Review Article

The Impact of Next-Generation Sequencing on Biotechnology: A Review of Current Applications

Abstract

Since sequencing was developed by Sanger and Gilbert in 1977, various sequencing methods have been introduced in biotechnology to sequence DNA. Next-generation sequencing is a novel high-tech method where the sequencing of thousands of DNA molecules can be done in parallel to generate enormous amounts of genetic data within one experiment. This novel generation of sequencing superseded conventional techniques like Sanger sequencing with parallelization to achieve rapid genome sequencing at minimum costs. In this chapter, we highlighted various applications of using NGS in different fields, such as genomic research, clinical diagnostics, metagenomics, water quality assessment, and food authenticity. Besides, NGS could identify genetic variations, diagnose diseases, and assess food quality with high accuracy, rapidity, and specificity.

Keywords: Next-generation sequencing (NGS), food authenticity, Cancer treatment, Antibiotic resistance

Introduction

Next-generation sequencing (NGS) is one of the most innovative technologies that science and technology have presented to us to sequence thousands of DNA molecules simultaneously. This innovative method is exceptionally beneficial for various applications since it produces enormous amounts of genetic data in a single experiment. In 1977, Frederick Sanger and Walter Gilbert pioneered accelerated sequencing techniques designed explicitly for lengthy DNA strands. This groundbreaking advancement marked the inaugural capability to decipher the complete nucleotide sequence of entire genes, ranging from 1,000 to 30,000 bases in length.

The Sanger sequencing method is the first generation sequencing platform for sequencing DNA and has been replaced by more sophisticated next-generation sequencing (NGS) methods. NGS is the recommended option for large-scale sequencing projects because it provides higher throughput, faster results, and lower costs, even though Sanger sequencing

established the foundation for genomic sequencing (Qin, 2019). However Sanger sequencing plays a major role in samall scale projects inorder to confirm the varientsidentified by NGS

Consequently, these newer technologies facilitated rapid and precise genome sequencing. With high-throughput sequencing, this approach integrates various contemporary sequencing methods and has developed into an invaluable resource for investigating genomes and epigenomics. This method allows researchers to identify nucleotides inside specific DNA or RNA regions and across the genome.In this era of extensive sequencing data, an NGS plot is a valuable tool for connecting genomic information with large datasets. In clinical settings, the primary application of NGS will likely involve resequencing genomic DNA. Whole genome sequencing (WGS) offers a comprehensive view of single nucleotide variations (SNVs), insertions and deletions (indels), complex structural changes, and variations in copy number, all in one analysis. While WGS reveals a wealth of genetic information that can be analyzed and applied, some may contain data of uncertain clinical significance or entirely new insights. WGS is the most thorough genomic examination (Pleasance *etal*., 2010). Data privacy has significant ethical issues, particularly since genetic information is sensitive and should be protected from unauthorized access. Additionally, unexpected genetic variants can impact patient health, leading to ethical dilemmas about whether to inform patients of these findings. Balancing the potential benefits of sharing this information with the risks and psychological effects on patients is critical.

As the need for more affordable sequencing techniques grew, high-throughput sequencing emerged, generating dozens or even millions of sequences at once. Beyond what is possible with conventional dye-terminator procedures, these technologies promise to reduce the cost of DNA sequencing significantly. Consequently, they have revolutionized the fields of genomics and molecular biology research by making it possible to sequence DNA and RNA more quickly and affordably than with previous techniques like Sanger sequencing. In molecular diagnostics, NGS methods are gradually replacing the Sanger approach. However, they have a similar ancestor in that they use DNA replication mechanisms that have evolved over millions of years. Among other techniques, these advancements include ion semiconductor sequencing (It is a next-generation sequencing (NGS) technology used for DNA sequencing. It works on the principle of detecting changes in pH when nucleotides are added to the grown DNA during the sequencing process), sequencing-by-synthesis (In this approach, the strand which is going to be sequenced is replicated by DNA polymerase along

with the fluorescent levelled dNTPs and thus detect the types of incorporated nucleotide base when the replication expand).

