Journal of Bio-X Research
A SCIENCE PARTNER JOURNAL

REVIEW ARTICLE

Genomic Sequencing: Techniques, Advancements, and the Path Ahead

Mohsina Patwekar[1,2*], Faheem Patwekar[1], A Venkata Badarinath[3], Abdul Ajeed Mohathasim Billah[4], Vamseekrishna Gorijavolu[5], Karthickeyan Krishnan[6], Palani Shanmugasundaram[6], P. Dharani Prasad[7], and A. A. Kazi[8]

[1]Luqman College of Pharmacy, Gulbarga, Karnataka, India. [2]Novel Global Community Educational Foundation, Hebersham, New South Wales, Australia. [3]Department of Pharmaceutics, Santhiram College of Pharmacy, Neravada, Nandyala 518112, Andhra Pradesh, India. [4]Department of Pharmacy Practice, Sri Ramachandra Faculty of Pharmacy, SRIHER (DU), Porur, Chennai, Tamil Nadu, India. [5]Department of Pharmaceutical Analysis, NRI College of Pharmacy, Pothavarappadu (v), Eluru (D.t), Andhra Pradesh 521212, India. [6]School of Pharmaceutical Sciences, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai 600 117, India. [7]Department of Pharmacology, Mohan Babu University, MB School of Pharmaceutical Sciences (Erstwhile Sree Vidyaniketan College of Pharmacy), Tirupati, India. [8]Department of Pharmaceutical Chemistry, N.B.S. Institute of Pharmacy, Ausa, Maharashtra, India.

*Address correspondence to: mohsina.patwekar@gmail.com

Citation: Patwekar M, Patwekar F, Badarinath AV, Billah AAM, Gorijavolu V, Krishnan K, Shanmugasundaram P, Prasad PD, Kazi AA. Genomic Sequencing: Techniques, Advancements, and the Path Ahead. *J. Bio-X Res.* 2025;8:Article 0046. https://doi.org/10.34133/jbioxresearch.0046

Submitted 11 January 2025
Revised 6 April 2025
Accepted 9 April 2025
Published 3 June 2025

The development of genomic sequencing technology, from conventional techniques to state-of-the-art inventions, has greatly improved our understanding of genetic material. This review examines important advancements in sequencing techniques and how they have revolutionized genomics research. High-throughput capabilities made possible by next-generation sequencing (NGS) have enabled quick and affordable genomic analysis. Digital gene expression profiling was made possible by methods such as serial analysis of gene expression (SAGE), whereas long-read capabilities without amplification were analyzed by single-molecule sequencing, as demonstrated by Oxford Nanopore's nanopore-based sequencing and PacBio's single-molecule real-time (SMRT) technology. Synthetic long-read sequencing is one example of a hybrid technique that enhances genome assembly. New techniques, such as epigenetic sequencing, have revealed that DNA alterations are essential for gene control, and spatial transcriptomics has connected gene expression to tissue-specific patterns. Target analysis and knowledge of microbial ecosystems were further enhanced via the use of sophisticated techniques, including metagenomics and CRISPR-Cas9-based sequencing. When combined, these techniques allow researchers to examine microbial communities, transcriptome diversity, genomic structure, and epigenetic changes with new clarity. For example, single-cell sequencing has shown molecular heterogeneity between cells, and long-read sequencing has revealed intricate isoform variants. Personalized medicine has advanced owing to spatial transcriptomics, which targets gene expression in specific organs. Digital sequencing has also improved the sensitivity of mutation identification, transforming the diagnosis of the disease. The convergence of sequencing technologies has ushered in a new era of genomic studies, opening the door to groundbreaking findings in ecology, biology, and medicine. Future developments will improve knowledge of human genetics by further improving sequencing accuracy, affordability, and applicability.

Introduction

Determining the arrangement of nucleotide bases (adenine, cytosine, guanine, and thymine) within a DNA molecule is the main goal of a crucial molecular biology technique known as gene sequencing. Over time, this technique has developed substantially, resulting in several related methodologies and applications. Technological developments, from Sanger sequencing to contemporary techniques, such as next-generation sequencing (NGS), have enhanced gene sequencing technology [1]. NGS-enabled massive parallel sequencing has revolutionized the field by enabling researchers to read whole genomes quickly and affordably. Gene sequencing is crucial for elucidating the genetic foundations of many characteristics and illnesses, examining genetic variations, and identifying pathogenic mutations. DNA sequencing enables the development of customized

medical therapies for the genetic profile of each individual. Sequencing has also made it easier to examine microbial populations and their functions and interactions in microbiology [2]. Genome sequencing of numerous species has shed light on relationships and adaptations that have evolved over time. NGS technologies, such as Illumina, Ion Torrent, and Pacific Biosciences (PacBio), have made it possible for researchers to sequence many DNA fragments simultaneously, leading to high-throughput sequencing with reduced costs and accelerated speed. These techniques require complex bioinformatics tools and algorithms to analyze enormous volumes of sequencing data. Researchers use these tools to compile, annotate, and interpret the data. Managing the enormous amount of data produced by NGS is difficult in terms of processing, storage, and analysis. NGS methods have intrinsic error rates, and discriminating between genuine genomic variants and errors can be difficult. The accessibility of genetic information to an individual presents issues with permission, privacy, and potential abuse [3]. The future perspective long-read sequencing technique aims to provide longer reads, allowing for a more precise assembly of complicated genomes. The single-cell sequencing method enables the sequencing of a single cell, providing insights into cellular heterogeneity. Advancements in sequencing technology, known as metagenomics, have enabled the examination of whole microbial populations from environmental samples [4].

Methodologies for Genome Sequencing

This review discusses the evolution of genome sequencing methodologies, focusing on human genome sequencing, and the ensuing advancements in sequencing technology. The reference human genome was finally completed after several years of using high-throughput, high-capacity DNA sequencing technology and related procedures. This process involved the use of large bacterial artificial chromosomes (BACs), each containing approximately 100 kb of the human genome. These BACs were used for physical mapping and overlapping mapping. BAC-based sequencing involves amplifying BAC clones in bacterial cultures, isolating them, and fragmenting them into size-selected pieces (2 to 3 kb). Subcloning these fragments into plasmid vectors leads to amplification, resulting in oversampling (coverage) [5–9]. A contiguous, high-quality sequence is produced after refinement, which includes gap closure and quality improvement, and the BAC insert sequence is reconstructed using computer-aided assembly. Following the completion of the Human Genome Project, whole-genome sequencing (WGS) emerged as the preferred method for genome sequencing, phasing out BAC-based techniques. The WGS process includes cutting genomic DNA into different sizes, putting it into subclones, and making paired-end sequencing reads for algorithms that combine the reads [10]. This approach enables faster and more efficient genome sequencing; nevertheless, genomes with important polymorphisms or repetitions may remain fragmented post-assembly. The advent of NGS technology rapidly transformed the field of genome sequencing. First developed in 2004, these technologies have substantially influenced numerous biological studies and inspired several peer-reviewed articles. Not only are NGS platforms revolutionizing genome sequencing methods, duration, and cost, they are also opening up new avenues for investigating ancient genomes, ecological diversity, and unknown disease causes [11]. This review outlines the development of WGS, the transition from BAC-based sequencing, and the completion of the reference human genome to the innovative aspects of NGS technology. These developments have revolutionized genome sequencing methods and opened new avenues for research [7–10]. Figure 1 presents an outline of next-generation DNA sequencing technologies focusing on 5 primary platforms: the Applied Biosystems SOLiD System, Helicos Heliscope, Roche/454 FLX, Illumina/Solexa Genome Analyzer, and single-molecule real-time (SMRT) devices from PacBio [12].

