HUMAN GENOME PROJECT

Interesting Facts about HGP

- Establishment of genetic engineering techniques made it possible to isolate and clone any piece of DNA.
- HGP (1990-2003) was a (13 years) first mega project for the sequencing of nucleotides and mapping of all the genes in Human Genome.
- The entire DNA in the haploid set of Chromosomes of an organism is called a Genome.
- Human Genome packed in 23 chromosomes containing about 3×10^9BP.
- Cost of project 9 billion USD (3$ per BP).
- If sequences to be stored in typed form in books, then 3300 books required contained 1000 pages, with 1000 letters on each page hence require high speed computational devices for data storage and retrieval, and analysis
- HGP was coordinated by US Department of Energy and the National Institute of Health. Partners are Wellcome Trust U.K. (major), Japan, France, Germany, China and others.

Goals of HGP

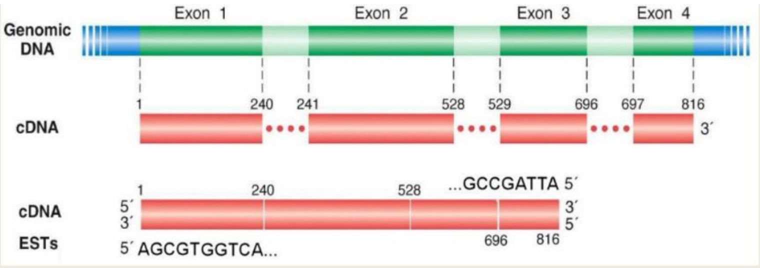
- Identify all the estimated gene (20,000-25,000) in Human DNA.
- Sequencing of 3 billion chemical BP of Human DNA.
- Store this information in database
- Improve tools for data analysis.
- Transfer related technologies to other sectors.
- Address the ethical, legal and social issues (ELSI) that may arise from the project.

Relative Information

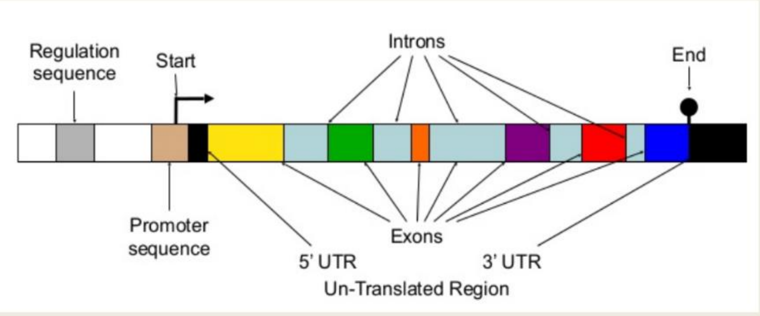
- Sanger has also developed method for sequencing of amino acids in proteins.
- DNA is converted to random fragments due to technical limitations in sequencing very long pieces.
- HGP was closely associated with Bioinformatics (Application of computer science & information technology to the field of biology & medicine.)
- Of 24 chromosomes (22 autosomes & X & Y) the last sequenced one is Chromosome 1st (May 2006).
- Genetic & physical maps on the genome were generated using information on polymorphism of restriction endonuclease recognition sites and some repetitive DNA sequence (microsatellites).
- Non-human organisms DNA sequences (Bacteria, yeast, Caenorhabditis elegans (a free-living non-pathogenic nematode/roundworm), Drosophila, plants (rice & Arabidopsis) etc) could contribute toward solving challenges in health care, agriculture, energy production, environmental remediation.

Methodologies

- **Expressed Sequence Tags (ETSs)**- Focus on identifying all the genes that are expressed as RNA.



- **Sequence Annotation**: Sequencing whole set of genomes containing all the coding and non-coding sequence and later assigning different regions in the sequence with functions.



Isolate total DNA from a cell

↓

Convert into random fragments

↓

Clone in a host (bacteria & yeast) using vector [Bacterial Artificial Chromosome (BAC) & Yeast Artificial Chromosome (YAC)] for amplification

↓

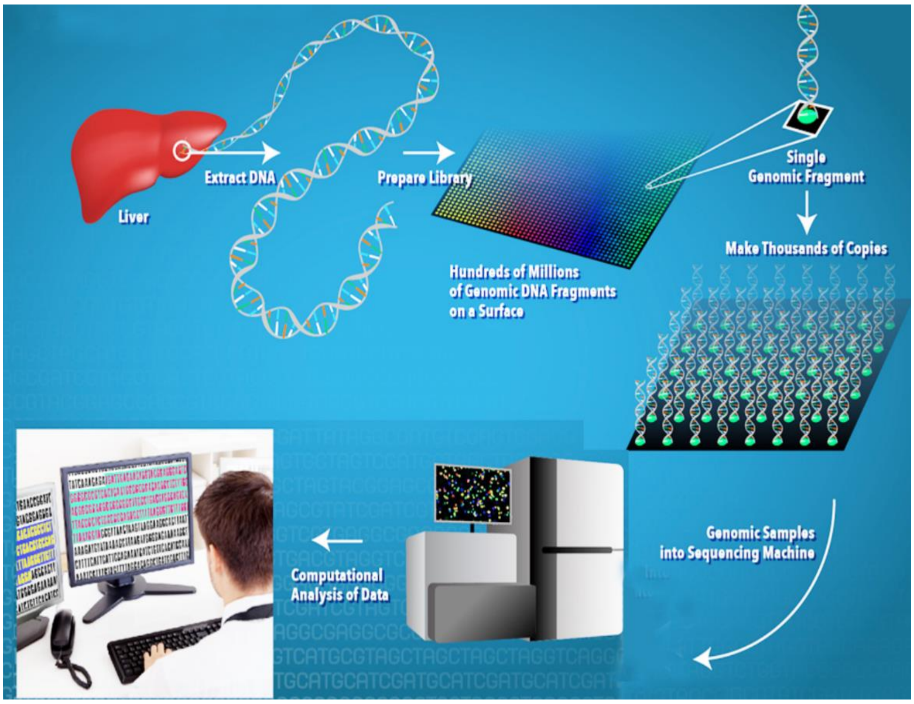
Sequencing of fragments using Automated DNA sequencer (Using Frederik Sanger Principle)

↓

Arrange the sequences based on overlapping regions using computer programs

↓

Annotated and assign to each chromosome



Salient Features of Human Genome

- HG contains 3164.7 million bp. 99.9% bp are same only 0.1% difference makes every individual unique.
- Total no. of gene is about 30,000. Over 50% gene function is unknown.
- Average gene consist of 3000 bases, but size vary, largest known human gene i.e. is **dystrophin on X-chromosome** containing about 2.4 million bases.
- Chromosome 1st has most gene (2968) and Y has the fewest (231).
- Less than 2% of the genome codes for protein.
- Very large portion of genome (DNA) is Repetitive (100-1000 times) & have no direct coding function. They shed light on Chromosome structure, dynamics & evolution.
- About 1.4 million locations have single-base DNA differences called SNPs (Single nucleotide polymorphism). This help to find chromosomal locations for disease-associated sequences and tracking human history.

Applications of HGP and Future Challenges

- Challenge is that it require the coordinated effort of scientists from various fields.
- Application:
    - Providing radically new approach in biological research.
    - Earlier, one or a few genes studied at a time. But with advance technology we can study all the genes in genome.
    - Can study how thousands of gene and protein work together.