

DOI: 10.31073/onehealthjournal2024-I-05

APPLICATION OF THE NEXT GENERATION SEQUENCING
IN BIOLOGY AND MEDICINE

Gerilovych A. P.^{1,2} (ORCID ID 0000-0002-3280-4172), **Sushko M. I.**¹ (ORCID ID 0000-0003-2723-9903), **Mandyhra S. S.**¹ (ORCID ID 0000-0002-3559-9258), **Rodyna N. S.**³ (ORCID ID 0000-0002-4338-8872), **Chechet O. M.**¹ (ORCID ID 0000-0001-5099-5577), **Romanko M. Ye.**¹ (ORCID ID 0000-0003-0285-5603), **Kuchinskiy M. V.**¹ (ORCID ID 0009-0003-2652-3360) and **Gerilovych I. O.**⁴ (ORCID ID 0000-0002-6742-2086)

1 – State Scientific and Research Institute for Laboratory Diagnostics and Veterinary and Sanitary Expertise, Kyiv, Ukraine, antger2011@gmail.com;

2 – One Health Institute, NGO, Kharkiv, Ukraine;

3 – SI “Kyiv Center for Diseases Control and Prevention” of MoH of Ukraine;

4 – National Scientific Center “Institute for Experimental and Clinical Veterinary Medicine”, Kharkiv, Ukraine

Abstract. *Next-Generation Sequencing (NGS), also known as high-throughput sequencing, refers to a set of modern DNA sequencing technologies that have revolutionized the field of genomics. Advantages of NGS techniques involving high speed (parallel sequencing is faster than traditional methods, allowing researchers to obtain results more quickly), cost-effectiveness (ability to sequence multiple fragments simultaneously reduces the cost per base compared to traditional sequencing), and scalability (platforms can be scaled to accommodate varying levels of throughput depending on experimental needs). NGS has significantly accelerated genomics research, enabling breakthroughs in fields such as personalized medicine, cancer genomics, and evolutionary biology. However, challenges such as data analysis complexity, error rates, and cost still exist and are areas of ongoing research and improvement within the field of sequencing technologies. Paper contains the brief explanation of the current NGS platforms and their features. NGS biomedical application is described with the main advantages and abilities of the analysed tools.*

Key words: DNA, RNA, genomics, metagenomes, Next-Generation Sequencing, diagnostics, biology, medicine,

Next-Generation Sequencing (NGS), also known as high-throughput sequencing, is th modern approach, that refers to a set of novel DNA and RNA sequencing technologies that support studies of genomes of various organisms at the genomics' level. Unlike traditional Sanger sequencing, which involves sequencing a single DNA fragment at a time, NGS allows for the simultaneous sequencing of millions to billions of DNA fragments, and to get a significant yeld of the genomic information (Levy, 2019). This parallel sequencing capability, coupled with high-throughput data generation, has significantly increased the speed and efficiency of DNA and RNA sequencing for biomedical purposes (McCombie, 2019, Kumar, 2019).

First popular sequencing method, known, as Sanger's dideoxynucleotide sequencing method was introduced in 1977. Also other enzymatic sequencing methods were developed and published at the same time, including Barnes' (1978) partial ribosubstitution sequencing, Sanger's plus and minus method (Sanger and Coulson 1975), and the chemical cleavage method (Maxam and Gilbert, 1977; McCombie, 2019).

Iniintially used radioactively labeled P- and S- nucleotides and film-based detection were replaced by Hood and his colleagues with four different fluorescent dyes that allow detection

during electrophoretic separation by the laser-induced fluorescent emission of each fragment on a special sequencing instrument (Smith et al., 1986).

Early sequencing projects based on Sanger's method were focused on determining sequences for single genes or very small genomes (~5000 bp at most). The development of a computational package by Rodger Staden, created in Medical Research Council (MRC) in collaboration with Sanger's laboratory, provided possibility to increase the length of sequenced DNA or genomes (Staden et al. 2000). The Staden package allowed to perform the small DNA-fragments to be randomly cut from a larger original DNA source, sequenced randomly, and computationally overlaid to reform the entire sequence of the original, larger input source. This package was created for the Unix OS and was widely distributed. It gave an ability to assemble sequences from input data and a viewer for visualisation of overlaps between fragments. It also displayed a six-frame translation of the identified sequences, after which the potential of open reading frames could be determined. The combination of Sanger sequencing and the Staden sequence assembly package made it possible to sequence genomes about 10 times larger than before.

However, the effort to sequence and annotate whole genomes by Sanger sequencing was still a significant and expensive. This scenario began to change shortly thereafter, with the introduction of the first massively parallel DNA sequencing technology in 2005, that produced NGS approach in the next stage (Margulies et al. 2005).

The main difference of NGS from Sanger sequencing, is that NGS instruments perform both the enzymology and data acquisition, sequence data generation from thousands to billions of templates simultaneously, that improves the productivity and reliability of the received results under NGS approach application (McCombie, 2019).

The NGS technique allows to perform the Parallel Sequencing. Its platforms conduct sequencing reactions on many fragments in parallel, allowing for the simultaneous analysis of multiple DNA sequences. This parallelization contributes to the rapid and cost-effective nature of NGS. Parallel sequencing, also known as massively parallel sequencing or high-throughput sequencing, is a DNA sequencing approach that allows the simultaneous sequencing of multiple DNA fragments. This method contrasts with traditional Sanger sequencing, which sequences one DNA fragment at a time. Parallel sequencing technologies have revolutionized genomics by significantly increasing the speed, throughput, and cost-effectiveness of DNA sequencing (Yang, 2023).

This allows to achieve the higher level of throughput. Traditional Sequencing is limited by the capacity to sequence one fragment at a time, Sanger sequencing has lower throughput. Massively the parallel sequencing technologies can generate millions to billions of sequence reads in a single run, greatly increasing throughput.