Overview of platforms

The Roche/454 FLX, Illumina/Solexa Genome Analyzer, and Applied Biosystems SOLiD systems are 3 prominent sequencing platforms. There are also 2 newer systems: PacBio SMRT and the Helicos Heliscope. These systems represent a complex combination of chemical, optical, hardware, and software engineering. Compared with conventional clone-based sequencing pipelines, these platforms provide simplified sample preparation, save time, and lower the requirement for expensive equipment [13].

Sample preparation and amplification

By attaching platform-specific connectors to blunt-ended pieces directly from the relevant genomic or DNA source, each platform generates fragment libraries. Owing to the inclusion of adapter sequences, polymerase chain reaction (PCR) amplification can preferentially amplify molecules, eliminating the need for bacterial cloning, as in conventional methods. Instruments from Helicos and PacBio are "single-molecule" sequencers that do not need to amplify DNA fragments first [14].

Run time and data yield

Depending on the platform and read type (single or paired ends), next-generation sequencers can operate for up to 10 d. Parallel sequencing operations require longer run times for imaging, unlike the periodic charged-coupled device (CCD) snapshots of capillary platforms. Next-generation systems generate many more sequence reads and total bases per instrument run than do capillary sequencers. Applications such as Biosystems and Illumina can generate tens of millions of readings, according to a previous report [15].

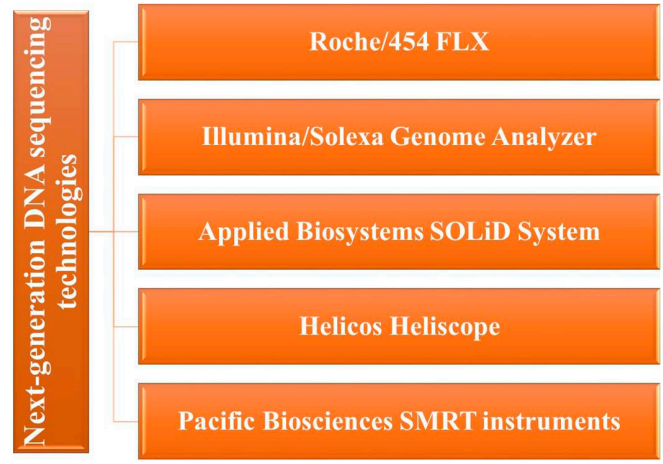

Fig. 1. Overview of advanced DNA sequencing methods.

Scalability and optimal sample preparation

The combination of prolonged run times and effective sample preparation allows a single operator to operate several NGS instruments effectively.

Roche/454 FLX Pyrosequencer

The R454 FLX (Roche Pyrosequencer) is a next-generation DNA sequencing platform that uses pyrosequencing technology. A summary of the key concepts is provided below. The Roche/454 FLX Sequencer Pyrosequencer, which uses pyrosequencing technology, was the first next-generation sequencer to be commercially launched in 2004. Pyrosequencing releases pyrophosphate when DNA polymerase incorporates nucleotides. When the luciferase enzyme is used, a sequence of processes initiated by pyrophosphate results in the generation of light [13–15]. The Roche/454 method involves combining library fragments with agarose beads containing oligonucleotides complementary to adapter sequences. One fragment is then connected to each bead. Each bead–fragment complex is separated into oil–water micelles containing PCR reactants, followed by emulsion PCR amplification. Thus, each bead replicates each DNA fragment approximately 1 million times. Single molecules amplified on the bead surfaces are sequenced simultaneously. Sequencing reactions for emitted light are observed on a picotiter plate (PTP), which comprises hundreds of thousands of single-well beads [16]. The PTP functions as a flow cell during the sequencing process, and pure nucleotide solutions are added progressively, with imaging occurring after each incorporation. The first 4 nucleotides on the adapter segment, which match the sequential flow of nucleotides, aid in base-calling calibration (Table 1). During base calling, homopolymer runs (repeated nucleotides) can result in insertion and deletion errors; however, substitution errors are uncommon because of nucleotide specificity. In an 8-h operation, the FLX instrument generates 100 flows per nucleotide, yielding an average read length of 250 nucleotides [17]. The 454 analysis tool processes the data and then applies quality filtering to generate a standard 100 Mb of high-quality data. Roche/454 FLX reads are suitable for assembling tiny genomes with high quality and contiguity, despite being shorter than the capillary sequencer reads. Because there is no bacterial cloning stage, cloning bias is reduced, leading to more thorough genome coverage. The Roche/454 FLX Pyrosequencer uses pyrosequencing technology for high-throughput DNA sequencing. The steps include PTP-based sequencing, emulsion PCR amplification, and data processing to obtain high-quality reads for combining and analyzing genomes [16–20].

Table 1. Characteristics of the Roche/454 FLX platform

Aspect	Description
Sequencing technology	Pyrosequencing
Read length	Extended read durations in comparison to some platforms
Applications	De novo genome assembly, repetitive sequences
Throughput	Moderate
Read accuracy	Restricted reading precision
Platform limitations	Inferior throughput and precision relative to more recent platforms
Historical Importance	An early NGS platform

Sequencing methodology and technology of the Illumina Genome Analyzer

The Illumina Genome Analyzer uses a sequencing-by-synthesis method and a single-molecule amplification phase. The main ideas are summarized below.

Using an adaptor library created especially for the Illumina Genome Analyzer, the sequencing procedure begins with flow cell-based single-molecule amplification. A device known as a cluster station is used to amplify the oligo-derivatized surface of an 8-channel flow cell. The flow cell facilitates bridge amplification of fragments, which produces DNA copy clusters [21,22]. Each cluster represents a single molecule that initiates amplification. Each cluster contains approximately 1 million copies of the original fragment, which provides the signal strength required for sequencing detection.

Approach to sequencing by synthesis

Illumina uses a sequencing-by-synthesis method in which all 4 nucleotides are simultaneously added to flow cell channels. Each inclusion is a distinct occurrence because each nucleotide has a fluorescent label that is specific to each base, and the 3'-OH group is chemically inhibited. An imaging stage follows each nucleotide incorporation, during which the instrument optics capture images of the clusters in three 100-tile segments. After imaging, the chemical blocking group is removed to obtain each strand ready for the subsequent integration phase by DNA polymerase [21,22]. User-defined instrument settings determine the number of cycles in cycle-based sequencing, which permits read lengths of 25 to 35 bases. DNA polymerase incorporates nucleotides one at a time, and imaging detects the fluorescent signals associated with these incorporations.

Data evaluation and quality assurance

Each read is assigned a sequence and quality grade using a base-calling algorithm. The data of each run are evaluated by a quality control procedure that identifies and eliminates sequences of lower quality [23]. The Illumina Genome Analyzer uses a sequencing-by-synthesis method that concurrently integrates and images nucleotides after single-molecule amplification in a flow cell. High-throughput sequencing with defined read lengths is possible using this technology, and the resulting data are subjected to stringent quality control procedures (Table 2) [24,25].

Applied Biosystems SOLiD Sequencer

The SOLiD Sequencer is produced by Applied Biosystems. This system emphasizes important traits and distinctions from other NGS systems. Like other NGS systems, SOLiD technology uses an adapter-ligated fragment library. Using tiny magnetic beads, emulsion PCR is used to amplify fragments for sequencing [23,26,27].

