DOI: 10.55522/jmpas.V15I3.7030 ISSN NO. 2320 – 7418

International peer reviewed open access journal

Journal of Medical Pharmaceutical and Allied Sciences

Journal homepage: www.jmpas.com CODEN: JMPACO

Research article

Role of next-generation sequencing in the discovery of mutational biomarkers

Anushka Jain[1], Prachi Srivastava[1], Prekshi Garg*[2]

[1] Amity Institute of Biotechnology, Amity University Lucknow, Uttar Pradesh, India
[2] Bioinfo Core Solutions (OPC) Pvt. Ltd., Lucknow, Uttar Pradesh, India

Corresponding author: Prekshi Garg, ✉ prekshigarg23@gmail.com, **Orcid Id**: https://orcid.org/0000-0001-9161-0768

Amity Institute of Biotechnology, Amity University Lucknow, Uttar Pradesh, India

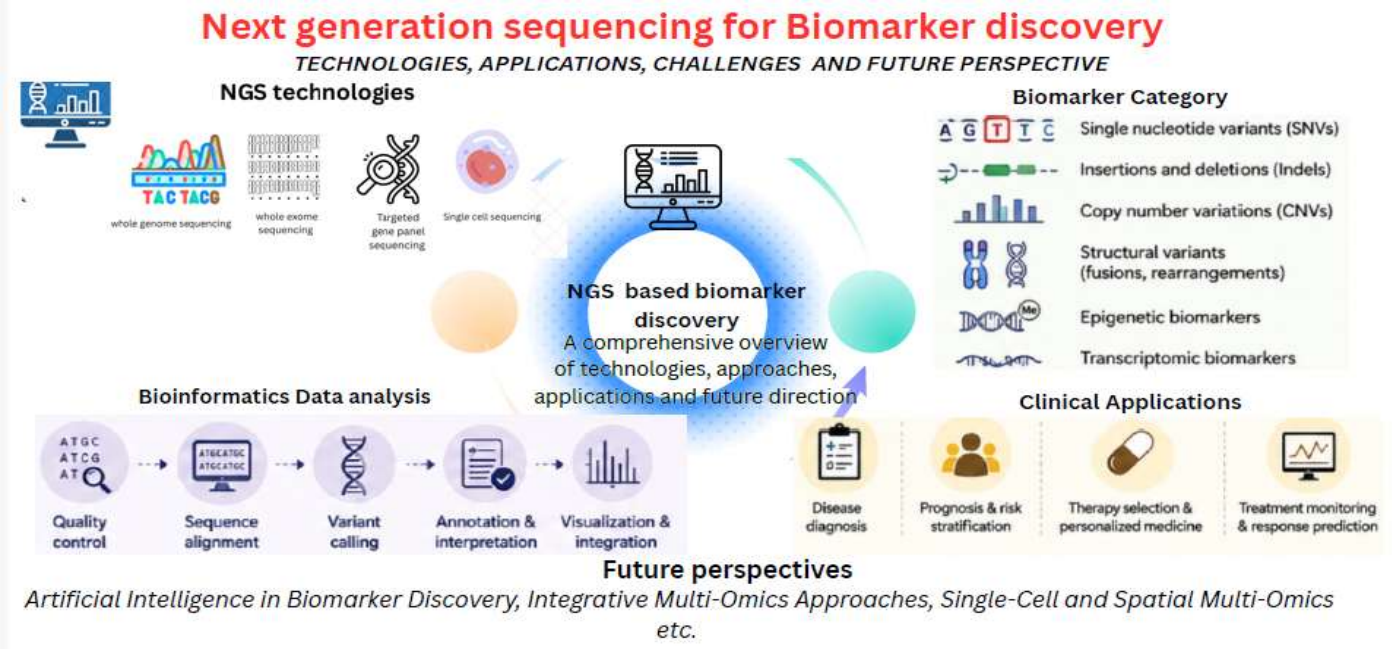
Received - 30-03-2026**, Revised** - 10-04-2026, **Accepted** - 15-05-2026 **(DD-MM-YYYY)**

Refer This Article

Anushka Jain, Prachi Srivastava, Prekshi Garg, 2026. Role of next-generation sequencing in the discovery of mutational biomarkers. Journal of medical pharmaceutical and allied sciences, V 15, I 3, Pages 17 – 22. Doi: https://doi.org/10.55522/jmpas.V15I3.7030.