Parallel sequencing generates massive data, enabling the sequencing of entire genomes, transcriptomes, or targeted regions comprehensively. Classical approach performs only small data amount generation.

Current most popular NGS-based platforms include:

- Illumina Sequencing, that utilizes reversible terminator chemistry, where fluorescently labeled nucleotides are added in parallel across millions of clusters on a flow cell.
- Ion Torrent Sequencing, that relies on semiconductor sequencing, with each incorporated nucleotide releasing a hydrogen ion that is detected in parallel.
- Nanopore Sequencing, that uses nanopores to sequence DNA or RNA strands as they pass through, enabling real-time sequencing of individual molecules.

Parallel sequencing is applied for whole genome sequencing (enables the rapid sequencing of entire genomes), exome sequencing (targets the protein-coding regions of the genome for efficient analysis), RNA sequencing (RNA-Seq, captures the transcriptome to study gene expression), ChIP-Seq (identifies protein-DNA interactions to study chromatin binding), and metagenomics analysis (analyses DNA from complex microbial communities in diverse environments).

Advantages of this approach involving high speed (parallel sequencing is faster than traditional methods, allowing researchers to obtain results more quickly), cost-effectiveness (ability to sequence multiple fragments simultaneously reduces the cost per base compared to traditional sequencing), and scalability (platforms can be scaled to accommodate varying levels of throughput depending on experimental needs).

Ongoing advancements in parallel sequencing technologies continually enhance speed, accuracy, and read length. Also the Third-Generation Sequencing trend is developing. It introduces longer read lengths and real-time sequencing capabilities.

Data Analysis process for NGS results are quite challengeable. It requires complicity of the bioinformatics approach. Managing and analysing the significant and huge amount of data generated by parallel sequencing requires sophisticated bioinformatics tools. Massive datasets also require necessitate efficient storage solutions and computational resources for their management and analysis (Budolwe, 2017).

Parallel sequencing has become the standard in genomics research and has facilitated breakthroughs in personalized medicine, functional genomics, and our understanding of genetic variation and diversity. The ability to sequence DNA in parallel has significantly advanced the field, making large-scale genomic projects and comprehensive analyses feasible (Roman, 2022).

Most NGS platforms produce relatively based on short DNA sequence reads, typically ranging from a few decades to a few hundred base pairs. Advances in technology have allowed for longer read lengths in some newer platforms. These fragments are managing by the bioinformatical software, providing full-scale data for further analysis using phylogenetical tools (DeKnijff, 2019).

NGS demonstrates the diversity of possible applications. Its versatility makes it applicable to a wide range of research areas.

Multiple platforms of NGS are available, each with its own technology and workflow. Illumina, Ion Torrent (Life Technologies), 454 (Roche), and Pacific Biosciences (PacBio) are examples of NGS platforms, each with unique sequencing chemistries and readout technologies.

Illumina is a leading biotechnology company that has pioneered Next-Generation Sequencing technologies. Illumina sequencing platforms have become widely adopted in genomics research, clinical diagnostics, and various applications due to their high-throughput, accuracy, and versatility.

Illumina Sequencing Platforms involve:

1. Illumina HiSeq Series is presented by HiSeq 2500, HiSeq 3000, HiSeq 4000 machines. These are the high-throughput platforms suitable for a wide range of applications, including whole-genome sequencing, exome sequencing, RNA sequencing (RNA-Seq), and more. HiSeq X series devices are specifically designed for large-scale human whole-genome sequencing projects.

2. Illumina NextSeq Series presented by NextSeq 500, and NextSeq 550, the mid-throughput platforms offering flexibility and a shorter turnaround time. Suitable for various applications, including targeted sequencing and small to mid-sized projects.

3. Illumina MiSeq series presented by MiSeq, and MiSeqDx. These benchtop sequencers designed for smaller-scale projects, targeted sequencing, and amplicon sequencing. They are often used for clinical applications and microbiome studies.

4. Illumina NovaSeq Series involves NovaSeq 5000, and NovaSeq 6000 DNA-analysers. These are the high-throughput platforms designed for scalability, cost-effectiveness, and the ability to handle large-scale projects. They are suitable for whole-genome sequencing and large-scale genomics initiatives (Yin, 2021; Pervez 2022).

Illumina approached sequencing is based on SBS chemistry, where fluorescently labeled nucleotides are added one at a time to a growing DNA strand. Each nucleotide incorporation is captured in parallel across millions of clusters on a flow cell.

It is known for its high accuracy, with low error rates in base calling. The platform is suitable for a wide range of applications, including whole-genome sequencing, exome sequencing, RNA-Seq, ChIP-Seq, and targeted sequencing.

Illumina platforms offer cost-effective sequencing, making large-scale genomic projects and population-scale studies feasible. While Illumina platforms typically produce shorter read lengths compared to some other technologies, advances in chemistry and technology have improved the accuracy and efficiency of short reads.

It's important to note that the field of NGS evolves rapidly, and Illumina may have introduced new platforms or technologies since my last update. Researchers and clinicians often choose Illumina platforms based on their specific project requirements, such as throughput, read length, and cost considerations (Pereira, 2020).

Another platform, Ion Torrent is a sequencing technology developed by Ion Torrent, a division of Thermo Fisher Scientific (formerly Life Technologies). The Ion Torrent sequencing approach is based on semiconductor sequencing, which detects changes in pH caused by the release of protons during DNA polymerization.

There are several Ion Torrent Sequencing based platforms, including Ion Personal Genome Machine (the Ion PGM is a benchtop sequencer designed for small to mid-sized projects. It is suitable for applications such as targeted sequencing, amplicon sequencing, and small genome sequencing), Ion Proton (the platform is designed for higher throughput compared to the Ion PGM. It is suitable for a range of applications, including exome sequencing, transcriptome sequencing, and targeted sequencing), Ion S5 and Ion S5 XL (part of the Ion S5 Series, these platforms offer scalability and higher throughput, making them suitable for projects with increased sequencing demands) (Hu, 2021).