Ligase-mediated sequencing approach

In contrast to other platforms, SOLiD uses DNA ligase in its sequencing process. Before the DNA fragments are inserted

Table 2. Sequencing process and technology of the Illumina Genome Analyzer

Aspect	Description
Single-molecule amplification on flow cell	Starts with a library adapter made especially for Illumina.
	Amplification is performed on an oligo-derivatized surface of an 8-channel flow cell using a Cluster Station.
	Bridge amplification of fragments occurs inside a flow cell, leading to the formation of DNA copy clusters.
	The original fragment is duplicated about 1 million times inside each cluster, guaranteeing enough signal intensity for detection.
Sequencing via synthesis methodology	Uses sequencing-by-synthesis, whereby flow cell channels are concurrently supplemented with each of the 4 nucleotides.
	The 3'-OH group is chemically obstructed for particular occurrences, and each nucleotide has a fluorescent marker unique to the base.
	Fluorescence signals are obtained using imaging after the incorporation of each nucleotide.
	DNA polymerase removes inhibiting groups from strands to make them ready for the next inclusion.
Cycle-based sequencing	Read lengths (25 to 35 bases) are used to determine the cycle count during sequencing, which is done in steps.
	While imaging records fluorescent signals, DNA polymerase adds nucleotides one at a time.
Data analysis and quality control	Each read is given a sequence and a quality value by the base-calling algorithm.
	A pipeline for quality checking analysis data and filters out low-quality sequences.

into a flow cell slide, magnetic beads are used to amplify the fragments. Ligase-mediated sequencing begins with the adapter sequences and primers joining together. Next, the fluorescently labeled 8-mers are ligated, and the encoded bases are shown by the fluorescent group. To prepare the primer for a further round of ligation, a regeneration step eliminates bases from the ligated 8-mers [28].

Two-base encoding

True single-base variants and base-calling errors are distinguished using SOLiD's innovative 2-base encoding technique. To identify errors, a ligated 8-mer fluorescent group identifies a pair of bases, which may then be aligned to a high-quality reference sequence [29–31].

Read length and data output

The user-specified SOLiD read lengths range from 2,535 base pairs (bp). Each sequencing run produces 2 to 4 Gb of DNA sequence data. Two-base encoding is used for additional quality assessment after base calling, quality filtering, and alignment to a reference genome [32].

Bioinformatics and IT considerations

The shorter read lengths and distinct error models of NGS technologies require modifications to bioinformatics analysis and processes. Implementing these platforms in a production-sequencing context requires an important amount of information technology (IT) infrastructure, data storage, and processing power. NGS has generated a paradigm change that impacts downstream bioinformatics and data production, necessitating new techniques and data visualization tools. The Applied Biosystems SOLiD Sequencer can sequence a large amount of DNA at once using ligase-mediated sequencing and 2-base encoding. The platform's distinct features and data output affect bioinformatics analyses and the IT setup in sequencing centers (Table 3) [23,26,30–35].

Chromatin immunoprecipitation sequencing

The interactions between DNA and proteins can be examined using chromatin immunoprecipitation sequencing (ChIP-seq), which highlights the progress that this technique has made owing to NGS technology. The DNA–protein interactions that are important in gene regulation and other biological processes can be assessed using the ChIP technique. The process includes chemically linking DNA and related proteins, separating nuclei, fragmenting DNA, using antibodies to pick out specific protein–DNA complexes, and then breaking down the crosslinks for further study [33,36]. ChIP-Chip and Restrictions using methods such as quantitative PCR (qPCR) and microarrays (ChIP-chip) enables the analysis of certain gene characteristics. The ChIP-chip has drawbacks such as a poor signal-to-noise ratio and the requirement for duplicates to confirm binding locations [37].

NGS enhancement

DNA sequences produced via ChIP have been advanced using the ChIP-seq NGS technology. A groundbreaking study used Roche technology to assess the position of nucleosomes in the genomic DNA of *Caenorhabditis elegans*. ChIP-seq has made it possible to profile DNA–protein interactions across the entire genome, providing accurate positional profiles of chromatin. Findings from ChIP-seq research of transcription factor binding locations in the human genome, such as NRSF and STAT1, were made possible using ChIP-seq on the Illumina platform. One study used lineage-committed cells and mouse embryonic stem cells to study chromatin architecture and variations in gene expression [38,39]. Researchers have connected lysine methylation marks to cellular lineage potential and gene expression states. ChIP-seq enables the mapping of genome-wide alterations in transcription factor binding and chromatin packing under various environmental conditions. These studies offer a useful method for assessing variations in gene expression and DNA–protein interactions in response to environmental

Table 3. Sequencing methodology and technology used by the Applied Biosystems SOLiD Sequencer

Aspect	Description
Adapter-ligated fragment library and emulsion polymerase chain reaction	Utilizes emulsion PCR with minute magnetic beads and an adapter-ligated fragment library.
Ligase-mediated sequencing method	Uses DNA ligase to facilitate sequencing.
	Magnetic beads with amplified DNA fragments are placed on a flow cell slide.
	Annealing a primer, ligating fluorescently labeled 8-mers, and renewing the primer are the steps in ligase-mediated sequencing.
Two-base encoding	Makes use of 2-base encoding to identify errors.
	On ligated 8-mers, fluorescent groups stand in for 2-base combinations.
Read length and data output	User-specified read lengths of 25 to 35 bp.
	2–4 Gb of DNA sequence data are generated by each sequencing run.
	Two-base encoding is used for quality assessment after base calling, quality filtering, and alignment.
Bioinformatics and IT considerations	Bioinformatics pipeline tweaks are necessary due to shorter read lengths and distinctive error models.
	Production sequencing necessitates a substantial investment in computing power, data storage, and IT infrastructure.
	New methods and data visualization are required due to a paradigm change in data creation and bioinformatics.

stimuli, which has implications for the cellular response. Delineating alterations in transcription factor binding and chromatin architecture can enhance understanding of cellular responses on the basis of gene expression. In summary, ChIP-seq, which integrates ChIP with NGS, has transformed the study of DNA–protein interactions. This approach enables the analysis of transcription factor binding and chromatin architecture across the entire genome, offering insights into gene regulation and cellular responses to environmental stimuli (Table 4) [40,41].

NGS

NGS is a novel technique for DNA sequencing that enables quick and efficient analysis of genetic data. In contrast to conventional Sanger sequencing, NGS methods can sequence millions of DNA fragments simultaneously, enabling quicker and more economical study. NGS, sometimes referred to as SGS or high-throughput sequencing, was a crucial development in DNA sequencing technology. NGS systems have transformed genomics research by enabling the rapid and comprehensive analysis of genetic material at unprecedented depths and scales. NGS platforms use cutting-edge techniques to parallelize and expedite the sequencing process, in contrast to the conventional Sanger sequencing method, which depends on gel electrophoresis and capillary sequencing [2–4,7,8,42].

Parallel sequencing

Massively parallel sequencing is a core characteristic of NGS. Sanger sequencing sequentially processes a single DNA fragment. Conversely, NGS systems process several DNA fragments simultaneously in a highly parallel fashion. This parallelization substantially boosts sequencing throughput, allowing researchers to examine large volumes of genomic material in a single run [43–46].

Sequencing-by-synthesis

The "sequencing-by-synthesis" method is used by the most popular NGS platforms. This method involves attaching DNA fragments to a solid surface or sequencing flow cell. One important advancement in DNA synthesis was the use of fluorescently labeled nucleotides. DNA polymerase releases a fluorescent signal when it adds a complementary nucleotide to the template strand, enabling the identification of the integrated base. Rapid and accurate sequencing is possible using real-time nucleotide incorporation detection [47–49].

Adapters and library preparation

The first step in the sequencing process is to transform the genetic material into a format compatible with the platform. To achieve this, DNA or RNA material is broken into smaller pieces, and adapters are affixed to their ends. Because the fragments contain sequences recognized by the sequencing platform, these adapters can be used to anchor the fragments to the flow cell or surface [50].

Cluster generation

After the adapters are attached, techniques such as bridge amplification and emulsion PCR are used to amplify the DNA fragments. This phase produces cluster-localized amplified collections of thousands of identical DNA fragments. All clusters are sequenced at the same time. Each cluster represents a single DNA molecule [51].

Data generation and output

The sensors of the platform capture the fluorescent signals produced by nucleotide incorporation as sequencing advances. The nucleotide sequence dictates the arrangement of the DNA fragment. In a single run, NGS technology generates millions to billions of short DNA sequences, referred to as "reads". For further analysis, these sequences are compiled and aligned to the reference genome [52–54].

