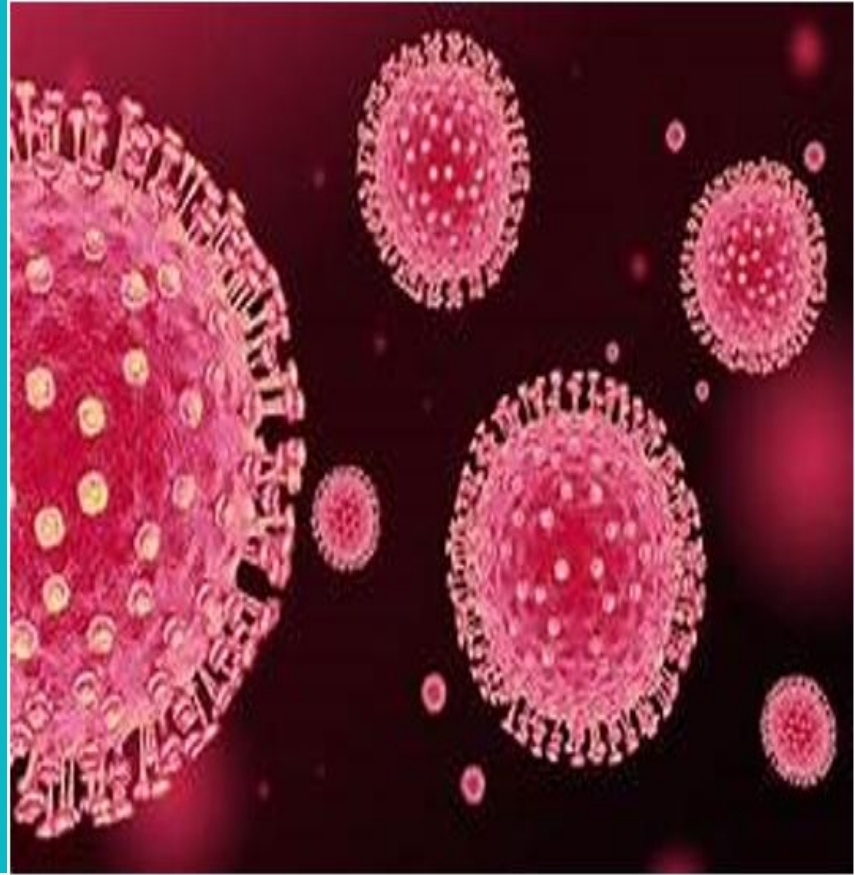


# INTRODUCTION TO BIOINFORMATICS

## Lecture – 1

Department of CSE, DIU



# CONTENTS

1. Basics
  - Cell
  - Nucleus
  - Chromosome
  - DNA
  - RNA
  - Gene
  - Genome
2. Brief History
3. Present and Future Scopes

# 1. Begin With Basics

Let's start with the very basics and learn step by step, TOGETHER!

