International Journal of Research Studies in Agricultural Sciences (IJRSAS)
Volume 7, Issue 8, 2021, PP 27-36
ISSN No. (Online) 2454–6224
DOI: http://dx.doi.org/10.20431/2454-6224.0708004
www.arcjournals.org

Review on Mechanism of DNA Sequencing Technologies in Crop Improvement

Zewdu Asrat *

Ethiopian Institute of Agricultural Research, Chiro National Sorghum Research and Training Center P. O. Box 190, Chiro, Ethiopia

Corresponding Authors:** ***Zewdu Asrat, *Ethiopian Institute of Agricultural Research, Chiro National Sorghum Research and Training Center P. O. Box 190, Chiro, Ethiopia*

Abstract: *The development of DNA sequencing strategies has been a high priority in genetics research since the discovery of the structure of DNA and the basic molecular mechanisms of heredity. However, it was not until the works by Maxam and Gilbert, and Sanger (Sanger et al, that the first practical sequencing methods were developed and implemented on a large scale. Currently there are five companies commercializing one or more NGS platforms. However, there are only three NGS technologies, Roche 454, Illumina and ABI SOLiD, that account for the vast majority of usage in plant research and are widely available in academic institutions, private research centers and service-providing companies*

Keywords: *NGS platforms, Sequencing, Genome Analyzer (GA).*

1. INTRODUCTION

The development of DNA sequencing strategies has been a high priority in genetics research since the discovery of the structure of DNA and the basic molecular mechanisms of heredity. However, it was not until the works by Maxam and Gilbert (1977), and Sanger (Sanger et al, 1977), that the first practical sequencing methods were developed and implemented on a large scale. The first isolation and sequencing of a plant cDNA by Bedbrook and colleagues a few years later initiated the field of Plant Molecular Genetics (Bedbrook et al, 1980). Plant biotechnology started shortly thereafter with the successful integration of recombinant DNA and sequencing techniques to generate the first transgenic plants using Agrobacterium (Fraley et al, 1983; Herrera Estrella et al, 1983). The first genetic map in plants based on restriction fragment length polymorphisms (RFLPs; Bernatzky & Tanksley, 1986) enabled the capture of genetic variation and started the era of molecular marker-assisted plant breeding. Since then, sequencing methodologies have been essential tools in plant research. They have allowed the characterization and modification of genes and metabolic pathways, as well as the use of genetic variation for studies in species diversity, marker-assisted selection (MAS), germplasm characterization and seed purity.

The determination of the reference genomes in Arabidopsis thaliana, rice and maize using Sanger sequencing strategies constituted major milestones that enabled the analysis of genome architecture and gene characterization in plants (The Arabidopsis Genome Initiative, 2001; International Rice Genome Project, 2005; Schnable et al, 2009). More recently, the development and increasing availability of multiple Next-Generation sequencing (NGS) technologies minimized research limitations and bottlenecks based on sequence information (Metzker, 2010; Glenn, 2011). It is difficult to overstate the influence that these massively parallel systems have had in our understanding of plant genomes and in the expansion, acceleration and diversification of breeding and biotechnology projects. At the same time, this influence tends to understate the importance that capillary Sanger sequencing still has in day-by-day research and development work.

This review provides a description of major sequencing technologies that are developed through time.

Kind of Sequencing Technologies

The development of recent sequencing technologies has generated a remarkable increase, by orders of magnitude, in sequencing throughput with a corresponding drop in cost per base. A simple exercise to comprehend the scale of acceleration in sequencing is to look back at the state of the art of sequencing in 1980. At that time, earlier improvements in Sanger and Maxam-Gilbert methodologies had initiated the wide use of sequencing in research laboratories around the world. Then, typical sequencing throughputs per slab gel run were under 10,000 bp. During the period from 1980 to 2005 sequencing platforms based on Sanger chemistry had a 500 to 1,000-fold increase, to more than 5 Mbp per run. The number of reads that could be processed, quality, read length and analysis all improved and were optimized, propelled by the development of the human genome project (Barnhart, 1989). While these technological advances were certainly impressive, they dwarf when compared to the acceleration in sequencing capacity after 2005. At that time, novel ultra-high throughput technologies started to become commercially available. From 2005 through the second half of 2011, the throughput per run had increased an additional 100,000-fold, or 5 orders of magnitude. This acceleration has been unprecedented in science and technology. It has outpaced Moore's law that famously predicted that the number of transistors in a computer processor would double every two years (Moore, 1965 Figure 1). This fast increase in sequencing capacity has had important consequences in analysis and logistics, and has changed expectations in all aspects of plant genetics, breeding and biotechnology.

