© Kamla-Raj 2019 Int J Hum Genet, 19(3): 123-133 (2019)
PRINT: ISSN 0972-3757 ONLINE: ISSN 2456-6330 DOI: 10.31901/24566330.2019/19.03.732

Sequencing Technologies: Introduction and Applications

Rishabh Gaur, Anupam Jyoti, Sanket Kaushik and Vijay Kumar Srivastava*

Amity Institute of Biotechnology, Amity University Rajasthan, Kant Kalwar, NH-11C, Jaipur-Delhi Highway, Jaipur 303 002, Rajasthan, India

KEYWORDS Base Sequence. High-throughput Nucleotide Sequencing. Next-generation Sequencing

ABSTRACT In the past decade, the expansion of next-generation sequencing (NGS) has smoothened the approach for whole-genome analysis in individuals. The sequencing technology until today has developed very fast. Many techniques of sequencing are present these days, but the challenge is cost effectiveness in sequencing technologies. The cost varies from technique to technique and the new and improved sequencing targets the region of determination and cost. That's why, the new and improved sequencing techniques have more effective methods in terms of run time, time is taken, cost per megabase, read length, error rate and output data run. The massive parallel DNA sequencing is revolutionary, changing the paradigm of genomics to deal with biological questions at a genome-wide scale. Sequencing techniques now entitle medical diagnostics and other forms of health care, as well as disease risk, therapeutic recognition, and prenatal testing. This review explores the current state, methodology, and comparison of first, second, third and fourth generation sequencing techniques.

INTRODUCTION

In recent years, techniques of molecular biology developing at a very fast speed have led to the improvement of Next Generation Sequencing (NGS) (Zhang et al. 2011). All the NGS techniques share a common feature of massively parallel sequencing strategy, where thousands of reads are achieved at the same time (Srinivasan and Batra 2014**)**. The DNA sequencing method subsists of three components: preparation of the sample, physical sequencing, and reassociation. In the first phase, the genome of interest is fragmented into multiple pieces. In the second phase it rely on the quantity of sample DNA, these fragments may be amplified into several replicas by means of a different molecular biology technique. In the third phase, the individual bases in each and every fragment are analyzed in a sequential manner to create an individual reads (Schatz et al. 2010). Sanger and Coulson in 1975 (the chain-termination method) (Sanger and Coulson 1975; Sanger et al. 1977) and simultaneously by Maxam and Gilbert in 1977 (chemical method), are known as founders of sequencing technology. These sequencing techniques co-relate genotype with phenotypic characters and directly help in genomics, transcriptomics, oncology, evolutionary biology, forensic science, etc. (Alvarez-Cubero et al. 2017). Sequencing technologies recently developed, are so fast which gives us a map of the genome of the simple virus to complicated human beings. Nowadays, new and advance sequencing technologies improve the sequencing platform. These advancements helped researchers to better analyze and understand the genome of different organisms to a very large extent (Alvarez-Cubero et al. 2017). Sequencing is mainly categorized into four generations: Sanger method and Maxam Gilbert's method in the first generation. More than twenty years among them, the Sanger method is a dominant methodology in industries.

Later on in the Second generation, sequencing includes De Novo whole genome sequencing, re-sequencing of genome, Roche 454 system- pyrosequencing, Illumina Hi sequencing system and Illumine Mi sequencing: sequence by synthesis, AB SOLiD 5500 system: sequence by ligation. After second-generation sequencing, new techniques that came into the picture in the third generation, comprise of Helioscope and Pac Bio. The fourth generation sequencing consists of many new techniques, among them Nanopore and DNA nanoball sequencing are mostly used (Alvarez-Cubero et al. 2017). Above mentioned techniques have a similar concert on throughput, precision and cost in contrast to Sanger and Maxam, after the success of these

*Address for correspondence:
E-mail: vksrivastava@jpr.amity.edu, vijaytechno@gmail.com

platforms, new machines and techniques were developed to fulfil the requirements of various laboratories that belong to the academics and industry (Levy and Myers 2016). However, the limited quantity of sample and its high grade of degradation is still the major hurdle for the NGS platforms in the forensic field application. This paper focuses on comparing and contrasting different commercial technologies most normally applied - NGS technologies, namely the Sanger, Maxam, Roche/454, ABI/SOLiD, Illumina, Ion PGM Nano-pore, DNA nanoball and presenting the applications areas of each in terms of basic science as well as in translational research areas such as clinical diagnostics, metagenomics, phylogenetic studies, plant biology and forensic science.

Objectives

The objective of this paper is a profound investigation of sequencing technologies and their application, particularly in various fields of research, including a conclusion and limitation about sequencing technologies.

MATERIAL AND METHODS

An efficient literature quest in PubMed, Scopus electronic databases and Science Direct was performed for the period of October 2012 to March 2019. After carefully examining all the literature out of 200 papers, the researchers took 60 articles to write the manuscript. All these papers were identified, and utilized techniques of sequencing and NGS platforms (Roche/454, ABI/SOLiD, and different techniques) in selected journals were identified.