# CELL

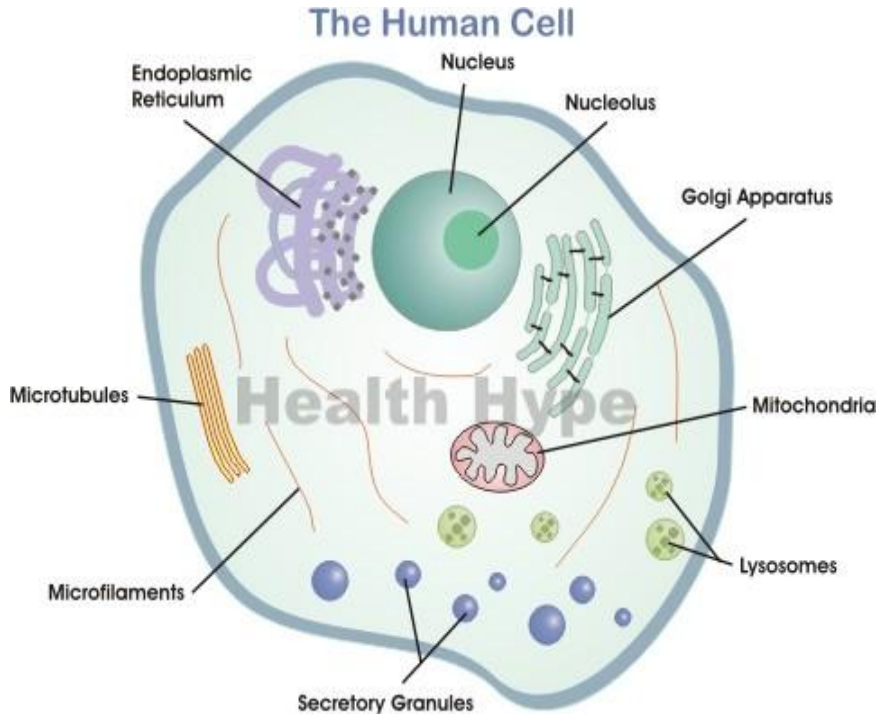


## ► Cell Types –

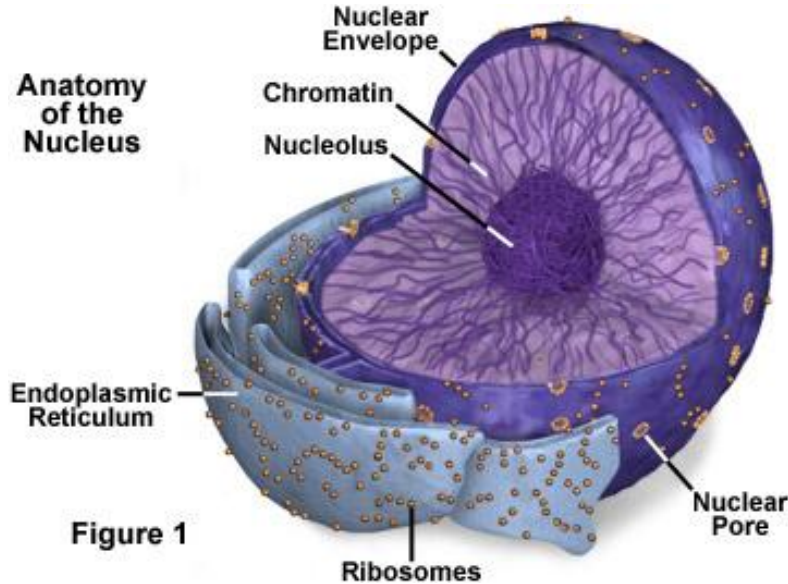
- Prokaryote – no nucleus or membrane-bounded organelles. Ex- Bacteria
- Eukaryote – have nucleus and other membrane-bounded organelles. Ex- Human

## ► Cell Contents –

- Mitochondria – Power House of Cell
- Lysosomes – Holds enzymes created by a cell
- Golgi Apparatus (body) – Processes protein and lipids
- Endoplasmic Reticulum – Involved in protein and lipid synthesis
- Nucleus – Brain of the cell
- Cytoplasm – liquid inside cell where organelles float

