Nanopore Sequencing: A Game-Changer for NGS

Gayathri S.S, Aswin Mohan, Shahanas Naisam.

Dated: 01th August, 2023
Keywords: *Galleria mellonella,* PE, Gut microbiota, Long-chain hydrocarbons, Salivary glands

INTRODUCTION

In this era of the scientific revolution and technological advancement, frontiers like genomic and DNA analysis experienced a significant transformation. Once the human genome was considered a mysterious code of life that is difficult to comprehend, on the other hand now it projects data with utmost clarity, offering insights into our biology, evolution, and susceptibility to diseases. Understanding the whole sequencing system is now being traced more keenly, and there also exists a high demand for more cost-effective techniques in DNA sequencing for the practical deployment of genome technologies. This communication discusses the intersection of two cutting-edge disciplines i.e.; Nanotechnology and Next-Generation Sequencing (NGS) for exploring the dynamic landscape of genomics.

A significant milestone in the field of genome sequencing is the completion of the Human Genome Project by scientists all over the world which started in 1990. They completed the massive sequencing of 3.2 billion base pairs in the human genome after 13 years of dedicated effort. It was for the first time that scientists could decode the entire genetic blueprint of humans, which paved the way for new exciting discoveries. Furthermore, this milestone points out the

complexities of our genetic code and the need to understand it rapidly, accurately, moreover cost-effectively.

Next Generation Sequencing (NGS) is a massively parallel sequencing technology that provides ultra-high-throughput sequencing of DNA and RNA. It has revolutionized the field of genomic research by sequencing the entire human genome within a single day which overcomes the limitations of traditional sequencing without compromising its speed, accuracy, and throughput of sequencing.

On the other hand, nanotechnology is the science of manipulating matter at the nanoscale, where materials show unique properties that extend their application in diverse fields such as medicine, electronics, etc. The technology has made and can make a significant impact in the field of life sciences, especially in genomics. Nanotechnology has opened up a lot of opportunities in the area of genomics and empowered researchers to approach DNA analysis even at the single-molecule level. The integration of NGS technologies with nanotechnologies i.e.; nanopores, nanomaterials, and nanostructures has led to remarkable breakthroughs. Nanoparticle-based library preparation, nanomaterial-assisted sample purification, and Nanopore sequencing are some of the few examples of nanotechnology-augmented capabilities of NGS.

This white paper initiates a comprehensive exploration of nanopore sequencing in Next-Generation Sequencing, It discusses the capabilities of nanopores in reading the individual DNA molecules in a fast and cost-effective way and also unravels the ability of nanoparticles in sequencing. Moreover, it will also go through the challenges and future directions offering a glimpse into the exciting future of genomics, diagnostics, and personalized medicine. empowered by nanopores in nanotechnology.

Nanopore sequencing

The concept of nanopore sequencing was started in the 1980s and significantly expanded since Oxford Nanopore Technologies (ONT) introduced a nanopore sequencer, MinION, in 2014. It was built on a nanoscale protein pore or nanopore, which acts as a biosensor and is placed in an electrically resistant polymer membrane. Within an electrolytic solution, a constant voltage is applied to generate an ionic current across the nanopore so that the negatively charged single-stranded DNA or RNA molecules propel through the nanopore from the 'cis' side (negatively charged) to the 'trans' side (positively charged). The translocation speed is regulated by a motor protein that will drive the nucleic acid molecule along the nanopore. Alterations in the ionic current in the course of translocation corresponding to a nucleotide sequence within the sensing region were decoded with the help of computational algorithms, enabling the real-time sequencing of DNA or RNA molecules.

.

Fig 1: MinION Device

Application of Nanopore Sequencing

One of the main applications of ONT sequencing is genome assembly for closing the gaps in the reference genome. For instance, ONT long reads are used to fill the 12 gaps in the human reference genome and to count the length of telomeric repeats 132. ONT reads have also been used to assemble the centromeric region of human Y chromosome 133. Other than the human reference genome, ONT long reads sequencing has been used to identify the repetitive regions in the Caenorhabditis elegans, in modeled organisms, and also in closely related species including 15 Drosophila species139, Escherichia coli 109, Arabidopsis thaliana138 and Saccharomyces cerevisiae137 precisely.