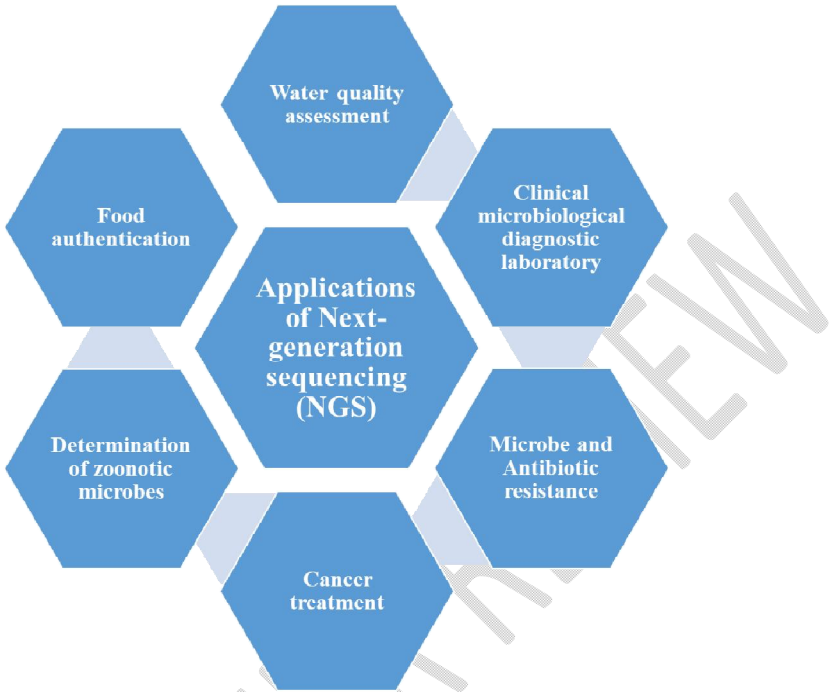

Fig.1. Next-generation sequencing (NGS) applications-an overview

Application ofNext-generation sequencing(NGS) across different Fields

1. Water quality assessment

The use of NGS-based techniques is growing in water quality and assessment research. Researchers are using NGS to learn more about its applicability and biases as an analytical tool. They evaluate biological risks and debate the benefits and drawbacks of using direct quantitative analysis to interpret NGS data. Research has looked at the breakdown processes of dangerous chemicals endangering water quality and the existence of toxin and antibiotic-resistance genes in water samples. NGS approaches have also generated bioindicators for sewage contamination and microbiological source tracing.

While Next-Generation Sequencing (NGS) holds promise for tracking the deterioration of water quality by identifying signs of human feces or sewage contamination, a

more straightforward method of assessing biological threats in water involves looking directly at the diversity and prevalence of pathogens. On the other hand, tracking each waterborne pathogen separately is expensive, necessitates technical know-how, and depends on determining which pathogens to target (Varela & Manaia, 2013). Potential danger sources can be identified by analyzing sequences that closely resemble those of aquatic pathogens using Next-Generation Sequencing (NGS). Furthermore, this approach might offer a ballpark estimate of their relative abundance in wastewater and comparable environments by evaluating variations in sample preparation (McLellan *et al.*, 2010; Cai *et al.*, 2014; Lu *et al.*, 2015). Earlier, researchers used the V6 region of the SSU rRNA gene, but due to its short sequence length (around 60 base pairs), it could not identify individual pathogen species. Differentiating between taxonomic levels beyond the family or genus was hampered by the short length of the V6 region. However, it is now possible to correctly categorize bacterial genomes at the species level thanks to advancements in Next-Generation Sequencing (NGS) technology, which allows for more extended sequence read lengths.

2. NGS for the standard clinical microbiological diagnostic laboratory

Identifying, characterizing, and detecting pathogens is crucial in creating treatment strategies that effectively address infectious diseases. A wide range of methods—both dependent and independent of culture are used in regular clinical microbiological diagnostic laboratories to investigate the causes of microbial infections. However, some organisms may not survive in laboratory settings due to the unique growth conditions required by many of them, making identification by culture-dependent approaches impractical (Lagier *et al.*, 2015). In addition, most methods that do not depend on cultivating microorganisms (like PCR) require prior knowledge of the microorganisms thought to be present in a particular clinical sample under investigation (Kulski, 2016). This requirement is essential for their identification. Therefore, these culture-independent methods may leave unexpected bacteria unidentified (Yang & Rothman, 2004). These variables have created a need for novel diagnostic assays independent of culture-based techniques. Enhancing the identification of agents responsible for infectious diseases is crucial. This development could result in better patient outcomes, more effective management of the use of antibiotics, improved disease outbreak identification and surveillance, the ability to identify viable but challenging-to-cultivate microorganisms (like VBNC), and the ability to research pathogens that have not yet been identified. Cliendo et al. (2013). These discoveries will ultimately lead to a better understanding of the pathogenesis of diseases, opening the door to a new era of molecular