ABSTRACT

Any measurable indicators of biological processes, disease states or therapeutic response is called biomarkers. They play a key role in modern clinical decision-making. Mutational biomarkers include single-nucleotide variants, insertions and deletions, copy-number alterations, and structural rearrangements. They are important as they provide stable genomic insights into disease susceptibility, progression, and treatment response. With the advent of next-generation sequencing (NGS) technology, this concept of mutational biomarker has been revolutionised. This is because NGS enables rapid, high-throughput, and cost-effective genome analysis compared with traditional sequencing technologies. This review outlines the principles, platforms, and experimental methodologies of NGS, including whole-genome, whole-exome, targeted panel, and single-cell sequencing, for identifying mutational biomarkers. The significance of bioinformatics analysis and the therapeutic utility of NGS in cancer precision medicine are also emphasised. Despite obstacles in interpretation and standardisation, NGS has great potential for improving early diagnosis, tailored therapy, and patient outcomes.



Keywords: Single-nucleotide variants, Next-generation sequencing, High-throughput, Whole-genome, Whole-exome, Targeted panel.

INTRODUCTION

Biomarkers are defined as characteristics that are measurable and act as indicators of normal biological processes, pathogenic processes, responses to exposure, and interventions. It also includes therapeutic interventions. The nature of these characteristics can vary; they could be either molecular, histologic, radiographic or even physiological [1]. Conventional biomarkers can be classified into many different categories, and all these categories are interrelated based upon their biological role, measurable characteristics and clinical applications. The broad classification of biomarkers is as Type0, Type1 and Type 2. Type 0 biomarkers are those that tell the natural history and progression of a disease. Type 1 biomarkers indicated biological responses to any treatment, Type 2, which are often called as surrogate biomarkers, predict the therapeutic responses before any clinical outcome occurs. Another classification of biomarkers is on the basis of biological nature as molecular, cellular and imaging biomarkers. The third criterion on which biomarkers can be classified is according to clinical applications as diagnostic, prognostic and therapeutic biomarkers [2, 3, 4].

In this article, we will be primarily focusing on mutational biomarkers. Mutational biomarkers are defined as specific, measurable genetic alterations in a genome. They include single-nucleotide variants, insertions and deletions, copy-number changes, and structural rearrangements. These mutations are very important as they may serve as molecular indicators that are stable and are capable of capturing individual variability at the level of DNA and provide insights into disease susceptibility, pathogenesis, progression, and response to interventions [5].

High-throughput sequencing technologies have enabled the identification and characterisation of mutational biomarkers in a more precise and accurate manner. The major advantage of tailoring therapies to an individual's genetic profile is that it contributes to the development of personalised medicine.

Diagnosis

According to the 2017 International League Against Epilepsy (ILAE) standards, seizure type should be categorised to help direct treatment **[10]**. Magnetic resonance imaging (MRI) and electroencephalogram (EEG) tests should be performed on patients who are suspected of having GRE **[11]**. The clinical data and imaging results should both be used to make the GRE diagnosis. Glioma is mostly diagnosed by MRI, which includes FLAIR, DWI, PWI, T2-weighed, and contrast – enhanced T1- weighted sequences. Differential diagnosis may benefit from the use of CT, MRS, and PET. The gold standard is still histopathological examination, which includes molecular testing according to the 2016 WHO classification [11].

METHODOLOGY

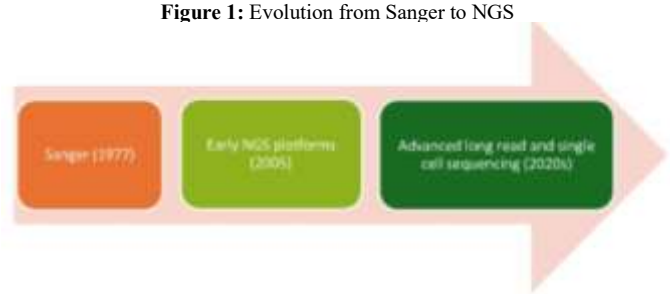
A comprehensive literature search was conducted to collect relevant studies. These were related to next-generation sequencing and mutational biomarker discovery. Scientific databases such as PubMed, Scopus, Google Scholar, and Web of Science were used to retrieve the articles. The search was carried out by adding keywords such as "next-generation sequencing," "mutational biomarkers," "whole-genome sequencing," "whole-exome sequencing," "targeted gene panel," and "single-cell sequencing." Only articles written in English were taken into account. This review consisted of recent studies, review articles, and research papers related to next-generation sequencing technologies. The articles that were not relevant to the subject or lacked adequate information were excluded. The articles selected were reviewed, and relevant information was obtained to provide a clear idea about next-generation sequencing technologies, methodologies, and their role in mutational biomarkers.