Moreover, it is capable of building a new reference genome of any non-model organisms. There are cases in which the first reference genome was built only using the ONT data alone 142. But in complicated cases, ONT long reads needed to be combined with other techniques such as 10x Genomics linked reads, Illumina short reads, optical mapping by Bionano Genomics, and spatial distance by Hi-C and PacBio long reads. Examples of such cases include the assembling of initial reference genomes of Eumeta variegata 149, Maniola jurtina 145, Panthera leo 148, Pavo cristatus 147, and Varanus komodoensis 146. During the SARS-CoV-2 pandemic 151, the full-length SARS-CoV-2 genome sequences are reconstructed through cDNA and direct RNA sequencing using ONT sequencing which provides valuable insight about the virus's biology, evolution, and pathogenicity.

A significant application of ONT long reads is for detecting Structural variants (SVs), especially from humans. Examples of such applications are the identification of SVs in breast cancer cell lines (HCC1187), individuals afflicted with acute myeloid leukemia 113, the construction of the initial haplotype-resolved structural variation profiles for two individuals with congenital abnormalities

112, and the discovery of 29,436 structural variations in an individual of Yoruban descent named NA19240 165.

The application of nanopore sequencing also includes the characterization of epigenetic marks. In early 2013, some reports revealed that methylated cytosines (5mC and 5hmC) can be differentiated from native cytosine by analyzing the distinct electrical current signals using the MspA nanopore technology 172, 173. Subsequent developments in bioinformatic tools identified three types of DNA modifications (5hmC, 5mC, and 6mA) from the ONT data.

Mapping DNA modifications is another application of ONT sequencing which with methyltransferase treatment results in the development of bioinformatics approaches such as MeSMLR-seq, which maps the nucleosome occupancy and chromatin accessibility even at single molecule level and at long scale 72 and other one is SMAC-seq, a similar approach, built by incorporating an additional exogenous modification to 6mA which will enhance the resolution of mapping nucleosome occupancy and chromatin accessibility.

A comprehensive evaluation of the feasibility of ONT cDNA sequencing, utilizing R7 and R9 nanopores, in transcriptome analyses has demonstrated its comparable performance to PacBio long reads in identifying gene isoforms, both of which outperform Illumina short reads 42. Even Though ONT data alone still faces challenges in accurately estimating gene/isoform abundance, detecting splice sites, and mapping alternative polyadenylation sites, recent enhancements in accuracy and throughput have made strides in improving these analyses. Furthermore, ONT cDNA sequencing has been applied to individual B cells from both mice and humans. Additionally, ONT direct RNA sequencing has been employed to measure the poly(A) tail length of native RNA molecules in humans, C. elegans, A. thaliana, and Locusta migratoria, revealing a negative correlation between poly(A) tail length and gene expression 67,168. Moreover, ONT

sequencing, in conjunction with rolling circle amplification, has facilitated the characterization of full-length isoforms of circular RNAs in humans 170,171.

ONT sequencing has also found application in various cancer types, including breast 33,176,193, colorectal 194, brain193, pancreatic 195, leukemia, and lung cancer 196, to identify multiple genomic variants of interest with improved sensitivity and time efficiency when compared to Sanger sequencing. Moreover, ONT whole-genome sequencing facilitates rapid detection of chromosomal translocation and the identification of breakpoints in an individual with acute myeloid leukemia 192. A combination of ONT sequencing and Cas9-assisted target enrichment successfully characterized a 200-kb region that extends over the breast cancer susceptibility gene BRCA1 and its flanking region 197. This study serves as a template for the analysis of complete variant profiles of disease-related genes in humans.

The application of ONT long reads also extends up to the characterization of genetic disorders. Examples include, the ONT sequencing of human genomes which provides an insight into the increased risk of Alzheimer's disease in association with the expansion of the tandem repeats in ABCA7 207. Moreover, it discovers a new 3.8-Mb duplication (in the intronic region of the F8 gene) in an individual affected with hemophilia 208. Likewise, detection of other disorders and abnormalities were also detected. MinION technology accelerated the identification of aneuploidy in prenatal and miscarriage samples, it reduced the detection time to 4 hours when compared to the conventional techniques (it takes 3-4 weeks).