pathology and tailored medicine. Harris and McCormick (2010). Next-generation sequencing (NGS) is a frequently used culture-independent technique in microbiology research. However, it is not routinely used in clinical microbiological diagnoses. This is completed in a single sequencing step and does not require conventional culture methods. Goodwin et al. (2016). However, the interpretation of NGS data is difficult to analyse and can make variations among different disease variants, like benign and non-benign tumours. Different techniques of NGS are used, such as ion semiconductor sequencing, Roche sequencing, Nanopore Sequencing, Single-Molecule Real-Time (SMRT), etc. It can lead to variability in results and make interpretation very difficult and time-consuming. However, tests using the NGS platform require proper permission and approval from authorized bodies like the FDA (Food and Drug Administration) or the European Medicines Agency (EMA).

Obligate anaerobes are known to cause serious infections, but they may be challenging to identify using standard specimen collection and detection methods in regular clinical microbiology labs. This is because particular safety measures must be taken to maintain an anaerobic environment when collecting and transporting specimens. Furthermore, culture-based detection techniques can require the laboratory to supply growth factors that still need to be identified (Brook I, 2002). Techniques such as Next-Generation Sequencing (NGS) hold considerable promise for improving the identification and diagnosis of anaerobic infections and illnesses brought on by hard-to-culture bacteria like fastidious or viable but non-culturable (VBNC) germs. NGS techniques are quite helpful in detecting VBNC bacteria, especially those impacted by antibiotics during antimicrobial therapy in patients. Pasquaroli et al. (2013). A thorough understanding of the microbiota in clinical samples is also crucial since it allows one to consider the entire microbial population when choosing therapeutic approaches. Hajishengallis et al. (2012).

3. Cancer treatment

Technological developments have made it easier to incorporate NGS-based assays into clinical practice, where they are being used more and more for carrier screening, fetal aneuploidy testing, finding rare illnesses, and determining the presence or risk of cancer. To sequence long and complex genes or several genes per tumor sample and identify the driver and targetable changes, NGS has now moved into clinical settings. Recent studies have

demonstrated that Next-Generation Sequencing (NGS) exhibits vital analytical accuracy when detecting the most common changes in clones. (Frampton *et al.*, 2013).

As cancer is primarily a genetic disease caused by genetic changes that can be inherited or acquired, the development of novel DNA sequencing techniques will have a significant influence on the diagnosis, prognosis, and course of therapy of the illness. NGS makes it possible to thoroughly profile the genomic landscape of several cancer genomes across different disease types. It is expected that all people's genetic data, including inherited characteristics and mutations linked to cancer, will soon be regularly sequenced. Furthermore, it is possible to monitor the progression of the cancer genome through numerous sequencing iterations, which will help with accurate diagnosis and focused molecular therapy implementation.

4. Microbe and Antibiotic resistance study by using NGS approach

The microbiome affects the host's immunity, affecting our body's immune system. Some bacteria, viruses, yeast, etc microbes can not be cultured in the laboratory. They are mainly found in the various parts of our body like the buccal cavity, nose, vagina, stomach and hair. The National Institute of Health started a human microbiome project in 2008 called the Human Microbiome Project (HMP) (Gupta and Verma 2009). Using NGS technology, they isolate and characterise microbes present in healthy and diseased persons (Qin *et al*., 2010).

According to various studies, the growth of Antimicrobial resistance pathogens is of great concern. Antibiotic resistance determinants (ARDs), which include antibiotic resistance genes (ARGs) and mobile genetic elements (MGEs) responsible for the transfer of these genes, are prevalent in environmental bacteria (Li *et al*., 2005). Microbial communities found in sediments near wastewater treatment plant effluents and in soil, water, sediments, and faecal samples from humans and animals contain ARDs. They are of great concern to prevent the spread and management of antibiotic resistance (Port *et al.*, 2014). Gaining a deeper understanding of the prevalence and characteristics of environmental reservoirs of ARDs will be crucial in addressing this issue (Bush et al., 2014). Total DNA is extracted from environmental samples, and sequencing is conducted randomly. The resulting sequence reads are then compared against a reference database containing known antibiotic resistance gene (ARG) sequences to assess the resistance potential derived from the metagenome (Allen *et al*., 2009).

