Long Read Sequencing and Y Chromosome

For Prelims: Long Read Sequencing, X and Y Chromosomes, DNA

For Mains: Significance of Long Read Sequencing

Source: TH

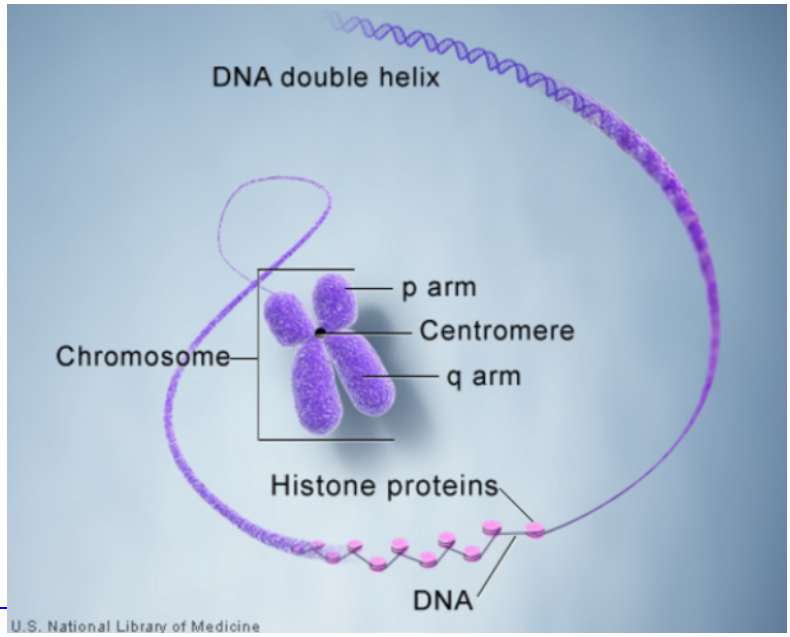
Why in News?

The new **"long-read" sequencing technique** has provided a reliable sequence from one end of the **Y chromosome** to the other.

- The findings published in *Nature* Journal provide information about the working of sex genes and sperm, the evolution of the Y chromosome, and its possible disappearance in a few million years.
- Earlier, some studies shed light on the **role of the Y chromosome in colorectal and bladder cancer,** revealing key genetic mechanisms that contribute to tumour progression, **immune response,** and clinical prognosis.

What is the Difference between DNA, Gene, and Chromosome?

- **DNA:**
  - **DNA is a long molecule that contains our unique genetic code.** DNA is composed of two strands that wrap around each other to form a double helix shape, like a spiral staircase.
  - Each strand of DNA is formed of four basic building blocks or **'bases'**: adenine (A), cytosine (C), guanine (G), and thymine (T).
- **Gene:**
  - **Genes are sections of DNA** that contain the set of instructions to produce one specific molecule in the body, usually a protein.
    - These proteins control how the body grows and works and are responsible for characteristics like eye colour, blood type, or height.
  - Each cell contains two sets of genes, one from your mother and one from your father. For ease of storage and access, the **genes are packaged up into 46 parcels called chromosomes.**
- **Chromosome:**
  - In the nucleus of each cell, the DNA molecule is packaged into thread-like structures called chromosomes.
  - Each chromosome is made up of DNA tightly coiled many times around proteins called **histones** that support its structure.
  - **Chromosomes are not visible in the cell's nucleus** - not even under a microscope.



What is the Y Chromosome and Its Related Findings?

- **About:**
  - The Y chromosome is a male-determining chromosome; it bears a gene called **SRY (Sex-Determining region Y)**, which directs the development of a testis in the embryo.
- **Y vs X:**
  - Y is very different from X and the 22 other chromosomes of the human genome; **it is small in size** and has a lot of DNA sequences that don't seem to contribute to traits (**aka "junk DNA"**) (only 27 compared to about 1,000 on the X) making it difficult to sequence the chromosome.
- **Disappearance of Y:**
  - About 150 mn years ago, the SRY evolved and defined a new proto-Y which degenerated fast enough (losing ~10 active genes per million years).
  - If this degradation continues, then in a few million years, the **whole human Y chromosome will disappear** (as it already has in some rodents).
- **Findings of Y Sequencing:**
  - The Y is the **last human chromosome to have been sequenced end-to-end**, or T2T (telomere-to-telomere).
    - Telomeres are structures made from DNA sequences and proteins **found at the ends of chromosomes.**
  - Some new genes have been discovered, but they are just extra copies of known genes.
  - The **centromere structure is now known**, and the repetitive sequences at the end of the Y have been read.
    - Centromere is a region of the chromosome that **pulls copies apart when the cell divides.**
    - The **location of the centromere on each chromosome gives the chromosome its characteristic shape** and can be used to help describe the location of specific genes.
  - The findings are significant for scientists all over the world. It will **help in examining the details of Y genes;** how SRY and the sperm genes are expressed or where and how repeated sequences originated.

What is Long Read Sequencing?

- Long-read sequencing, also called **third-generation sequencing**, is a DNA sequencing technique that enables the **sequencing of much longer DNA fragments** than traditional short-read

sequencing methods.
    - One of the most basic forms of DNA sequencing is **Sanger sequencing** which can sequence relatively small fragments of DNA (up to 900 base pairs).
    - The more modern forms of DNA sequencing are called **next-generation sequencing** which can efficiently determine longer DNA sequences compared to Sanger sequencing.
- Over the past decade, long-read, single-molecule DNA sequencing technologies have emerged as powerful players in genomics. It can read the DNA **sequence of much longer DNA fragments (**normal range: 10,000 - 100,000 base pairs).
    - While short reads can capture the majority of genetic variation, **long-read sequencing allows the detection of complex structural variants** that may be difficult to detect with short reads.

UPSC Civil Services Examination, Previous Year Questions (PYQs)

Q. What is Cas9 protein that is often mentioned in the news? (2019)

(a) A molecular scissors used in targeted gene editing

(b) A biosensor used in the accurate detection of pathogens in patients

(c) A gene that makes plants pest-resistant

(d) A herbicidal substance synthesized in genetically modified crops

Ans: (a)

Q. With reference to agriculture in India, how can the technique of 'genome sequencing', often seen in the news, be used in the immediate future? (2017)

1. Genome sequencing can be used to identify genetic markers for disease resistance and drought tolerance in various crop plants.
2. This technique helps in reducing the time required to develop new varieties of crop plants.
3. It can be used to decipher the host-pathogen relationships in crops.

Select the correct answer using the code given below:

(a) 1 only

(b) 2 and 3 only

(c) 1 and 3 only,

(d) 1, 2 and 3

Ans: (d)

Mains

Q. What are the research and developmental achievements in applied biotechnology? How will these achievements help to uplift the poorer sections of society? **(2021)**