Table 4. ChIP-seq for investigating DNA–protein interactions and improvements via NGS technologies

Aspect	Description
Chromatin immunoprecipitation (ChIP)	Technique for analyzing DNA–protein interactions in the regulation of genes and cellular activities.
	Involves isolation of nuclei, cross-linking DNA–protein complexes, antibody immunoprecipitation, and downstream analysis.
ChIP-chip and limitations	Early research examined particular gene areas using qPCR and microarrays (ChIP-chip).
	Low signal-to-noise ratio and the necessity for replications are 2 drawbacks of ChIP-chip technology.
ChIP-seq: next-gene sequencing enhancement	DNA sequences from ChIP platforms have been improved.
	The Roche platform was used to examine *Caenorhabditis elegans*'s nucleosome location.
	Provides chromatin positioning profiles and enables genome-wide profiling of DNA–protein interactions.
Insights from ChIP-seq studies	The Illumina technology was used to elucidate the binding sites of transcription factors in the human genome, namely, NRSF and STAT1.
	A study on mouse cells examined chromatin organization and changes in gene expression.
	Lysine methylation marks associated with cellular lineage and gene expression.
	Genome-wide mapping of chromatin packing and transcription factor binding under various circumstances.
Implications for cellular responses	ChIP-seq sheds light on how DNA–protein interactions and variations in gene expression occur in response to environmental signals.
	Understanding of biological reactions is improved by mapping transcription factor binding and chromatin structure.

Traditional methods for gene expression analysis

Historically, methods such as microarrays or qPCR have been used to evaluate gene expression levels. qPCR is sensitive and economical for genome-wide gene expression surveys, but it is limited in terms of finding new alternative splicing isoforms. Traditional methods for gene expression analysis have largely relied on understanding how genes control and contribute to diverse biological processes [55]. These methods measure the amount of RNA or proteins produced by particular genes to determine their level of activity. There are 3 basic conventional techniques used for assessing gene expression, and these are discussed below (Fig. 2).

Northern blotting is used to identify and measure specific RNA molecules in a sample. This technique involves size-based separation of RNA molecules using gel electrophoresis, membrane transfer, and hybridization with labeled DNA or RNA probes. This technique provides details regarding the size and quantity of particular RNA transcripts [56,57].

Reverse transcription PCR (RT-PCR) is a common technique for assessing the expression of particular RNA transcripts. Reverse transcriptase enzymes are required to transform RNA into complementary DNA (cDNA). cDNA is subsequently amplified by PCR using gene-specific primers. RT-PCR provides quantitative data regarding the initial concentration of RNA in samples [52,58].

qPCR, an improvement over RT-PCR, enables accurate measurement of RNA molecules. It uses labeled probes or fluorescent dyes that bind to DNA to track the amplification of DNA in real time during PCR. The initial concentration of RNA in the sample is estimated by measuring the cycle at which the fluorescence signal reaches a threshold level.

Serial analysis of gene expression (SAGE) measures the amount of gene expression by sequencing short tags from mRNA transcripts. These tags are made into a library that is subsequently sequenced. The frequency of each tag in the library corresponds to how common its related transcript is in the sample [59,60].

Microarrays are a high-throughput equipment that allows for the simultaneous analysis of the expression of hundreds of genes in a single experiment. Labeled cDNA or RNA from the sample is hybridized to immobilized DNA or RNA probes on a solid surface. The expression level of each gene is determined by hybridization [61].

Protein-based methods are also used since protein abundance does not always directly correlate with mRNA levels. Common techniques for measuring protein expression include Western blotting, immunohistochemistry, and enzyme-linked immunosorbent assay [62].

Although these conventional techniques provide useful information on gene expression patterns, they frequently require considerable time, effort, and sample input. They may also exhibit throughput, dynamic range, and sensitivity restrictions. Our understanding of the transcriptome and its regulation has improved as a result of the introduction of NGS and high-throughput technologies, which have produced more comprehensive and scalable approaches for investigating gene expression (Table 5) [60–62].

SAGE

SAGE and its successors digitally represent gene expression levels using DNA sequencing. SAGE is effective in detecting gene expression levels below microarray detection thresholds. However, the cost of DNA sequencing restricts its application. SAGE is a molecular biology method used for characterizing and measuring gene expression patterns in large quantities (Table 6). Developed in the early 1990s, SAGE provides a comprehensive view of the transcriptome by generating short-sequence tags that represent

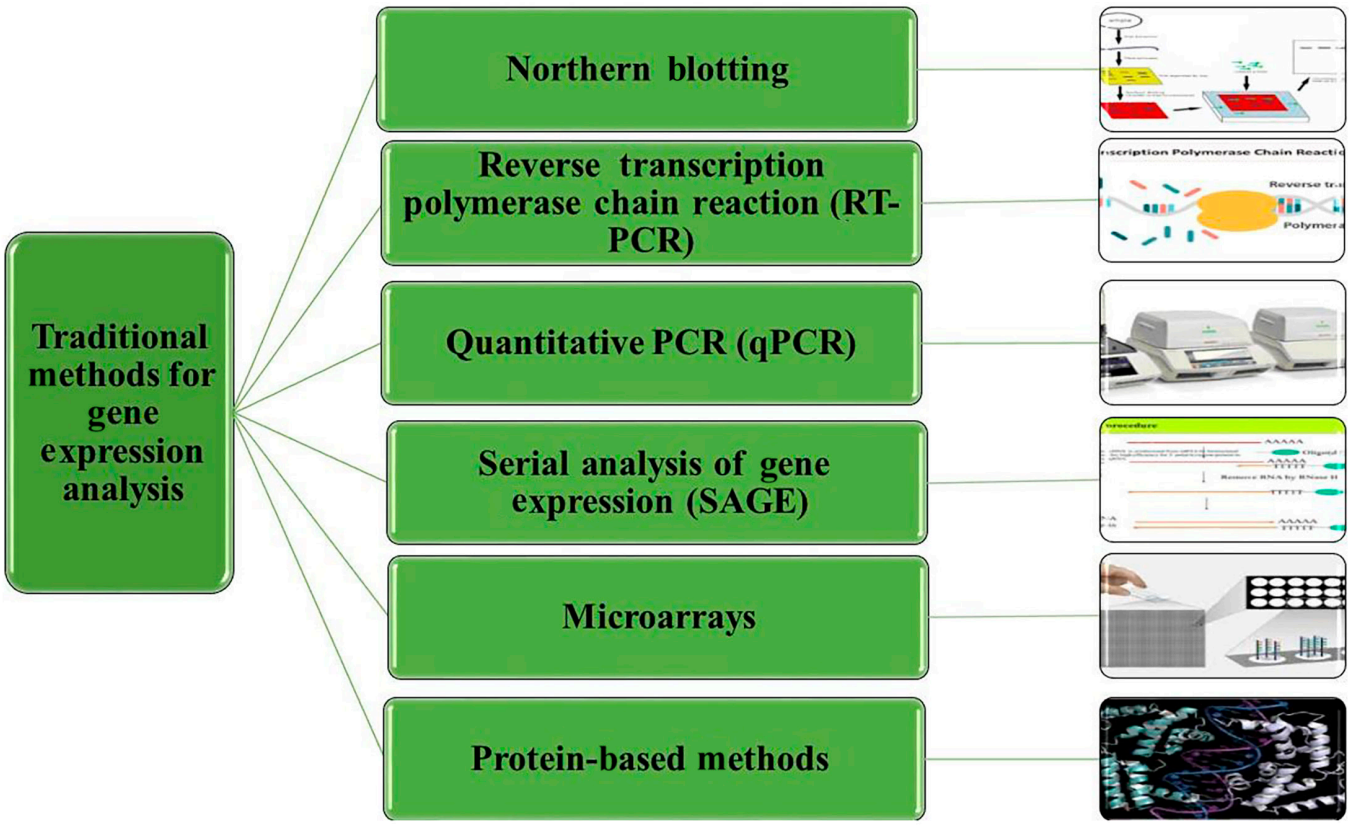

Fig. 2. Traditional methods for gene expression analysis.