Ion Torrent sequencing is based on the detection of hydrogen ions (protons) released during the natural DNA polymerization process. As each nucleotide is added to the growing DNA strand, a proton is released, and the change in pH is detected by a semiconductor sensor.

The Ion Torrent platforms have multiple advantages in use.

These platforms offer relatively rapid run times, making them suitable for applications that require quick results. Ion Torrent platforms are scalable, allowing researchers to choose the platform that best fits the scale of their sequencing project. This approach is often considered cost-effective, making it suitable for labs with varying budget constraints (Ringlander, 2022).

The Ion Torrent workflow is known for its simplicity, involving fewer steps compared to some other sequencing technologies.

It is also particularly suitable for targeted sequencing applications, including panels designed for specific genes or genomic regions. Ion Torrent allows amplification-free library preparation in certain applications, reducing the risk of bias introduced during PCR amplification.

Ion Torrent platforms are used for a variety of applications, including targeted sequencing, exome sequencing, transcriptome sequencing, and microbial genomics. Ion AmpliSeq technology is often used in conjunction with Ion Torrent platforms. It involves a highly multiplexed PCR approach for targeted amplification of specific genomic regions.

This DNA-analysis approach is utilized in clinical research and diagnostics, including cancer genomics for detecting somatic mutations (Chen, 2020).

It's important to note that the choice of a sequencing platform depends on various factors, including the scale of the project, required throughput, read length, and specific applications.

Here are several other approaches used in the practice.

The next NGS platform is Pacific Biosciences (PacBio) Sequencers, presented by PacBio Sequel II, and Sequel I. These tools use single-molecule, real-time (SMRT) sequencing technology. They are known for producing longer read lengths, making them valuable for applications requiring *de novo* sequencing and resolving complex genomic regions (Rhoads, 2015).

The cheapest and most perspective, and extremely high scale developing NGS platform is based on the Oxford Nanopore Sequencing (Wang, 2021).

Oxford Nanopore Technologies is a biotechnology company that has developed a distinctive sequencing technology known as nanopore sequencing. This technology is implemented in their series of Next-Generation Sequencing (NGS) platforms. Oxford Nanopore sequencing allows the direct and real-time analysis of single DNA or RNA molecules as they pass through nanopores, providing advantages such as long read lengths and the ability to sequence native DNA without amplification.

Oxford Nanopore Sequencing Platforms include: MinION (the MinION is a portable sequencer that connects to a computer via USB. It is designed for smaller-scale projects, fieldwork, and rapid sequencing), GridION X5 (a benchtop sequencer that can process up to five MinION flow cells simultaneously, that offers higher throughput compared to the MinION), PromethION (a high-throughput sequencer designed for large-scale genomic projects. It features multiple flow cells and offers scalability).

Nanopore sequencing involves passing a single-stranded DNA or RNA molecule through a nanopore embedded in a membrane. Changes in electrical current as the molecule translocates through the nanopore are used to identify the sequence of nucleotides.

One of the primary advantages of Oxford Nanopore sequencing is the potential for long read lengths. This is especially beneficial for resolving complex genomic regions and studying structural variations. Sequencing occurs in real-time, allowing for immediate data analysis during the sequencing run. This feature is valuable for dynamic experiments and real-time monitoring. Oxford Nanopore sequencing allows for the direct sequencing of native DNA or RNA without the need for amplification. This can reduce bias and artifacts introduced during amplification. The technology supports a wide range of applications, including whole-genome sequencing, RNA sequencing, metagenomics, epigenetics, and more.

The MinION is a portable device that can be used for on-the-go sequencing or in settings where a compact instrument is preferred. The GridION X5 and PromethION offer higher throughput and are suitable for larger-scale projects.

Oxford Nanopore has embraced a community-driven model, encouraging user feedback and collaborations. This has led to an active user community and the development of various analysis tools (Gilpatrick, 2020).

The real-time nature of Oxford Nanopore sequencing allows for relatively rapid turnaround times, enabling quick access to sequencing results. Sequencing is performed at the level of individual molecules, providing a more direct and accurate representation of the genetic material.

Oxford Nanopore sequencing can detect base modifications, providing information about epigenetic modifications in DNA or RNA (Sheka, 2021).

Other approach is presented by Roche 454 Sequencers: GS FLX+ and GS Junior. These platforms use pyrosequencing technology. While not as widely used as some other platforms, they have been valuable for applications requiring longer read lengths (Harrington, 2013; Wieckhusen, 2015).

BGISEQ Sequencers (BGI) model line is presented by BGISEQ-500, and BGISEQ-50. BGISEQ platforms are developed by BGI and use DNA nanoball technology. They are known for their cost-effectiveness and have been used in large-scale genomics projects.

Complete Genomics platforms, presented by MGISEQ-T7 and MGISEQ-2000, now owned by BGI, utilize combinatorial probe-anchor synthesis (cPAS) technology. They are designed for high-throughput whole-genome sequencing (McCarthy, 2013).

The key fields for application of NGS-technologies in biological and ecological practice include:

- Whole Genome Sequencing (WGS) – the approach is commonly used for sequencing entire genomes, providing a comprehensive view of an organism's genetic makeup.

- RNA Sequencing (RNA-Seq) – can be applied to study gene expression profiles by sequencing RNA molecules. This is valuable for understanding transcriptional activity and identifying differentially expressed genes.

- Targeted sequencing – allows researchers to selectively sequence specific regions of interest, which is useful for targeted genetic studies or clinical applications.

- Metagenomics – study microbial communities in environmental samples, human microbiomes, or other complex ecosystems (Zhong, 2022).