Overview of next-generation sequencing technologies

Evolution from Sanger sequencing to next-generation sequencing

As is known, the first sequencing technology that was developed was Sanger sequencing technology in the late 1970s, often known as first-generation sequencing technology, where the sequencing method is based upon dideoxy chain termination, the DNA synthesis stops by the incorporation of dd NTPs, and the sequence is later determined through gel electrophoresis. For decades, this sequencing technology has served as a gold standard due to its high accuracy and its important role in early gene studies and the Human Genome Project. However, there were certain limitations of this which were later overcome by Next generation sequencing technologies, including its high cost, low throughput and its slow processing, which limits its suitability for large-scale genome analysis [6]. Next-generation sequencing technology in the early 2000s as massively parallel sequencing platform that was capable of processing millions of fragments of DNA simultaneously [7]. This revolutionised genomics by enabling high-throughput sequencing, a very large reduction in cost per base, and turnaround time was fast, and thus this permitted genome-wide assessments which were previously considered impractical [8].

Figure 1: Evolution from Sanger to NGS



Principles of next-generation sequencing

The core concept of next-generation sequencing is based on massively parallel sequencing. As mentioned above, in massively

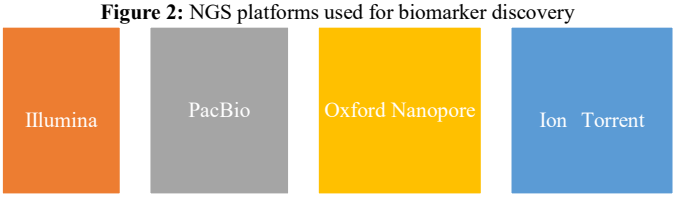
parallel sequencing, millions of DNA fragments are sequenced simultaneously rather than individually as in the case of first-generation sequencing. In the workflow, the first step is DNA fragmentation, which is followed by library preparation along with ligation of the adapter, clonal amplification of fragments, cyclic sequencing of the reactions, and detection of high-throughput signals. To reconstruct the sequence information, all these steps are followed by computational analysis [9].

There are different NGS platforms that use different sequencing methods, such as sequencing by synthesis, detecting semiconductors or real-time sequencing of single molecules. Although the basic idea of all these methods is the same, i.e. reading many DNA fragments at the same time and using computer analysis for processing the large amount of data that is generated. This framework of technology enables the rapid, large-scale genome analysis with very high sensitivity for detecting the genetic variants, hence this supports the applications of NGS in large-scale genome analysis and in the detection of genetic variants [10].

NGS platforms used in biomarker discovery

There are various platforms of next-generation sequencing which are used in the discovery of mutational biomarkers, like the Illumina, Ion Torrent, PacBio and Oxford Nanopore. Now, we will look at each of them in detail. Illumina is based on sequencing- by-synthesis chemistry, where fragmented DNA is amplified clonally on a flow cell, and nucleotides that are fluorescently labelled are incorporated for sequence determination [11]. There are various applications of Illumina sequencing, and comparative studies have even demonstrated its effectiveness for the identification of even plant viruses and viroids using this technology [12]. In contrast, in Ion Torrent sequencing, the synthesis of DNA is measured using Hydrogen ions, which are released during the incorporation of nucleotide, rather than using fluorescent signals. Hence, it is a semiconductor-based approach that enables rapid sequencing [13]. When we talk about the third-generation sequencing technology that includes the PacBio sequencing, it performs single-molecule real-time sequencing called SMRT, which allows for reading the DNA molecules without prior amplification. It generates long continuous reads which are used for resolving the structural variants, repetitive regions and those complex genome assemblies that are difficult to analyse with the short-read platforms [14]. Oxford nanopore sequencing determines DNA sequence by measuring changes in ionic current as nucleic acids pass through a biological nanopore, i.e. it is a single-molecule sequencing technology. This enables real-time sequencing of very long DNA fragments without the need for prior amplification or complex detection systems [15].