individual mRNA molecules. This technology goes beyond some of the drawbacks of conventional techniques, such as Northern blotting and microarrays, which may be biased toward known sequences or have a constrained dynamic range [63]. In SAGE, the initial step involves building a cDNA library from the target mRNA population. Reverse transcriptase is used to extract mRNA from cells or tissues, which are subsequently transformed into cDNA. Each cDNA molecule is given a particular "anchoring enzyme" recognition site. Sequencing and tag extraction are performed using a website. The anchoring enzyme then breaks the cDNA molecules into fragments. The anchoring enzyme assigns a distinct 10- to 14-nucleotide tag to each fragment, which corresponds to a specific location in the original mRNA. Each cDNA molecule has its own concatenated tag sequence because of the ligation of these tags, which act as molecular barcodes. A SAGE library is created by cloning concatenated tag sequences into a vector. The quantity of each clone in the library corresponds to the degree of expression of the relevant gene, and each clone represents a distinct mRNA molecule. SAGE libraries are sequenced using conventional DNA sequencing techniques [64]. The identities and frequencies of the sequenced tags are then analyzed. The abundance of each tag is inversely correlated with the concentration of the original mRNA molecules in the sample. In gene expression profiling, researchers can calculate the level of gene expression by counting the instances of each tag. Bioinformatic tools and databases link tag sequences to known genes, enabling gene identification and annotation. SAGE offers an unbiased view of the transcriptome because it does not rely on predesigned probes or arrays. The tag frequencies represent the gene expression levels of the sample. Accurate quantification is made possible by counting each tag individually. SAGE locates new regulatory components and transcripts. Building and sequencing SAGE libraries is expensive and technically difficult. For highly expressed genes, short tags may clash and lead to erroneous quantification. NGS platforms have largely replaced SAGE despite their groundbreaking nature. Owing to their increased throughput, longer read lengths, and enhanced sensitivity, NGS techniques offer more thorough and effective gene expression analysis [65,66].

Integration of NGS

Studies that integrate NGS with SAGE tagging or cDNA sequencing use Roche/454 technology [8,17,51,59]. Equipment from Applied Biosystems and Illumina may also be used for shorter read durations. Next-generation technologies enable individual sequencing of SAGE tags without the need for tag concatenation. Merging coisolated and sequenced mRNA populations from the same cells using ChIP-derived DNA sequencing may be a possibility. Future technologies will enable investigations with high sensitivity and low biomolecule input. In summary, NGS offers a potent and reasonably priced method for analyzing transcriptomes and gene expression. By enabling the sequencing of individual tags, it enhances established methodologies such as SAGE and opens new avenues for integrated research that integrates DNA and mRNA data [67–69].

NGS plays a crucial role in the identification and analysis of noncoding RNAs (ncRNAs). In recent years, a fascinating new field of biological research has focused on the identification and functional examination of ncRNA systems in numerous species. ncRNAs, which were first identified in plants, are now known to be essential for regulating gene expression in animal systems. NGS techniques have had a crucial impact on the identification of novel ncRNAs from a variety of classes in numerous animal

Table 5. Comparison of traditional methods for gene expression analysis

Method	Principle	Advantages	Limitations
Northern blotting	RNA can be separated using gel electrophoresis and hybridization	Provides size and abundance information	Limited sensitivity and low throughput
RT-PCR	RNA to cDNA conversion and PCR amplification	Quantitative, specific, relatively fast	Confined, target-specific multiplexing
Quantitative PCR (qPCR)	Fluorescence-based real-time monitoring of DNA amplification	High sensitivity, quantitative, real-time analysis	Limited multiplexing, adequate controls are needed
Serial analysis of gene expression	Short tags from mRNA transcripts are sequenced	Comprehensive, quantitative	Sequencing requirements and tag-based restrictions
Microarrays	Labeled cDNA/RNA hybridization to immobilized probes	High-throughput, simultaneous gene analysis	Dynamic range limitation and cross-hybridization
Protein-based methods	Detection and measurement of protein concentrations	Direct measurement of protein abundance	Possibly not related to mRNA levels

Table 6. SAGE

Aspect	Description
Technique type	Molecular biology technique for gene expression analysis
Development period	Early 1990s
Purpose	High-throughput quantification and profile gene expression patterns
mRNA extraction	mRNA is removed from the cells or tissues.
cDNA synthesis	Reverse transcriptase is used to transform mRNA into cDNA.
Anchoring enzyme site addition	A particular anchoring enzyme recognition site is added to the cDNA. The anchoring enzyme digests the cDNA, producing small fragments.
Tag generation	A SAGE library is created by separating 10–14 unique nucleotide tags from fragments, ligating the tags together, and cloning the concatenated tag sequences into a vector.
Tag extraction and concatenation	High-throughput quantification and profile gene expression patterns
Cloning	mRNA is removed from the cells or tissues
Sequencing	High-throughput quantification and profile gene expression patterns
Tag analysis	mRNA is removed from the cells or tissues.
Gene mapping and annotation	Reverse transcriptase is used to transform mRNA into cDNA.
Expression quantification	A particular anchoring enzyme recognition site is added to the cDNA. The anchoring enzyme digests the cDNA, producing small fragments.
Advantages	A SAGE library is created by separating 10–14 unique nucleotide tags from fragments, ligating the tags together, and cloning the concatenated tag sequences into a vector.
Limitations	High-throughput quantification and profile gene expression patterns
Current relevance	mRNA is removed from the cells or tissues.

taxa. Owing to NGS, new ncRNA classes have been discovered in organisms that were previously unknown to possess them. Owing to the diversity of ncRNA gene sequences, it is difficult to anticipate their sequences using computational methods alone, making sequencing an essential tool for their discovery. Converting processed ncRNAs into NGS libraries is difficult because of their distinctive architectures. Despite these difficulties, ncRNA molecule characterization has advanced substantially. The capacity and affordability of NGS platforms have enabled rapid and continuous ncRNA discovery. The numerical results from NGS make it possible to identify changes in the levels of ncRNA expression that are linked to environmental factors and medical conditions. Characterization and discovery of ncRNAs improves genome annotation for sequenced organisms. This improved annotation aids in assessing the effects of mutations across the genome, particularly in model organisms and humans [70,71] (Fig. 3). In conclusion, NGS is essential for the identification and research of ncRNAs because it sheds light on