Managing and interpreting the massive datasets generated by NGS requires sophisticated bioinformatics tools. This includes aligning sequences to a reference genome, variant calling, and functional annotation.

Third-Generation Sequencing and Single-Cell Sequencing are the novel trends of NGS biomedical application. Technologies like PacBio and Oxford Nanopore represent third-generation sequencing, offering longer read lengths and the ability to sequence longer DNA fragments. NGS can be applied at the single-cell level, allowing the study of individual cells' genomic or transcriptomic profiles (Petersen, 2019; Zhu, 2020; Searle, 2023).

NGS has significantly accelerated genomics research, enabling breakthroughs in fields such as personalized medicine, cancer genomics, and evolutionary biology. However, challenges such as data analysis complexity, error rates, and cost still exist and are areas of ongoing research and improvement within the field of sequencing technologies.

The key fields of NGS application in biomedical practice are:

- a) WGS, that is valuable for identifying micro- and macroorganism origin and species belonging, determine the genetic variations, understanding genomic diversity, and evolution study (Zhao, 2019).

- b) Exome Sequencing: sequencing of the exome, the protein-coding regions of the genome, is cost-effective and focuses on regions most likely to harbor disease-causing mutations. It is widely used in identifying rare genetic disorders (Jelin, 2018).

- c) *De novo* Sequencing: NGS facilitates the assembly of genomes without the need for a reference sequence, enabling the study of non-model organisms and uncovering novel genetic information (Lu, 2016).

- d) Transcriptomics: NGS is used to analyze the transcriptome, providing insights into gene expression levels, alternative splicing, and novel transcripts. This is crucial for understanding cellular processes, development, and disease. Single-Cell RNA Sequencing (scRNA-Seq) using NGS at the single-cell level enables the study of gene expression profiles in individual cells, providing a more detailed understanding of cellular heterogeneity and dynamics (Jain, 2023).

- e) Epigenomics: NGS allows the mapping of DNA methylation patterns across the genome (DNA Methylation Profiling), providing insights into epigenetic modifications and their role in gene regulation, development, and diseases such as cancer. NGS is also used to identify protein-DNA interactions, revealing information about transcription factor binding, histone modifications, and chromatin structure (Armand, 2021).

- f) Metagenomics (Microbiome/virome studies): NGS enables the characterization of microbial communities in diverse environments, including the human gut, soil, water, and air. This has implications for understanding the role of microbiota in health and disease (Purushothaman, 2022).

- g) Clinical Diagnostics:

- *Cancer Genomics*. NGS is extensively used in cancer research and diagnostics to identify somatic mutations, copy number variations, and gene fusions. This information aids in personalized treatment strategies (Berger, 2018).

- *Infectious Disease Diagnostics*: NGS allows rapid and accurate identification of pathogens, facilitating the diagnosis and tracking of infectious diseases, including *de novo* sequencing based detection of novel and atypical pathogens (Zhong, 2021).

- *Non-Invasive Prenatal Testing (NIPT)*: NGS is employed for the detection of fetal chromosomal abnormalities from maternal blood samples, offering a safer alternative to traditional invasive methods (Carbone, 2020).

h) *Pharmacogenomics* (identification of the drug targets): NGS aids in identifying genetic variations that may influence drug response, allowing for the development of targeted therapies and personalized medicine (Morganti, 2019).

i) *Functional Genomics (CRISPR Screening)*: NGS is used to analyse the outcomes of CRISPR-based genetic screens, helping researchers identify genes associated with specific phenotypes or diseases (Shi, 2023).

j) *Evolutionary Biology*: NGS facilitates the study of genetic variation within populations, contributing to our understanding of evolutionary processes and adaptation (Piantadosi, 2021).

NGS has transformed biological and medical research by providing unprecedented insights into genetic information. As technology continues to advance, NGS is likely to play an even more significant role in advancing our understanding of genetics and improving clinical diagnostics and treatment strategies.

NGS is also the applicable tool for forensic genomics. NGS technologies are applied in forensic science for the analysis of DNA samples, helping in criminal investigations and paternity testing.

Biodiversity monitoring using NGS allows to assess biodiversity by analyzing DNA from environmental samples, such as soil or water, providing insights into the diversity of species.

The versatility and scalability of NGS technologies have made them indispensable tools across a wide range of scientific disciplines, contributing to advances in research, diagnostics, and personalized medicine. Ongoing technological innovations continue to expand the applications and impact of NGS in diverse fields.

Biodiversity monitoring using Next-Generation Sequencing (NGS) has emerged as a powerful and efficient approach for assessing and understanding the diversity of species within various ecosystems. NGS technologies enable researchers to analyze environmental DNA (eDNA), which includes genetic material shed by organisms into their surrounding environment (water, soil, air) (Wensel, 2022).

NGS allows the analysis of water samples for eDNA, providing insights into aquatic biodiversity (Antich, 2021). Researchers can detect the presence of various organisms, including fish, amphibians, and invertebrates, without the need to observe or capture them directly.

Soil samples can be also analyzed by the NGS to assess the microbial diversity and the presence of various organisms, including fungi, bacteria, and small invertebrates. NGS can be also applied to air samples to capture genetic material from airborne organisms, such as pollen, spores, and microorganisms (Köhler, 2021).

NGS facilitates the use of DNA barcoding, where specific genomic regions are sequenced to identify and distinguish different species. This is particularly useful for identifying organisms that may be challenging to observe visually.

This type of sequencing allows the simultaneous sequencing of multiple DNA barcodes from diverse organisms in a mixed sample. Metabarcoding provides a comprehensive overview of biodiversity within an ecosystem (Giampaoli, 2020).

NGS can be used to detect the presence of rare or endangered species by analyzing eDNA samples. This is crucial for monitoring the status and distribution of species at risk. It helps in detecting invasive species by analyzing environmental samples for the presence of non-native genetic material. Early detection is essential for managing and controlling invasive species.

