MOLECULAR BASIS OF INHERITANCE

HUMAN GENOME PROJECT

HUMAN GENOME PROJECT

Human Genome project (HGP) :

HGP is a mega project started by **U.S. Department of Energy and national Institute of Health** for sequencing human genomein 1990. **Welcome Trust (UK)** joined the project as a **major partner.** Later on japan, France, Germany, China and some other countries also joined it.

Aims of HGP :

(i) To determine the sequence of 3.2 billion base pairs of human genome

(ii) To Identify all the genes of human genome and determined their functions

(iii) To Identify those genes that are responsible for genetic disorders.

(iv) To Store this information in data bases.

(v) The project may result in many ethical, legal and social issue which must be addressed and solved.

Methodology : Two approaches have been recognized for analysing the human genome.

(i) ESTs or expressed sequence tags : To Identify all the genes that are expresed as RNA.

(ii) Sequence annotation : Sequencing both coding and noncoding regions of whole genome and assigning the different regions with functions. The project was completed for sequencing in 2003. However, chromosome I was last to be sequenced in May 2006.

SALIENT FEATURES OF HUMAN GENOME :

(1) Human genome contains **30,000 genes** which are much lower than previous estimate of 80,000 – 10,0000.

(2) Human genome contains **3.1647 billion** nucleotide base.

(3) Less than 2% of the genome shows structural genes that code for proteins.

(4) Average gene size is 3000 base pairs but size vary greatly, with the largest known human gene being dystrophin at 2·4 million bases.

(5) Chromosome I contains **2968 genes (maximum gene)** while **Y-chromosome** bears **231 genes (minimum genes)** in human chromosome.

(6) 99.9% of the nucleotide bases are **exactly similar** in all human beings.

(7) About 1.4 million locations have been reported where single nucleotide differences or **SNPs (snips) or single nucleotide polymorphism** are found. They have the potential to helping and finding chromosomal locations for disease associated sequences and tracing human history.

(8) Repeated sequences makeup very large portion of the human genome.

(9) Repeated sequences are stretches of DNA sequnces that are repeated hundred to thousand times. They are thought to have no direct coding functions but they shed light on chromosome structure. dynamics & evolution. Many non human organisms like, Bacteria (E. coli), yeast, Caenorhabditis elegans, Drosophila, plants (Rice and Arabidopsis) etc have also been sequenced.