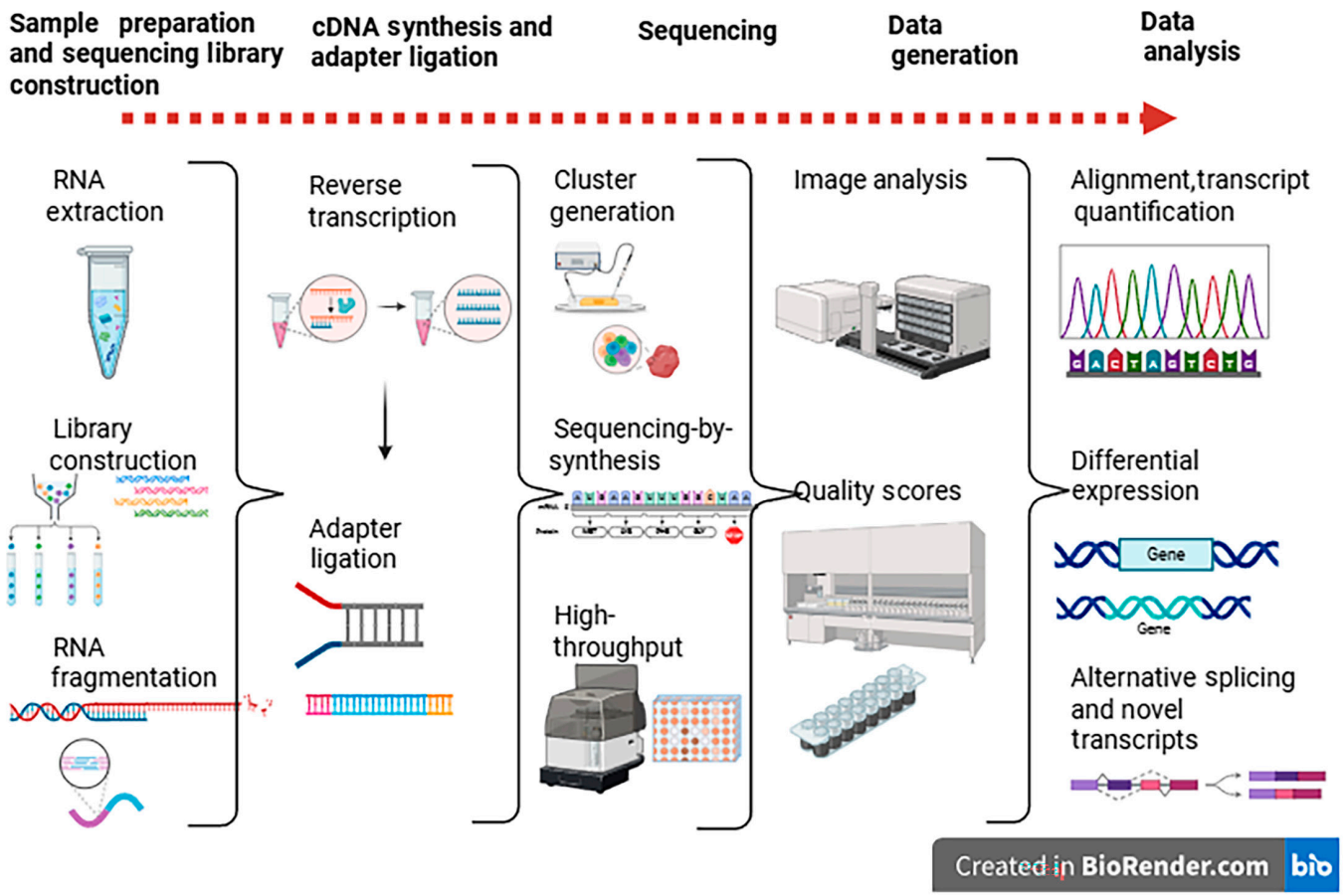

Fig. 3. Incorporation of NGS in gene expression analysis.

the control of genes, causes of disease, and genome annotation. Traditionally, studies on ancient genomes have been limited to mitochondrial DNA sequencing and PCR amplification of specific regions. Direct nuclear genome sampling and sequencing from prehistoric remains, such as those of cave bears, mammoths, and Neanderthals, has been made possible by NGS. Despite technical difficulties and contamination problems, important characterizations have been made, including the sequencing of a million nucleotides of the Neanderthal genome. Metagenomics characterizes biodiversity on Earth by obtaining a direct genetic sample of the surrounding environment. Analysis of the 16S rRNA gene and subcloning are 2 traditional approaches; however, NGS has revolutionized metagenomic research by enabling the deep sequencing of many habitats [72]. NGS-enabled metagenomic studies have investigated numerous habitats, including the ocean, acid mine sites, soil, coral reefs, and human microbiome. Microbial populations that live in different areas of the human body are referred to as human microbiomes. The human microbiome, especially that of the lower intestine, has been well characterized by NGS, which has demonstrated intricate interactions between human hosts and their microbial symbionts. The Human Microbiome Initiative was founded to learn more about how the human microbiome affects health. In summary, the study of ancient genomes, metagenomics, and the human microbiome has been made possible by NGS technology, which has revolutionized genomic research. As a result of these developments, our knowledge of genetic variety, ecological interactions, and their effects on human health has increased. The revolutionary impact of NGS on genomics has enabled techniques such as transcriptomics, epigenomics, metagenomics, and WGS. Understanding genetic variation, disease mechanisms, and intricate biological processes has advanced owing to their high-throughput capabilities. NGS has substantially advanced research but also poses difficulties for data analysis, storage, and interpretation. Processing and making sense of the enormous volumes of data produced by NGS requires computer power and bioinformatics pipelines. From fundamental research to clinical diagnostics, NGS has had a profound influence on several sectors. This has made it possible to practice personalized medicine, in which a patient's genetic profile can inform treatment choices (Table 7) [68,70–72].

Novel method of gene sequencing

Single-molecule sequencing

Advanced platforms, such as Oxford Nanopore Technologies (ONT) and PacBio, use single-molecule sequencing. The PacBio method involves monitoring DNA polymerase in real time, as it incorporates fluorescently labeled nucleotides and enables longer read lengths. The Oxford Nanopore uses nanopores built into membranes that allow DNA strands to pass through while producing distinctive electrical signals that match the base order. An innovative technique in genomics, known as single-molecule sequencing, permits direct sequencing of individual DNA or RNA molecules without the need for amplification or cloning. In contrast to conventional sequencing approaches that depend on PCR or other amplification techniques, single-molecule

Table 7. Integration of NGS in gene expression analysis

Sample preparation and sequencing library construction	
RNA extraction	Total RNA is extracted from cells or tissues while preserving the mRNA fraction.
Library construction	RNA-seq or mRNA-seq methods are used to build cDNA libraries from the extracted RNA.
RNA fragmentation	Techniques like chemical or ultrasonic fragmentation are frequently used to break apart RNA molecules.
cDNA synthesis and adapter ligation	
Reverse transcription	The fragmented RNA is reverse transcribed into complementary DNA (cDNA) using reverse transcriptase and random primers.
Adapter ligation	Adapters with unique barcodes are attached to the cDNA fragments to multiplex samples in a single sequencing run.
Sequencing	
Cluster generation	After the sequencing library is put onto a flow cell, clusters of DNA fragments are produced using bridge amplification.
Sequencing-by-synthesis	Fluorescently tagged nucleotides are added to the flow cell during each incorporation cycle, and the light signals that are released are captured.
High-throughput	Sequencing millions of clusters simultaneously generates large amounts of short-sequence reads.
Data generation	
Image analysis	The flow cell is photographed at each cycle, and the light signals it emits are converted into nucleotide calls.
Quality scores	A quality score that reflects the dependability of the called base is assigned to each base call.
Data analysis	
Read alignment	Sequence readings are aligned to a reference genome using alignment techniques.
Transcript quantification	The amount of reads that have been mapped to each gene is used to determine its degree of expression.
Differential expression	NGS data can be used to identify genes that show differential expression in different contexts.
Alternative splicing and novel transcripts	Alternative splicing events and new transcript isoforms can be found using NGS.
Advantages of NGS in gene expression analysis	
Comprehensive	NGS discovers rare transcripts and scans the entire genome.
Digital nature	Accurate comparisons are made possible by digital quantitative data.
Higher dynamic range	Genes with high and low expression levels in the same sample can be distinguished using NGS.

sequencing attempts to overcome the limitations caused by amplification biases and errors. Advances in nanotechnology, fluorescence detection, and real-time monitoring have made this procedure practical (Table 8) [73,74].