The new chemistries and platforms, broadly described as Next-Generation sequencing (NGS) technologies, take advantage of diverse chemistries and detection approaches. While some of these technologies appear to have little in common with each other, they share key characteristics. All NGS technologies are massively parallel systems relying on the immobilization of millions of, up to billions of DNA templates in a solid surface. They do not use electrophoresis, relying instead on in situ base detection and extension. With the exception of one system, developed by Helicos, NGS platforms need to amplify the templates and use one of several PCR-based approaches. One additional characteristic of NGS systems that took more than one early-adopting institution unprepared, is the increased need for computer power and storage necessary to process and retain the massive data produced.

Currently there are five companies commercializing one or more NGS platforms. However, there are only three NGS technologies, Roche 454, Illumina and ABI SOLiD, that account for the vast majority of usage in plant research and are widely available in academic institutions, private research centers and service-providing companies. As it will be emphasized in the next sections, these platforms have different input and output characteristics that make them more or less advantageous to specific applications. Finally, one 'Third Generation' sequencing platform has recently become commercially available from Pacific Biosciences. Third generation technologies are also massively parallel systems although they use single-molecule DNA templates, real-time detection and are able to generate longer reads faster. The expectation is that third generation machines will eventually produce large numbers of high-quality reads with an average of several kilobase-pairs from a single molecule.

2. THE FIRST GENERATION OF SEQUENCING

Sanger and Maxam-Gilbert sequencing technologies were classified as the First Generation Sequencing Technology [Thudi et al, 2012] who initiated the field of DNA sequencing with their publication in 1977.

Sanger sequence analyzers/Chain Termination Method

Sanger Sequencing is known as the chain termination method or the dideoxy nucleotide method or the sequencing by synthesis method. It consists in using one strand of the double stranded DNA as template to be sequenced. this sequencing is made using chemically modified nucleotides called dideoxy-nucleotides (dNTPs). Hese dNTPs are marked for each DNA bases by ddG, ddA, ddT, and ddC. The dideoxy nucleotides are used dNTPs are used for elongation of nucleotide, once incorporated into the DNA strand they prevent the further elongation and the elongation is complete. The DNA fragments ended by a dNTP with different sizes. The fragments are separated according to their size using gel slab where the resultant bands corresponding to DNA fragments can be visualized by an imaging system (X-ray or UV light) [Masudi-Nejda et al, 2013].

The first genomes sequenced by the Sanger sequencing are phiX174 genome with size of 5374 bp and in 1980 the bacteriophage λ genome with length of 48501 bp [Sanger et al, 1980]. After years of

improvement, Applied Biosystems is the first company that has automated Sanger sequencing. Applied Biosystems has built in 1995 an automatic sequencing machine called ABI Prism 370 based on capillary electrophoresis allowing fast an accurate sequencing. The Sanger sequencing was used in several sequencing projects of different plant species such as Arabidopsis [Goff et al, 2002], and soybean [Schmotz J et al, 2010] and the most emblematic achievement of this sequencing technology is the decoding of the first human genome [Durban RM, 2010].

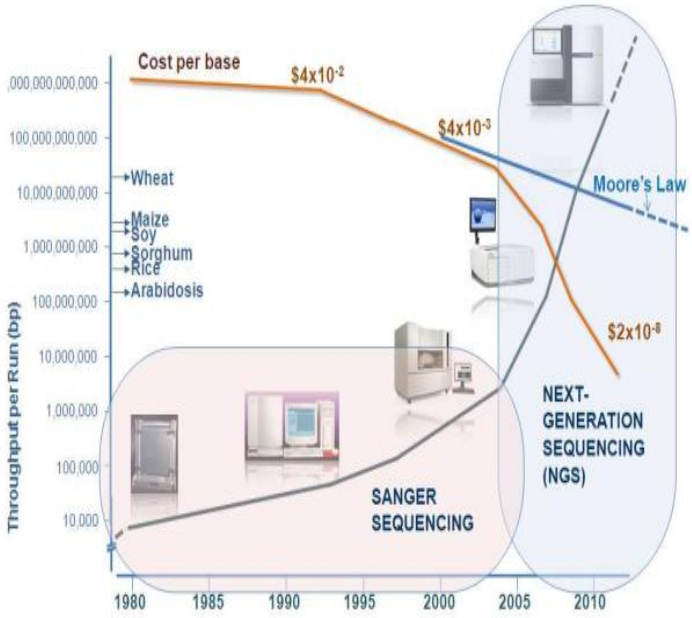

Figure1. *Increase in maximum throughput per run in sequencing platforms from 1980 to 2011. Stratton et al (2009) and Glenn (2011).*

The sanger sequencing was widely used for three decades and even today for single or low-throughput DNA sequencing, however, it is difficult to further improve the speed of analysis that does not allow the sequencing of complex genomes such as the plant species genomes and the sequencing was still extremely expensive and time consuming.