Figure 2: NGS platforms used for biomarker discovery



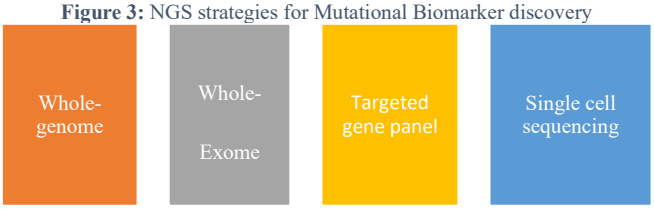
Comparison of short-read and long-read sequencing

Short read sequencing generates very large volumes of short DNA reads that are highly accurate. These are good for variant selection and taxonomic profiling, but they can't resolve repetitive regions and complex genome structures. Whereas long-read sequencing can produce contiguous reads that are longer. They improve the assembly of genomes, identification of structural variants, and they can resolve repetitive and highly similar sequences. Long read and hybrid strategies are known to enhance assembly continuity and completeness as compared to the short-read approaches when comparative analysis was done in genome-resolved metagenomics [16]. Moreover, long read sequencing technology minimize sequencing errors by reducing the dependency on complex computational assembly algorithms.

NGS strategies for mutational biomarker discovery

There are various experimental strategies that are facilitated by NGS for the discovery of mutational biomarkers.

Figure 3: NGS strategies for Mutational Biomarker discovery



Whole-Genome sequencing (WGS)

Using WGS, we can sequence the complete DNA sequence of an organism, which covers both the coding and the non-coding regions. The major advantage is that it allows unbiased detection of diverse genetic variations, like SNPs, insertions and deletions, copy number changes and structural rearrangements [17]. WGS provides high-resolution insight into the mechanisms of complex diseases and also the genetic architecture. Also, it can generate huge genomic datasets and, when integrated with omics data, can enhance precision medicine and aid in mutational biomarker discovery [18].

Whole-exome sequencing (WES)

WES, in contrast to WGS, is a targeted next-generation sequencing approach; this means that it focuses on the protein-coding regions of the genome. These protein coding regions are the ones that that contains majority of known disease-causing variants. Since it selectively captures and sequences exons, it's a cost-effective and efficient alternative to WGS [19]. The major application of whole genome sequencing is in the diagnosis of rare and heterogeneous diseases, and this significantly improves molecular diagnostic yield across diverse clinical interventions [20].

Targeted gene panel sequencing

This technique is used when we want to analyse or sequence a selected set of genes or a genomic region that is associated with a particular disease. This technique is highly sensitive and provides sequencing depth, hence it enables accurate detection of pathogenic variants. This sequencing technique is cost-effective, faster and easier to interpret because it limits sequencing to clinically relevant loci. It is also applied to the field of oncology to identify novel pathogenic mutations and clinically actionable variants in patients with certain cancers [21].

Single-cell sequencing approaches

It is an advanced genomic approach where analysis of DNA or RNA is done from individual cells, rather than from bulk populations. This method is high resolution and provides insights into cellular heterogeneity. When genomic and transcriptomic information is captured at the single-cell level, it is easier to understand disease mechanisms and cellular diversity [22].

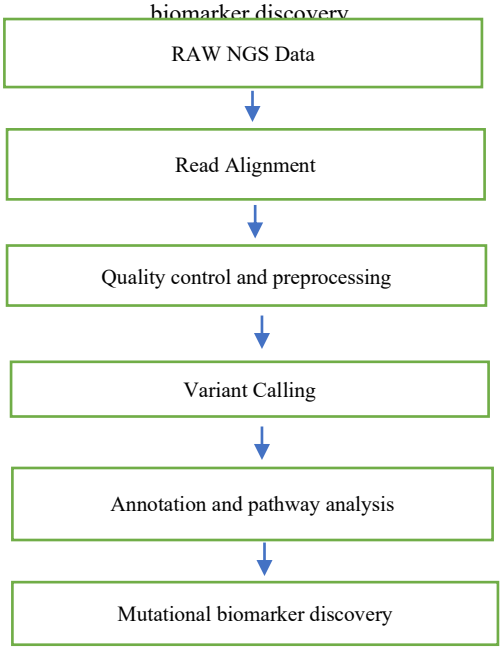
Identification of mutation types using NGS