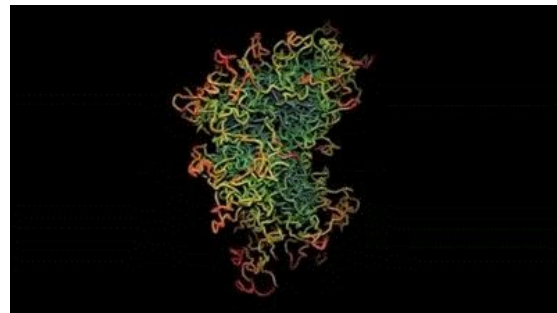
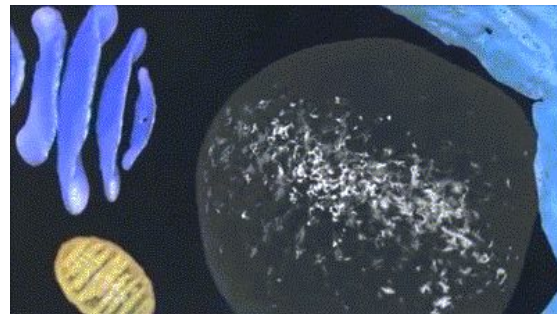
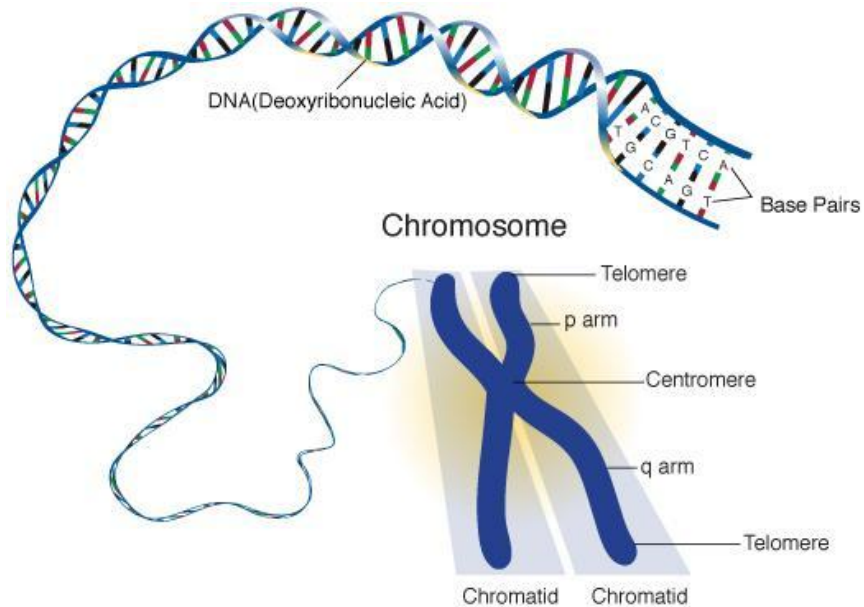


# NUCLEUS

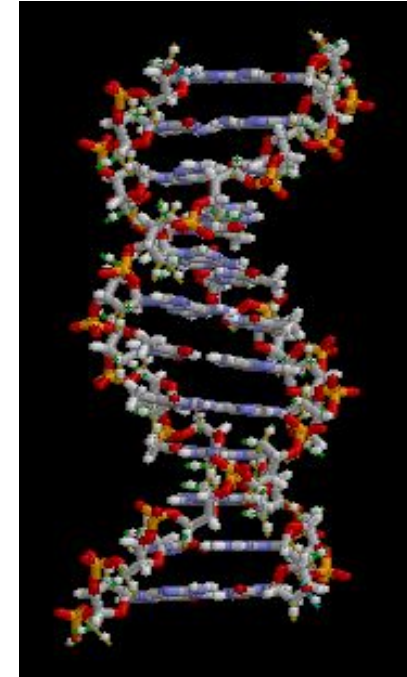
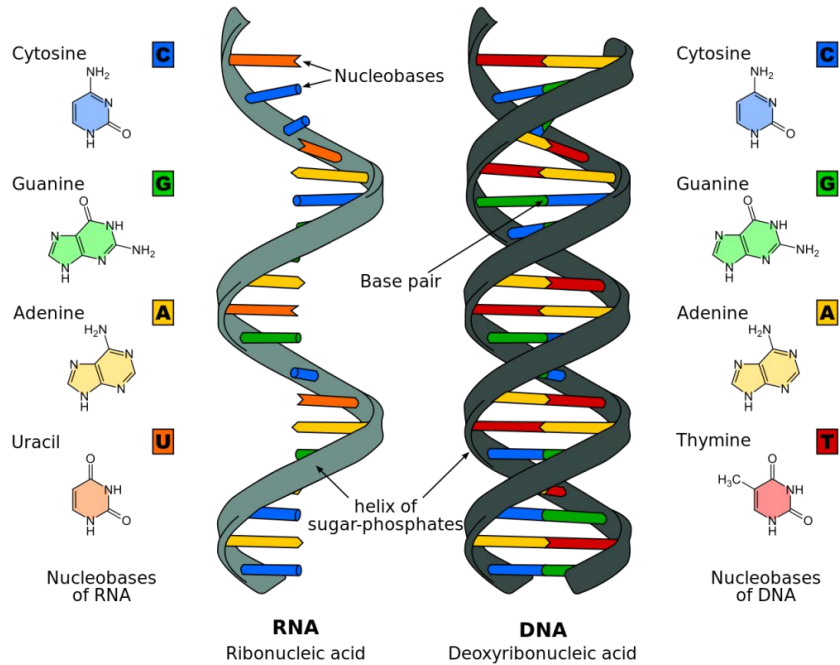