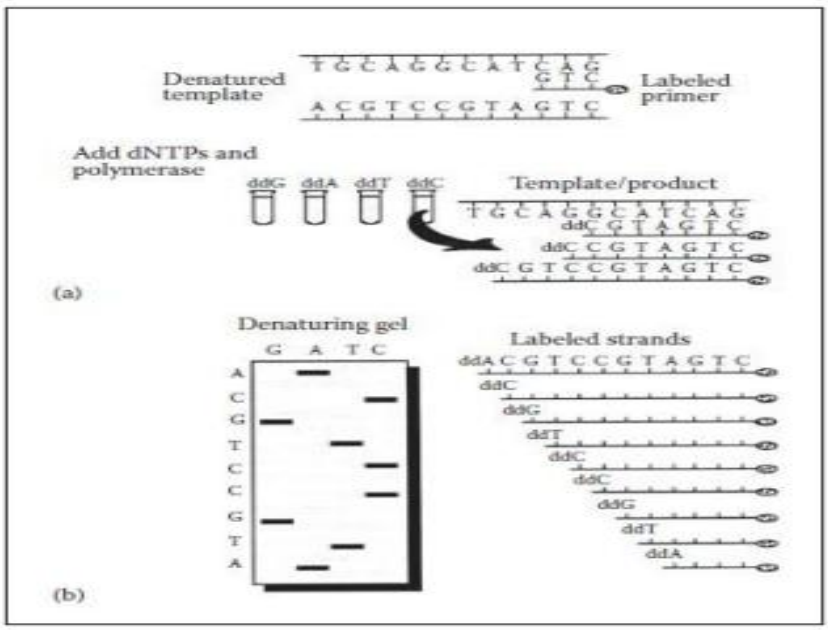

Figure 2: *Sanger sequencing technology* ***(a)*** *The sequencing reaction is performed by the presence of denatured DNA template, radioactively labeled primer, DNA polymerase, and dNTPs.* ***(b)*** *Illustrates the separation of these DNA fragments in a denaturing gel by electrophoresis(P Moran, and RS Waples, 1994).*

Maxam and Gilbert/Chemical cleavage Method

Maxam-Gilbert is another sequencing belonging to the first generation of sequencing known as the chemical degradation method. Relies on the cleaving of nucleotides by chemicals and is most effective with small nucleotides polymers. Chemical treatment generates breaks at a small proportion of one or two of the four nucleotide bases in each of the four reactions (C, T+C, G, A+G). This reaction leads to a series of marked fragments that can be separated according to their size by electrophoresis [Masudi-Nejad et al, 2013].

3. THE SECOND GENERATION OF SEQUENCING

The first generation of sequencing was dominant for three decades especially Sanger sequencing, however, the cost and time was a major stumbling block. In 2005 and in subsequent years, have marked the emergence of a new generation of sequencers to break the limitations of the frst generation. The basic characteristics of second generation sequencing technology are: (1) The generation of many millions of short reads in parallel, (2) The speed up of sequencing the process compared to the first generation, (3) The low cost of sequencing and (4) The sequencing output is directly detected without the need for electrophoresis. Short read sequencing approaches divided under two wide approaches: sequencing by ligation (SBL) and sequencing by synthesis (SBS), (more details for these sequencing categories are presented in [Goodwins et al, 2016]) and are mainly classifed into three major sequencing platforms: Roche/454 launched in 2005, Illumina/Solexa in 2006 and in 2007 the ABI/SOLiD. We will briefly describe these commonly utilized sequencing platforms.

Roche/454 sequencing

Roche/454 sequencing appeared on the market in 2005, using Pyro sequencing technique which is based on the detection of pyrophosphate released aier each nucleotide incorporation in the new synthetic DNA strand (http://www.454.com). The pyrosequencing technique is a sequencing-by-synthesis approach.

DNA samples are randomly fragmented and each fragment is attached to a bead whose surface carries primers that have oligonucleotides complementary to the DNA fragments so each bead is associated with a single fragment (Figure 3A). Then, each bead is isolated and amplifed using PCR emulsion which produces about one million copies of each DNA fragment on the surface of the bead (Figure 3B). The beads are then transferred to a plate containing many wells called picotiter plate (PTP) and the pyrosequencing technique is applied which consists in activating of a series of downstream reactions producing light at each incorporation of nucleotide. By detecting the light emission after each incorporation of nucleotide, the sequence of the DNA fragment is deduced (Figure 3C) [Mordis ER, 2008]. The use of the picotiter plate allows hundreds of thousands of reactions occur in parallel, considerably increasing sequencing throughput [Vezzi F, 2012]. The latest instrument launched by Roche/454 called GS FLX+ that generates reads with lengths of up to 1000 bp and can produce ~1Million reads per run (454.com GS FLX+Systems http//454.com/products/gs-flxsystem/index.asp). Other characteristics of Roche/454 instruments are listed in [Litul et al, 2012].