NGS allows researchers to assess microbial diversity in different environments, providing insights into the health and functioning of ecosystems.

Changes in the composition of eDNA over time can reveal shifts in species composition and community dynamics, providing information on ecosystem stability and resilience.

NGS can be applied to monitor the impact of pollutants on biodiversity by assessing changes in the microbial community and the presence of indicator species.

Biodiversity monitoring with NGS enables the assessment of habitat quality by analyzing eDNA samples to determine the presence or absence of key species (Salazar, 2020).

NGS technologies support large-scale biodiversity surveys by enabling the collection and analysis of genetic data from diverse ecosystems around the world.

NGS empowers citizen scientists to contribute to biodiversity monitoring by collecting environmental samples and providing valuable data for research projects.

Biodiversity monitoring with NGS offers a transformative approach to understanding and conserving the rich tapestry of life on Earth. As technology advances and costs decrease, NGS is likely to play an increasingly central role in global biodiversity assessments and conservation efforts (Afouda, 2020).

It's essential to check for the latest developments and updates in the field, as new sequencing technologies and platforms are regularly introduced. Additionally, the choice of a sequencing platform depends on the specific requirements of the experiment, such as read length, throughput, and cost considerations.

NGS is used in the human medicine to identify genetic disorders/mutations associated with hereditary disorders, allowing for the diagnosis of conditions such as cystic fibrosis, muscular dystrophy, and various rare genetic diseases.

The versatility of NGS technologies has transformed medical research and clinical practice, facilitating precision medicine approaches that consider individual genetic variations for more effective and personalized patient care. Ongoing advancements in NGS continue to expand its applications and impact in the medical field.

Bacterial DNA and RNA sequencing has been used for decades to identify casual pathogens and resistance genes in clinical isolates and even before the onset of NGS, could yield rapid results with high specificity (Kapur, 1995). However, with the advance of NGS technologies, clinicians and laboratory professionals have seen tremendous growth and opportunity for using sequencing as a front-line diagnostic tool. Below, we summarize and briefly highlight the current applications of NGS for infectious diseases including a brief comparison of NGS methods, epidemiological surveillance, identification of pathogens and their resistance markers for diagnosis and treatment, and the importance of the host microbiome.

Targeted NGS uses panels of known pathogen sequences to screen clinical isolates. The panels can be specific for or target multiple types of pathogens including bacteria, viruses and eukaryotic organisms (Briese, 2015; Allicock, 2018). These panels can also target pathogens causing particular diseases. The advantages of these panels are their high specificity, sensitivity, rapid in time, and ability to sequence directly from a host isolate. However, the downsides include their limited scope and inability to identify novel pathogens or antibiotic resistance markers (Nimmo, 2017).

In the case of bacterial samples, WGS allows to characterize entire pathogen genome including plasmids. This broad sequencing allows for the identification of antibiotic resistance profiles, that can be used for first-line drug choice. WGS datasets accurately define known drug resistance markers, the discovery of novel mutations and their effects on phenotype bring added uncertainty to the test (Ellington, 2017).

Metagenomic NGS (mNGS) can analyse the direct specimen from the patients and amplify the sequences of all organisms in the sample, including host sequences. Furthermore, mNGS can detect pathogen sequences that comprise a very small portion of the overall sequenced reads; such low-level sequences can easily be missed by other methodologies (Mongkolrattanothai, 2017).

NGS has been playing an increasingly critical role in rapidly track the transmission, spread, and evolution of pathogens. Using a MinION sequencer (Oxford Nanopore Technologies), researchers were able to sequence over 140 Ebola virus positive samples

directly in the field (Quick, 2016). Sequencing was performed for SARS-CoV-2 detection and characterization, facilitated disease diagnosis and surveillance, and the development of effective drugs for disease treatment and vaccines for prevention (Zhou, 2020).

Numerous examples in the literature describe the use of NGS to identify bacterial, viral, or eukaryotic pathogens in a wide range of sample types including synovial fluid, feces, corneal tissue, blood and plasma, as well as nasopharyngeal swabs (Zhou, 2016).

Commonly used NGS technologies have also numerous limitations such as short reads and relying on clonal PCR to generate enough signals for detection. New technologies with long reads and single molecular sequencing (e.g., Pacific Biosystems and Oxford Nanopore) would theoretically be better and require less starting material. However, their high error rate prevents them from becoming the method of choice. In oncology, cancer heterogeneity is a challenge for sampling, variant detection, variant interpretation, and treatment recommendations. New methods, such as single cell sequencing and liquid biopsy, are promising for addressing this issue (Schrader, 2016).

Conclusion. NGS provides the new opportunities for molecular diagnostics. Many clinical laboratories have applied NGS technology to identify causative variants for the diagnosis of constitutional disorders, genomic profiling for precision oncology, and pathogen detection for infectious diseases. These novel technologies together with bioinformatics tools will continue to evolve and become the primary diagnostic tool and standard of care for genomic analysis to meet the ever-increasing demands of precision medicine.

