BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024,
Tamil Nadu, India.

Programme : M.Sc., Biomedical Science

Course Title : Bioinformatics

Course Code : BM35S1BI

Unit-IV

TOPIC: HUMAN GENOME PROJECT

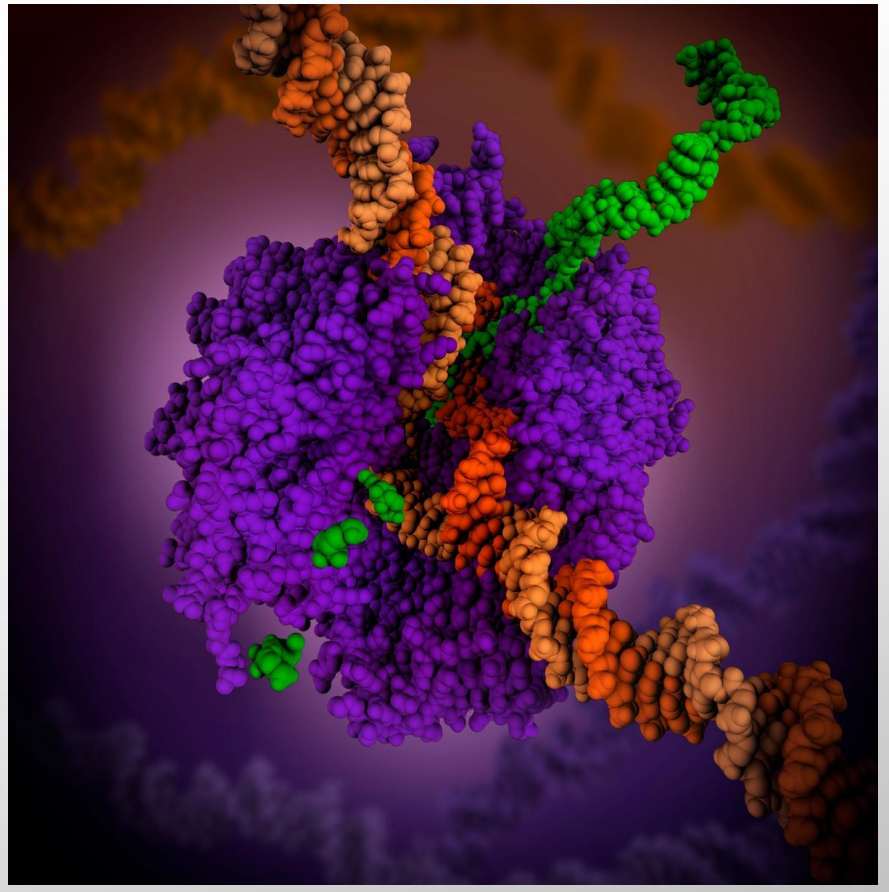
Dr. P. JEGANATHAN
Guest Lecturer
Department of Biomedical Science

HUMAN GENOME PROJECT

OBJECTIVES

- Techniques that enabled genome sequencing
- Process of Human Genome Sequence
- Importance of Genome Sequencing

- **HUMAN GENOME PROJECT (HGP)**, AN INTERNATIONAL COLLABORATION THAT SUCCESSFULLY DETERMINED, STORED, AND RENDERED SEQUENCES OF ALMOST ALL THE GENETIC CONTENT OF THE CHROMOSOMES OF THE HUMANS OTHERWISE KNOWN AS THE HUMAN GENOME.



HUMAN GENOME PROJECT

- THE HUMAN GENOME PROJECT (HGP) WAS AN INTERNATIONAL SCIENTIFIC RESEARCH PROJECT WITH THE GOAL OF DETERMINING THE BASE PAIRS THAT MAKE UP HUMAN DNA, AND OF IDENTIFYING, MAPPING AND SEQUENCING ALL OF THE GENES OF THE HUMAN GENOME FROM BOTH A PHYSICAL AND A FUNCTIONAL STANDPOINT.
- IT STARTED IN 1990 AND WAS COMPLETED IN 2003.

- THE **HUMAN GENOME PROJECT (HGP)** WAS INITIATED TO DETERMINE THE COMPLETE DNA SEQUENCE OF THE HUMAN GENOME AND ALL OF THE GENES IT ENCODES.
- THE MOST IMMEDIATE BENEFIT OF THIS INFORMATION WAS TO FACILITATE **DISEASE GENE RESEARCH.**

ESTABLISHMENT OF HGP

- HUMAN GENOME PROJECT (HGP) REMAINS THE WORLD'S LARGEST COLLABORATIVE BIOLOGICAL PROJECT.
- PLANNING FOR THE PROJECT STARTED AFTER IT WAS ADOPTED IN 1984 BY THE US GOVERNMENT, AND IT OFFICIALLY LAUNCHED IN 1990.
- IT WAS DECLARED COMPLETE ON APRIL 14, 2003, AND INCLUDED ABOUT 92% OF THE GENOME.

FOUNDER

FRANCIS COLLINS IS THE FATHER OF THE HUMAN GENOME PROJECT.

COLLINS LED THE HUMAN GENOME PROJECT AND OTHER GENOMICS RESEARCH INITIATIVES AS DIRECTOR OF THE NATIONAL HUMAN GENOME RESEARCH INSTITUTE (NHGRI), ONE OF THE 27 INSTITUTES AND CENTERS AT NIH.

