Hello, everyone. Welcome to this tutorial. Here, we are going to have a walkthrough of sequencing where we shall explore Sanger sequencing and Next-Generation Sequencing (NGS).

Sanger sequencing. It was first developed in the 1970s by Sir Frederick Sanger. It is the gold standard of sequencing. It is still used today for routine applications and to validate NGS data. Sanger sequencing is also known as sequencing by chain termination.

It uses chain-terminating dideoxyribonucleotide Triphosphates, ddNTPs, that are tagged with fluorochromes. A ddNTP is a deoxyribonucleotide Triphosphate, dNTP, without the free 3 prime OH group. This means that they are incapable of chain extension in case they are inserted into a growing chain by DNA polymerase.

In classical Sanger sequencing, Sanger set up four different reactions, with each having all the basic PCR components, including a primer, DNA template, a polymerase, dNTPs. He also added ddNTPs in very small amounts. For more than Sanger sequencing, all these reactions are combined and run as a multiplex.

The resulting reaction produces extension products of varying length, terminated by ddNTPs at the 3 prime end. These fragments are then separated by capillary electrophoresis. The extension products are injected by an electrical field or current into a capillary tube filled with a gel polymer.

The extension products move from the negatively charged end to the positively charged end, and the speed at which they move is inversely proportional to their molecular weight. This process separates the extension products by size at a resolution of one base. A laser excites the dye-labelled ends of the fragments as they pass through the capillary tube.

The excited dye emits a signature light that is detected by the sensor. The detected light is translated into a base call by a programme. At the end of the process, the Sanger sequencer gives a series of chromatographs together with the called bases as a single .ab1 file.

Next-generation sequencing. Also known as massively parallel sequencing, is a powerful technique that has enabled advances in personalised medicine, genetic disease research, and clinical diagnosis. Examples of next-generation sequencing platforms include the MiSeq, the HiSeq, 454, SOLiD, ion proton.

Although many NGS technologies have been recently developed over the years, they all share the same common features, and these include the nucleic acid extraction, sample preparation, sequencing, and data analysis. On nucleic acid extraction, cells, viruses are lysed to release the DNA and RNA.

The RNA or DNA is then purified by various methods and kits. The RNA or DNA is then eluted and concentrated. Quality control of the extracted DNA or RNA is crucial.

Sample preparation. The goal is to add sequencing adaptors to the DNA or RNA to be sequenced. It starts with target selection, where DNA or RNA are subjected to fragmentation or PCR. Short oligo motifs, also known as adapters, are added to either ends of each fragments or amplicons.

The end product of sample preparation is what we call a sequencing library, where all fragments or amplicons are tagged with sequencing adapters. The library is then quality controlled and loaded onto the sequencer.

Sequencing. The DNA with adapters is fast amplified into clones or clusters. After a series of chemical reactions that either produce change in pH, luminescence, or fluorescence, signals are captured by cameras and computers, which are then translated into base calls.

Data analysis. Data from the sequencer contains the base calls and their quality scores. This is quality controlled. The data is then aligned or assembled de novo to generate consensus sequences that are used to answer biological questions.