The Rche/454 is able to generate relatively long reads which are easier to map to a reference genome. The main errors detected of sequencing are insertions and deletions due to the presence of homopolymer regions [Margules et al, 2005]. Indeed, the identification of the size of homopolymers should be determined by the intensity of the light emitted by pyro sequencing. Signals with too high or too low intensity lead to under or overestimation of the number of nucleotides which causes errors of nucleotides identifcation. Ion torrent sequencing

Life Technologies commercialized the Ion Torrent semiconductor sequencing technology in 2010 (https//www.thermofsher.com/us/en/home/brands/ion-torrent.html).Itis similar to 454pyrosequencing technology but it does not use fluorescent labeled nucleotides like other second-generation technologies. It is based on the detection of the hydrogen ion released during the sequencing process. Specifically, Ion Torrent uses a chip that contains a set of micro wells and each has a bead with several identical fragments. The incorporation of each nucleotide with a fragment in the pearl, a hydrogen ion is released which change the pH of the solution. This change is detected by a sensor attached to the bottom of the micro well and converted into a voltage signal which is proportional to the number of nucleotides incorporated (Figure 4).

The Ion Torrent sequencers are capable of producing reads lengths of 200 bp, 400 bp and 600 bp with throughput that can reach 10 Gb for ion proton sequencer. the major advantages of this sequencing technology are focused on read lengths which are longer to other SGS sequencers and fast sequencing time between 2 and 8 hours. The major disadvantage is the difficulty of interpreting the homopolymer sequences (more than 6 bp) [Reuter JA et al, 2015] which causes insertion and deletion (indel) error with a rate about ~1%.

