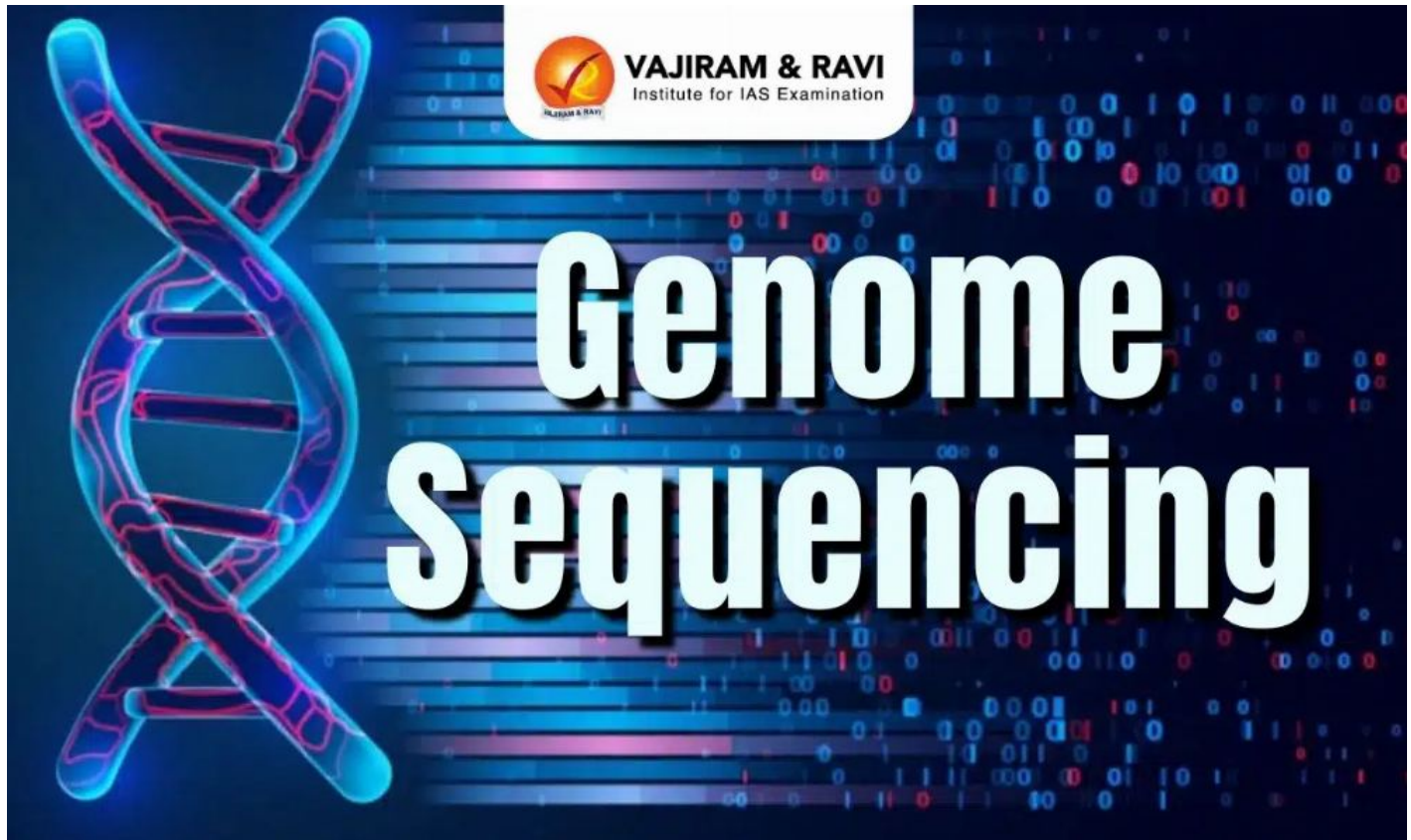


Genome Sequencing, Applications, Significance, Limitations

July 10, 2026 • vajiramandravi.com



Genome sequencing is the process of determining the **DNA** sequence of an organism's genome. A genome is a complete set of DNA that contains all of the genes of an organism. Genome sequencing involves figuring out the order of bases in an **organism's entire genome**. It is supported by **automated DNA sequencing** methods and computer software to assemble the massive **sequence data**.

The Genome India Project is a significant initiative aimed at sequencing the genomes of 10,000 individuals from diverse populations across India. Launched by the Department of Biotechnology, this project seeks to create a comprehensive catalogue of genetic variations specific to India's population.

Genome Overview

The genome is the complete set of DNA instructions present in a cell. The human genome is made up of a tiny chromosome in the **cell's mitochondria** and 23 pairs of chromosomes that are found in the nucleus of the cell. A genome contains all of the information required for a person to develop and function.