5. Determination of zoonotic microbes transmission by using NGS

Zoonotic diseases are a significant threat to public health. This concern has only recently gained recognition despite the increasing transmission of zoonotic microbial agents (infections spread from animals to humans). According to studies, it was found that out of four pathogens, three pathogens originate from zoonotic sources (Taylor *et al*., 2001). So, It's a global concern and can lead to human disaster, including Ebola virus infections, bird flu (highly pathogenic avian influenza), and severe acute respiratory syndrome (SARS), among others (Heymann and Dar 2014). These are emerging public health issues in our healthcare systems.

6. Food testing authentication using NGS

Regarding guaranteeing food authenticity, regulatory bodies, food producers, and processors have the utmost regard for consumer safety and confidence. It has been shown that false food labeling and other dishonest business tactics undermine consumer confidence and, in some instances, endanger consumer safety. (Barnett, *et al.*, 2016). So, food product authenticity is a crucial issue for consumers, authorities, farmers, and processors because dishonest business practices can have a negative impact on consumer safety and confidence as well as the business models of reliable organizations. DNA sequencing is spreading quickly, which has caused databases that contain sequences to grow quickly in response. In terms of accuracy, creating current, trustworthy databases might be advantageous for the successful application of NGS. Strong quality control standards and proficiency testing techniques, both of which are currently being developed, will be required for these uses. NGS can be used to check the authencity of sea food in order to prevent sea fraud and microbial communities in food products.

Several methods, including spectroscopy, chromatography, proteomics, and Polymerase Chain Reaction (PCR) techniques, have been employed to detect food fraud in cases of food adulteration. (Primrose *et al*., 2010). However, since fraud often involves replacing one ingredient with one of a different race or species and DNA testing is convenient and reliable across various food processing methods, DNA-based methods are the most accurate and sensitive way to address this problem. (Catalano *et al.*, 2016, Pardo, 2015). Standardized gene snippets are used in techniques such as DNA Barcoding (Hebert *et al*., 2003) and Forensically Informative Nucleotide Sequencing (FINS) (Bartlett & Davidson, 1992) to differentiate between various products within taxonomic levels. These methods

depend on specific DNA sequences that vary among species, making precise identification and differentiation possible. In addition, various techniques have been created to find and amplify particular oligonucleotide primer products. These techniques include digital PCR, real-time PCR, oligonucleotide ligation assay (OLA), and LAMP. These techniques are used to detect adulteration in food goods. However, instead of offering a thorough breakdown of the food's composition, these approaches usually concentrate on specific ingredients, which can worry customers. On the contrary, next-generation sequencing (NGS) provides several capabilities to overcome these drawbacks. It can simultaneously screen out the different genomic regions of multiple genes, which made the platform to identify plants, animals, and food ingredients derived from fungi and microalgae.

Conclusion

High throughput sequencing has sped up research in several biological and applied science domains. NGS is a sophisticated, realistic, and more accurate platform that can produce massive amounts of data. It becomes a preferable toolfor discovering SNPs and other genome mutations. If costs decrease and turnaround times increase, it might become a standard diagnostic tool for clinical routine. The concept of cancer treatment is derived from the detailed characterization of the numerous genetic changes discovered in tumours. The ability to identify every gene linked to a disease inside a genome opens the door to auto-immune disease detection and diagnosis. Work on food testing and authentication, water quality assessment, disease diagnosis, standard clinical, microbiological diagnostic laboratory, etc., has already begun through NGS, which has positive results and indicates that NGS can be used for multiple purposes in multi-directional ways. This invaluable invention will have a greater scope in the advanced biotechnology field.

Disclaimer (Artificial intelligence)

We declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

References

Allen, H.K., Moe, L.A., Rodbumrer, J., Gaarder, A. and Handelsman, J., 2009. Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. *The ISME journal*, *3*(2), pp.243-251.

Barnett, J., Begen, F., Howes, S., Regan, A., McConnon, A., Marcu, A., ... & Verbeke, W. (2016). Consumers' confidence, reflections and response strategies following the horsemeat incident. *Food Control*, *59*, 721-730.