PROJECT AIM

- NOTABLY, THE PROJECT WAS NOT ABLE TO SEQUENCE ALL OF THE DNA FOUND IN HUMAN CELLS; RATHER, THE AIM WAS TO SEQUENCE ONLY *EUCHROMATIC* REGIONS OF THE NUCLEAR GENOME, WHICH MAKE UP 92.1% OF THE HUMAN GENOME.
- THE REMAINING 7.9% EXISTS SCATTERED *HETEROCHROMATIC* REGIONS SUCH AS THOSE FOUND IN CENTROMERES AND TELOMERES.
- THESE REGIONS BY THEIR NATURE ARE GENERALLY MORE DIFFICULT TO SEQUENCE AND SO WERE NOT INCLUDED AS PART OF THE PROJECT'S ORIGINAL PLANS.

GOALS OF THE HUMAN GENOME PROJECT

GOALS OF THE HUMAN GENOME PROJECT INCLUDE:

- OPTIMIZATION OF THE DATA ANALYSIS.
- SEQUENCING THE ENTIRE GENOME.
- IDENTIFICATION OF THE COMPLETE HUMAN GENOME.
- CREATING GENOME SEQUENCE DATABASES TO STORE THE DATA.
- TAKING CARE OF THE LEGAL, ETHICAL AND SOCIAL ISSUES THAT THE PROJECT MAY POSE

FOCUS OF THE HUMAN GENOME PROJECT

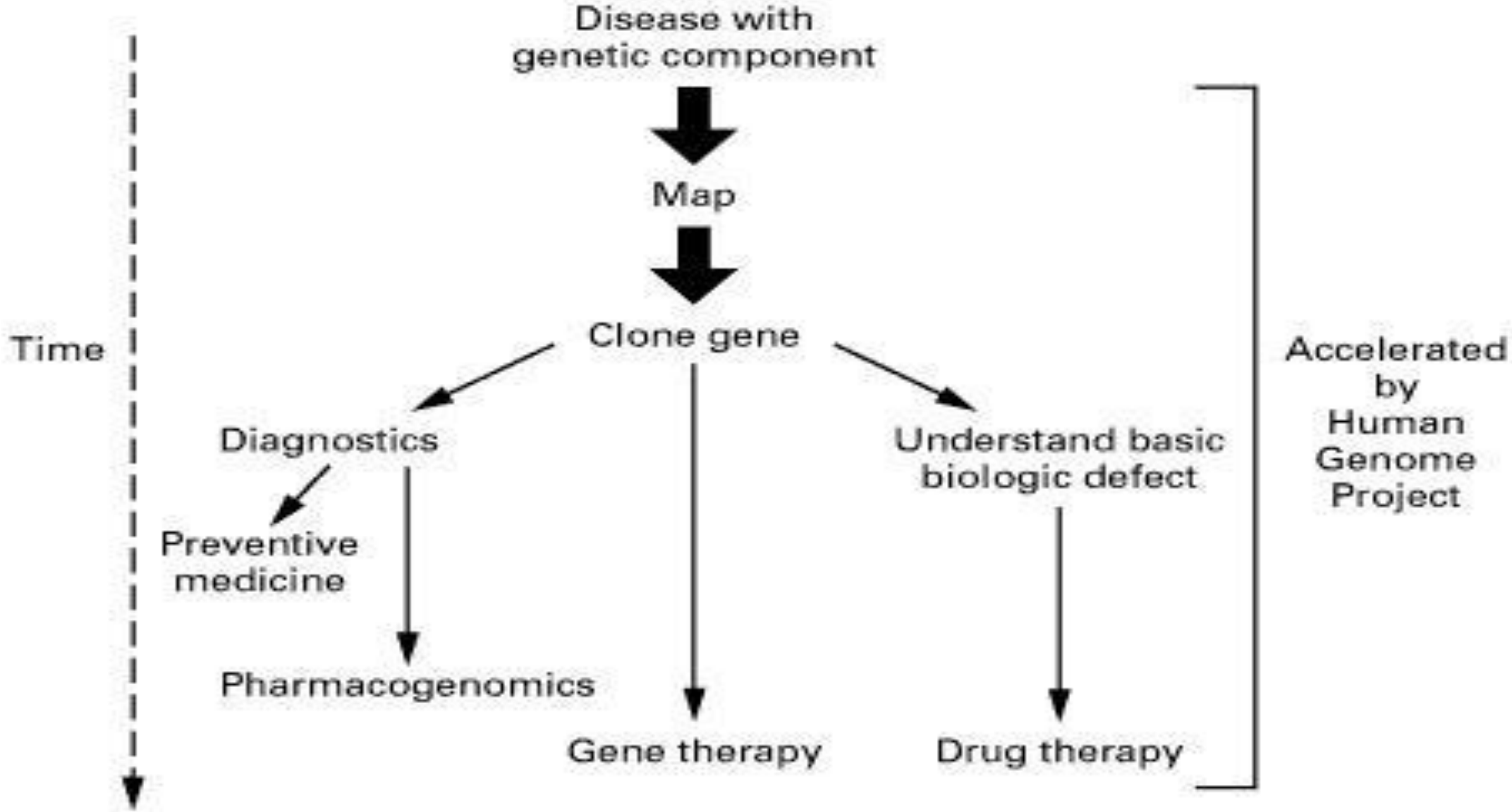
(*)THE PRIMARY WORK OF THE HUMAN GENOME PROJECT HAS BEEN TO PRODUCE THREE MAIN RESEARCH TOOLS THAT WILL ALLOW INVESTIGATORS **TO IDENTIFY GENES INVOLVED IN NORMAL BIOLOGY AS WELL AS IN BOTH RARE AND COMMON DISEASES.**

(*) THE TOOLS ARE KNOWN AS POSITIONAL CLONING.