- **Deoxyribonucleic acid (DNA):** It is the chemical substance that holds the instructions required for regulating the growth and development of almost all living things. The two twisted, linked strands that makeup DNA molecules are frequently referred to as a "**double helix.**"
- **Gene:** A gene is a unit of DNA that contains the instructions for making a specific protein or set of proteins.
- **Sequencing:** The process of determining the exact order of the bases in a strand of **DNA** is known as **sequencing**. Because bases exist in pairs and the identity of one of the bases in the pair determines the identity of the other member of the pair, researchers are not required to report both bases of the pair.

What is Genome Sequencing?

The sequence of base pairs is identical in all humans, there are differences in the genome of every human being that make them unique. The process of deciphering the order of base pairs, to decode the genetic fingerprint of a human is called genome sequencing.

Whole Genome Sequencing

All organisms (bacteria, plants, and mammals) have their **own genetic code**, or genome, which is made up of nucleotide bases. Whole genome sequencing is a laboratory procedure used to determine the order of bases in the genome of an organism in a single step.

Steps Involved in Whole Genome Sequencing	
DNA shearing	- Scientists start by using molecular scissors to cut the DNA of any organism (for example bacteria), which is made up of millions of bases, into pieces small enough for the sequencing machine to read.
DNA barcoding	- Small pieces of DNA tags, or bar codes, are added by scientists to identify which piece of sheared DNA belongs to which organism.
DNA sequencing	- Multiple bacteria's bar-coded DNA is combined and placed in a DNA sequencer. - Each bacterial sequence's A, C, T, and G bases are identified by the sequencer. - The sequencer uses the bar code to keep track of which bases belong to which bacteria.
Data analysis	- Scientists use bioinformatic tools to compare the sequences from different bacteria and identify differences. - The number of differences between bacteria can indicate how closely related they are.

Global efforts towards Genome Sequencing

Various international projects are working together to map the genomes of all plants, animals, fungi, and other eukaryotic life on Earth.

Human Genome Project

The Human Genome Project was a significant **global** scientific endeavour whose primary goal was to create the **first sequence of the human genome**.

- The Project formally began in **1990** and was completed in 2003, to discover all the estimated **20,000-25,000** human genes and make them accessible for further biological study.
- **Objectives:** To create three research tools that will enable scientists to identify genes that are involved in both rare and common diseases.
 - To investigate and educate the public about the ethical, legal, and social implications of new genetic technologies.

Encyclopedia of DNA Elements (ENCODE) Project

The Encyclopedia Of DNA Elements (ENCODE) project **began in 2003** by the **US National Human Genome Research Institute (NHGRI)**.

- It is an international research effort funded by the **National Human Genome Research Institute (NHGRI)** that aims to identify all functional elements (FE) in the human genome.
 - **FE include** protein-coding regions, regulatory elements such as promoters, silencers, or enhancers, and sequences that are important for chromosomal structure.
- **Objective:** To create a comprehensive list of functional elements in the human genome, including elements that act at the protein and **RNA** levels, as well as regulatory elements that control cells and the conditions under which a gene is active.

Earth Biogenome Project

The Earth BioGenome Project (EBP), a **biology moonshot**, aims to sequence, catalogue, and characterize the genomes of all **eukaryotic biodiversity** on Earth over a ten-year period.

- The project was officially launched in 2018 and is expected to take approximately ten years to complete.
- **Objectives:** Creating a digital library of all known **eukaryotic life's DNA** sequences can aid in the development of effective tools for preventing biodiversity loss and pathogen spread, monitoring and protecting ecosystems, and improving **ecosystem services**.

Genome Sequencing India

Today, genomic data has enormous potential to improve healthcare strategies in a variety of ways, including disease prevention, improved diagnosis, and optimized treatment. Genome sequencing technologies are now more widely available to the general public in India.

- However, information on variants linked to a number of major diseases is not readily available in publicly accessible databases.
- In order to fill the gap of whole genome sequences from different populations in India, the **Government of India** has launched many programmes.

IndiGen Programme

The IndiGen programme launched in **2019 aims to** undertake whole genome sequencing of thousands of individuals representing diverse ethnic groups from India.

- The **goal is to** enable genetic epidemiology and develop public health technology applications using population genome data.
- IndiGen is endorsed by the **Council for Scientific and Industrial Research (CSIR)**.
- The outcomes of the **IndiGen** will be utilized towards understanding the genetic diversity on a population scale, making available genetic variant frequencies for clinical applications, and enabling genetic epidemiology of diseases.

Indian Initiative on Earth Bio-Genome Sequencing

The project was launched in 2020 and is part of the **Earth Biogenome Project**. The project will allow for the collection and preservation of endangered and economically significant species.