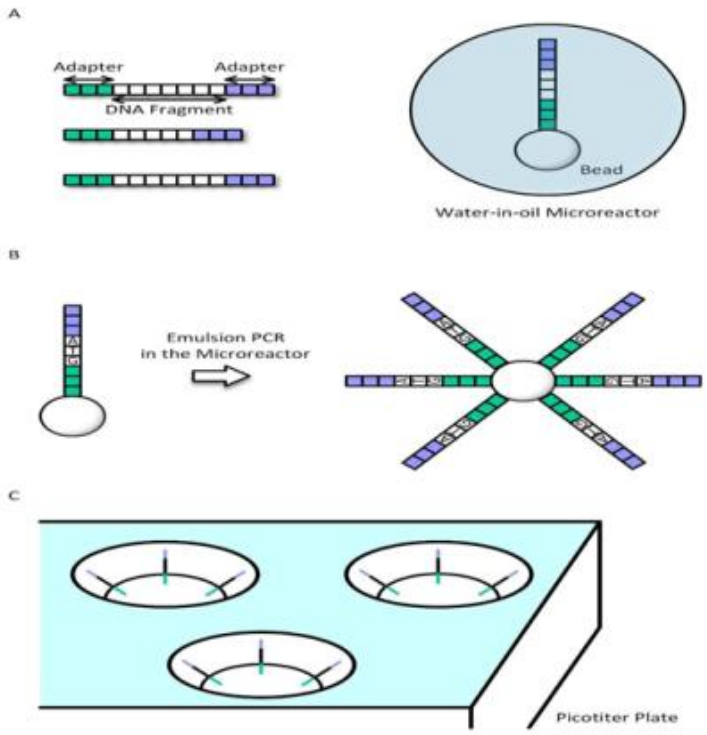

Figure3: *Roche/454 sequencing technology [Heo Yun, 2015].*

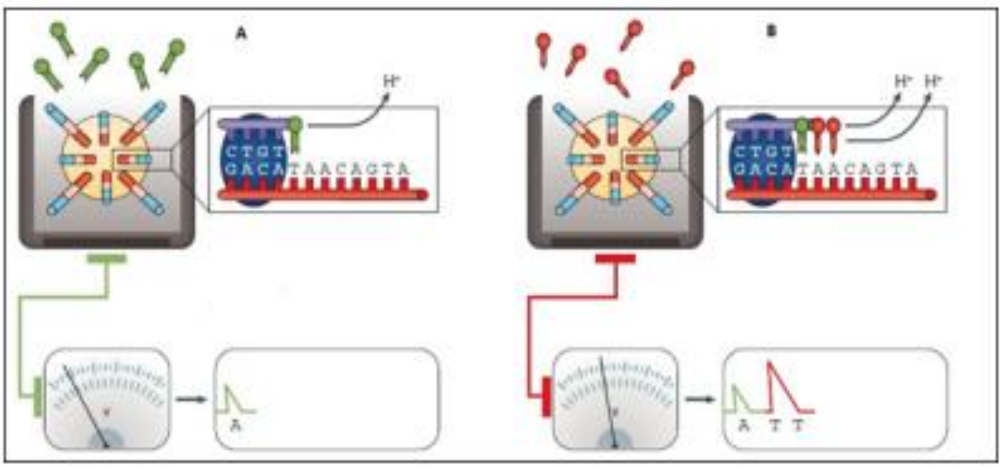

Figure4: *Ion torrent sequencing technology*

Illumina/Solexa sequencing

The Solexa company has developed a new method of sequencing Illumina company (http://www.illumina.com) purchased Solexa that started to commercialize the sequencer Ilumina/Solexa Genome Analyzer (GA) [Balasubramanian S et al., 2008]. Illumina technology is sequencing by synthesis approach and is currently the most used technology in the NGS market. The sequencing process is shown in Figure 4. During the first step, the DNA samples are randomly fragmented into sequences and adapters are ligated to both ends of each sequence. Then, these adapters are fixed themselves to the respective complementary adapters, the latter are hooked on a slide with many variants of adapters (complementary) placed on a solid plate (Figure 4A).

During the second step, each attached sequence to the solid plate is amplified by "PCR bridge amplifcation" that creates several identical copies of each sequence; a set of sequences made from the same original sequence is called a cluster. Each cluster contains approximately one million copies of the same original sequence (Figure 4B). The last step is to determine each nucleotide in the sequences, Illumina uses the sequencing by synthesis approach that employs reversible terminators [Bently DR et al., 2008] in which the four modified nucleotides, sequencing primers and DNA polymerases are added as a mix, and the primers are hybridized to the sequences. Then, polymerases are used to extend the primers using the modified nucleotides. Each type of nucleotide is labeled with a fluorescent specific in order for each type to be unique. The nucleotides have an inactive 3'-hydroxyl group which ensures that only one nucleotide is incorporated. Clusters are excited by laser for emitting a light signal specific to each nucleotide, which will be detected by a coupled-charge device (CCD) camera and Computer programs will translate these signals into a nucleotide sequence (Figure 4C). The process continues with the elimination of the terminator with the fluorescent label and the starting of a new cycle with a new incorporation [Heo Yun, 2015].

The first sequencers Illumina/Solexa GA has been able to produce very short reads ~35 bp and they had an advantage in that they could produce paired-end (PE) short reads, in which the sequence at both ends of each DNA cluster is recorded. The output data of the last Illumina sequencers is currently higher than 600 Gpb and lengths of short reads are about 125 bp. Details on Illumina sequencers

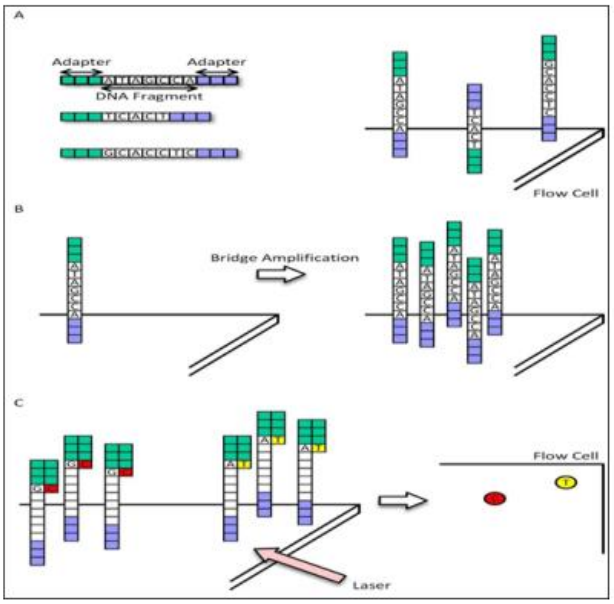

Figure5: *Illumina sequencing technology [Heo Yun,2015].*

ABI/SOLiD sequencing

Supported Oligonucleotide Ligation and Detection (SOLiD) is a NGS sequencer Marketed by Life Technologies (http:// www.lifetechnologies.com). In 2007, Applied Biosystems (ABI) has acquired SOLiD and developed ABI/SOLID sequencing technology that adopts by ligation (SBL) approach. The ABI/SOLiD process consists of multiple sequencing rounds. It starts by attaching adapters to the DNA fragments, fixed on beads and cloned by PCR emulsion. These beads are then placed on a glass slide and the 8-mer with a fluorescent label at the end are sequentially ligated to DNA fragments, and the color emitted by the label is recorded (Figure 5A). Then, the output format is color space which is the encoded form of the nucleotide where four fluorescent colors are used to represent 16 possible combinations of two bases. The sequencer repeats this ligation cycle and each cycle the complementary strand is removed and a new sequencing cycle starts at the position n-1 of the template. The cycle is repeated until each base is sequenced twice (Figure 5B). The recovered data from the color space can be translated to letters of DNA bases and the sequence of the DNA fragment can be deduced [Mardis ER, 2008**].**

ABI/SOLiD launched the first sequencer that produce short reads with length 35 bp and output of 3 Gb/run and continued to improve their sequencing which increased the length of reads to 75 bp with

an output up to 30 Gb/run [Goodwins et al, 2016]. The strength of ABI/SOLiD platform is high accuracy because each base is read twice while the drawback is the relatively short reads and long run times. The errors of sequencing in this technology is due to noise during the ligation cycle which causes error identification of bases.

