

2. Genome India Project

Prelims Syllabus: Bio-Technology

Mains Syllabus: GS-III Awareness in the fields of IT, Space, Computers, robotics, nano-technology, bio-technology and issues relating to intellectual property rights.



Why in News?

- The Genome India Project, a Centre-backed initiative to sequence 10,000 Indian human genomes and create a database, is about two-thirds completed and will be 100% complete by year-end.

About the News:

- Of the 7,000 genomes sequenced about 3,000 are already available for public access (as per the Department of Biotechnology).

About Human Genome Project:

- The Human Genome Project (1990 to 2003) 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.

What is Whole Genome Sequencing?

- All organisms have a unique genetic code, or genome, that is composed of nucleotide bases- Adenine (A), Thymine (T), Cytosine (C) and Guanine (G).
- The unique Deoxyribonucleic Acid (DNA) fingerprint, or pattern can be identified by knowing the sequence of the bases in an organism.

- Determining the order of bases is called sequencing.
- Whole genome sequencing is a laboratory procedure that determines the order of bases in the genome of an organism in one process.

Methodology:

- **DNA Shearing:**
 - ✓ Scientists begin by using molecular scissors to cut the DNA, which is composed of millions of bases (A's, C's, T's and G's), into pieces that are small enough for the sequencing machine to read.
- **DNA Bar Coding:**
 - ✓ Scientists add small pieces of DNA tags, or bar codes, to identify which piece of sheared DNA belongs to which bacteria.
 - ✓ This is similar to how a bar code identifies a product at a grocery store.
- **DNA Sequencing:**
 - ✓ The bar-coded DNA from multiple bacteria is combined and put in a DNA sequencer.
 - ✓ The sequencer identifies the A's, C's, T's, and G's, or bases, that make up each bacterial sequence.
 - ✓ The sequencer uses the bar code to keep track of which bases belong to which bacteria.
- **Data Analysis:**
 - ✓ Scientists use computer analysis tools to compare sequences from multiple bacteria and identify differences. The number of differences can tell the scientists how closely related the bacteria are, and how likely it is that they are part of the same outbreak.

Advantages of Genome Sequencing:

- Provides a high-resolution, base-by-base view of the genome
- Captures both large and small variants that might be missed with targeted approaches
- Identifies potential causative variants for further follow-up studies of gene expression and regulation mechanisms
- Delivers large volumes of data in a short amount of time to support assembly of novel genomes

Significance of Genome Sequencing:

- Genomic information has been instrumental in identifying inherited disorders, characterizing the mutations that drive cancer progression, and tracking disease outbreaks.

- It is beneficial for sequencing agriculturally important livestock, plants, or disease-related microbes.

What is Genome?

- A genome refers to all of the genetic material in an organism, and the human genome is mostly the same in all people, but a very small part of the DNA does vary between one individual and another.
- Every organism’s genetic code is contained in its DNA, the building blocks of life.
- The discovery that DNA is structured as a “double helix” by James Watson and Francis Crick in 1953, started the quest for understanding how genes dictate life, its traits, and what causes diseases.
- Each genome contains all of the information needed to build and maintain that organism.
- In humans, a copy of the entire genome contains more than 3 billion DNA base pairs.

What is the Difference between Genome and Gene?

**GENE
VERSUS
GENOME**

| | |
|--|---|
| A gene is a part of a DNA molecule | The genome is the total DNA in a cell |
| Hereditary element of genetic information | All set of nuclear DNA |
| Encodes protein synthesis | Encodes both proteins and regulatory elements for protein synthesis |
| Length is about a few hundreds of bases | Length of the genome of a higher organism is about billion base pairs |
| A higher organism has about thousands of genes | Each organism has only one genome |
| Variations of the gene named alleles can be naturally selected | Horizontal gene transfer & duplication cause large variations in the genome |