The MinION is a powerful, portable, real-time DNA and RNA sequencing device built on nanopore sequencing technology. It is a rapid real-time sequencing device used for rapid pathogen identification such as diagnosis of pneumonia, bacterial lower respiratory tract infection, infective endocarditis, and infection in prosthetic joints. There was a real-time case in which MiniON was able to complete the 16s amplicon sequencing within 10 minutes highlighting its ability in the early detection of bacterial infections and the administration of antibiotics. Likewise, MinION was

faster and showed higher sensitivity during the clinical diagnosis of bacterial lower respiratory tract infection than the current culture-based 'gold standard methods. Besides its role in pathogen detection, it can also accelerate the profiling of antimicrobial/antibiotic resistance in bacteria and other microorganisms. In addition to it, ONT sequencing is useful for the detection of specific species and strains from microbiome samples 57,206.

The portable MinION device also enables on-site, real-time genomic surveillance of emerging infectious diseases, the characterization of its evolution rate, transmission rate, and response to treatment. During the Ebola outbreak in April 2015, the portable MinIOn devices were shipped to Guinea for real-time surveillance and it took only 15-30 minutes for sequencing (Per samples). Moreover, there were also other reports pointing out the advantages of MiniOn which includes genomic surveillance in less than 2h during the outbreak of salmonella in hospital and in the outbreak of Dengue virus, Yellow fever virus, and Zika virus in Brazil. As the ONT sequencing technology continues to improve its throughput capabilities, real-time genomic surveillance has been employed for larger genome pathogens across the years, ranging from a few kilobases to fungal pathogens with a genome range of greater than 10Mb in humans.

Other on-site applications

ONT devices have found application in on-site metagenomics research such as in the characterization of virulence genes, and pathogenic microbes and in the detection of antibiotic resistance markers in polluted river water. Its application also includes early and rapid detection of pests and viruses in farms. A new portable method called 'MinION sketching ' was initiated by forensic research to detect human DNA with a time of 3m of sequencing, which provides a faster and better solution to cell authentication in the course of cell or tissue culture. The MinION system is portable and consists of compact-sized MinION, mobile DNA extraction devices like

Bento Lab and VolTRAX, and a real-time onboard base calling with Guppy and other offline bioinformatics tools. It is very helpful in a situation like field research which is challenging to culture or preserve samples or where rapid genomic data acquisition is required.

Limitations and Future Direction

Even Though nanopore sequencing opens up many possibilities in biomedical research by enabling the real-time sequencing of single sequences to ultralong reads it has a lot of limitations in terms of error rates and the requirement of a high amount of sample. Overcoming these limitations will need further developments in bioinformatic software, nanopore technology, and molecular experiments. The major concern of nanopore sequencing is its error rate which is higher than that of Illumina sequencing. The error rates are tolerable in the case of sequencing large amounts of the same data because the user can detect multiple copies of the same sequence and allow them to eliminate the errors. But, it will be a big challenge in case of small or rare sequences. These error rates have been dramatically decreasing over time according to the changes in proteins involved in the sequencing process and it is anticipated to have further improvements in accuracy in the future. Another limitation includes the limited shelf life of nanopore membranes; they need to be replaced after specific periods. It will be overcome in the future by the introduction of solid-state nanopores which are in their developmental stage. It will use pores directly made from synthetic materials instead of protein channels. This could increase the durability of membranes, enabling the membranes to be preserved for significantly longer durations and will also accommodate a large number of nanopores into an array, which leads to a further reduction in cost.

Nanopore sequencing machines offer a major advantage which includes their compact size, usage of minimal reagent, and the need for little power in comparison with other sequencing technologies which require bulk machines. Nowadays nanopore sequencers can conveniently

fit in your pocket and can access it through a USB connection to a laptop. This advancement has empowered the scientific community to establish mobile laboratories and will allow the collection of organisms and DNA sequencing directly within the field. For instance, in 2016 the International Space Station integrated nanopore sequencing into its scientific capabilities. While the specific targets and locations for future sequencing endeavors remain uncertain, one thing is clear: the next frontier of sequencing is going to take us to uncharted territories.

CONCLUSION

The combination of nanotechnology and Next Generation Sequencing (NGS) brings a transformative era in genomic sequencing in which Oxford Nanopore Technologies put forward a wide range of applications in various scientific disciplines. It ranges from closing gaps in the reference genome to identifying structural variation and detection of epigenetic marks. It facilitates the profiling of antimicrobial resistance, rapid detection, and onsite surveillance of diseases during their outbreaks. The compatibility and portability of the device empowered the real-time genetic analysis, in field research even in remote areas. There are many challenges in the field of nanopore sequencing which include error rate and the requirement of more samples. Further improvement in the field of nanopores technology and bioinformatics opens a door for exciting scientific exploration and application that will benefit society as a whole.