Real-time monitoring

Real-time monitoring of the integration of nucleotides into a developing DNA strand is performed using single-molecule sequencing systems. This phenomenon suggests that DNA polymerase generates sequence information by adding individual nucleotides. Nanopore sequencing is a prominent example of single-molecule sequencing. This technique involves embedding a membrane with a nanopore and a hole with a nanometer size. When DNA strands thread through the nanopore, variations in electrical current as nucleotides traverse the pore determine the sequence. Another method uses fluorescently labeled nucleotides for detection. When each nucleotide is added to the expanding strand, the fluorescent signal is recognized. The DNA molecule sequence is identified by observing the timing of the fluorescent signal. Applications include WGS, targeted sequencing of specific genes or regions, epigenetic studies, and RNA sequencing (RNA-seq) [75]. Single-molecule sequencing has several applications, including the study of intricate genomic regions, structural changes, and repetitive sequences. Compared with several other NGS technologies, single-molecule sequencing platforms often result in extended read lengths. This technique is useful for the resolution of challenging genomic areas and detection of structural changes. The amplification bias and errors produced by PCR amplification are minimized because individual molecules can be sequenced without amplification. Direct sequencing preserves true biological information by allowing direct sequencing of native DNA or RNA molecules without the requirement for cloning or PCR amplification. Single-molecule sequencing techniques may have higher error rates than other sequencing techniques. However, advancements in data analysis pipelines and base-calling algorithms have addressed this problem. Current single-molecule sequencing platforms may impact the

Table 8. Novel methods of gene sequencing

Method	Description
Single-molecule sequencing	Uses nanopores or real-time DNA polymerase monitoring for extended read lengths.
Third-generation sequencing	Larger read lengths are provided, assisting in complex genomic analysis and structural variances.
Synthetic long-read sequencing	Generates synthetic long reads for structural insights by combining short-read sequencing data.
Spatial transcriptomics	Allows for the study of gene expression in tissue sections with spatial resolution.
Epigenetic sequencing	Elucidates protein–DNA interactions and DNA methylation (bisulfite sequencing and ChIP-seq).
Long-read isoform sequencing	Focuses on understanding alternative splicing by sequencing full-length RNA isoforms.
CRISPR-Cas9-based sequencing	Uses CRISPR-Cas9 to target particular genomic areas of interest before sequencing.
Metagenomic sequencing	Uses direct DNA sequencing to analyze microbial populations from environmental samples.
Single-cell sequencing	Profiling for information on cellular dynamics and heterogeneity at the single-cell level.
Digital sequencing	Enables precise analysis by providing absolute quantification of DNA or RNA molecules.

cost-effectiveness of large-scale projects because of their lower throughput compared with other technologies [76]. DNA is threaded through a nanopore using the nanopore sequencing method (ONT), which includes monitoring changes in ionic current as nucleotides travel through. Real-time sequencing with extended read length can be performed. Real-time sequencing (PacBio SMRT sequencing) can be performed when a DNA polymerase attached to a chip is used in PacBio SMRT sequencing to observe DNA synthesis in real time and facilitate extended readings. Single-molecule sequencing has advanced genomic research by enabling longer read lengths, direct sequencing, and examination of challenging genomic regions. Compared to other NGS technologies, single-molecule sequencing platforms frequently have longer read lengths, which makes them useful for the resolution of challenging genomic areas and detection of structural changes. The amplification bias and errors produced by PCR amplification are minimized because individual molecules can be sequenced without amplification. Direct sequencing preserves true biological information by allowing direct sequencing of native DNA or RNA molecules without the requirement for cloning or PCR amplification. Single-molecule sequencing techniques may have higher error rates than other sequencing techniques. However, advancements in data analysis pipelines and base-calling algorithms have addressed this problem. The cost-effectiveness of large-scale projects may be affected by the throughput of current single-molecule sequencing platforms, which may be less than that of other technologies [69]. For example, DNA is threaded through a nanopore using the nanopore sequencing method (ONT), which includes monitoring changes in the ionic current as nucleotides travel through, and real-time sequencing of long-read lengths is possible. Real-time sequencing (PacBio SMRT sequencing) can be performed when a DNA polymerase immobilized on a chip is used in PacBio SMRT sequencing to monitor DNA synthesis in real time and enable lengthy readings. By enabling larger read lengths, direct sequencing, and analysis of difficult genomic areas, single-molecule sequencing has expanded the field of genomics research (Table 9). Despite its challenges, continuous advancements have enhanced the technology and expanded its applications [75,76].

Third-generation sequencing

Third-generation sequencing (TGS) technologies such as PacBio and Oxford Nanopore have the advantage of extended read lengths, which allow the sequencing of large regions of DNA without the need for assembly. This approach is particularly helpful for analyzing structural variants in complicated genomes. In DNA sequencing technology, TGS, also known as long-read sequencing, provides important advancements over second-generation sequencing (SGS) methods. Compared with SGS technology, TGS platforms promise to yield substantially longer read lengths, enabling the direct sequencing of longer DNA segments. This function is particularly useful for overcoming obstacles caused by repeated sequences, structural changes, and complicated areas of the genome. TGS technologies that focus on single-molecule sequencing do not require amplification to sequence individual DNA or RNA molecules. This method eliminates the biases and errors caused by amplification. One of the distinguishing characteristics of TGS platforms is their capacity to provide read lengths noticeably longer than those of SGS systems. These lengthy readings are useful for resolving intricate genomic architectures and locating widespread variants. TGS platforms frequently allow real-time monitoring of the sequencing process, similar to single-molecule sequencing techniques [77]. This phenomenon suggests that DNA polymerase generates sequence information by adding individual nucleotides. Numerous genomic investigations, including de novo genome assembly, structural variation detection, haplotype phasing, and epigenetic modification studies, have utilized TGS. TGS works especially effectively for genomes with intricate structures, such as plant genomes, and specific alterations linked to illness. One of the main benefits of TGS is the production of long reads that frequently cover thousands to tens of thousands of bases. This allows for direct sequencing of massive structural differences, repetitive sequences, and complex genomic areas. Like single-molecule sequencing, TGS technologies do not require PCR amplification, which minimizes the biases and errors caused by amplification. Compared with SGS platforms, some TGS technologies may have greater per-base error rates. Nevertheless, this problem has been addressed through error-correction techniques and advancements in base-calling algorithms. The throughput of TGS platforms, which may be less than that of

Table 9. Quantitative performance metrics for short-read (Illumina) and long-read (PacBio/ONT) sequencing technologies

Metric	Short-read sequencing (Illumina)	Long-read sequencing (PacBio/ONT)
Read length	Typically, 50–300 bp	Thousands of base pairs (kb) or more
Error rate (%)	~0.1% (Illumina HiSeq, NovaSeq)	5–15% (PacBio HiFi <1%, ONT ~5–15%)
Throughput (Gb/run)	100–600 Gb (NovaSeq)	10–100 Gb (PacBio Sequel II, ONT PromethION)
Cost per Gb	$5–100	$30–500
Time to completion	1–2 d	6–48 h (ONT faster, PacBio longer for HiFi)
Consensus accuracy (%)	>97.5% (after read correction)	~99.9% (PacBio HiFi), ~98–99% (ONT with polishing)

high-throughput SGS platforms, could affect the time and expense of sequencing projects [78]. SMRT sequencing by PacBio uses DNA polymerase immobilized on a chip to track DNA synthesis in real time. Compared with SGS technologies, it provides extended readings but with greater per-base error rates. To detect changes in ionic current as nucleotides move through a nanopore, ONT (ONT) uses a nanopore sequencing technique to thread DNA through a nanopore. Real-time sequencing yields long reads and is portable for field use. The direct sequencing of lengthy DNA fragments made possible by TGS has completely changed genomic research. Given that this has enabled improvements in complicated genomic region knowledge, structural variation identification, and genome assembly, it is a crucial tool in contemporary genomics [77,78].