(*) **POSITIONAL CLONING** IS A TECHNIQUE THAT IDENTIFIES A TRAIT-ASSOCIATED GENE BASED ON ITS LOCATION IN THE GENOME AND INVOLVES METHODS SUCH AS LINKAGE ANALYSIS, ASSOCIATION MAPPING, AND BIOINFORMATICS.

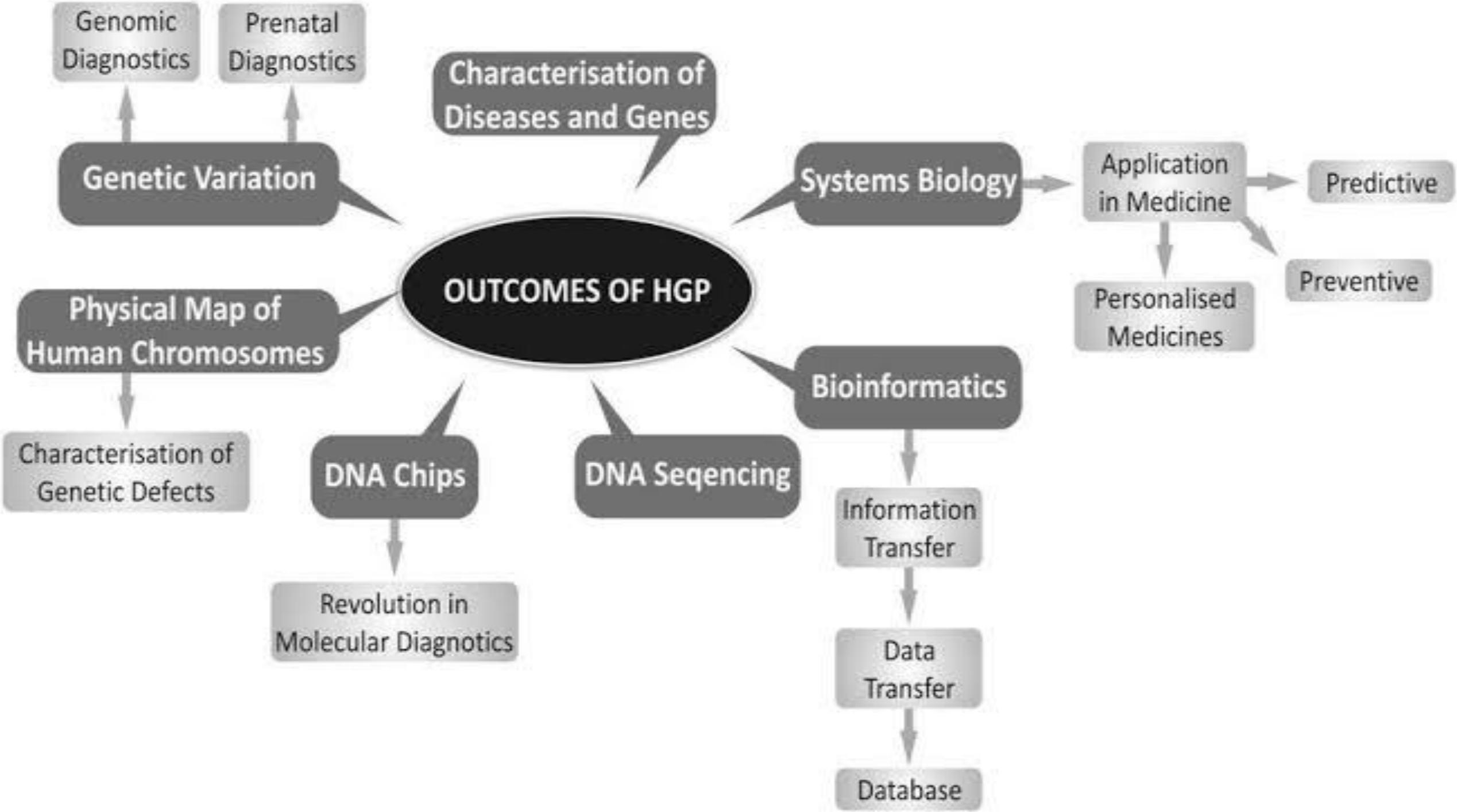
(*) THESE ADVANCED TECHNIQUES ENABLE RESEARCHERS TO SEARCH FOR **DISEASE-LINKED GENES** DIRECTLY IN THE GENOME WITHOUT FIRST HAVING TO IDENTIFY THE GENE'S PROTEIN PRODUCT OR FUNCTION.

REFERENCE: (ARTICLE BY GOATE, PP. 217–220.)