There are different types of mutations that are identified by next-generation sequencing data analysis, like single-nucleotide variants, Insertions and deletions, Copy number alterations, structural variants and gene rearrangements. SNVs are alterations in the genome that involve a change in a single nucleotide at a specific position in the genome [23]. In Insertion and deletion mutations, the addition or loss of one or more nucleotides within a DNA sequence is seen. These alterations may have variation in size, and they may occur in either coding or non-coding regions [24]. In copy number alterations, segments of DNA are either duplicated or deleted, and this leads to changes in the copies of a specific gene or of a chromosomal region [25]. Apart from all these, structural Variants are genomic changes that are on a large scale, that is, typically DNA segment which are greater than 1 kilobase in size. A major application of them is in disease susceptibility and studying the genetic diversity [26]. Identification of all these biomarkers makes next-generation sequencing a powerful tool in the discovery of biomarkers.

Bioinformatics pipeline for the discovery of mutational biomarkers

As we know, bioinformatics plays a very crucial role in converting NGS data to get clinically useful results. The first step of this process is data preprocessing and quality control, which is followed by sequence alignment and then finally variant calling, i.e. a Vcf file is prepared [27]. The most important step is quality control and preprocessing of data, as here we are removing any low-quality reads or omitting the sequencing errors (if any present). For all the steps shown in the flowchart different bioinformatics tools, or coding-based approach can be used. The results are usually compiled in the form of graphs for easy understanding and interpretation.

Flowchart showing the bioinformatics pipeline used for mutational biomarker discovery



Role of NGS in cancer mutational biomarker discovery

Due to its comprehensive high-throughput detection of genetic alterations across tumour genes, next-generation sequencing has transformed cancer biomarker discovery. NGS can be used to identify single-nucleotide variants, insertions, deletions and also copy number changes, and gene fusion across tumour genomes. This allows for the identification of diagnostic, prognostic and predictive biomarkers [28]. There are certain biomarkers like EGFR, KRAS, ALK and BRAF which are critical for selecting targeted therapies and treatment of cancer [29]. In recent studies, NGS has been used in the discovery of mutational biomarkers, especially in the case of lung cancer [30].

DISCUSSION

Next-generation sequencing has greatly changed the way mutational biomarkers are discovered. By using this technique, we can easily study the genetic changes in a fast and detailed way, on a large scale. In this review article we have discussed different NGS methods. All of these methods have their own advantage and disadvantages, and the choice depends upon the user preferences. Producing large amounts of data, a lack of uniformity in results, and the expertise required for identifying the most important genetic changes associated with the disease of interest are some challenges associated with using NGS. Another difficulty is in designing standard pipelines for data analysis due to the large number of tools and software available for performing the same task. However, this has enhanced the research in cases of certain rare diseases and also in certain types of cancers, which were thereby difficult to treat.

The future of Next Generation sequencing in the discovery of mutational biomarkers looks promising. It is going to revolutionise the field of research and give major useful contributions

in clinical practice. The integration of NGS with Artificial intelligence (AI) and multiomics is going to provide a complete understanding of disease mechanisms in detail. This would also lead to the generation of an automated bioinformatics pipeline for mutational biomarker discovery. When we see from a clinical perspective, then next-generation sequencing is going to play a key role in the field of personalised medicine. Overall, rapid development in technology along with enhancements in analysis methods, will increase the role of NGS in the discovery of novel biomarkers and hence would aid in better clinical outcomes in the future.

CONCLUSION

Therefore, it is evident that NGS is a powerful tool that has actually transformed the field of research due to its ability to discover novel biomarkers by high-throughput analysis of alterations across diverse biological contexts. In contrast to traditional sequencing approaches, next-generation sequencing has allowed for the detection of single-nucleotide variants, insertions and deletions, copy number alterations and even structural variants, through different types of sequencing strategies like WGS, WES, targeted panel and even single-cell sequencing approaches. This has enabled the diagnosis of certain diseases to be more precise and the selection of the treatment methods or therapies to be more selective. Despite there are certain challenges which are related to data interpretation, cost and standardisation.