REFERENCES

1. Levy, S. E., & Boone, B. E. (2019). Next-Generation Sequencing Strategies. *Cold Spring Harbor perspectives in medicine*, 9(7):a025791. doi:10.1101/cshperspect.a025791.
2. McCombie, W. R., McPherson, J. D., & Mardis, E. R. (2019). Next-Generation Sequencing Technologies. *Cold Spring Harbor perspectives in medicine*, 9(11):a036798. doi:10.1101/cshperspect.a036798.
3. Kumar, K. R., Cowley, M. J., & Davis, R. L. (2019). Next-Generation Sequencing and Emerging Technologies. *Seminars in thrombosis and hemostasis*, 45(7):661–673. doi:10.1055/s-0039-1688446.
4. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SB, Hood LE. 1986. Fluorescence detection in automated DNA sequence analysis. *Nature*; 321:674–679. doi:10.1038/321674a0.
5. Staden R, Beal KF, Bonfield JK. 2000. The Staden Package, 1998. *Methods Mol Biol*; 132:115–130.
6. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*; 437:376–380. doi:10.1038/nature03959.
7. Yang, T., Li, W., Huang, T., & Zhou, J. (2023). Genetic Testing Enhances the Precision Diagnosis and Treatment of Breast Cancer. *International journal of molecular sciences*; 24(23):16607. doi:10.3390/ijms242316607.
8. Budowle, B., Schmedes, S. E., & Wendt, F. R. (2017). Increasing the reach of forensic genetics with massively parallel sequencing. *Forensic science, medicine, and pathology*; 13(3):342–349. doi:10.1007/s12024-017-9882-5.
9. Roman, M. G., Gutierrez, R., & Houston, R. (2022). Massively parallel sequencing of Cannabis sativa chloroplast hotspots for forensic typing. *Journal of cannabis research*; 4(1):13. doi:10.1186/s42238-022-00123-2.
10. de Knijff P. (2019). From next generation sequencing to now generation sequencing in forensics. *Forensic science international. Genetics*, 38:175–180. doi:10.1016/j.fsigen.2018.10.017.

11. Yin, Y., Butler, C., & Zhang, Q. (2021). Challenges in the application of NGS in the clinical laboratory. *Human immunology*, 82(11):812–819. doi:10.1016/j.humimm.2021.03.011.
12. Pervez, M. T., Hasnain, M. J. U., Abbas, S. H., Moustafa, M. F., Aslam, N., & Shah, S. S. M. (2022). A Comprehensive Review of Performance of Next-Generation Sequencing Platforms. *BioMed research international*; 2022:3457806. doi:10.1155/2022/3457806.
13. Pereira, R., Oliveira, J., & Sousa, M. (2020). Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *Journal of clinical medicine*; 9(1):132. doi:10.3390/jcm9010132.
14. Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human immunology*; 82(11):801–811. doi:10.1016/j.humimm.2021.02.012.
15. Ringlander, J., Andersson, M. E., Prakash, K., Larsson, S. B., & Lindh, M. (2022). Deep sequencing of hepatitis B virus using Ion Torrent fusion primer method. *Journal of virological methods*; 299:114315. doi:10.1016/j.jviromet.2021.114315.
16. Chen, P., Yu, X., Huang, H., Zeng, W., He, X., Liu, M., & Huang, B. (2020). Evaluation of Ion Torrent next-generation sequencing for thalassemia diagnosis. *The Journal of international medical research*; 48(12):300060520967778. doi:10.1177/0300060520967778.
17. Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. *Genomics, proteomics & bioinformatics*; 13(5):278–289. doi:10.1016/j.gpb.2015.08.002
18. Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature biotechnology*; 39(11):1348–1365. doi:10.1038/s41587-021-01108-x.
19. Sheka, D., Alabi, N., & Gordon, P. M. K. (2021). Oxford nanopore sequencing in clinical microbiology and infection diagnostics. *Briefings in bioinformatics*; 22(5):bbaa403. doi:10.1093/bib/bbaa403.
20. Gilpatrick, T., Lee, I., Graham, J. E., Raimondeau, E., Bowen, R., Heron, A., Downs, B., Sukumar, S., Sedlazeck, F. J., & Timp, W. (2020). Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nature biotechnology*; 38(4):433–438. doi:10.1038/s41587-020-0407-5.
21. Harrington, C. T., Lin, E. I., Olson, M. T., & Eshleman, J. R. (2013). Fundamentals of pyrosequencing. *Archives of pathology & laboratory medicine*; 137(9):1296–1303. doi:10.5858/arpa.2012-0463-RA.
22. Wieckhusen, C., & Bugert, P. (2015). 454-Sequencing™ for the KEL, JR, and LAN Blood Groups. *Methods in molecular biology (Clifton, N.J.)*; 1310:123–133. doi:10.1007/978-1-4939-2690-9_11.
23. McCarthy A. (2013). BGI Americas: commercializing next-generation sequencing. *Chemistry & biology*; 20(6):743–744. doi:10.1016/j.chembiol.2013.06.003.
24. Zhong, Y., Xu, F., Wu, J., Schubert, J., & Li, M. M. (2021). Application of Next Generation Sequencing in Laboratory Medicine. *Annals of laboratory medicine*; 41(1):25–43. doi:10.3343/alm.2021.41.1.25.
25. Searle, B., Müller, M., Carell, T., & Kellett, A. (2023). Third-Generation Sequencing of Epigenetic DNA. *Angewandte Chemie (International ed. in English)*; 62(14):e202215704. doi:10.1002/anie.202215704.
26. Petersen, L. M., Martin, I. W., Moschetti, W. E., Kershaw, C. M., & Tsongalis, G. J. (2019). Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *Journal of clinical microbiology*; 58(1):e01315-19. doi:10.1128/JCM.01315-19.