Disease with genetic component
Map
Clone gene
Time
Diagnostics
Preventive medicine
Pharmacogenomics
Gene therapy
Understand basic biologic defect
Drug therapy
Accelerated by Human Genome Project

APPLICATIONS

- THE SEQUENCING OF THE HUMAN GENOME HOLDS BENEFITS FOR MANY FIELDS, FROM MOLECULAR MEDICINE TO **HUMAN EVOLUTION** .
- THE HUMAN GENOME PROJECT, THROUGH ITS SEQUENCING OF THE DNA, CAN HELP TO UNDERSTAND DISEASES INCLUDING:
- GENOTYPING OF SPECIFIC VIRUSES TO DIRECT APPROPRIATE TREATMENT,
- IDENTIFICATION OF MUTATIONS LINKED TO DIFFERENT FORMS OF CANCER,
- THE DESIGN OF MEDICATION AND MORE ACCURATE PREDICTION OF THEIR EFFECTS.


OUTCOMES OF HGP
Genetic Variation
Genomic Diagnostics
Prenatal Diagnostics
Characterisation of Diseases and Genes
Systems Biology
Application in Medicine
Predictive
Preventive
Personalised Medicines
Physical Map of Human Chromosomes
Characterisation of Genetic Defects
DNA Chips
Revolution in Molecular Diagnotics
DNA Seqencing
Bioinformatics
Information Transfer
Data Transfer
Database

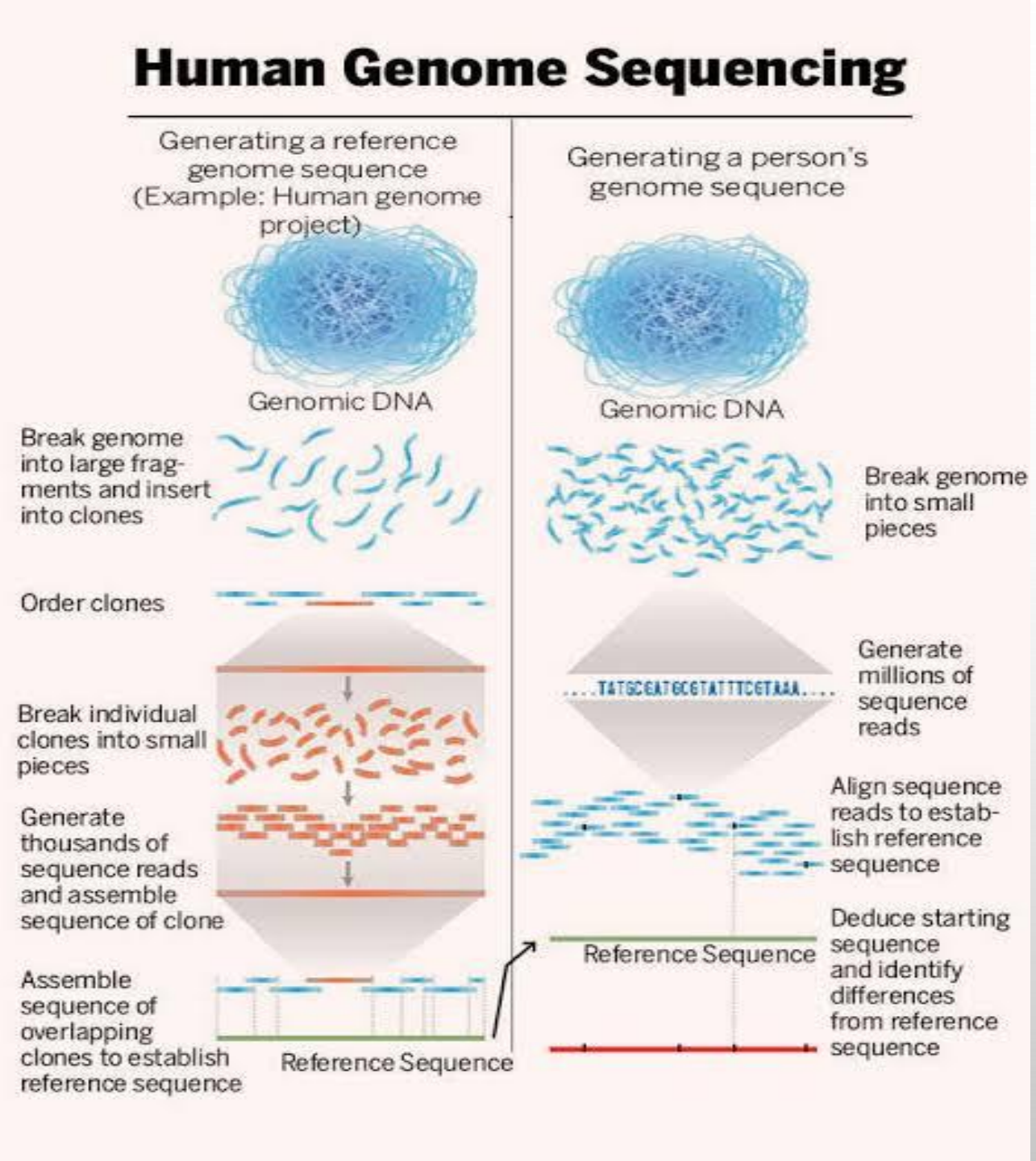
- THE SEQUENCE OF THE DNA IS STORED IN DATABASES AVAILABLE TO ANYONE ON THE INTERNET.