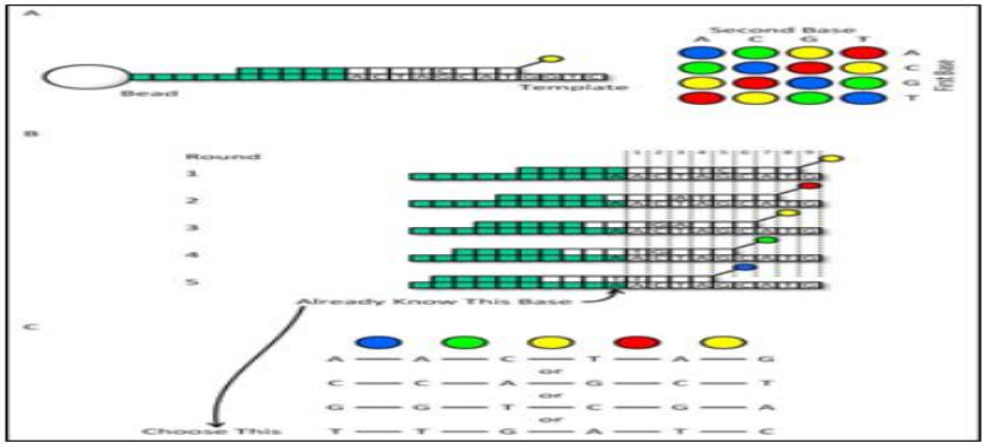

Figure6: *ABI/SOLID sequencing technology [Heo Yun., 2015].*

4. Third Generation of Sequencing

The second-generation of sequencing technologies previously discussed have revolutionized the analysis of DNA and have been the most widely used compared to the first generation of sequencing technologies. However, the SGS technologies generally require PCR amplification step which is a long procedure in execution time and expansive in sequencing price. Also, it became clear that the genomes are very complex with many repetitive areas that SGS technologies are incapable to solve them and the relatively short reads made genome assembly more difficult. To remedy the problems caused by SGS technologies, scientists have developed a new generation of sequencing called "third generation sequencing". These third generations of sequencing have the ability to offer a low sequencing cost and easy sample preparation without the need PCR amplification in an execution time significantly faster than SGS technologies. In addition, TGS are able to produce long reads exceeding several kilobases for the resolution of the assembly problem and repetitive regions of complex genomes**.**

There are two main approaches that characterize TGS : The single molecule real time sequencing approach (SMRT) [38] that was developed by Quake laboratory [Haris Td et al,2008] and the synthetic approach that rely on existing short reads technologies used by Illumina (Moleculo) [43] and 10xGenomics (https://www.10xgenomics.com) to construct long reads. The most widely used TGS technology approach is SMRT and the sequencers that have used this approach are Pacifc Biosciences and Oxford Nanopore sequencing (specifcally the MinION sequencer).

Pacific biosciences SMRT sequencing

Pacifc Biosciences (http//www.pacifcbiosciences.com/) developed the frst genomic sequencer using SMRT approach and it's the most widely used third-generation sequencing technology. Pacifc Biosciences uses the same fluorescent labelling as the other technologies, but instead of executing cycles of amplifcation nucleotide, it detects the signals in real time, as they are emitted whennthe incorporations occur. It uses a structure composed of many SMRT cells, each cell contains microfabricated nanostructures called zeromode waveguides (ZMWs) which are wells of tens of nanometers in diameter microfabricated in a metal flm which is in turn deposited onto a glass substrate [McCoy RC et al,2014].

These ZMWs exploit the properties of light passing through openings with a diameter less than its wavelength, so light cannot be propagated. Due to their small diameter, the light intensity decreases along the wells and the bottom of the wells illuminated (Figure 7A). Each ZMW contains a DNA polymerase attached to their bottom and the target DNA fragment for sequencing. During the sequencing reaction, the DNA fragment is incorporated by the DNA polymerase with fluorescent labeled nucleotides (with different colors). Whenever a nucleotide is incorporated, it releases a

luminous signal that is recorded by sensors (Figure 7B). He detection of the labeled nucleotides makes it possible to determine the DNA sequence**.**