Brook I (2002) Clinical review: bacteremia caused by anaerobic bacteria in children. Crit Care 6:205–211.

Kulski, J. K. (2016). Next-generation sequencing—an overview of the history, tools, and "Omic" applications. *Next generation sequencing-advances, applications and challenges*, *10*, 61964.

Bush, K., Courvalin, P., Dantas, G., Davies, J., Eisenstein, B., Huovinen, P., Jacoby, G.A., Kishony, R., Kreiswirth, B.N., Kutter, E. and Lerner, S.A., 2011. Tackling antibiotic resistance. *Nature Reviews Microbiology*, *9*(12), pp.894-896.

Cai, L., Ju, F., and Zhang, T. (2014). Tracking human sewage microbiome in a municipal wastewater treatment plant. Appl. Microbiol. Biotechnol. 98, 3317–3326. doi: 10.1007/s00253-013-5402-z

Cliendo AM, Gilbert DN, Ginocchio CC, Hanson KE, May L, Quinn TC et al (2013) Better tests, better care: improved diagnostics for infectious diseases. Clin Infect Dis 57(Suppl 3): S139–S170.

Frampton, G. M., Fichtenholtz, A., Otto, G. A., Wang, K., Downing, S. R., He, J., ... &Yelensky, R. (2013). Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nature biotechnology, 31(11), 1023-1031

Gupta, N. and Verma, V.K., 2019. Next-generation sequencing and its application: empowering in public health beyond reality. *Microbial Technology for the Welfare of society*, pp.313-341.

Hajishengallis G, Darveau RP, Curtis MA (2012) The keystone-pathogen hypothesis. Nat Rev Microbiol 10:717–725.

Harris, T. J., & McCormick, F. (2010). The molecular pathology of cancer. *Nature reviews Clinical oncology*, *7*(5), 251-265.

Heymann, D.L. and Dar, O.A., 2014. Prevention is better than cure for emerging infectious diseases. *Bmj*, *348*.

Lagier JC, Hugon P, Khelaifia S, Fournies PE, La Scola B, Raoult D (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 28:237–264.

Li B., Yang Y., Ma L., Ju F., Guo F., Tiedje J. M., et al. (2015). Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J.* 10.1038/ismej.2015.59

Lu, X., Zhang, X.-X., Wang, Z., Huang, K., Wang, Y., Liang, W., et al. (2015). Bacterial pathogens and community composition in advanced sewage treatment systems revealed by metagenomics analysis based on high-throughput sequencing. PLoS ONE 10:e0125549. doi: 10.1371/journal.pone.0125549

McLellan, S. L., Huse, S. M., Mueller-Spitz, S. R., Andreishcheva, E. N., and Sogin, M. L. (2010). Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. Environ. Microbiol. 12, 378–392. doi: 10.1111/j.1462-2920.2009.02075.x

Pasquaroli S, Zandri G, Vignaroli C, Vuotto C, Donelli G, Biavasco F (2013) Antibiotic pressure can induce the viable but non-culturable state in Staphylococcus aureus growing in biofilms. J AntimicrobChemother 68:1812–1817.

Pleasance, E. D., Cheetham, R. K., Stephens, P. J., McBride, D. J., Humphray, S. J., Greenman, C. D., ... & Stratton, M. R. (2010). A comprehensive catalogue of somatic mutations from a human cancer genome. Nature, 463(7278), 191-196.

Port J. A., Cullen A. C., Wallace J. C., Smith M. N., Faustman E. M. (2014). Metagenomic frameworks for monitoring antibiotic resistance in aquatic environments. *Environ. Health Perspect.* 122 222–228. 10.1289/ehp.1307009

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T. and Mende, D.R., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *nature*, *464*(7285), pp.59-65.

Qin, D. (2019). Next-generation sequencing and its clinical application. *Cancer biology & medicine*, *16*(1), 4.

Taylor LH, Latham SM, Woolhouse ME. Risk factors for human disease emergence. *Philos Trans R Soc Lond Ser B Biol Sci.* 2001;356:983–989.

Varela, A. R., and Manaia, C. M. (2013). Human health implications of clinically relevant bacteria in wastewater habitats. Environ. Sci. Pollut. Res. Int. 20, 3550–3569. doi: 10.1007/s11356-013-1594-0.

Yang S, Rothman RE (2004) PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. Lancet Infect Dis 4:337–348.