But overall, the utility of NGS in the discovery of mutational biomarkers in research and diagnostic healthcare represents advancements in the field of precision medicine, and this can further enhance early diagnosis and better patient outcomes in the near future.

ACKNOWLEDGEMENT

I would like to acknowledge Bioinfocore Solutions for providing me with guidance on writing a review paper.

Author's contributions statement

All authors contributed equally to the manuscript.

Conflict of interest

Authors declare no conflict of interest, financial or otherwise, amongst each other.

REFERENCES

1. Califf R M, 2018. Biomarker definitions and their applications. Experimental Biology and Medicine. 243(3), Pages 213–221. Doi: https://doi.org/10.1177/1535370217750088.

2. Bodaghi A, Fattahi N, Ramazani A, 2023. Biomarkers: Promising and valuable tools towards diagnosis, prognosis and treatment of COVID-19 and other diseases. Heliyon. 9(2), Pages e13323. Doi: https://doi.org/10.1016/j.heliyon.2023.e13323.

3. Wishart D S, Bartok B, Oler E, et al, 2021. MarkerDB: an online database of molecular biomarkers. Nucleic Acids Research. 49(D1), Pages D1259–D1267. Doi: https://doi.org/10.1093/nar/gkaa1067.

4. Dhama K, Latheef S K, Dadar M, et al, 2019. Biomarkers in Stress Related Diseases/Disorders: Diagnostic, Prognostic, and Therapeutic Values. Frontiers in Molecular Biosciences. 6, Doi: https://doi.org/10.3389/fmolb.2019.00091.

5. Novelli G, Cinzia Ciccacci, Borgiani P, et al, 2026. Genetic tests and genomic biomarkers: regulation, qualification and validation. Clinical Cases in Mineral and Bone Metabolism. 5(2), 149.

6. Gomes A, Korf B, 2017. Genetic Testing Techniques. Elsevier EBooks. Pages 47–64. Doi: https://doi.org/10.1016/b978-0-323-48555-5.00005-3.

7. Valerie J. Carabetta, 2025. High Throughput Gene Screening. SpringerLink. Doi: https://doi.org/10.1007-978-1-0716-4192-7.

8. Hu T, Chitnis N, Monos D, et al, 2021. Next-generation sequencing technologies: An overview. Human Immunology. 82(11), Pages 801–811. https://doi.org/10.1016/j.humimm.2021.02.012.

9. Reis-Filho J S, 2009. Next-generation sequencing. Breast Cancer Research. 11(S3), Pages S12–S12. Doi: https://doi.org/10.1186/bcr2431.

10. Alexandru Isaic, Nadica Motofelea, Hoinoiu T, et al, 2025. Next-Generation Sequencing: A Review of Its Transformative Impact on Cancer Diagnosis, Treatment, and Resistance Management. Diagnostics. 15(19), Pages 2425–2425. Doi: https://doi.org/10.3390/diagnostics15192425.

11. Bogaerts B, Van den Bossche A, Verhaegen B, et al, 2024. Closing the gap: Oxford Nanopore Technologies R10 sequencing allows comparable results to Illumina sequencing for SNP-based outbreak investigation of bacterial pathogens. Journal of Clinical Microbiology. 62(5), Doi: https://doi.org/10.1128/jcm.01576-23.

12. Pecman A, Adams I, Ion Gutiérrez-Aguirre, et al, 2022. Systematic Comparison of Nanopore and Illumina Sequencing for the Detection of Plant Viruses and Viroids Using Total RNA Sequencing Approach. Frontiers in Microbiology. 13, Pages 883921. Doi: https://doi.org/10.3389/fmicb.2022.883921.

13. Damerla R R, Chatterjee B, Li Y, et al, 2013. Ion Torrent sequencing for conducting genome-wide scans for mutation mapping analysis. Mammalian Genome. 25(3-4), Pages 120–128. Doi: https://doi.org/10.1007/s00335-013-9494-7.

14. Rhoads A, Au K F, 2015. PacBio Sequencing and its Applications. Genomics, Proteomics & Bioinformatics. 13(5), Pages 278–289. Doi: https://doi.org/10.1016/j.gpb.2015.08.002.