OBSERVATIONS AND DISCUSSION

First Generation Sequencing Technologies

Sanger Method

Previously in 1975, the Sanger proposed the concept of DNA sequencing technique in his pioneered Croonian lecture (Sanger and Coulson 1975) and afterward, published a technique for identifying sequences in genetic material by primed synthesis with DNA polymerase enzyme (Sanger and Coulson 1975). It is also referred to as di-deoxy sequencing or chain termination. It is based on de-deoxy nucleotides along with normal nucleotides formed in DNA. DNA has to be denatured into single strand before sequencing and a primer is annealed. Either primer or one of nucleotide is radioactively or fluorescently labelled for detection of a final product or a gel. After annealing of primer, the solution is divided into four tubes containing different di-deoxy nucleotide triphosphate present. To synthesize the DNA, nucleotides and DNA polymerase are added (Sanger et al. 1977). In Sanger sequencing, all reactions start from similar nucleotides and end with specific bases and it was the main tool to complete the human genome project in the year 2001 (Fig.1). Due to the high cost of sequencing, it accelerated the development of next-generation sequencing (Alvarez-Cubero et al. 2017). In Parallel, Maxam and Walter Gilbert gave the chemical degradation method of DNA sequencing in which terminally labelled DNA fragments were chemically cleaved at specific bases and separated by using gel electrophoresis technique (Maxam and Gillbert 1977).

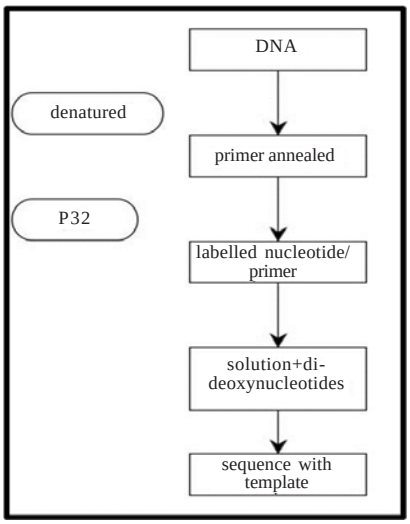

Fig. 1. Simplified representation of a Sanger sequencing *Source:* Author

Maxam and Gilbert Method

Maxam and Gilbert develop a new advanced technique for sequencing. In this technique, the

determination of nucleotides sequence of the terminally labelled DNA molecule was done by breaking adenine, guanine, cytosine, and thymine by the chemical reagent. Set of radioactive fragments expand from the labelled end to each of the position due to partial cleavage at each base. Polyacrylamide gel electrophoresis resolves this single strand fragment. Four different chemical cleavages produce the autoradiograph of a gel, each specific for a base (Fig. 2). Thereafter the sequence can be read directly by analyzing the band's pattern (Fig. 3) (Maxam and Gillbert 1977).

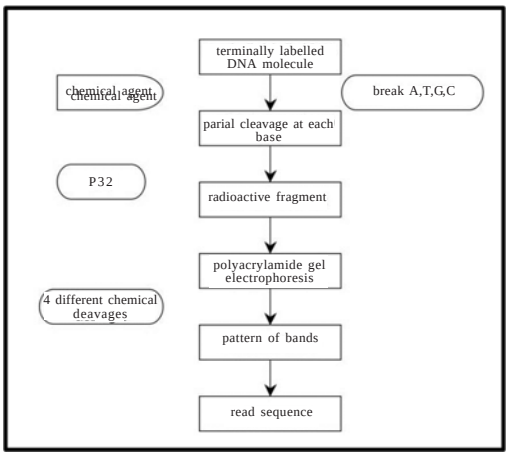

Fig. 2. The diagram showing the flow chart for Maxam and Gilbert method
Source: Author

Second Generation Sequencing (SGS) Technologies

Commercial SGS was developed in 2005, in reaction to the high cost and the low throughput of the first generation sequencing technology (Liu et al. 2012). To determine the issue, NGS tools accomplish much higher throughput by sequencing a large number of DNA molecules side by side. There are different methods developed for SGS which are as follows:

Roche 454 System

Pyrosequencing technology is the first substitute to the Sanger method. It was conceived for de novo sequencing (Dudley et al. 2012). It provides chains of enzymes to precisely identify a nucleic acid sequence. The sequencing primer is hybridized to single strand DNA and mixed with four enzymes; DNA polymerase, ATP Sulfurase, luciferase, and apyrase (Ronaghi et al. 1996). Roche 454 system was the first commercially achieved next-generation system, used the biggest libraries among all the other platforms. Initially, the read length is 100 to 150 bp (Alvarez-Cubero et al. 2017; Kochling et al. 2017). In this sequencing method, a single DNA fragment is hybridized to beads array and emulsified with reagent necessary to PCR bound template. The beads are deposited in the pico-titer plate along with DNA polymerase, primers and some enzymes essential to produce fluorescence through consumption of inorganic phosphates produced during sequencing (Medini et al. 2008). In this method, the pyrophosphates are identified by an enzymatic luminometric inorganic pyrophosphate detection assay (ELIDA) through the generation of a light signal due to change of pyrophosphate into adenosine triphosphate (ATP) (Kumar et al. 2012) (Fig. 4).