THE U.S. NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION (AND SISTER ORGANIZATIONS IN EUROPE AND JAPAN) HOUSE (STORE) THE GENE SEQUENCE IN A DATABASE KNOWN AS GENBANK, ALONG WITH SEQUENCES OF KNOWN AND HYPOTHETICAL GENES AND PROTEINS.

ENSEMBL PRESENT ADDITIONAL DATA AND ANNOTATION AND POWERFUL TOOLS FOR VISUALIZING AND SEARCHING GENOME SEQUENCE.

TECHNIQUES AND ANALYSIS

- THE PROCESS OF IDENTIFYING THE DNA SEQUENCE IS CALLED GENOME ANNOTATION AND IS THE DOMAIN OF BIOINFORMATICS.
- DNA SEQUENCING IS THE PROCESS OF DETERMINING THE SPECIFIC ORDER AND IDENTITY OF THE **THREE BILLION BASE PAIRS** IN THE GENOME,WITH THE ULTIMATE GOAL OF IDENTIFYING **DISCRETE DNA MOLECULES** OF KNOWN POSITION ON A CHROMOSOME,WHICH ARE THEN USED FOR SEQUENCING.



DNA EXTRACTION

(RESTRICTION ENDUNUCLEASE)

DNA FRAGMENTATION

BAC,YAC –BACTERIA ,YEAST

DNA CLONING

Sanger/SHOT GUN sequencing

DNA SEQUENCING

How ?

2 Methods

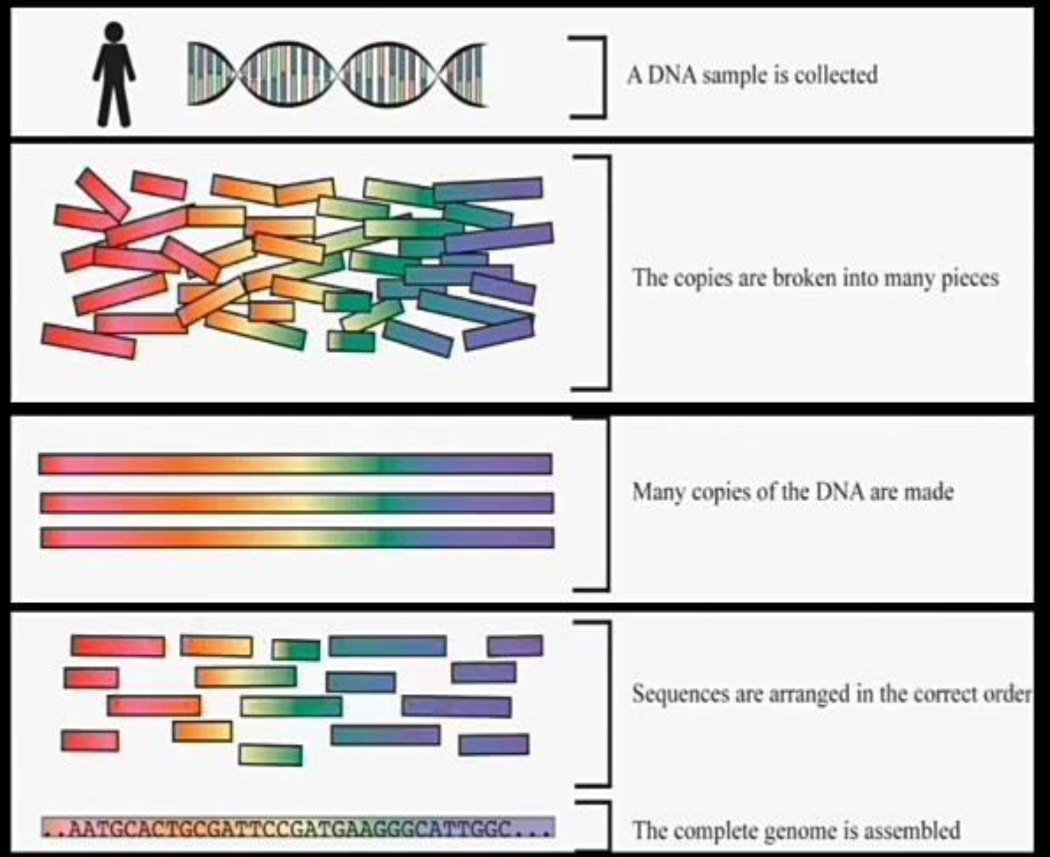
- Expressed Sequence Tags
- Sequence Annotation

- **Expressed Sequence Tags** – Only the Exons (Coding Part which produce proteins)
- **Sequence Annotation** – Whole entire DNA Sequence