- Nuclear membrane – 2 layered membrane
- Nucleolus – dense object inside nucleus
- Chromatin – Packed inside the nucleus of every human cell is nearly 6 feet of DNA, which is divided into 46 individual molecules, one for each chromosome and each about 1.5 inches long. 23 pairs of chromosome, 22 pairs of common autosome, 1 pair sex chromosome
- Nucleoplasm – dense liquid inside nucleus

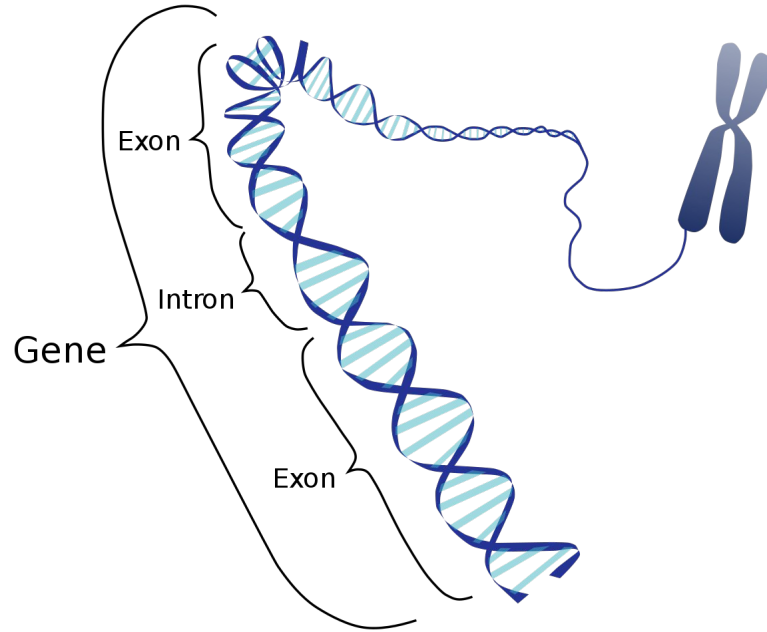
# Chromosome



# NUCLEIC ACID (DNA & RNA)



# GENE



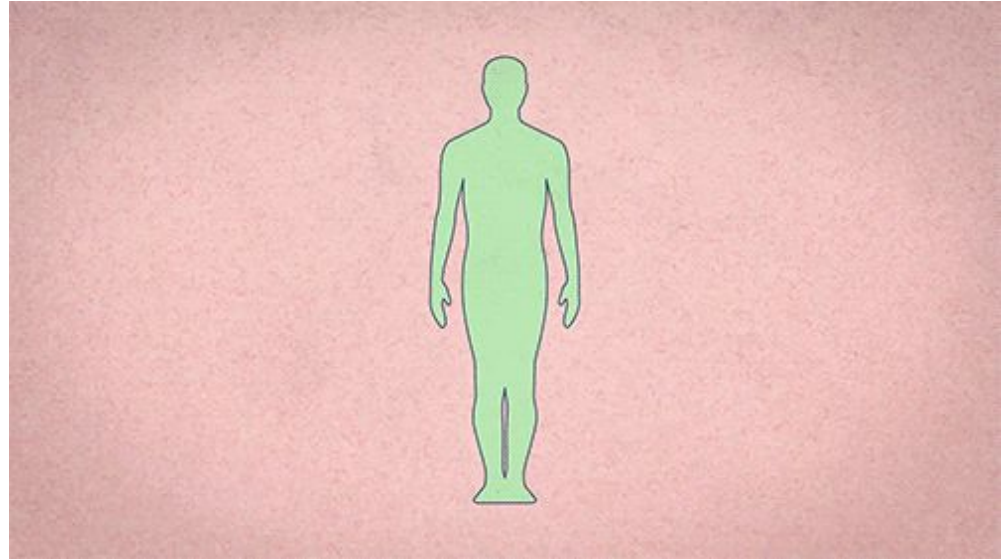
- A gene is the basic physical and functional unit of heredity.
- Portion of DNA which acts as instruction to make proteins.
- Human gene can have few hundred to 2 million base pairs.
- The Human Genome Project has estimated that humans have between 20,000 and 25,000 genes.
- Every person has two copies of each gene, one inherited from each parent.
- Most genes are the same in all people, but a small number of genes (less than 1 percent of the total) are slightly different between people (called allele).
- These small differences contribute to each person's unique physical features.