Illumina Hi Seq System and Illumina Mi Seq System

Both the systems are based on the technique of sequencing by synthesis using reversible

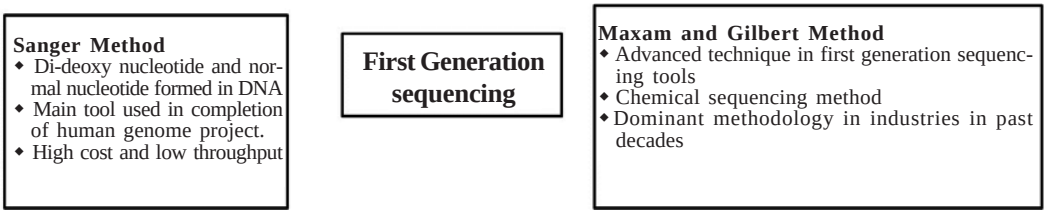

Fig. 3. Specific features of the first generation sequencing technologies
Source: Author

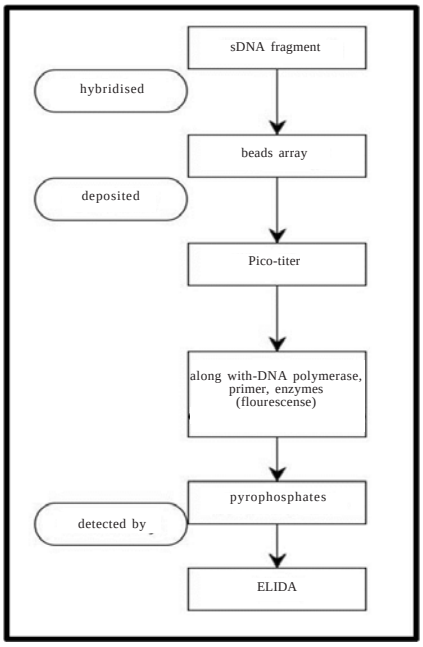

Fig. 4. Overview of the Roche 454 workflow
Source: Author

dye-terminators (Alvarez-Cubero et al. 2017). Both of the techniques usually generate 50-60 million reads, average read length is 40-50 bases per run. The ligation of the adapters occurs at one and the other ends of the fragmented DNA, and then loaded on glass slides. The DNA fragment hybridized to - oligo1, is covalently attached to the surface. The free opposite end is hybridized with complementary oligo2 on a slide to form the bridge. By adding polymerase enzyme DNA and dNTP's, the ssDNA is the bridge, amplified as synthesis starts in the 5' to 3' from oligo2 (Caporaso et al. 2012). For the next cycle, newly synthesized complementary strand and template strand denatured to again start the amplification. In each channel of the flow cell, millions of clusters of duplex DNA are generated, amplified fragments are annealed, denatured with primer and synthesis begins with the incorporation of dNTPs. Thus, the Illumina system has the capacity to produce large measure data in less cost and time as compared to Sanger sequencing (Aradhya and Nussbaum 2018; Meyer and Kircher 2010; Kozich et al. 2013).

AB SOLiD 5500 System

This sequencing is based on the technology of two base sequencing that is based on ligation sequencing. It consists of a series of hybridization-ligation steps. The di-base probe is octamer containing two bases of known position, three degenerate bases, and three more degenerate bases attached to the fluorophore. Sixteen combinations of di-base probes are used, with each base represented by one of the four fluorescent dyes. The read length of SOLiD initially was 35 bp, and the output is 3G data per run. Then, SOLiD 5500 XL improved read length and also improved accuracy, output to 85 bp. By using wildfire technology new update, 5500 xl-w, increased throughput and lowered the cost (Alvarez-Cubero et al. 2017). The system was developed by the medical school of Harvard and Howard Hugs Medical Institute. The clonal beads are made inside the microreactor in the presence of DNA template, beads, primers and PCR components. The beads are allowed to attach to the glass slide due to the modification at their 3' end by DNA template. A primer annealing was done to an adapter on the DNA template and flow cell are filled with fluorescently tagged oligonucleotides. Once the oligonucleotides match with the template sequence, it is ligated onto the primer and unincorporated nucleotides are washed away. The different colors attached to the primer are captured by a charged coupled device (CCD) camera. Each fluorescence wavelength coincides to dinucleotide combinations (Kumar et al. 2012; Mardis 2008a, 2008b, 2008c; Van dijk et al. 2014) (Fig. 5).

Ion PGM (Personal Genome Machine) and Ion Proton

PGM (Personal Genome Machine) is based on semiconductor sequencing by synthesis technology (Millat et al. 2014). It involves natural nucleotides that are not modified. It does not require laser, optics, camera, fluorescence and fast recognition to sequence extension. In this technology, a DNA fragment is flanked by specific ion torrent adapters and it is ideal for small genome, set of a gene, gene expression profiling. A novel system is established for the WGS in similar run time as PGM ion proton sequencer (Bakker et al. 2014). Chips for ion proton sequencer provide high throughput to 1000 folds higher

