Computation Analysis of Indian Genes

Why in News

Recently, the results from the computation analysis of the 1029 sequenced genomes from India were published in the scientific journal, Nucleic Acid Research.

Key Points

- The analysis has found that out of **55,898,122 single nucleotide variants** in India 18,016,257 **(32.23%) variants are unique and found only in the samples sequenced from India.**
  This emphasizes the need for an India centric population genomic initiative.
- The **analysis was carried out by CSIR** constituent labs, Institute of Genomics and Integrative Biology (IGIB), Delhi and Centre for Cellular and Molecular Biology (CCMB), Hyderabad.

IndiGenomes Resource Database

- It encompasses the **genomic data from over 1000 whole genome sequences sequenced** from across India as part of the **IndiGen programme** and represents diverse geographies and ethnicities.
  - The **IndiGen Programme** was launched in **April 2019,** under which **genome sequencing** of 1029 self-declared healthy Indians drawn from across the country has been completed.
  - Its **objective** was to enable genetic epidemiology and develop public health technologies applications using population genome data.
- The resource provides **access to over 55 million genetic variants representing the contemporary Indian population** with an objective to classify variants involved in **mendelian disorders** and **improve precision medicine outcomes.**
  Mendelian disorder is a type of genetic disorder, resulting due to alterations in a gene or due to abnormalities in the genome.

- It enables the **identification of markers for carrier screening, variations causing genetic diseases, prevention of adverse events** and provides better diagnosis and optimal therapy **through mining data.**
- This resource can provide useful insights for clinicians and researchers in comprehending genetics not only at the population level but at the individual level.

Background

- Despite having this rich genetic diversity, India has been under-represented in global genome studies. Further, the population architecture of India has resulted in high prevalence of recessive alleles.
- In the absence of large-scale whole genome studies from India, the population-specific genetic variants are not adequately captured and catalogued in global medical literature.
- In order to fill the gap of whole genome sequences from different populations in India, CSIR initiated the **IndiGen Programme.**

Genome

- A genome is all the **genetic matter** in an organism. It is defined as an organism's complete set of **Deoxyribose Nucleic Acid (DNA),** including all of its genes.
  Every organism's genetic code is contained in its **DNA,** the building blocks of life.
- 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.

Genome sequencing

- Genome sequencing is figuring out the **order of DNA nucleotides, or bases,** in a genome—the order of **Adenine, Cytosine, Guanines, and Thymine** that make up an organism's DNA.
- **Application:**
  - Genome sequencing can play a crucial role for **new advancements in medical science** (like predictive diagnosis and precision medicine, genomic information) and in **disease management.**
  - Through genome sequencing methodology, researchers and clinicians can **easily detect the disease related to genetic disorder.**

- **Importance of Genome Sequencing:**
    - The study of entire genome sequences will help understand **how the genome as a whole works—**how genes work together to direct the growth, development and maintenance of an entire organism.
    - The genes account for less than 25% of the DNA in the genome, and so knowing the entire genome sequence will **help scientists study the parts of the genome outside the genes.**

Source: PIB