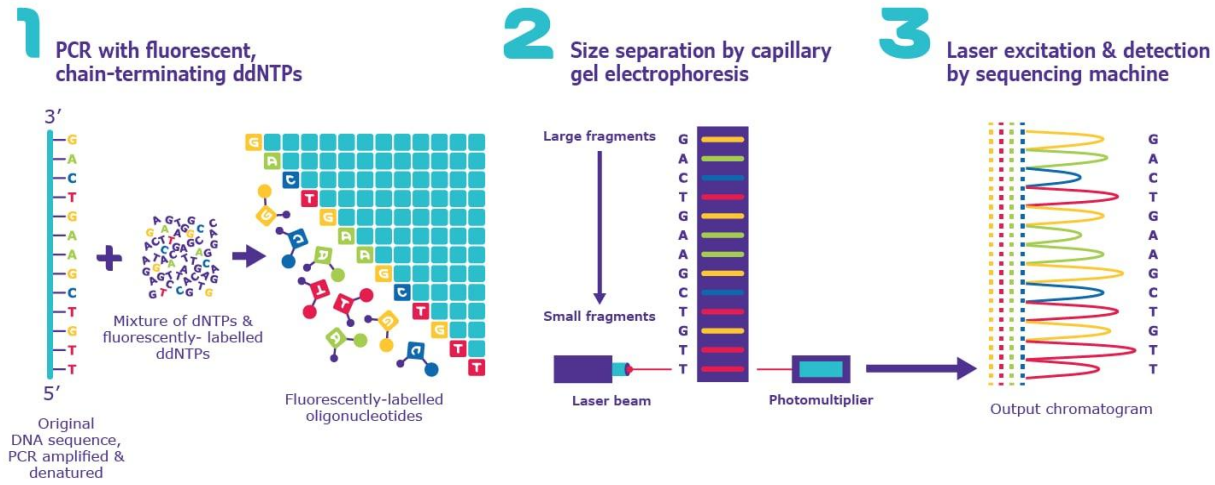
Shotgun sequencing method: Shotgun sequencing is a laboratory technique for determining the DNA sequence of an organism's genome. The method involves randomly breaking up the genome into small DNA fragments that are sequenced individually. A computer program looks for overlaps in the DNA sequences, using them to reassemble the fragments in their correct order to reconstitute the genome.



Sanger sequencing method: In Sanger sequencing, a DNA primer complementary to the template DNA (the DNA to be sequenced) is used to be a starting point for DNA synthesis. In the presence of the four deoxynucleotide triphosphates (dNTPs: A, G, C, and T), the polymerase extends the primer by adding the complementary dNTP to the template DNA strand. To determine which nucleotide is incorporated into the chain of nucleotides, four dideoxynucleotide triphosphates (ddNTPs: ddATP, ddGTP, ddCTP, and ddTTP) labeled with

JUST AGRICULTURE
multidisciplinary e-Newsletter

a distinct fluorescent dye are used to terminate the synthesis reaction. Compared to dNTPs, ddNTPs has an oxygen atom removed from the ribonucleotide, hence cannot form a link with the next nucleotide. Following synthesis, the reaction products are loaded into four lanes of a single gel depending on the diverse chain-terminating nucleotide and subjected to gel electrophoresis. According to their sizes, the sequence of the DNA is thus determined.



Outcomes of HGP:

- Human genome consists the information of 24 chromosomes (22 autosomes+ X chromosome+ one Y chromosome); in *Homo sapiens* 2n= 2x= 46.
- Human genome contains 3164.7m (3.1b) base pairs
- Average gene consists of 3000 bases
- Largest human gene- dystrophin (2.4m base)
- Total number of genes- 30000
- Repeated sequences make large portions
- Less than 2% genome codes for proteins
- 1.4 million locations have SNP's
- The functions are unknown for over 50% of discovered genes.

Conclusion:

The HGP has transformed biology through its interdisciplinary approach to deciphering a reference human genome sequence. The HGP exemplifies the power, necessity, and success of large, integrated, cross-disciplinary efforts (*e.g.*, engineering, computer science, math, and biology) directed toward complex major objectives. This ambitious

JUST AGRICULTURE
multidisciplinary e-Newsletter

endeavor led to the development of novel technologies and analytical tools. Importantly, the outcomes of the project have been established as open and accessible to all. Finally, the HGP has inspired several other exciting projects that promise to open new avenues in biology, medicine, and psychology