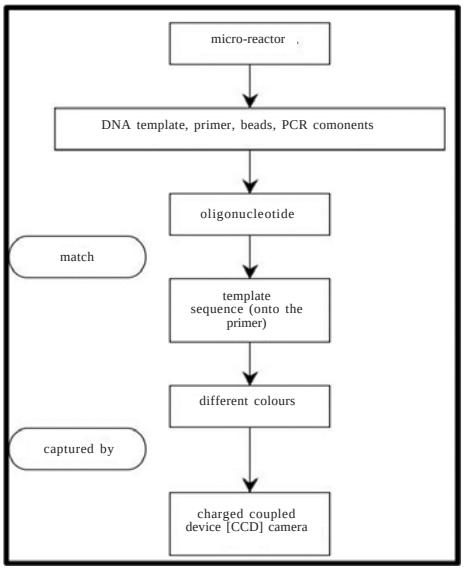

Fig. 5. Flow chart explaining procedure of the AB SOLiD system
Source: Author

than ion PGM, in simply two hours (Alvarez-Cubero et al. 2017). Sequencing templates are generated on sphere via emulsion PCR (an oil-water emulsion is used to partition small reaction vessel). One of the complementary primers is present in solution while others bound to the sphere. Finally, sphere containing amplified strands are deposited into sequencing chips. There are flow compartments and solid-state pH sensors micro arrayed well in the ion torrent chips (Rougemont and Naef 2012). The detection of bases during sequencing is based on the release of H^+ during nucleotide extension. The sensor detects the change in pH when H^+ is released (Fig. 6) (Buermans and Dunnen 2014).

Third Generation Sequencing or Single Molecule Sequencing

Advantage of TGS technology is the capability to sequence a single molecule of DNA. The data obtained or generated by TGS is of a different type as compared to the other sequencing techniques.

Helioscope

To sequence a single molecule of DNA, the method was brought into notice by Braslavsky et al. (2003). It depends on the principle of single-molecule fluorescent sequencing. It is a plat-

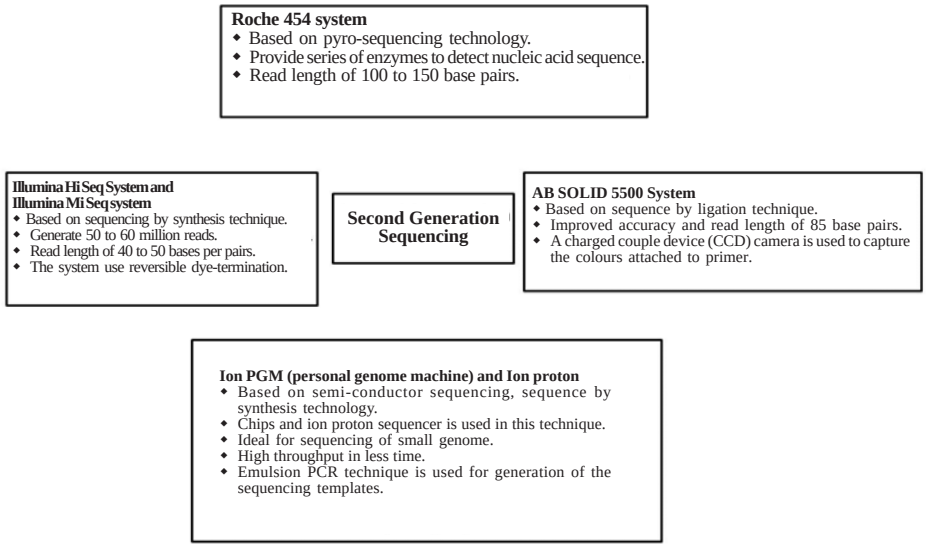

Fig. 6. Characteristics of four leading second generation sequencing
Source: Author

form for first commercial, high throughput for single molecule sequencing. The read length is 25 bp and 35 G of sequence per run in output (Alvarez-Cubero et al. 2017). Sequencing involves library preparations by adding a poly-A tail to the shared DNA fragment, attached to flow cells through poly-T anchors (Ozsolak and Milos 2011a, b). Heliscope sequencer detects only one fluorescent tagged nucleotide out of four by DNA extension (Kumar et al. 2012).

Pac Bio

Single molecule run time (SMRT) is a basis for the Pacific bioscience (Schadt et al. 2010). This type of sequencing requires four to six hours and does not require polymerase chain reaction, direct observation to enzyme reaction in real time and read length is approximately 1300 bp (Pareek et al. 2011). The DNA polymerase is affixed at the base of the zero mode waveguide (ZMW). DNA fragments along with anchored DNA polymerase and nucleotides (each labelled to different color fluorophore) are added into the ZMW chamber. Once the nucleotide circulates all the way through the ZMW, incorporated nucleotide remains at the bottom. The fluorophore of the integrated nucleotides is excited by the laser placed below the ZMW (Kumar et al. 2012).

Single-molecule Real-time Sequencing

Its data is obtained by real-time sequencing engine. The synthesis of data is depending upon continuous template-directed synthesis using four recognizable fluorescently labelled dNTP's, performed by DNA polymerase. The enzymatic integration into a growing strand of DNA with ZMW, provide a determination of thousands of single molecule sequencing reaction. The fluorophore associated to the terminal phosphate to the dNTP's allows observation of DNA synthesis without steric hindrance over thousands of bases (Fig. 7) (Levene et al. 2003; Eid et al. 2009; Flusberg et al. 2010).

Fourth Generation Sequencing

Over the years, advances in sequencing technologies have enhanced vastly. One of the significant accomplishment of NGS is the nanopore sequencing, Nanopore DNA sequencing with MspA and DNA Nano-ball sequencing.

