Genome Sequencing in India

[UPSC Notes]

What is Genome Sequencing Meaning?

A Genome Sequence involves determining the order of base pillars in an individual's genome. As a result of sequencing, the Genome is deciphered. The cost of Genome Sequencing differs from each other, depending on how the Genome is being read and how accurately it is detected. For reference, there are more than 3 billion genetic letters in a Human Genome. It is not possible to sequence the entire genome at once due to the limitations of available DNA Sequencing methods.

Therefore, only DNA Sequencing can handle a small stretch of DNA at a time. The virus genome can be made from either DNA or RNA, while the human genome is made of DNA. Every organism has a different set of genomes. For example, Corona Virus is made of RNA. In genome sequencing, the genetic information is read and interpreted from RNA or DNA.

Approaches to Genome Sequencing

- Genome cutting and reassembling can be accomplished in two ways.
- The clone by clone is the method where the genome is broken into approximately 150000 base pair chunks, called clones.
- The location of each clone in the genome is determined by genome mapping techniques.
- To Sequence about 500 base pairs per clone, the existing clones are cut into smaller pieces that overlap each other.
- After reassembling the pieces, they are sequenced and reassembled in the whole sequence.
- By breaking up the genome into pieces, and sequencing each one, the whole genome shotgun method creates a complete genome sequence.

Human Genome

- Approximately 3 billion DNA base pairs make up the Human Genome which consists of 23 chromosome pairs.
- Besides the 22 autosomal chromosomes, there are also the X and Y chromosomes that determine the sex of the baby.
- According to their sizes, the chromosomes 1-22 are numbered roughly from largest to smallest.
- Chromosomes 1-22 are duplicated in each somatic cell from both the parents and X or Y chromosomes along with 1 X chromosome from the mother get duplicated from the father to make a total of 46.

- Human Protein Coding genes number between 20000 to 25000.
- In recent years as the quality of genome sequences has improved and gene discovery methods of finding jeans have become more accurate the estimate of the number of human genes has been consistently revised.

Human Genome Project

Genome Sequencing and Gene Identification were the goals of the human genome project, which was an international research project. This project was coordinated by the US Department of Energy and the National Institute of Health.

This was a public-funded project that was Initiated in 1990, which aimed to determine the complete euchromatic human genome sequence within 15 years.
Undoubtedly, this project was a success, but in addition to it, this project also led to significant advancements in DNA Sequencing technology. The goal of personalized medicine is to develop better and safer drugs by using information about a person's genes, including their nucleotide sequencing.

Importance of DNA Sequencing

- There are Genetic tests that are available commercially to analyze a portion of the genome known to contain aberrant genes associated with diseases.
- To properly represent India's diverse human population in genomic data, the country must develop the indigenous capabilities for creating, maintaining, utilizing, and analyzing the large genome data on a scalable basis.
- Since 2003, the sequence of the human genome has opened up new insights into the link between diseases and genetics.
- With the help of genome Sequencing, it was found that there are about 10000 diseases associated with a single gene malfunctioning, including cystic fibrosis and thalassemia.
- Genomic Sequencing has shown that, in addition to certain organ diseases, cancer can also be understood from the genetic point of view, rather than the perspective of certain organ diseases.
- A few years back, the cost of Sequencing was about $10000 but now has fallen to tenth.
- In the future, scientists will be able to find the genes more quickly and easily with the help of the genome sequence.
- Even though scientists are just learning how to interpret genome sequences, they contain some clues about where genes are located.
- Scientists can easily find out the genes much quicker with the help of Genome Sequencing.
- Even though scientists are still researching how to interpret genome sequences, they still contain slight knowledge about their locations.
- It is also believed that studying Genome Sequencing allows scientists to understand the interactions of genes for controlling the growth and development of an organism.

- As the genome stores less than 20% DNA of the genes, the entire genome sequence will aid scientists in studying parts of the genome other than genes in a detailed manner.
- As well as the 'nonsense' or 'junk' DNA, which hasn't been defined, there are regulatory regions that control gene expressions.

Applications of Genome Sequencing

Although the Genome Sequencing process is widely used in medical-related streams and fields, there are several other applications of Genome Sequencing-

Medical Application

The immunity stimulating oral plant vaccines are created by DNA and transgenes. With precision medicine, the type of treatment received by a patient is determined by his or her genetic makeup.

Biotechnology Application

There are many uses of Genome Sequencing in the field of biotechnology and bioengineering. The genes of Mycoplasma Genitalium were synthesized in mycoplasma laboratories.

Social Science Applications

In order to determine the factors involved in species conservation. Genome Sequencing is used. Genetic diversity of a population, for example, is used to predict a species' health and conservation of species. A population's genetic patterns can be assessed in order to analyze the consequences of the evolutionary process. It may be possible to devise methods to preserve the species through the analysis of the patterns of these populations.

Advantages and Disadvantages of Genome Sequencing

Following are some of the advantages and disadvantages of Genome Sequencing-

Advantages

1. **Technical accuracy**
   As part of the Genome Sequencing process and data interpretation, the highest standards are compiled with high-quality levels available to the public. There are continually discoveries being made in improving the data quality of Human Genome Sequencing which is the state of the art technology.
2. **Protection of Information**
   Merogenomics DNA tests are clinical tests, and therefore they need to be signed duly by ordering the physician before the test can be conducted. In compliance

with HIPAA regulations, the service providers safeguard patients' information and work with clinics and hospitals worldwide.

3. **Lifetime Use**
   Sequencing of a genome differs from testing genotypes or sequencing the genes individually. In whole genome sequencing, the entire DNA is analyzed while in genotyping or sequencing a panel of desired genes only examines the selected fragments.
   Using such fragments can lead to a situation where important information may get missed which is needed to analyze the physical trait effectively. Therefore, another advantage of sequencing is the discovery of rare or novel variants.
4. **Information for the Future Generations**
   Obtaining genome sequences for additional members of a family increases the power of genomic information especially when it occurs between the generations. It is possible to infer inheritance patterns from a direct comparison of genome sequences of relatives who share the DNA variations. The interpretation of genomic variation can be made easier with each additional genome sequence.
5. **Psychological Benefits**
   If a family history indicates inheritance of a disease then some individuals might experience distress over perceived disease risk. Direct-to-consumer testing can determine an individual's genetic condition even if they are deemed a non-high risk for a specific disease. Removing the uncertainty about disease risk elevates anxiety associated with it- according to the studies.

Disadvantages

1. **Structural Variants**
   The majority of technologies available are capable of detection of so-called structural variants only, even though they are highly accurate at determining the DNA sequence. The DNA alterations such as duplication, deletion, and inversion effect larger segments at once.
2. **Test Failure**
   The quality of genome sequences and their interpretation can be negatively affected by factors outside the control of the service provider tasked with the isolation and sequencing of DNA. An example of it can be the low quantity, high bacterial contamination, or degradation of DNA samples provided for analysis.
3. **Discriminations**
   Genetic discrimination is banned in many countries but still is rarely protected in life insurance policies by such laws. Due to the slow adoption of new technologies and difficulties in interpreting new laws, there is a fear of the public regarding discrimination. This issue can be resolved by appealing in the judicial system, which may help in the future.
4. **Psychological Impact**
   The results of genome sequencing can reveal data that may have negative psychological implications for the client and their family. For example- some pathogenic variant being discovered, or getting indecisive results in case of

serious medical conditions, or in the worse case, it could be that the genome sequence doesn't belong to the family biologically.