But finally SEQUENCE ANNOTATIONS were used

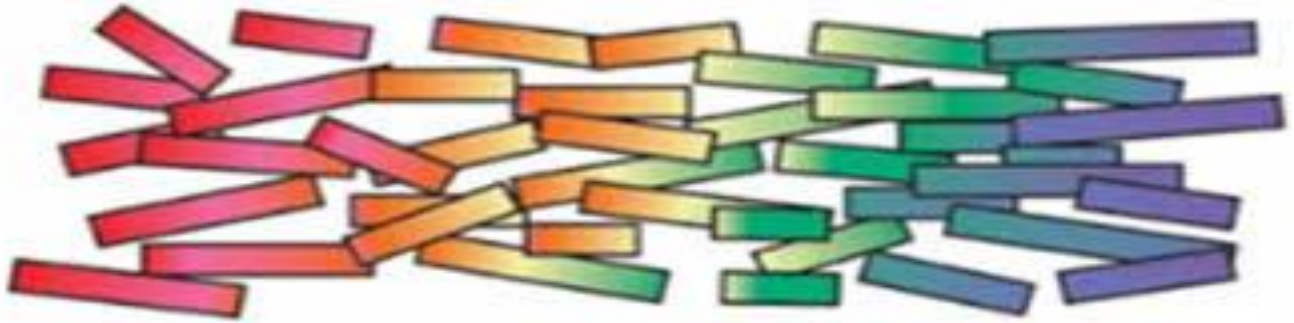
Procedure :

- DNA Extraction
  - (Restriction Endonuclease)
- DNA Fragmentation
  - Host and Vector (Bacteria and BAC) (Yeast and YAC)
- DNA Cloning
  - Sanger Sequencing (Shot gun sequencing
- DNA Sequencing