NanoPore Based Sequencing Technology

Nano-pore sequencing has the potential to rapidly or consistently sequence the entire human genome for the cost between a hundred US dollars to thousand US dollars (Maitra et al. 2012), depending on single molecule technique. Nanopore technologies are characterized by two types: biological and single state. Hagan Bayley and Gordon Sanghera founded Oxford Nano-pore technology (ONT) for commercial use of Nanopore based sequencing system (Fig. 8) (Feng et al. 2015; Schadt et al. 2010). In Nano-pore technology, throughout the sequencing procedure,

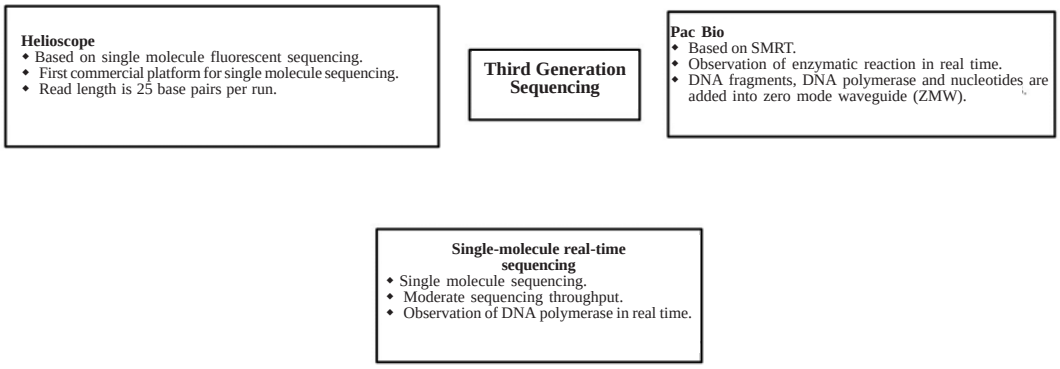

Fig. 7. Technical advances of three leading third generation sequencing
Source: Author

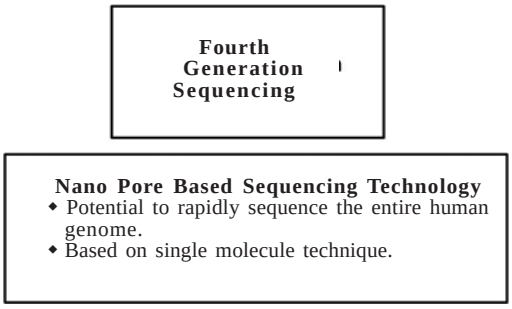

Fig. 8. Advanced technological features of the fourth generation sequencing
Source: Author

an ion current passes through the pore that is blocked by the nucleotide and it is devoid of nucleotide labelling and detection. During the completion of the method, different features and parameters, such as diameter, length, and confirmation of the molecule are identified (Astier et al. 2006; Rhee and Burns 2006).

Nanopore DNA Sequencing with *Mycobacterium smegmatis* Porin A (MspA)

In this method, MspA protein has a better resolution because of the shorter blockade region. It is used as the pore and the consequence of a linear molecule of single-stranded DNA (ssDNA) on the current transiting the pore is calculated (Derrington et al. 2010). A region of double-stranded DNA (dsDNA) is implemented to sluggish down the transit of the ssDNA via the pore so that it allows detecting the individual bases as they disrupt current transiting the pore (Ku and Roukos 2013).

DNA Nano-ball Sequencing

It involves separating DNA that is to be sequenced, shearing it into small 100 – 350 base pair (bp) fragments, they are ligated with adapter sequences, and further on the fragment is circularized. The rolling circle application is used to copy that result in many single-stranded copies of each fragment. The copies of the DNA concatenates head to tail in a long strand, and are packed together into a DNA nanoball and later on are adsorbed on to a sequencing flow cell. At each interrogated position the color of the fluorescence is documented all the way through a camera having high resolution. The created single-stranded circular DNA template containing sample DNA ligated to the adapter sequences and later on long string of DNA is formed by the amplification of the full-length sequence. It is accomplished by binding of Phi 29 DNA polymerase to DNA template by rolling circle replication. The synthesized strand does not have the circular template which result in nanoparticle self-assembles, that is compacted ball of 300 nm. Nanoballs repel each other due to being negatively charged. In order to determine the sequence of the DNA, the DNA nanoballs have adhered to a patterned arrangement flow cell. The flow cell is made up of positively-charged aminosilane on which the DNA nanoballs are attached to in a well-organized array, allowing a very high density of DNA nanoballs to be sequenced (Xu et al. 2019).

Applications of Next Generation Sequencing

Next-generation sequencing has so many applications in many areas such as WGS, targeted sequencing, de novo sequencing, single RNA sequencing and the sequencing of DNA extracted by antibodies targeting DNA binding protein (Jensen et al. 2013; Koboldt et al. 2013).

All these applications are used for studies in the number of areas:

Genetic Diagnosis of Rare Disease

NGS technology is widely used by researchers working on many human diseases (Metzker 2010). The whole genome sequencing of certain species and organism is easily done by NGS, which helps in analyzing the total sequence and gives information about the different disease (Van et al. 2016; Boycott et al. 2013; Choi et al. 2009).