27. Zhu, Y., Huang, Y., Tan, Y., Zhao, W., & Tian, Q. (2020). Single-Cell RNA Sequencing in Hematological Diseases. *Proteomics*; 20(13):e1900228. doi:10.1002/pmic.201900228.
28. Zhao, E. Y., Jones, M., & Jones, S. J. M. (2019). Whole-Genome Sequencing in Cancer. *Cold Spring Harbor perspectives in medicine*; 9(3):a034579. doi:10.1101/cshperspect.a034579.
29. Jelin, A. C., & Vora, N. (2018). Whole Exome Sequencing: Applications in Prenatal Genetics. *Obstetrics and gynecology clinics of North America*; 45(1):69–81. doi:10.1016/j.ogc.2017.10.003.
30. Lu, H., Giordano, F., & Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, proteomics & bioinformatics*; 14(5):265–279. doi:10.1016/j.gpb.2016.05.004.
31. Jain, S., Rick, J. W., Joshi, R. S., Beniwal, A., Spatz, J., Gill, S., Chang, A. C., Choudhary, N., Nguyen, A. T., Sudhir, S., Chalif, E. J., Chen, J. S., Chandra, A., Haddad, A. F., Wadhwa, H., Shah, S. S., Choi, S., Hayes, J. L., Wang, L., Yagnik, G., ... Aghi, M. K. (2023). Single-cell RNA sequencing and spatial transcriptomics reveal cancer-associated fibroblasts in glioblastoma with protumoral effects. *The Journal of clinical investigation*; 133(5):e147087. doi:10.1172/JCI147087.
32. Armand, E. J., Li, J., Xie, F., Luo, C., & Mukamel, E. A. (2021). Single-Cell Sequencing of Brain Cell Transcriptomes and Epigenomes. *Neuron*; 109(1):11–26. doi:10.1016/j.neuron.2020.12.010.
33. Purushothaman, S., Meola, M., & Egli, A. (2022). Combination of Whole Genome Sequencing and Metagenomics for Microbiological Diagnostics. *International journal of molecular sciences*; 23(17):9834. doi:10.3390/ijms23179834.
34. Berger, M. F., & Mardis, E. R. (2018). The emerging clinical relevance of genomics in cancer medicine. *Nature reviews. Clinical oncology*; 15(6):353–365. doi:10.1038/s41571-018-0002-6.
35. Carbone, L., Cariati, F., Sarno, L., Conforti, A., Bagnulo, F., Strina, I., Pastore, L., Maruotti, G. M., & Alviggi, C. (2020). Non-Invasive Prenatal Testing: Current Perspectives and Future Challenges. *Genes*; 12(1):15. doi:10.3390/genes12010015.
36. Morganti, S., Tarantino, P., Ferraro, E., D'Amico, P., Duso, B. A., & Curigliano, G. (2019). Next Generation Sequencing (NGS): A Revolutionary Technology in Pharmacogenomics and Personalized Medicine in Cancer. *Advances in experimental medicine and biology*; 1168:9–30. doi:10.1007/978-3-030-24100-1_2.
37. Shi, H., Doench, J. G., & Chi, H. (2023). CRISPR screens for functional interrogation of immunity. *Nature reviews. Immunology*; 23(6):363–380. doi:10.1038/s41577-022-00802-4.
38. Piantadosi, A., Mukerji, S. S., Ye, S., Leone, M. J., Freimark, L. M., Park, D., Adams, G., Lemieux, J., Kanjilal, S., Solomon, I. H., Ahmed, A. A., Goldstein, R., Ganesh, V., Ostrem, B., Cummins, K. C., Thon, J. M., Kinsella, C. M., Rosenberg, E., Frosch, M. P., Goldberg, M. B., ... Sabeti, P. (2021). Enhanced Virus Detection and Metagenomic Sequencing in Patients with Meningitis and Encephalitis. *mBio*; 12(4):e0114321. doi:10.1128/mBio.01143-21.
39. Wensel, C. R., Pluznick, J. L., Salzberg, S. L., & Sears, C. L. (2022). Next-generation sequencing: insights to advance clinical investigations of the microbiome. *The Journal of clinical investigation*; 132(7):e154944. doi:10.1172/JCI154944.
40. Antich, A., Palacín, C., Cebrian, E., Golo, R., Wangenstein, O. S., & Turon, X. (2021). Marine biomonitoring with eDNA: Can metabarcoding of water samples cut it as a tool for surveying benthic communities? *Molecular ecology*; 30(13):3175–3188. doi:10.1111/mec.15641.
41. Köhler, J. M., Beetz, N., Günther, P. M., Möller, F., & Cao, J. (2021). Extremophiles in Soil Communities of Former Copper Mining Sites of the East Harz Region (Germany)

Reflected by Re-Analyzed 16S rRNA Data. *Microorganisms*; 9(7):1422. doi:10.3390/microorganisms9071422.

42. Giampaoli, S., De Vittori, E., Frajese, G. V., Paytuví, A., Sanseverino, W., Anselmo, A., Barni, F., & Berti, A. (2020). A semi-automated protocol for NGS metabarcoding and fungal analysis in forensic. *Forensic science international*; 306:110052. doi:10.1016/j.forsciint.2019.110052.

43. Salazar, N., González, S., Nogacka, A. M., Rios-Covián, D., Arbolea, S., Gueimonde, M., & Reyes-Gavilán, C. G. L. (2020). Microbiome: Effects of Ageing and Diet. *Current issues in molecular biology*; 36:33–62. doi:10.21775/cimb.036.033.

44. Afouda, P., Dubourg, G., & Raoult, D. (2020). Archeomicrobiology applied to environmental samples. *Microbial pathogenesis*; 143:104140. doi:10.1016/j.micpath.2020.104140.

45. Kapur, V., Li, L. L., Hamrick, M. R., Plikaytis, B. B., Shinnick, T. M., Telenti, A., Jacobs, W. R., Jr, Banerjee, A., Cole, S., & Yuen, K. Y. (1995). Rapid Mycobacterium species assignment and unambiguous identification of mutations associated with antimicrobial resistance in Mycobacterium tuberculosis by automated DNA sequencing. *Archives of pathology & laboratory medicine*; 119(2):131–138.

46. Allicock, O. M., Guo, C., Uhlemann, A. C., Whittier, S., Chauhan, L. V., Garcia, J., Price, A., Morse, S. S., Mishra, N., Briesse, T., & Lipkin, W. I. (2018). BacCapSeq: a Platform for Diagnosis and Characterization of Bacterial Infections. *mBio*; 9(5):e02007-18. doi:10.1128/mBio.02007-18.

