# THE NEXT GENERATION SEQUENCING TECHNOLOGIES

**MOHAMAD, O.[1] & HO, W. S.[2]**
*[1]Kulliyyah of Science, IIUM, Kuantan, Pahang*
*[2]Dept. of Molecular Biology, FRST, UNIMAS, Kota Samarahan, Sarawak*

Sanger sequencing has been the major method in directly sequencing DNA, and has dominated the DNA sequencing market for nearly past 30 years (Varshney et al., 2009). Along with PCR, we cannot underestimate how important this technology has been to research in various fields of molecular biology. It has revolutionized genetics by allowing us to gain unprecedented insights into the workings of different organisms.

Sanger method is considered the first generation technology. The technology has been considered rather static with the main change being in the increase in number of samples that could be analyzed during any one machine run. However, the length of reads per run remains about the same. Its limitations include maximum read length usually < 1000 bp, extremely high cost, long experimental set up times, high DNA concentrations needed and some regions are unsequenceable. If we only use this technology to sequence human DNA, it would take 30 years and cost more than RM25 million (Veeramah, 2009). Major accomplishment of the technology is the human genome sequence through the much-publicised Human Genome Project which took 13 years and about USD3.0 billion to accomplish.

## NGST PLATFORMS

The limitations of Sanger sequencing has led to the development of newer technology platforms. Research nowadays has a niche for developing higher throughput technology. Newer technologies, a.k.a. next generation sequencing technologies (NGST), have been developed to substantially increase the productivity, decrease the cost of DNA sequencing, and offer fast and accurate genome information. These technologies are able to produce billions of data up to 1 Gb of short reads per instrument run in a high-throughput and cost-effective fashion in contrast to first generation technology (conventional method) which only produces Mb of data.

The bulk of these technology platforms that are currently available share common elements which can be best described by the term sequencing by synthesis. In these platforms, each single DNA molecule is replicated many 1,000 times, made single-stranded, fixed to a specific location, and then sequenced by using a DNA polymerase-based method. The latter is much the same way as Sanger method; however, newer technologies differ on how DNA is fixed, and what particular next generation sequencing (NGS) chemistry is used to get the required reaction. **Table 1** describes the major next generation sequencing technologies that are commercially available and have been taken up by many institutions. Examples of commercially available NGS technologies include Illumina/Solexa GAII, Roche 454 GS FLX Titanium, LifeTech/APG SOLiD 3 System, Helicos BioSciences HeliScope, Dover Systems Polonator G.007 and Pacific Biosciences DNA sequencing instrument. Illumina/Solexa is currently the most widely used in the field. Many have already demonstrated their potential to circumvent the limiting factors of Sanger sequencing.

**Table 1 :** Major next generation sequencing technologies that are commercially available and have been taken up by many institutions

| Platform | NGS Chemistry | Read length (bases) | Run time (days) | Gb per run | Machine cost (RM equivalent) | Pros | Cons |
|---|---|---|---|---|---|---|---|
| Roche 454 GS FLX Titanium | Pyrosequencing (PS) | 330 | 0.35 | 0.45 | 1,500,000 | Longer read, fast run times | High reagent cost |
| Illumina/Solexa GAII | Reversible terminators (RTs) | 75-100 | 4-9 | 18-35 | 1,620,000 | Currently widely used | Low multiplexing capability |
| LifeTech/APG SOLiD 3 System | Cleavable probe Sequencing by ligation (SBL) | 50 | 7-14 | 30-50 | 1,785,000 | Inherent error correction | Long run times |
| Dover Systems Polonator G.007 | Non-cleavable probe SBL | 26 | 5 | 12 | 510,000 | Least expensive | Shortest read lengths |
| Helicos Biosciences HeliScope | RTs | 32 | 8 | 37 | 2,997,000 | No bias representation of templates | High error rates compared with other RTs |
| Pacific Biosciences DNA sequencing instrument | Realtime | 964 | - | - | - | Has potential to read > 1 kb | Highest error rates compared with other NGS technologies |