Next Generation Sequencing

Anamaria Hodivoianu

April 11, 2025

1   Introduction

Next Generation Sequencing, or NGS, is a method of DNA and RNA sequencing. What makes it a "next generation" and break-through technology is the fact that it can sequence many fragments of DNA in parallel, which makes it much faster and cheaper than the traditional Sanger sequencing method. NGS can sequence a whole genome in a day, while the Sanger method would take months or even years.

2   The Human Genome Project

The human genome was first sequenced as part of the Human Genome Project, which relied on the Sanger method. It began in 1990, and by 2003 it was 92% complete and was officially declared complete. It reached its final 100% completion in 2022. The project took 13 years. Although NGS is much faster as it can sequence billions of DNA strands at the same time, while the Sanger method can only sequence one strand at a time, NGS only works because the Human Genome Project created a reference human genome.

3   How NGS works

NGS works by breaking the DNA into small fragments, which are then sequenced in parallel. The sequencing process involves several steps:

3.1   Library preparation

The first step is to prepare the DNA library. A DNA library is just a collection of small DNA fragments taken from a larger DNA sequence. If RNA is sequenced instead of DNA, then the RNA must be first reverse transcribed into DNA. In order to split the DNA into smaller fragments, it is cut using enzymes. Then, short DNA sequences called adapters are added to the ends of the fragments. These adapters are important because they contain relevant information for the sequencing process, as well as an index.

3.2   Sequencing by synthesis

The next step is to sequence the DNA fragments. This is done using a method called sequencing by synthesis. It takes place in a flow cell, which has a glass surface. On this surface are small DNA segments called oligonucleotides, which are complementary to the adapters on the DNA fragments. Then, the DNA fragments are denaturated in order to become one strand. The single-stranded DNA fragments are then added to the flow cell, where the adapters attach to the oligonucleotides. These are called the forward strands. Then the reverse strands are made, and the forward ones are discarded.
Before the strands are actually sequenced, they need to be amplified, because otherwise the sequencing signal would be too weak. This is done using PCR. The flow cell solution is changed, and the strands attach to a second oligonucleotide, forming a bridge. Then the strand is copied, making it a double strand, which is then denaturated, forming two single strands. This process is repeated, and the DNA fragments are amplified. The reverse strands are then discarded, and sequencing can begin.
A sequencing primer is attached to the strands, and then fluorescent nucleotides, A, T, C and G, and DNA polymerase are added to the flow cell. Only one nulceotide can get sequenced at a time. The complementary nucleotides attach to the strands, emitting a fluorescent signal that is

captured by a camera. Then the nuceluotides are washed away, and the process is repeated until the sequencing is complete.

3.3 Filtering and alignment

First, any faulty reads are removed. Then the reads are demultiplexed, which means that the reads are separated based on their indexes. The reads are then aligned to the reference genome with some overlap betweeen them. This overlap in connected to the read depth, an important metric of sequencing, which is the number of reads for a nucleotide. The average read depth depends on the purpose of the sequencing. For example, for whole genome sequencing, the average read depth is 30x, while for cancer mutations, it is 1500x. Another important metric in sequencing is coverage, which means that no DNA segment should be left uncovered.

4 Applications

NGS has many applications in medicine. It can be used to identify genetic mutations, which can help diagnose diseases. For example, it is very useful in cancer reserach, where it can identify mutations that cause cancer, but also in treatement, since any tumor could have more than one mutation. NGS can also be used to identify pathogens in infectious diseases, and it is also used in personalized medicine, where the treatment is tailored to the patient's genetic makeup.

5 Conclusion

In conclusion, NGS is a powerful technology that has revolutionized genetics. With NGS, DNA sequencing is now much faster and cheaper than before. It has many applications in medicine, and it will continue to be an important tool in genetics research.