

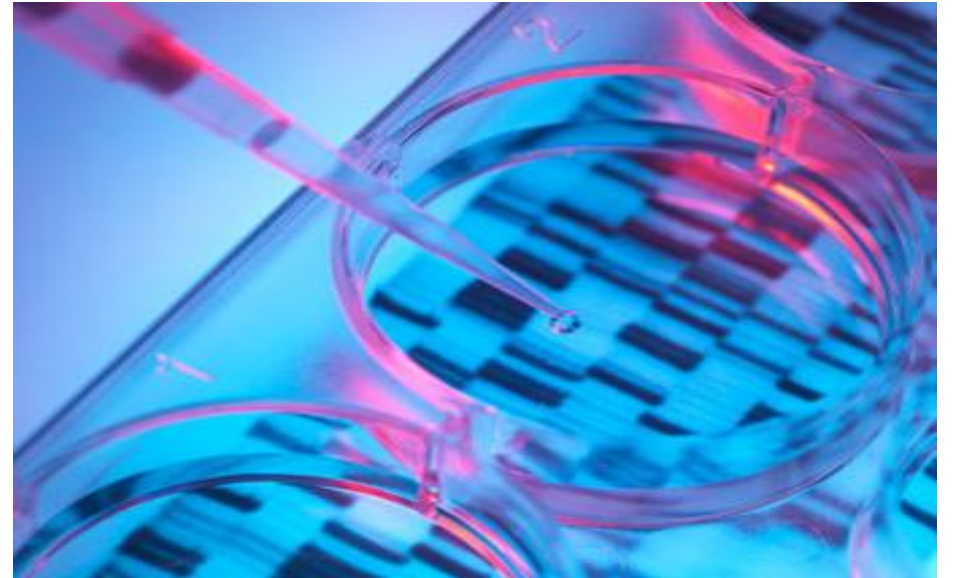
Genetics Engineering (Zoo-455)

Human Genome Project

Lecture-7

Introduction to the human genome project:

- ❑ The Human Genome Project (HGP) was the international, collaborative research program whose goal was the complete mapping and understanding of all the genes of human beings.
- ❑ HGP has **three major goals**:
 1. Determine the nucleotide sequences of all DNA in the human genome.
 2. Identify the location and sequence of every human gene.
 3. Producing linkage maps, through which inherited traits (such as those for genetic disease) can be tracked over generations.
 4. comparisons of genomes within and across species.



Time line of the human genome project:



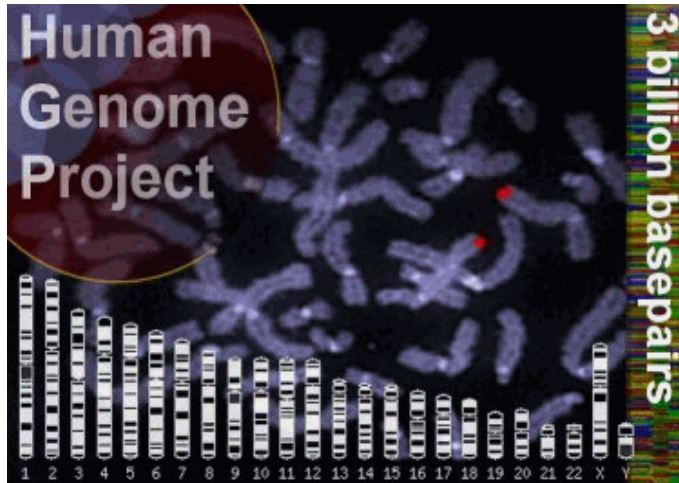
- ❑ The Human Genome Project officially **begins** in 1990 and **completes** in 2003.

Results of the HGP:

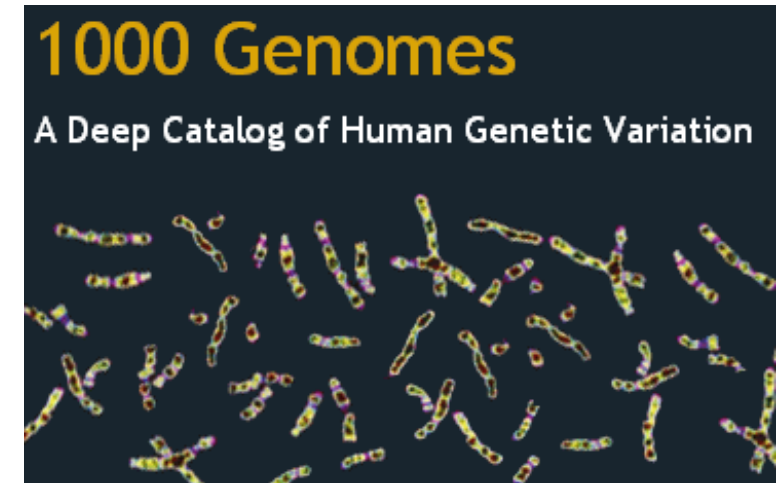
- ❑ HGP has sequenced all of the DNA base pairs of human chromosomes.
- ❑ HGP has sequenced **3 billion** base pairs (**ATGC**).
- ❑ HGP covers about 99% of the human genome's gene-containing regions.
- ❑ The identification of approximately **20.000** protein-coding genes in human.
- ❑ Many of protein-coding genes produce more than one protein product by alternative splicing of the primary transcript of the gene. Each **ORFs** produce more than 2 to 3 different proteins.
- ❑ The identification of more than 3 million human genetic variations, called single nucleotide polymorphisms (SNPs).
- ❑ The generation of full-length complementary DNAs (cDNAs) for more than 70% of known human and mouse genes.
- ❑ The identification of the mouse genome sequence, published in December 2002.
- ❑ The identification of the rat genome sequence, produced in November 2002.

