Genetic Mapping of Indians

Mains Syllabus: GS III - Science and Technology- developments and their applications and effects in everyday life.

Why in News?

Recently , the preliminary findings of the GenomeIndia project, were published in the journal Nature Genetics.

What is the GenomeIndia Project?

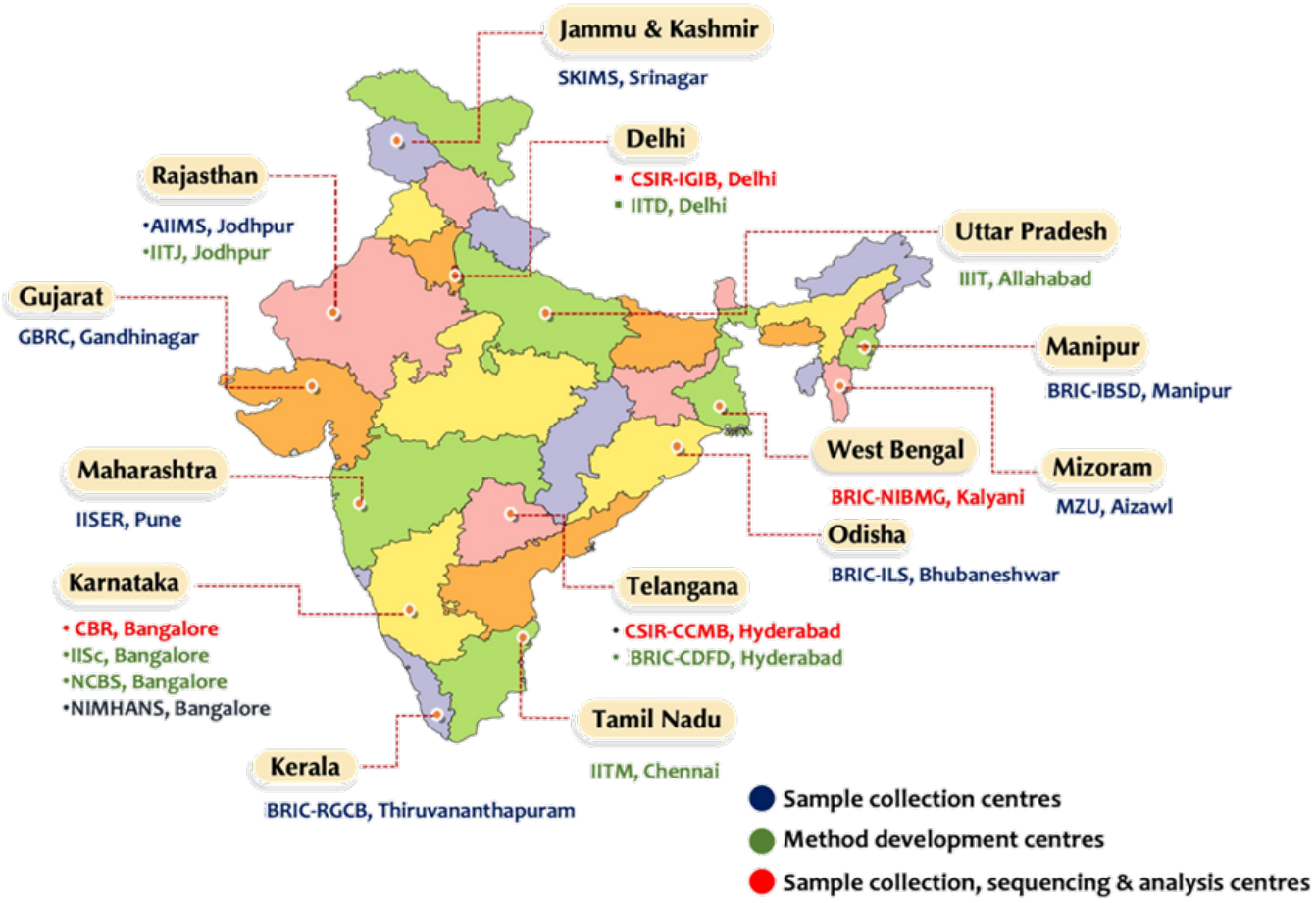
A genome is the complete set of genetic information for an organism, encoded in DNA or RNA. It contains all the instructions for an organism's development and maintenance.

- **Genome India** – It is a is a pioneering scientific project funded by the Department of Biotechnology, Ministry of Science and Technology, Government of India.