# GENOME



- A genome is an organism's complete set of DNA, including all of its genes.
- Each genome contains all of the information needed to build and maintain that organism.
- In humans, a copy of the entire genome—more than 3 billion DNA base pairs—is contained in all cells that have a nucleus.



# Overview of organizations of life

- **Nucleus = library**
  - **Chromosomes = bookshelves**
  - **Genes = books**
- 
- Almost every cell in an organism contains the same libraries and the same sets of books.
  - Books represent all the information (DNA) that every cell in the body needs so it can grow and carry out its various functions.
  - Moreover, more recent discoveries suggest that the books, bookshelves and libraries are not passive waiting to be read but are, sometimes, rewriting and rewiring themselves!

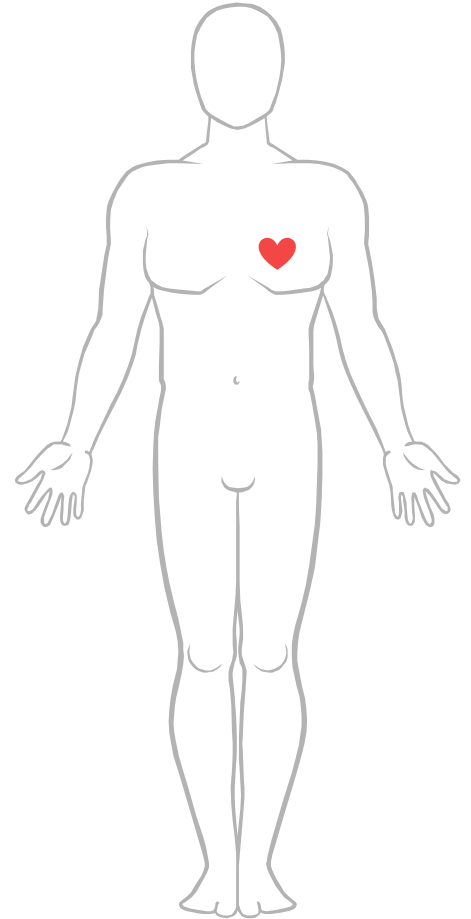


# 2. HISTORY

A Journey from Biology to Bioinformatics

# WHAT IS BIOINFORMATICS?

- Bioinformatics is the discipline that uses computers to store, retrieve, manipulate and distribute information related to biological macromolecules such as RNA, DNA and proteins
- Bringing biological themes to computers



The field of science in which biology, computer science and information technology merge into a single discipline

### **Biologists**

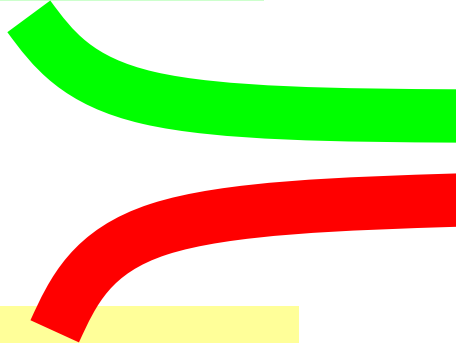
collect molecular data:  
DNA & Protein sequences,  
gene expression, etc.