Cancer Diagnosis

NGS used in molecular diagnostic laboratories for the recognition of DNA mutation in tumor tissue. Many medical centers and universities recently implemented NGS for clinical diagnosis. NGS is a multiplex prognostic test which examines a limited area of tumor DNA for deviation that can be used as a molecular target for

therapy (Todenhofer et al. 2018). NGS recover the efficiency of therapies, but NGS testing are costly and the profit is not yet achieved in relation to these costs (Sholl 2017; Nikiforov et al. 2014; Serrati et al. 2016).

Microbiological Studies

The major use of NGS in microbial technology is to alter the conservative characterization of a pathogen by staining, metabolism and morphology along with genomic criteria of pathogen (Mayo et al. 2014). The genomic properties of the pathogens define information about drug sensitivity and relation between different pathogen which can be used to draw infection outburst. Recently, NGS was used to trace an infection occurrence of methicillin-resistant *Staphylococcus aureus* on a neonatal intensive care unit (Behjati and Tarpey 2013; Roh et al. 2010).

Transcriptomics

The development of NGS technology led to the whole transcriptome amplification to enable sequencing of full-length mRNA from many individual cells (Van et al. 2014; Mutz et al. 2013).

Phylogenetic Studies

The human genome has all the information embedded in it, like ancestral information, phylogenetic relationship between other species, ethnicity, age and physical and physiological characteristics (Yang et al. 2014).

Forensic Studies/Research

NGS technology provides new chances in the biomedical field of forensic science. NGS technology can be put in forensic science to altogether examine many loci of forensic concern in diverse genetic circumstances, such as autosomes, mitochondrial and sex chromosomes (Marshall et al. 2017; Strobl et al. 2018).

Plant Biology

The NGS sequencing is a revolutionary technique in the sequencing of plant genomes. The recent achievement in the field is the sequencing of plant species *Arabidopsis thaliana* (Koornneef and Meinke 2010). To determine the whole genome analysis/sequencing of the plants, most of the researchers commonly prefer NGS technology (Hodzic et al. 2017; Egan et al. 2012; Brautigam and Gowik 2010).

Immunoinformatics

The NGS development has strengthened the way for WGS in persons (Chaitankar et al. 2016), with the help of NGS technology, research on human leukocyte antigen (HLA) gives very useful information involved in immunity. The HLA region has been associated with 100 or more than 100 diseases, and most of them are an autoimmune disease (Hosomichi et al. 2015; Gough and Simmonds 2007).

Limitations of the Next Generation Sequencing

The major disadvantage of NGS technology in the biomedical field is to put necessary things, computer facility, its storage, and an expert researcher to analyze and identify the data of different samples (Gullapalli et al. 2012; Pareek et al. 2011). One thousand dollars can read about 150,000,000 reads in the NGS, on the other hand, a single Sanger method read, typically costs approximately one dollar (Mayo et al. 2014). However, to run the NGS with less cost, one would have large batches of a sample with supra-regional centralization that need to be further investigated.

The identification of technical properties can be useful for suitable selection of a platform for the sequencing. To interpret the repetitive sequences longer reads are preferred so the Sanger method would be the most suitable. However, it is avoided due to high cost, time and labor requirements. The NGS platform is used in combination with Sanger sequencing, for genome sequencing and sequences that are improperly sequenced by NGS. Alvarez-Cubero et al. (2017) suggest the Roche/454, covering the longest read-length capacity and highest speed among NGS platforms, for the sequencing of the large genome. The third and fourth generation sequencing is the techniques that are mostly used to sequence the large genome of the plant and

animals with low cost and efficient accuracy. Still, there is a need in the near future to come up with the sequencing technologies that would be easily available to the users with the least cost.

CONCLUSION

In the fast-rising HT-NGS technologies, the major question is to handle the analysis of vast fabrication of sequencing database through advanced *In-silico* tools. Each day new innovations are developed in the field of sequencing and have a greater chance to achieve cheap, fast and accurate sequencing data. The sequencing is important in the diagnosis of many diseases, sequence analysis of organisms, drug designing, etc. and the sequencing techniques are improving day-by-day with new concepts still developing. Sequencing techniques based on different generation are categorized by run time, cost per megabase, error rate and time taken. In recent years, there has been a tremendous advancement in cost-effectiveness and rapid sequencing. In general, sequencing has been restricted to small fragments of around thousands of bases due to maintaining accuracy and high sequencing quality. In the Sanger sequencing method, sequencing is higher in cost and takes so much time, although it was very successful technology in most of the commercial sectors in the past. NGS comes with a very good machinery work along with a number of techniques providing low cost and high runtime sequencing technology which is mostly used by companies recently. On the other hand, the third generation sequencing offers low cost, high throughput, improved quantitative accuracy but the data obtained is of a different kind. In recently developed sequencing technology (fourth generation sequencing) the Nanopore sequencing technology is a very new concept working on lesser cost and higher run time. This study and these recent challenges hold the significance of performing normal evaluations of the rapidly improving hardware and software for identifying different genomic sequences. Hope the sequencing world discovers a new technology soon with more efficiency, accuracy and least cost for human welfare.

ABBREVIATIONS

NGS – Next Generation Sequencing
Ion PGM – ion personal genome machine
TGS – third generation sequencing
SMRT – single molecule real time
ZMW – zero mode waveguide

ACKNOWLEDGEMENTS

We extend our appreciation towards Amity University Rajasthan, Jaipur, India for its valuable support throughout the work.