Exploring the human genome:

2003



2008

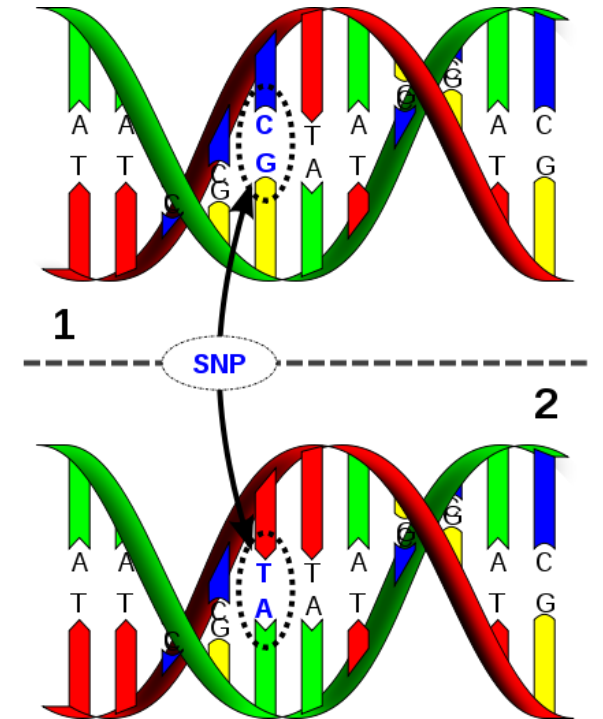


❑ Identification of common genetic variation in 270 individuals from 4 populations (Europe, China, Japan, and Nigeria)

❑ Whole genome sequencing and complete description of human genetic diversity in >1000 individuals from multiple world populations.

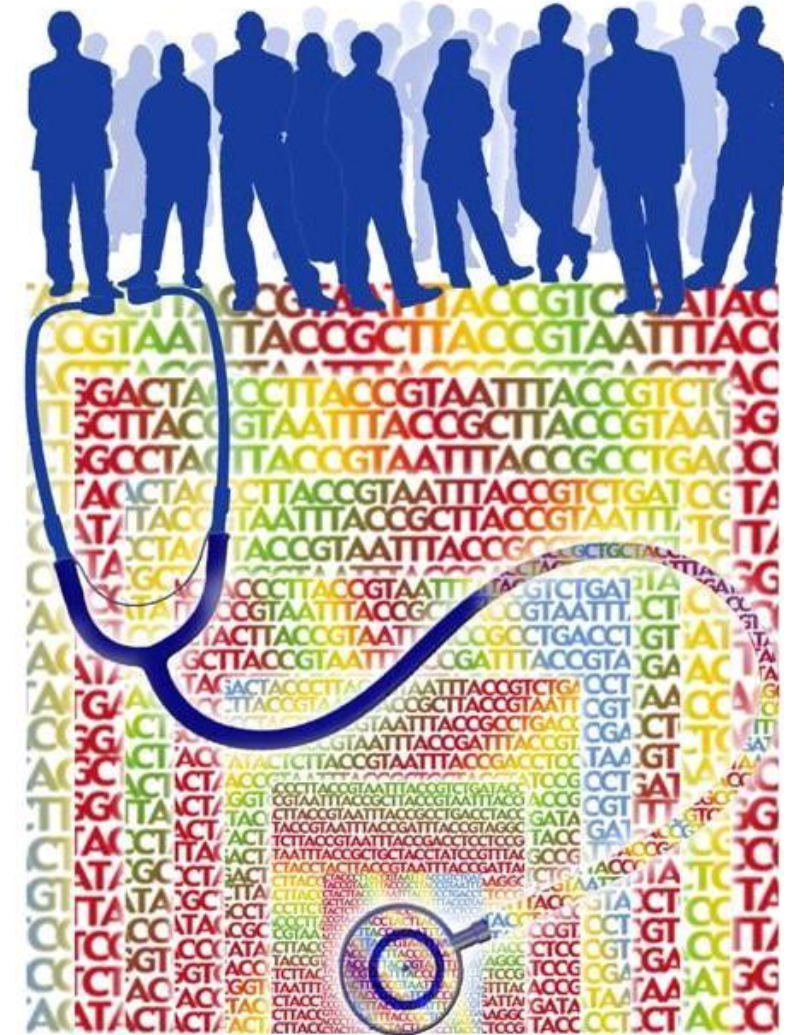
What is a genome?

- ❑ **Genome:** All the DNA for a human.
- ❑ **Gene:** Specific sequences on DNA that can be expressed into proteins.
- ❑ **Human genome contains:**
 - **Nucleus:** 3 billion base pairs (ATGC) packaged in 46 chromosome.
 - **Mitochondrion:** 16600 base pairs (ATGC).
 - 20.000 protein-coding genes.
 - 4 millions DNA variants/individual.
 - 50% is high copy number repeats.
 - About 10% is transcribed (made into RNA).
 - Only 1.5% codes for protein (98.5% non-coding DNA).
- ❑ **SNP = Single nucleotide polymorphism:** DNA sequence variation in which a single nucleotide A, T, C or G differs between members of the same species.
- ❑ SNP occurs commonly within a population (> 1%).



How to read the genome?

1. **Genotyping:** Process of determining genetic differences between individuals by using a set of markers
 2. **Sequencing:** Process of determining the full nucleotide order of a DNA sequence
- ❑ Technology allows the study and comparison of both genes and proteins.
 - ❑ **Bioinformatics:** is the use of computer databases to organize and analyze biological data.
 - ❑ **Proteomics:** is the study and comparison of proteins





Questions?