A DNA sample is collected

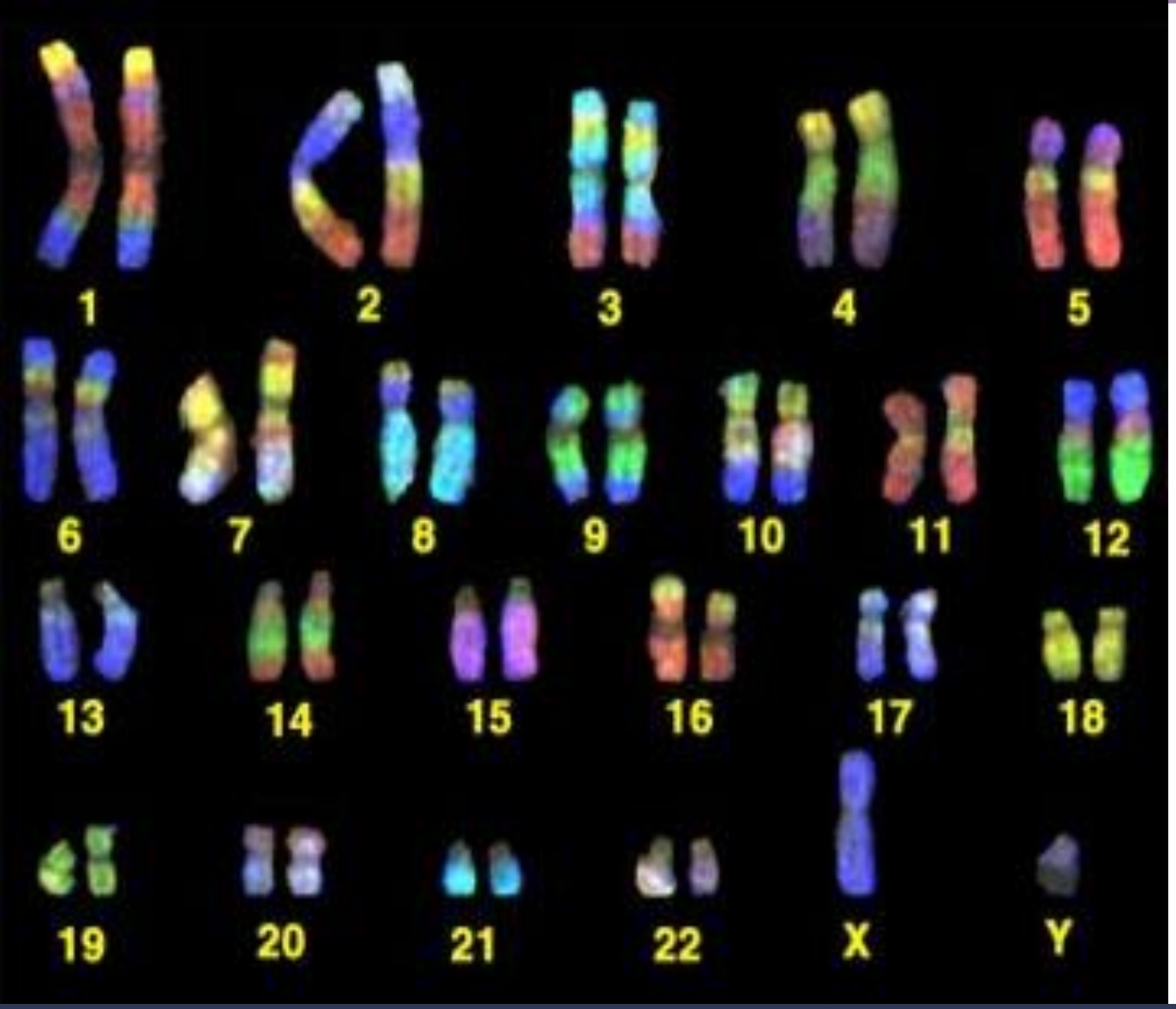
Many copies of the DNA are made

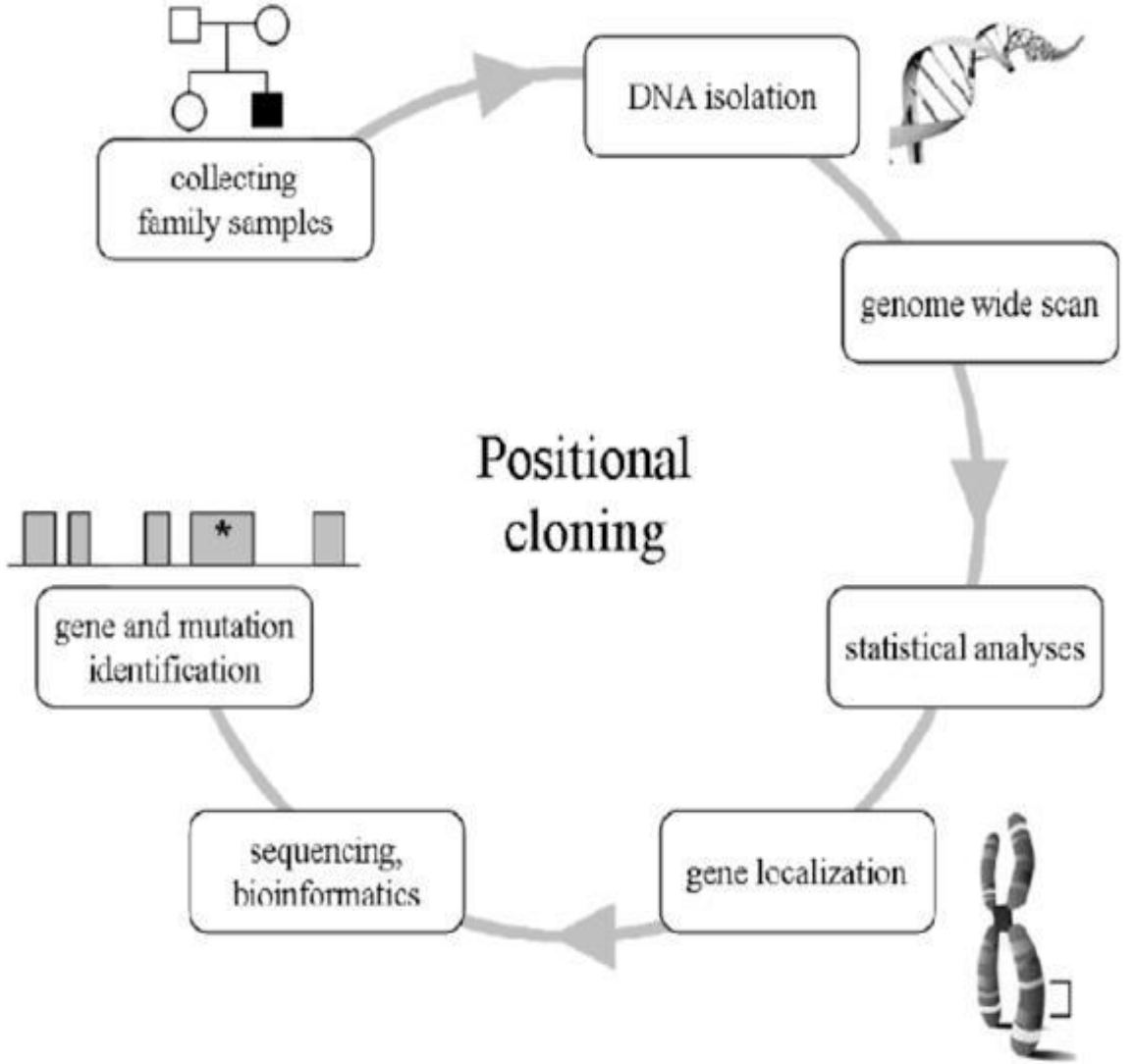


The copies are broken into many pieces

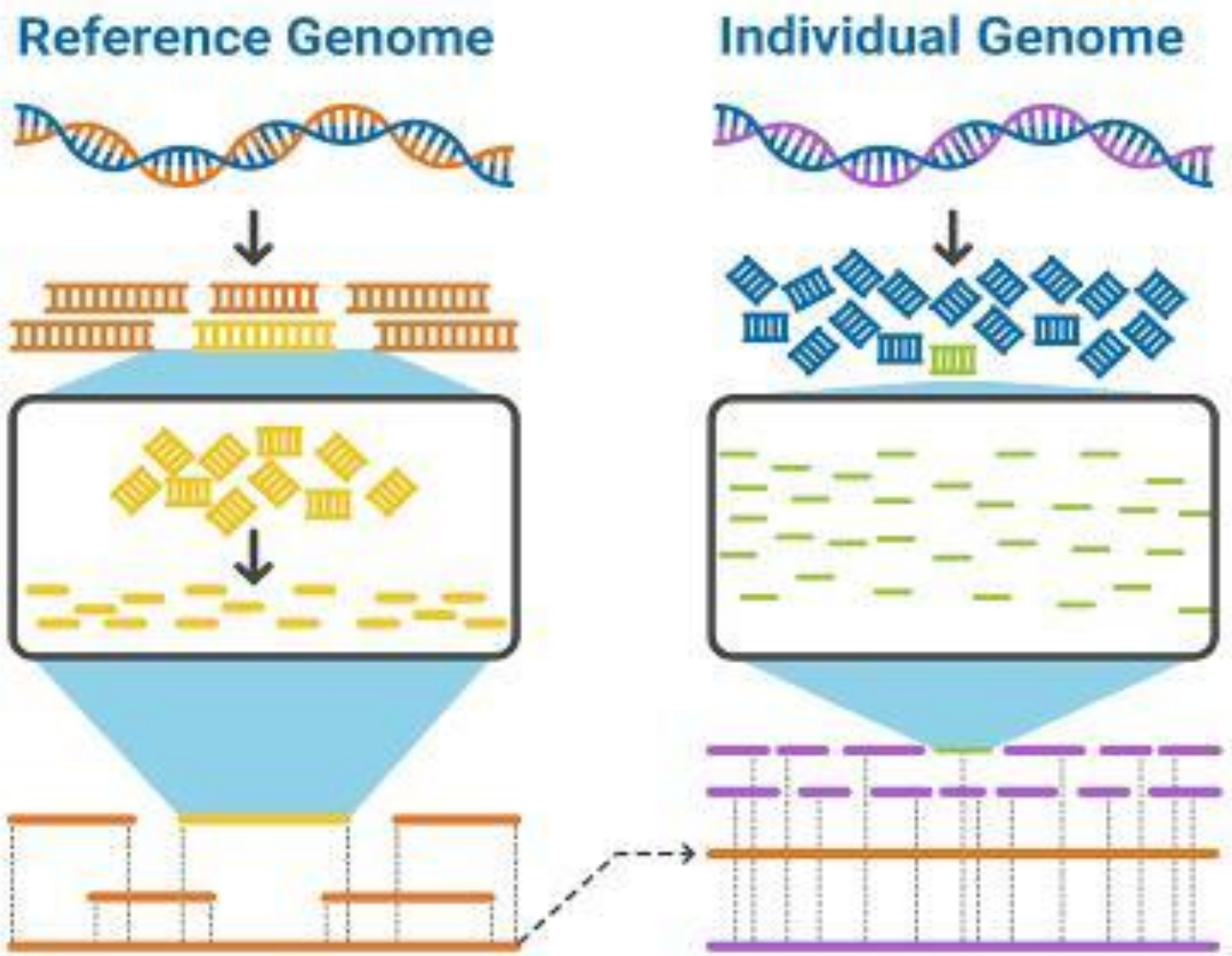
Sequences are arranged in the correct order

The complete genome is assembled


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y


collecting family samples
DNA isolation
genome wide scan
statistical analyses
gene localization
sequencing, bioinformatics
gene and mutation identification
Positional cloning

WHOLE GENOME SEQUENCING



1. Break genome into large fragments and clone
2. Break individual clone into small fragments
3. Generate thousands of sequence reads
4. Assemble sequence reads for each clone

Reference genome

1. Break genome into small fragments
2. Generate millions of sequence reads
3. Align sequence reads into a reference genome

Individual genome

ADVANCES BASED ON THE HGP

- ADVANCES IN GENETICS AND GENOMICS CONTINUE TO EMERGE.
- THE INTERNATIONAL HAPMAP PROJECT AND THE INITIATION OF LARGE-SCALE COMPARATIVE GENOMICS STUDIES DEVELOPED DUE TO HGP.
- THE INTERNATIONAL **HAPMAP** PROJECT COMPARES GENOMIC SEQUENCES WITHIN ONE SPECIES