REFERENCES

Alvarez-Cubero MJ, Saiz M, Martinez-Garcia B, Sayalero SM, Entrala C, Lorente JA, Martinez-Gonzalez LJ 2017. Next generation sequencing: An application in forensic sciences. *Annals of Human Biology,* 44: 581-592.

Aradhya S, Nussbaum RL 2018. Genetics in mainstream medicine: Finally within grasp to influence healthcare globally. *Molecular Genetics & Genomic Medicine,* 6: 473-480.

Astier Y, Braha O, Bayley H 2006. Toward single molecule DNA sequencing: Direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *Journal of the American Chemical Society,* 128: 1705-1710.

Behjati S, Tarpey PS 2013. What is next generation sequencing? *Archives of Disease in Childhood*, 98: 236-238.

Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE 2013. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nature Reviews,* 14: 681-691.

Braslavsky I, Hebert B, Kartalov E, Quake SR 2003. Sequence information can be obtained from single DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 100: 3960-3964.

Brautigam A, Gowik U 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology*, 12: 831-841.

Buermans HP, den Dunnen JT 2014. Next generation sequencing technology: Advances and applications. *Biochimica Et Biophysica Acta*, 1842: 1932-1941.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal,* 6: 1621-1624.

Chaitankar V, Karakulah G, Ratnapriya R, Giuste FO, Brooks MJ, Swaroop A 2016. Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. *Progress in Retinal and Eye Research,* 55: 1-31.

Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP 2009. Ge-

netic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 106: 19096-19101.

den Bakker HC, Allard MW, Bopp D, Brown EW, Fontana J, Iqbal Z, Kinney A, Limberger R, Musser KA, Shudt M, Strain E, Wiedmann M, Wolfgang WJ 2014. Rapid whole-genome sequencing for surveillance of Salmonella enterica serovar enteritidis. *Emerging Infectious Diseases*, 20: 1306-1314.

Derrington IM, Butler TZ, Collins M.D, Manrao E, Pavlenok M, Niederweis M, Gundlach JH 2010. Nanopore DNA sequencing with Msp. A. *Proceedings of the National Academy of Sciences of the United States of America*, 107: 16060–16065.

Dudley DM, Chin EN, Bimber BN, Sanabani SS, Tarosso LF, Costa PR, Sauer MM, Kallas EG, O'Connor DH 2012. Low-cost ultra-wide genotyping using Roche/454 pyrosequencing for surveillance of HIV drug resistance. *PloS One*, 7: e36494.

Egan AN, Schlueter J, Spooner DM 2012. Applications of next-generation sequencing in plant biology. *American Journal of Botany*, 99: 175-185.

Eid J, Fehr A, Gray J, Luong K, et al. 2009. Real-time DNA sequencing from single polymerase molecules. *Science*, 323: 133-138.

Feng Y, Zhang Y, Ying C, Wang D, Du C 2015. Nanopore-based fourth-generation DNA sequencing technology. *Genomics, Proteomics & Bioinformatics*, 13: 4-16.

Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods*, 7: 461-465.

Gough SC, Simmonds MJ 2007. The HLA region and autoimmune disease: Associations and mechanisms of action. *Current Genomics*, 8: 453-465.

Gullapalli RR, Desai KV, Santana-Santos L, Kant JA, Becich MJ 2012. Next generation sequencing in clinical medicine: Challenges and lessons for pathology and biomedical informatics. *Journal of Pathology Informatics*, 3: 40.

Hodzic J, Gurbeta L, Omanovic-Miklicanin E, Badnjevic 2017. An overview of next-generation sequencing platforms used in published draft plant genomes in light of genotypization of Immortelle Plant (*Helichrysium Arenarium*). *Medical Archives*, 71: 288-292.

Hosomichi K, Shiina T, Tajima A, Inoue I 2015. The impact of next-generation sequencing technologies on HLA research. *Journal of Human Genetics*, 60: 665-673.

Jensen TH, Jacquier A, Libri D 2013. Dealing with pervasive transcription. *Molecular Cell*, 52: 473-484.

Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER 2013. The next-generation sequencing revolution and its impact on genomics. *Cell*, 155: 27-38.

Kochling T, Ferraz AD, Jr., Florencio L, Kato MT, Gavazza S 2017. 454-Pyrosequencing analysis of highly adapted azo dye-degrading microbial communities in a two-stage anaerobic-aerobic bioreactor treating textile effluent. *Environmental Technology*, 38: 687-693.

Koornneef M, Meinke D 2010. The development of Arabidopsis as a model plant. *Plant J*, 61: 909-921.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79: 5112-5120.

Ku CS, Roukos DH 2013. From next-generation sequencing to nanopore sequencing technology: Paving the way to personalized genomic medicine. *Expert Review of Medical Devices*, 10: 1-6.

Kumar S, Banks TW, Cloutier S 2012. SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics*, 2012: 831460.

Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science*, 299: 682-686.

Levy SE, Myers RM 2016. Advancements in Next-Generation Sequencing. *Annual Review of Genomics and Human Genetics*, 17: 95-115.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M 2012. Comparison of next-generation sequencing systems. *Journal of Biomedicine & Biotechnology*, 2012: 251364.