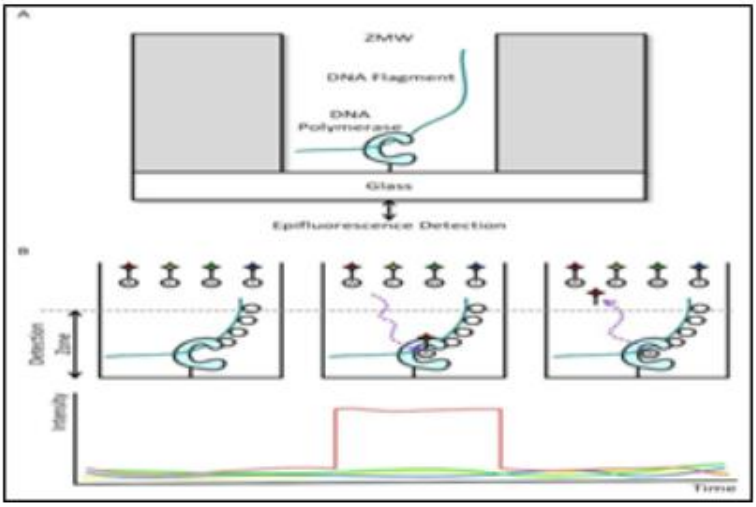

Figure7: *Pacifc biosciences sequencing technology [Heo Yun,2015].*

Oxford nanopore sequencing

The Oxford Nanopore sequencing (ONT) was developed as a technique to determine the order of nucleotides in a DNA sequence. In 2014, Oxford Nanopore Technologies released the MinION [Mikheyev As,2014] device that promises to generate longer reads that will ensure a better resolution structural genomic variants and repeat content [Laehnemann D et al,2015]. It's a mobile single-molecule Nanopore sequencing measures four inches in length and is connected by a USB 3.0 port of a laptop computer. This device has been released for testing by a community of users as part of the MinION Access Program (MAP) to examine the performance of the MinION sequencer [Laver T et al 2015].

In this sequencing technology, the frst strand of a DNA molecule is linked by a hairpin to its complementary strand. The DNA fragment is passed through a protein nanopore (a nanopore is a nanoscale hole made of proteins or synthetic materials [Heo Yun,2015].

When the DNA fragment is translated through the pore by the action of a motor protein attached to the pore, it generates a variation of an ionic current caused by di;erences in the moving nucleotides occupying the pore (Figure 8A). This variation of ionic current is recorded progressively on a graphic model and then interpreted to identify the sequence (Figure 8B). The sequencing is made on the direct strand generating the "template read" and then the hairpin structure is read followed by the inverse strand generating the "complement read", these reads is called "1D". If the "temple" and "complement" reads are combined, resulting consensus sequence called "two direction read" or "2D".

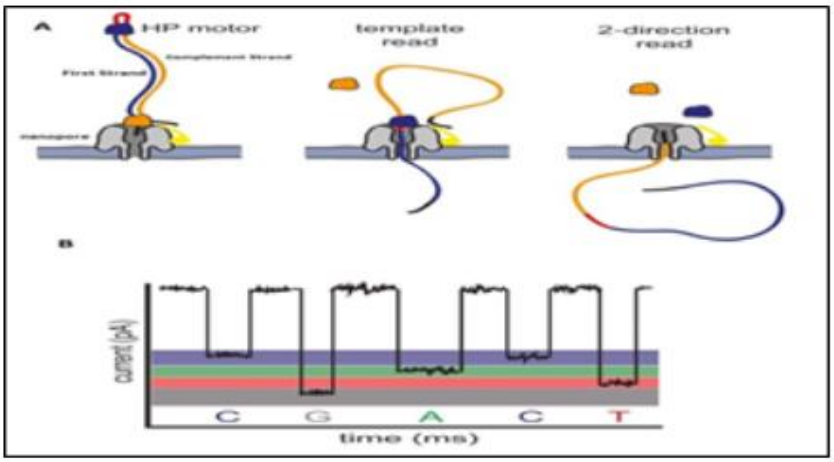

Figure7: *Oxford nanopore MinION sequencing [Reuter JA et al,2015].*

5. CONCLUSION

The first method of sequencing came about half a century ago, and since then, sequencing technologies have continued to evolve especially after the appearance of the first sequencers from NGS technology. These technologies are characterized by their high throughput which gives the opportunity to produce millions of reads with inexpensive sequencing. NGS technologies are now the starting point for several areas of research based on the study and analysis of biological sequences. In this review, I presented a concise overview of the generations of sequencing technologies by beginning with the first-generation sequencing history followed by the main commonly used NGS platforms. Nevertheless, there are significant challenges in NGS technologies, including the difficulty of storing and analyzing the data generated by these technologies. This is mainly due to the production of a high number of reads. In the coming years, new sequencing platforms will appear producing a larger amount of data (in Terabyte) which requires the development of new approaches and applications capable of analyzing this large amount of data.