- The genetic information that has been decoded will also be useful in preventing biopiracy.
- The initial phase of **IIEBS** will include the whole genome sequencing of 1,000 plant and animal species, which will take five years to complete.
- The **Jawaharlal Nehru Tropical Botanic Garden and Research Institute (JNTBGRI)** plays a key role in a nationwide project to decode the genetic information of all known species of plants and animals in the country.

Genome India Project

The **Genome India Project** is a significant initiative aimed at sequencing the genomes of 10,000 individuals from diverse populations across India. Launched by the Department of Biotechnology, Government of India, this project seeks to create a comprehensive catalogue of genetic variations specific to India's population.

- The project completed in January 2025, and the data is now available at the **Indian Biological Data Centre (IBDC)**.
- This data will help scientists and researchers better understand India's genetic landscape, aiding in the development of personalized healthcare solutions and informing policy decisions
- **Significance:** Creating an Indian genome database allows researchers to learn about genetic variants unique to India's population groups and use this information to **tailor drugs and therapies**.
 - **Objective:** To "**unravel the genetic underpinnings of chronic diseases** currently on the rise in India.

Significance of Genome Sequencing

Genome sequencing is crucial in pharmacogenomics, clinical diagnosis, and vaccine development. It provides precise data for outbreak detection, bacterial characterization, and mutation identification. Advancements have lowered costs, aiding in inherited disorder analysis and cancer research.

- Whole genome sequencing provides detailed and precise data for identifying outbreaks sooner.
 - Additionally, whole genome sequencing is used to characterize bacteria as well as **track outbreaks**.
- Dramatic advancements in **DNA-sequencing** technologies have massively reduced the time and cost required to sequence an entire human genome.
- Genomic information has been instrumental in identifying **inherited disorders** and characterizing the mutations that drive **cancer progression**.

Applications of Genome Sequencing

Genome sequencing has transformed various fields by unlocking insights into genetic information. Its applications span biological research, forensics, diagnostics, drug efficacy, vaccine development, population studies, and agriculture, revolutionizing healthcare, food security, and disease prevention.

- **Biological research:** The ability to read genetic sequences is extremely useful in biological research because the base sequence contains information for making proteins as well as **regulating gene functions**.
- **Forensics:** Sequencing has proven to be a powerful tool in forensics. Because differences in **DNA and RNA** sequences can differentiate organisms down to species and individual levels, it can help to classify diseases, identify **therapeutic targets**, and customize treatments.
- **Diagnostics:**
- **Pre-natal screening:** It has also been used in prenatal screening to determine whether the foetus has any **genetic disorders or anomalies**.
- **Evaluate disorders:** Genome sequencing has been used to assess rare disorders, preconditions for disorders, and even cancer from a genetic viewpoint, rather than as diseases of specific organs.
- **Drug Efficacy: Genome sequencing can provide information about drug efficacy or adverse drug effects.**
- The relationship between drugs and the genome is called **pharmacogenomics**.
- **Vaccine development:** Sequencing of viruses (e.g. **ebola, coronavirus**) and bacteria has led to the development of vaccines against them, once knowing their variants or strains.
 - Genomic data of pathogens could reveal hidden pathways of transmission.
- **Population Studies:Advanced analytics and AI** could be applied to critical datasets generated by collecting genomic profiles across the population, allowing for a better understanding of disease causation and potential treatments.
 - This is especially relevant for **rare genetic diseases**, where large datasets are required to find statistically significant correlations.
- **Agriculture and food security:** Genome sequencing has the power to revolutionize food security and sustainable agriculture by reducing the risks from disease outbreaks, and improving agriculture through effective plant and animal breeding, detecting multiple pathogens, etc.

Limitations of Genome Sequencing

Genome sequencing has revolutionized genetic research, offering profound insights into DNA structures and functions. However, its implementation faces several limitations, including data analysis challenges, incomplete

research, structural variant detection, ethical concerns, and technical constraints with larger genomes.

- **Data Analysis:** A vast amount of data is generated, which requires extensive analysis and interpretation.
- **Structural variants:** While technologies used to sequence DNA are highly accurate at deciphering the sequence, the majority of available technologies have limited scope in being able to determine so-called structural variants.
 - These are alterations that affect large segments of DNA at a time, such as duplications, deletions, and inversions.
- **Incomplete research and irrelevant data:** Despite growing knowledge in genomics, many genes still have unknown roles and a large number of genomic variants have not been identified as benign or pathogenic.
- **Ethical concerns:** Storing the large amount of data generated by WGS poses challenges related to capacity, cost, and privacy concerns, including potential ethical dilemmas with insurance companies and family members.
- **Not suitable for larger genomes:** Despite being a faster method, whole genome sequencing is not suitable for larger genomes because they have a number of repetitive DNA sequences for which assembling processes is sometimes challenging.

Source: <https://vajiramandravi.com/upsc-exam/genome-sequencing/>

© Vajiram & Ravi. For personal use only.