47. Briesse, T., Kapoor, A., Mishra, N., Jain, K., Kumar, A., Jabado, O. J., & Lipkin, W. I. (2015). Virome Capture Sequencing Enables Sensitive Viral Diagnosis and Comprehensive Virome Analysis. *mBio*; 6(5):e01491-15. doi:10.1128/mBio.01491-15.

48. Nimmo, C., Doyle, R., Burgess, C., Williams, R., Gorton, R., McHugh, T. D., Brown, M., Morris-Jones, S., Booth, H., & Breuer, J. (2017). Rapid identification of a Mycobacterium tuberculosis full genetic drug resistance profile through whole genome sequencing directly from sputum. *International journal of infectious diseases: IJID : official publication of the International Society for Infectious Diseases*; 62:44–46. doi:10.1016/j.ijid.2017.07.007.

49. Ellington, M. J., Ekelund, O., Aarestrup, F. M., Canton, R., Doumith, M., Giske, C., Grundman, H., Hasman, H., Holden, M. T. G., Hopkins, K. L., Iredell, J., Kahlmeter, G., Köser, C. U., MacGowan, A., Mevius, D., Mulvey, M., Naas, T., Peto, T., Rolain, J. M., Samuelsen, Ø., ... Woodford, N. (2017). The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*; 23(1):2–22. doi:10.1016/j.cmi.2016.11.012.

50. Mongkolrattanothai, K., Naccache, S. N., Bender, J. M., Samayoa, E., Pham, E., Yu, G., Dien Bard, J., Miller, S., Aldrovandi, G., & Chiu, C. Y. (2017). Neurobrucellosis: Unexpected Answer From Metagenomic Next-Generation Sequencing. *Journal of the Pediatric Infectious Diseases Society*; 6(4):393–398. doi:10.1093/jpids/piw066.

51. Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., Bore, J. A., Koundouno, R., Dudas, G., Mikhail, A., Ouédraogo, N., Afrough, B., Bah, A., Baum, J. H., Becker-Ziaja, B., Boettcher, J. P., Cabeza-Cabrerizo, M., Camino-Sanchez, A., Carter, L. L., Doerrbecker, J., ... Carroll, M. W. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*; 530(7589):228–232. doi:10.1038/nature16996.

52. Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. D., Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., Zheng, X. S., ... Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*; 579(7798):270–273. doi:10.1038/s41586-020-2012-7.

53. Zhou, Y., Wylie, K. M., El Feghaly, R. E., Mihindukulasuriya, K. A., Elward, A., Haslam, D. B., Storch, G. A., & Weinstock, G. M. (2016). Metagenomic Approach for Identification of the Pathogens Associated with Diarrhea in Stool Specimens. *Journal of clinical microbiology*; 54(2):368–375. doi:10.1128/JCM.01965-15.

54. Schrader, K. A., Cheng, D. T., Joseph, V., Prasad, M., Walsh, M., Zehir, A., Ni, A., Thomas, T., Benayed, R., Ashraf, A., Lincoln, A., Arcila, M., Stadler, Z., Solit, D., Hyman, D. M., Zhang, L., Klimstra, D., Ladanyi, M., Offit, K., Berger, M., ... Robson, M. (2016). Germline Variants in Targeted Tumor Sequencing Using Matched Normal DNA. *JAMA oncology*; 2(1):104–111. doi:10.1001/jamaoncol.2015.5208.

ЗАСТОСУВАННЯ СЕКВЕНУВАННЯ НОВОГО ПОКОЛІННЯ В БІОЛОГІЇ ТА МЕДИЦИНІ

Герілович А.П.^{1,2} (ORCID ID 0000-0002-3280-4172), **Сушко М.І.**¹ (ORCID ID 0000-0003-2723-9903), **Мандигра С.С.**¹ (ORCID ID 0000-0002-3559-9258), **Родина Н.С.**³ (ORCID ID 0000-0002-4338-8872), **Чечет О.М.**¹ (ORCID ID 0000-0001-5099-5577), **Романко М.Є.**¹ (ORCID ID 0000-0003-0285-5603), **Кучинський М. В.**¹ (ORCID ID 0009-0003-2652-3360), **Герілович І.О.**⁴ (ORCID ID 0000-0002-6742-2086)

1 – Державний науково-дослідний інститут з лабораторної діагностики та ветеринарно-санітарної експертизи, м. Київ, Україна, antger2011@gmail.com;

2 – ГО «Інститут єдиного здоров'я», м. Харків, Україна;

3 – ДУ «Київський центр контролю та профілактики захворювань» МОЗ України,

4 – ННЦ «Інститут експериментальної та клінічної ветеринарної медицини», м. Харків, Україна

Резюме. Секвенування нового покоління (NGS), також відоме як високопродуктивне секвенування, відноситься до низки сучасних технологій секвенування ДНК, які зробили революцію в галузі геноміки. Переваги методів NGS включають високу швидкість (паралельне секвенування є швидшим за традиційні методи, що дозволяє дослідникам швидше отримувати результати), економічність (можливість секвенувати кілька фрагментів одночасно зменшує вартість бази порівняно з традиційним секвенуванням) і масштабованість (платформи можна масштабувати, щоб відповідати різним рівням пропускну здатності залежно від експериментальних потреб). NGS суттєво прискорило геномні дослідження, зробивши можливим прорив у таких галузях, як персоналізована медицина, геноміка раку та еволюційна біологія. Однак такі проблеми, як складність аналізу даних, рівень помилок і вартість, все ще існують і є областями постійних досліджень і вдосконалень у сфері технологій секвенування. Стаття містить коротке пояснення поточних платформ NGS та їх особливості. Описано біомедичне застосування NGS з основними перевагами та можливостями аналізованих інструментів.

Ключові слова: ДНК, РНК, геноміка, метагеноми, секвенування нового покоління, діагностика, біологія, медицина

DOI: 10.31073/onehealthjournal2024-I-05