15. Eisenstein M, 2012. Oxford Nanopore announcement sets sequencing sector abuzz. Nature Biotechnology. 30(4), Pages 295–296. Doi: https://doi.org/10.1038/nbt0412-295.

16. Eisenhofer R, Nesme J, Santos-Bay L, et al, 2024. A comparison of short-read, HiFi long-read, and hybrid strategies for genome-resolved metagenomics. Microbiology Spectrum. 12(4). Doi: https://doi.org/10.1128/spectrum.03590-23.

17. Park S T, Kim J, 2016. Trends in Next-Generation Sequencing and a New Era for Whole Genome Sequencing. International Neurourology Journal. 20(Suppl 2), Pages S76-83. Doi: https://doi.org/10.5213/inj.1632742.371.

18. Ng P C, Kirkness E F, 2010. Whole Genome Sequencing. Methods in Molecular Biology. Pages 215–226. Doi: https://doi.org/10.1007/978-1-60327-367-1_12.

19. Foo JN, Liu JJ, Tan EK, 2012. Whole-genome and whole-exome sequencing in neurological diseases. Nature Reviews Neurology. 8(9), Pages 508–517. Doi: https://doi.org/10.1038/nrneurol.2012.148

20. Retterer K, Juusola J, Cho M T, et al, 2016. Clinical application of whole-exome sequencing across clinical indications. Genetics in Medicine. 18(7), Pag es696–704. Doi: https://doi.org/10.1038/gim.2015.148

21. Kartti S, Bouricha E M, Zarrik O, et al, 2023. Targeted Gene Panel Sequencing Unveiled New Pathogenic Mutations in Patients with Breast Cancer. Bioinformatics and Biology Insights. 17, Doi: https://doi.org/10.1177/11779322231182054.

22. Nawy T, 2013. Single-cell sequencing. Nature Methods. 11(1), Pages 18–18. Doi: https://doi.org/10.1038/nmeth.2771.

23. Zou H, Wu L X, Tan L, et al, 2020. Significance of Single-Nucleotide Variants in Long Intergenic Non-Protein-Coding RNAs. Frontiers in Cell and Developmental Biology. 8. Doi: https://doi.org/10.3389/fcell.2020.00347.

24. Rogozin I B, Saura A, Poliakov E, et al, 2024. Properties and Mechanisms of Deletions, Insertions, and Substitutions in the Evolutionary History of SARS-CoV-2. International Journal of Molecular Sciences. 25(7), Pages 3696. Doi: https://doi.org/10.3390/ijms25073696.

25. Tang Y C, Amon A, 2013. Gene Copy-Number Alterations: A Cost-Benefit Analysis. Cell. 152(3), Pages 394–405. Doi: https://doi.org/10.1016/j.cell.2012.11.043

26. Feuk L, Marshall CR, Wintle RF, et al, 2006. Structural variants: changing the landscape of chromosomes and design of disease studies. Human Molecular Genetics. 15(suppl_1), Pages R57–R66. Doi: https://doi.org/10.1093/hmg/ddl057.

27. Zolotovskaia M, Sorokin M, Garazha A, et al, (2019). Molecular Pathway Analysis of Mutation Data for Biomarkers Discovery and Scoring of Target Cancer Drugs. Methods in Molecular Biology. Pages 207–234. Doi: https://doi.org/10.1007/978-1-0716-0138-9_16.

28. Obeagu E I, 2026. Integrative bioinformatics approaches for early detection biomarkers in ovarian cancer. Annals of Medicine & Surgery. 88(2), Pages 1687–1694. Doi: https://doi.org/10.1097/ms9.0000000000004739.

29. Desai AN, Jere A, 2015. Next-Generation Sequencing for Cancer Biomarker Discovery. Next Generation Sequencing in Cancer Research. 2, Pages 103–125. Doi: https://doi.org/10.1007/978-3-319-15811-2_7.

30. Mehta A, Vasudevan S, Sharma, et al, 2020. Biomarker testing for advanced lung cancer by next-generation sequencing; a valid method to achieve a comprehensive glimpse at mutational landscape. Applied Cancer Research. 40(1), Doi: https://doi.org/10.1186/s41241-020-00089-8.