### **Bioinformaticians**

Study biological questions by  
analyzing molecular data

### **Computer scientists**

(+Mathematicians,  
Statisticians, etc.)  
Develop tools, softwares,  
algorithms  
to store and analyze the data.



# SYNONYMS

*Molecular Bioinformatics, Computational Biology,  
Biocomputing*

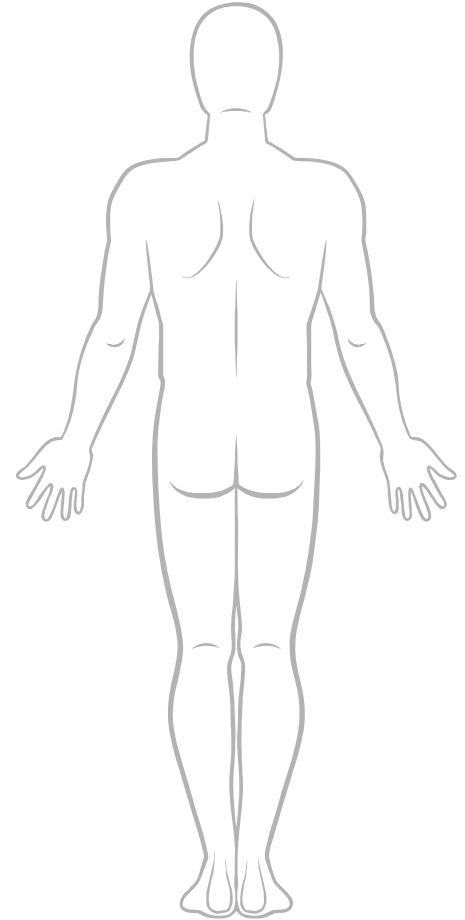


# FROM BIOLOGY TO BIOINFORMATICS

- ▶ From Academic Interest to Commercial Interest
- ▶ From Knowledge Discovery to Industrial Development
- ▶ Added Pursuit for Longer Life and Cure for Diseases
- ▶ Tremendous recent progress in Biology and Engineering

# HUMAN GENOME PROJECT

- ▶ **Completed between 1986 to April 2003**
- ▶ Identify all genes of Human DNA
- ▶ Around 3 Billion Nucleotide Bases (A, T, G, C) in a human body
- ▶ Average gene consists 3000 bases, varies in size (Largest Gene – 2.4 Million Bases)
- ▶ Total Number of Genes around 30,000
- ▶ Functions are unknown of almost 50% discovered genes





# 3. BIOINFORMATICS – PRESENT & FUTURE SCOPES

What's happening now and what might happen in the upcoming years?

# SCOPE OF RESEARCH

## **Sequence Alignment**

Arranging DNA/RNA/Protein Sequences to identify regions of similarity.

## **Genome Assembly**

Merging long/short reads by aligning and produce original ancestral gene.

## **Read Mapping**

Mapping generated long/short reads to a reference sequence.

## **Evolution Phylogenetics**

Studying evolutionary relationships among biological entities.

## **Statistical Phylogenetics**

Performing statistical analysis for finding evolutionary relationships between biological entities.

## **Population Genetics**

Study of genetic variation within populations.

# SCOPE OF RESEARCH (CONT.)

## **Genome Annotation**

Process of identifying location of genes and all other coding regions in a genome.

## **Gene Expression Analysis**

Gene expression analysis is the process by which information from a gene is used in the synthesis of a functional gene product.

## **Computer Aided Drug (CAD) Design**

Designing new medications with knowledge of a biological target and with help of computer tools.

## **Structure Prediction (RNA, Protein)**

Predicting 3D structures of RNA and Protein.