- **COMPARATIVE GENOMICS** IS THE STUDY OF SIMILARITIES AND DIFFERENCES BETWEEN DIFFERENT SPECIES.

*COMPARING DNA SEQUENCES IS OFTEN DONE USING A SOFTWARE TOOL CALLED **BLAST (BASIC LOCAL ALIGNMENT SEARCH TOOL)**

- THROUGH THIS RESEARCHERS ARE ABLE TO IDENTIFY DEGREES OF SIMILARITY AND DIVERGENCE BETWEEN THE GENES AND GENOMES OF RELATED OR DISPARATE SPECIES.

IMPACTS OF THE HGP

- **IMPACT ON MEDICINE:**

THE PUBLIC AVAILABILITY OF A COMPLETE HUMAN GENOME SEQUENCE REPRESENTED A DEFINING MOMENT FOR BOTH THE BIOMEDICAL COMMUNITY AND FOR SOCIETY.

IN THE YEARS SINCE COMPLETION OF THE HGP, THE HUMAN GENOME DATABASE, TOGETHER WITH OTHER PUBLICLY AVAILABLE RESOURCES SUCH AS THE HAPMAP DATABASE, HAS ENABLED THE IDENTIFICATION OF A VARIETY OF GENES THAT ARE ASSOCIATED WITH DISEASE.

REFERENCES

(*) https://www.genome.gov/human-genome-project&ved

(*) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6875757/

(*) Understanding The Human Genome Project by **Michael A. Palladino, 2002**

(*) The Book of Man: The Human Genome Project and the Quest to Discover Our Genetic Heritage-Walter Bodmer, 1994

THANK YOU