Genetic variation refers to the differences in DNA sequences between individuals within a species or population.

These differences are often expressed as variations in genes, which are then passed down to offspring, leading to a range of traits within a population.

- **Objective** - To build a comprehensive catalogue of genetic variations that reflect the unique diversity of the Indian population.
- **Collaborative work** - The project marks a landmark collaboration of 20 academic and research institutions to drive a genomics-based health revolution for India.



- **Sample collection** - A total of 20,000 samples have been collected from 83 diverse populations(30 tribal and 53 non-tribal) , cutting across the length and breadth of India.
- **Data collection** - Blood samples and associated phenotype data such as weight, height, hip circumference, waist circumference and blood pressure were collected.
- **Sample diversity** - Samples were collected from inhabitants of over 100 distinct geographical locations to estimate the relatively rare mutations that are important to understand complex diseases.
- **Genomes of five tribes across India** - Tibeto-Burman tribe, Indo-European tribe, Dravidian tribe, Austro-Asiatic tribe, and a continentally admixed outgroup — were sequenced.
- **Genomes of three non-tribes** — Tibeto-Burman non-tribe, Indo-European non-tribe, and Dravidian non-tribe — were also sequenced.
- **Unrelated individuals** - The samples were taken from unrelated individuals to ensure accurate estimation of mutation frequencies across groups.
- **Parent-child pairs** - Three to six parent-child pairs were included in each population group to uncover de novo mutations (mutations that occur randomly in a child but not seen in parents).

IndiGen is the genome sequencing project of Council of Scientific and Industrial Research (CSIR) for conducting a "whole-genome sequence" of a 1,008 Indians.

What do the preliminary findings reveal?

- **Sample sequencing** - Whole genome sequencing has been completed for 10,000 samples and the data for 10,000 individuals have been archived at the Indian Biological Data Centre (IBDC).
- After excluding two populations, the published findings are based on the genetic information of 9,772 individuals — 4,696 male participants and 5,076 female participants.
- **Mutations** - In total, 180 million mutations have been found from the individuals sequenced.

A mutation is a change in the DNA sequence of an organism resulting from errors in DNA replication or from exposure to environmental factors like radiation or chemicals. Mutations can be beneficial, harmful, or have no effect.

- While 130 million variations are in the non-sex chromosomes (22 pairs of autosomes), 50 million mutations are in the sex chromosomes X and Y.
- A large number of the 180 million variants found in the sequenced genomes of 9,772 individuals are very likely to be present in the non-coding regions.

Non-coding DNA refers to the regions of a genome that do not directly code for proteins.

While previously thought to have no function, non-coding DNA plays vital roles in various cellular processes and is crucial for gene regulation, genome structure, and evolution.

- **Tracing evolutionary history** - Polymorphisms or variations in the non-coding regions of the human genome, particularly the mutations that are evolutionarily conserved, will help in tracing evolutionary history.

Many of the “contemporary Indian populations have originated from a few founding groups and have maintained distinct identities through centuries of endogamy.

- **Prevalence of endogamy** - It is highly prevalent in all the 83 population groups under study, though the degree varies.

Endogamy is the cultural practice of mating within a specific social group, religious denomination, caste, or ethnic group.

- **Effect of endogamy** - As a result of the centuries-long practice of endogamy, population-specific unique variations, including distinct disease-causing mutations with amplified frequencies, are likely to be seen within specific groups.

What are the significances of these studies?

- **Recognising India's genomic landscape**- While the global genomic landscape is predominantly Eurocentric, and India has been severely underrepresented in these studies.
- The study is therefore important for having captured the genetic diversity of "one of the highly underrepresented populations in the global genomics landscape".
- **Public policy development** - Genetic mutations found associated with endogamous population-specific diseases will help the government to come up with targeted public health policies.
- **Spur further studies** - The 130 million variations identified are expected to spur studies that aim to determine the possible roles of population-specific genetic mutations in various diseases.
- **Medical implications** - Understanding genetic variations can pave the way for precision medicine, ensuring treatments and interventions tailor-made for Indian genetic profiles.
- The data on variants associated with diseases will enable the development of affordable, genomics-based diagnostic tools, facilitating early detection, and prevention and management of diseases in India.

Reference

The Hindu | Genetic mapping of Indians

SHANKAR IAS PARLIAMENT
Information is Empowering