REFERENCES

[1] Alic AS, Ruzafa D, Dopazo J, Blanquer I., 2016. Objective review of de novo stand-alone error correction methods for NGS data. In: WIREs Computational Molecular Science.

[2] Balasubramanian S., 2015. Solexa sequencing: Decoding genomes on a population scale. ClinChem 61: 21-24.

[3] Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, et al., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456: 53–59.

[4] Durbin RM., 2010. A map of human genome variation from population scale sequencing. Nature 467: 1061-73.

[5] El-Metwally S, Ouda OM, Helmy M. (2014). Next generation sequencing technologies and challenges in sequence assembly. Springer.

[6] Goff,SA, Ricke D, Lan TH, Presting G, Wang R, et al. (2002). A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92-100.

[7] Goodwin S, McPherson JD, Richard Mc Combie W. (2016). Coming of age: Ten years of next-generation sequencing technologies. Nature Reviews Genetics 17: 333–351.

[8] Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, et al. (2008). Single molecule DNA sequencing of a viral genome. Science 320: 106-9.

[9] Heo, Yun. (2015). Improving quality of high-throughput sequencing reads.

[10] Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biol 8: R143.

[11] Ip CL, Loose M, Tyson JR, de Cesare M, Brown BL, et al. (2015). MinION analysis and reference consortium: Phase 1 data release and analysis. F1000Research 4: 1075.

[12] Jain M, Hugh EO, Paten B, Akeson M. (2016). oxford nanopore MinION: Delivery of nanopore sequencing to the genomics community. Genome Biology 17: 239.

[13] Kulski JK. (2016). Next-generation sequencing-An overview of the history, tools, and "Omic" applications, next generation sequencing-advances, applications and challenges.

[14] Laehnemann D, Borkhardt A, McHardy AC. (2015). Denoising DNA deep sequencing data—high-throughput sequencing errors and their correction. Brief Bioinformatics 17: 154-79.

[15] Laver T, Harrisona J, O'Neill PA, Moorea K, Farbosa A, et al. (2015). Studholme. Assessing the performance of the oxford nanopore technologies MinION. Biomolecular Detection and Quantifcation 3: 1-8.

[16] Liu L, Li Y, Li S, Hu N, He Y, et al. (2012). Comparison of next-generation sequencing Systems. Journal of Biomedicine and Biotechnology.

[17] Lu H, Giordano F, Ning Z. (2016). Oxford nanopore MinION sequencing and genome assembly. Genomics Proteomics Bioinformatics 14: 265-279.

[18] Mardis ER. (2008). Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 9: 387-402.

[19] Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376-80.

[20] Masoudi-Nejad A, Narimani Z, Hosseinkhan N. (2013). Next generation sequencing and sequence assembly. Methodologies and algorithms. Springer.

[21] McCoy RC, Taylor RW, Blauwkamp TA, Kelley JL, Kertesz M, et al. (2014). Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. PLoS ONE 9:e106689.

[22] Mikheyev AS, Tin MMY. (2014). A frst look at the oxford nanopore MinION sequencer. Molecular Ecology Resources 14: 1097–1102.

[23] Myllykangas S, Buenrostro J, Ji HP. (2012). Overview of sequencing technology platforms, bioinformatics for high throughput sequencing.

[24] Reuter JA, Spacek DV, Snyder MP. (2015). High-throughput sequencing technologies. Molecular Cell 58: 586-597.

[25] Rhoads A, Au KF. (2015). PacBio sequencing and its applications.Genomics, Proteomics & Bioinformatics 13: 178-289.

[26] Sanger F, Coulson AR. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94:441-448.

[27] Sanger F, Coulson AR, Barrell BG, Smith AJ, Roe BA. (1980). Cloning in single stranded bacteriophage as an aid to rapid dna sequencing. J Mol Biol 143: 161-178.

[28] Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, et al. (2010). Genome sequence of the paleopolyploid soybean. Nature 463: 178-83.

[29] Shendure J, Ji H. (2008). Next-generation DNA sequencing. Nature Biotechnology 26: 135–1145.

[30] The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796-815.

[31] Thudi M, Li Y, Jackson SA, May GD, Varshney RK. (2012). Current stateof-art of sequencing technologies for plant genomics research. Brief Funct Genomics 11: 3-11.

[32] Vezzi F. (2012). Next generation sequencing revolution challenges: Search, assemble, and validate genomes. Ph.D, Universita degli Studi di Udine, Italy.

Citation: *Zewdu Asrat, (2021). "Review on Mechanism of DNA Sequencing Technologies in Crop Improvement", International Journal of Research Studies in Agricultural Sciences (IJRSAS), 7(8), pp. 27-36 DOI: http://dx.doi.org/10.20431/2454-6224.0708004*