Maitra RD, Kim J, Dunbar WB 2012. Recent advances in nanopore sequencing. *Electrophoresis*, 32: 3418-3428.

Mardis ER 2008a. A decade's perspective on DNA sequencing technology. *Nature*, 470: 198-203.

Mardis ER 2008b. The impact of next-generation sequencing technology on genetics. *Trends Genet*, 24: 133-141.

Mardis ER 2008c. Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, 9: 387-402.

Marshall C, Sturk-Andreaggi K, Daniels-Higginbotham J, Oliver RS, Barritt-Ross S, McMahon TP 2017. Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: Implications for the use of a positive control in a next-generation sequencing procedure. *Forensic Science International*, 31: 198-206.

Maxam AM, Gilbert W 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74: 560-564.

Mayo B, Rachid CT, Alegria A, Leite AM, Peixoto RS, Delgado S 2014. Impact of next generation sequencing techniques in food microbiology. *Current Genomics*, 15: 293-309.

Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R 2008. Microbiology in the post-genomic era. *Nat Rev Microbiol*, 6: 419-430.

Metzker ML 2010. Sequencing technologies - the next generation. *Nature Reviews*, 11: 31-46.
Meyer M, Kircher M 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6): pdb.prot5448. doi: 10.1101/pdb.prot5448.
Millat G, Chanavat V, Rousson R 2014. Evaluation of a new NGS method based on a custom AmpliSeq library and Ion Torrent PGM sequencing for the fast detection of genetic variations in cardiomyopathies. *Clinica Chimica Acta: International Journal of Clinical Chemistry*, 433: 266-271.
Mutz KO, Heilkenbrinker A, Lonne M, Walter JG, Stahl F 2013. Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology,* 24: 22-30.
Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, Hodak SP, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN 2014. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer,* 120: 3627-3634.
Ozsolak F, Milos PM 2011a. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews,* 12: 87-98.
Ozsolak F, Milos PM 2011b. Transcriptome profiling using single-molecule direct RNA sequencing. *Methods in Molecular Biology*, 733: 51-61.
Pareek CS, Smoczynski R, Tretyn A 2011. Sequencing technologies and genome sequencing. *Journal of Applied Genetics,* 52: 413-435.
Rhee M, Burns MA 2006. Nanopore sequencing technology: Research trends and applications. *Trends in Biotechnology,* 24: 580-586.
Roh SW, Abell GC, Kim KH, Nam YD, Bae JW 2010. Comparing microarrays and next-generation sequencing technologies for microbial ecology research. *Trends in Biotechnology*, 28: 291-299.
Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry,* 242: 84-89.
Rougemont J, Naef F 2012. Computational analysis of protein-DNA interactions from ChIP-seq data. *Methods in Molecular Biology*, 786: 263-273.
Sanger F, Coulson AR 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology,* 94: 441-448.
Sanger F, Nicklen S, Coulson AR 1977. DNA sequencing with chain-terminating inhibitors. *Biotechnology,* 24: 104-108.
Schadt EE, Turner S, Kasarskis A 2010. A window into third-generation sequencing. *Human Molecular Genetics,* 19: 227-240.
Schatz MC, Delcher AL, Salzberg SL 2010. Assembly of large genomes using second-generation sequencing. *Genome Research*, 20: 1165-1173.
Serrati S, De Summa S, Pilato B, Petriella D, Lacalamita R, Tommasi S, Pinto R 2016. Next-generation sequencing: Advances and applications in cancer diagnosis. *OncoTargets and Therapy,* 9: 7355-7365.
Sholl L 2017. Molecular diagnostics of lung cancer in the clinic. *Translational Lung Cancer Research,* 6: 560-569.
Srinivasan S, Batra J 2014. Four generations of sequencing- Is it ready for the clinic yet. *Journal of Next Generation Sequencing & Applications,* 1: 8.
Strobl C, Eduardoff M, Bus MM, Allen M, Parson W 2018. Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. *Forensic Science International*, 35: 21-25.
Todenhofer T, Struss WJ, Seiler R, Wyatt AW, Black PC 2018. Liquid biopsy-Analysis of Circulating Tumor DNA (ctDNA) in bladder cancer. *Bladder Cancer*, 4: 19-29.
Van Amerongen RA, Retel VP, Coupe VM, Nederlof PM, Vogel MJ, van Harten WH 2016. Next-generation sequencing in NSCLC and melanoma patients: A cost and budget impact analysis. *Ecancermedical Science*, 10: 684.
Van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C 2014. Ten years of next-generation sequencing technology. *Trends Genet,* 30: 418-426.
Yang Y, Xie B, Yan J 2014. Application of next-generation sequencing technology in forensic science. *Genomics, Proteomics & Bioinformatics,* 12: 190-197.
Xu Y, Lin Z, Tang C, Tang Y, Cai Y, Zhong H, Wang X, Zhang W, Xu C, Wang J, Wang J, Yang H, Yang L, Gao Q 2019. A new massively parallel nanoball sequencing platform for whole exome research. *BMC Bioinformatics,* 20: 153.
Zhang J, Chiodini R, Badr A, Zhang G 2011. The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, 38: 95-109.

Paper received for publication in March, 2019
Paper accepted for publication in June, 2019