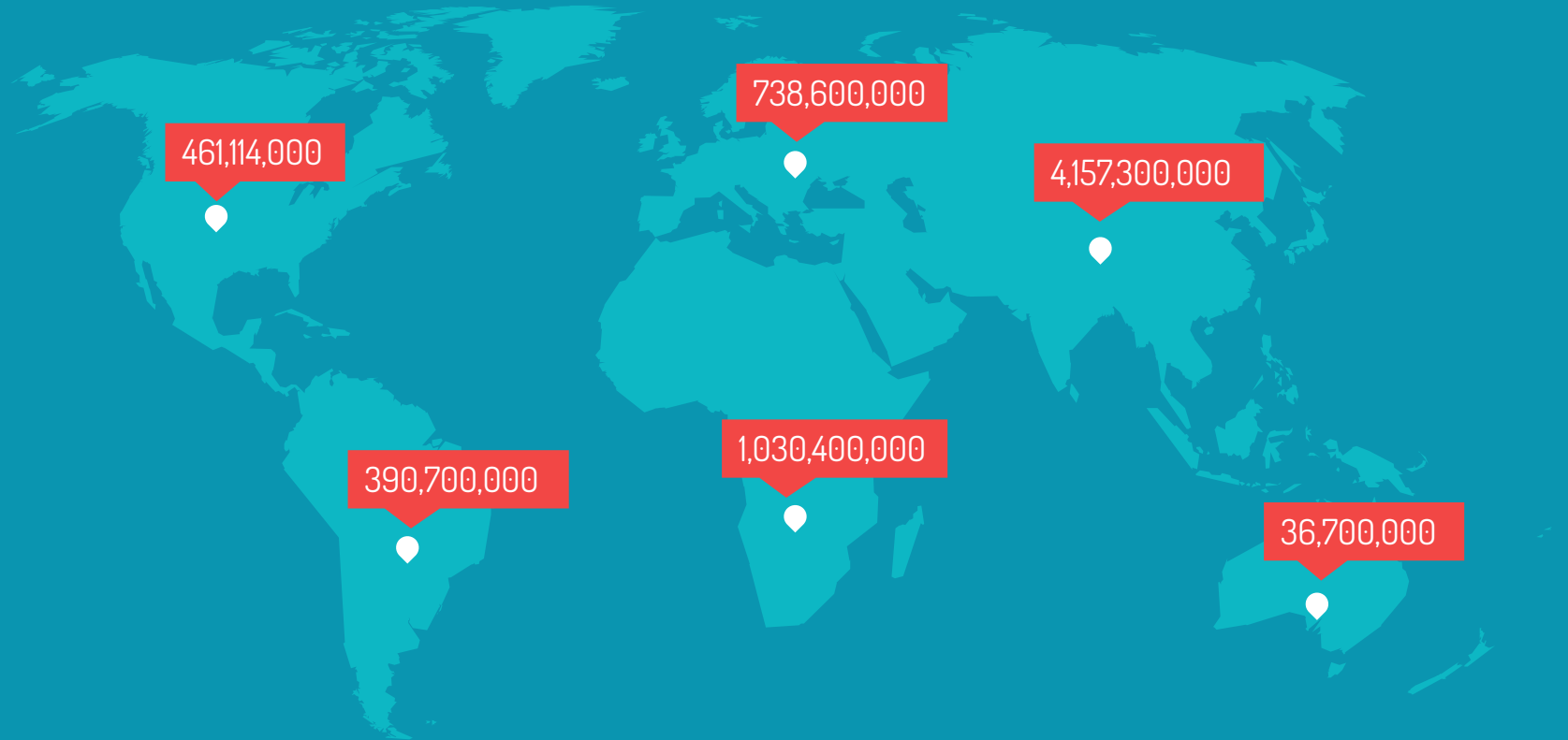
## **Association Mapping**

Linkage Analysis and disequilibrium mapping. Basically links phenotypes with genotypes.

## **Systems Biology**

Study of systems of biological components.

# CURRENT POPULATION



# 70 Miles per Hour

Nerve Impulses run to and from Brain

# 5x Times

Brain's storage capacity compared to  
Encyclopedia Britannica

# 20% Oxygen

Brain uses 20% of Oxygen entered into the blood

# 80% Water

Human brain has 80% water