TGS technologies

PacBio and Oxford Nanopore have the advantage of extended read lengths, which allow for the sequencing of large regions of DNA without the need for assembly. This approach is particularly helpful for analyzing structural variants in complicated genomes. In DNA sequencing technology, TGS, also known as long-read sequencing, provides important advancements over SGS methods. Unlike genuine long-read sequencing methods that directly generate extensive reads, synthetic long reads (SLRS) approaches create synthetic long reads by amalgamating data from many short-read sequences. This approach facilitates the analysis of complex genomic regions and structural variations by bridging the gaps between short- and long-read sequencing [78]. The creation of short-read sequences is the first step in the SLRS. High-throughput sequencing technologies, such as Illumina and Ion Torrent, can be used. Typically, these brief reads contain a few hundred base pairs. These small reads are assembled computationally to obtain the real power of the SLRS. Synthetic long reads are constructed by aligning and overlapping short reads using sophisticated bioinformatics tools and algorithms. Computational approaches construct longer sequences that span thousands of base pairs by deliberately overlapping short reads that span common regions. Complex genomic characteristics, such as repeated regions and structural changes, can be captured by these synthetic long reads. SLRS is used in a number of genomic investigations, such as de novo genome assembly, structural variation characterization, haplotype phasing, and the discovery of genetic variations in disease-related areas. Compared with specialized long-read sequencing methods, SLRS is more affordable because it makes use of existing short-read sequencing platforms. The computational nature of SLRS allows for flexibility in study design by modifying the parameters and algorithms for different genomic areas [79]. The quality of short-read sequences and assembly algorithms affects the accuracy of synthetic long reads. Short-read errors can result in inaccurate reconstruction. SLRS can address various issues in intricate genomic regions but may still be limited in the cases of highly repeated sections and intricate structural changes. With the 10x Genomics Linked-Read Sequencing method, large DNA molecules may be broken into smaller pieces for short-read sequencing while retaining their spatial information. The longer sequences are then reconstructed using computational techniques. With the aid of computational techniques and the chromatin conformation capture technique Hi-C, chromosome-level structures can be recreated by learning more about the spatial arrangements of genomic areas. Synthetic long-read sequencing is a flexible method that combines the effectiveness of short-read sequencing and computer assembly to produce useful synthetic long reads. This method has created new opportunities for investigating intricate genomic features and has enhanced our knowledge of genome structure and variation.

New developments in long-read sequencing technologies

Technologies for strand-seq high-resolution haplotype phasing and structural variation detection are made possible by the single-cell DNA sequencing technique known as "strand-seq", which maintains the orientation of DNA strands. Strand-seq has recently been combined with long-read sequencing technologies, such as the Oxford Nanopore and PacBio, enabling better structural variant characterization and genome assembly. It is a crucial tool for resolving intricate genomic regions, including massive inversions and repetitive sequences. 10x Genomics invented linked-read sequencing by reconstructing long-range genetic information using barcoded short reads. Despite 10x Genomics discontinuing this technology, new methods are being developed. The most recent developments have focused on combining linked reads with hybrid sequencing techniques and single-cell multiomics platforms. These developments have enhanced the discovery of rare variants, the phasing of complicated genomic variants, and the highly accurate de novo assembly of genomes. By maintaining the spatial arrangement of chromatin, Hi-C 2.0, an advancement in chromosomal conformation capture technology, improves the precision of genome scaffolding. The resolution of genome assemblies is greatly improved by this technique, especially for large and complex organisms, such as plants and mammals. Hi-C 2.0 is a useful technique for de novo genome assembly and chromatin architecture investigations because it enhances the assembly of repeated areas and structural variants when paired with long-read sequencing [80,81].

Spatial transcriptomic

Spatial transcriptomics is a cutting-edge method that combines high-throughput sequencing with spatial information to map gene expression patterns within tissue sections. Unlike traditional bulk RNA-seq, this method allows for spatially resolved transcriptomics, enabling researchers to understand gene expression within tissue sections [82].

Key features and workflow

The first step in spatial transcriptomics is to prepare tissue sections from a biological sample, such as a tissue biopsy. For further analysis, slices are mounted on a specific slide or substrate. A specialized matrix containing distinct spatial barcodes and capture probes processes the tissue sections in spatial transcriptomics. These capture probes allow for spatial tagging of polyadenylated mRNA molecules by binding to them only in the tissue. After capture, the mRNA molecules undergo reverse transcription to produce cDNA, which was used for cDNA synthesis and sequencing library preparation. Both the sequencing data and the spatial barcode data are obtained from cDNA. A sequencing library is constructed by adding sequencing adapters to the cDNA [83]. The sequencing library is subjected to this process, producing millions of short reads that represent the transcriptome profile of the tissue section. Each cDNA fragment has geographical barcodes that offer a spatial context. To decode the geographical barcodes and link them to particular transcriptome profiles, the sequenced data are subjected to computer analysis. This enables the reconstruction of spatially specific gene expression patterns. Spatial transcriptomics provides single-cell resolution data, allowing researchers to analyze gene expression patterns in specific cells within their native tissue context. Spatial visualization makes it possible to determine how genes are expressed in the context of the anatomy of tissue sections, which sheds light on how cells are arranged and interact. Spatial transcriptomics enables the discovery of new cell types or subpopulations by revealing unique gene expression patterns in specific tissue locations. Because the precise alignment of sequence reads to physical locations is needed, processing and analyzing spatial transcriptomic data can be challenging. High spatial resolution may be difficult to achieve because of the size of the tissue sections and the spatial precision of the capture probes. Spatial transcriptomics can reveal how different gene expression patterns change within intricate tissue contexts, thereby revealing heterogeneity and functional connections. This method can assist in identifying the spatial variations in gene expression linked to diseases, offering information on the development of the disease and potential therapeutic targets. Researchers have used spatial transcriptomics to study gene expression dynamics during organogenesis, tissue regeneration, and embryonic development. A popular spatial transcriptomics technique is the Visium platform from 10x Genomics [84]. A combination of capture probes, spatial barcoding, and high-throughput sequencing is used to produce geographically resolved transcriptome maps. By enabling examination of the diversity and interconnections of cells within complex tissues, spatial transcriptomics has revolutionized the understanding of how genes are expressed. It is used in a variety of disciplines, from fundamental research to clinical studies, and contributes to a more thorough characterization of biological systems [82–84].

Epigenetic sequencing

Epigenetic modifications substantially affect gene regulation. Techniques, such as ChIP-seq (to study protein–DNA interactions) and bisulfite sequencing (to study DNA methylation), can help elucidate epigenetic control. The aim of epigenetic sequencing is to clarify epigenetic modifications, heritable shifts in gene expression that occur without changing the DNA sequence. Histone abnormalities and DNA methylation are 2 examples of alterations essential for gene control during both development and disease. Understanding the epigenetic landscape of the genome using epigenetic sequencing facilitates the comprehension of the ways in which epigenetic modifications affect cellular processes. Bisulfite sequencing is commonly used to map DNA methylation patterns. Bisulfite treatment converts unmethylated cytosine in DNA into uracil while keeping methylated cytosine unchanged [85]. Sequencing identifies the converted cytosine as thymine, enabling differentiation between methylated and unmethylated cytosine. ChIP and high-throughput sequencing are used to identify histone modifications and DNA-binding protein interaction sites. Antibodies targeting specific histone modifications or proteins of interest are used to amplify chromatin fragments, which are subsequently sequenced to identify their genetic origins. To discover regions of open chromatin, ATAC-Seq (assay for transposase-accessible chromatin sequencing) uses a transposase enzyme to introduce sequencing adapters into accessible chromatin sites. Elucidating chromatin accessibility and regulatory components is possible by using this method [86]. DNA methylation (bisulfite sequencing and ChIP-seq) is a comprehensive technique for analyzing DNA methylation with single-base precision throughout the genome. High-throughput sequencing is used after bisulfite treatment to enable genome-wide characterization of DNA methylation patterns. Using certain antibodies, methylated DNA fragments are collected via MeDIP-Seq (methylated DNA immunoprecipitation sequencing) before sequencing. Compared with bisulfite sequencing, this technique offers information regarding DNA methylation patterns at a lower resolution [87]. Epigenetic sequencing methods elucidate intricate regulatory mechanisms governed by DNA methylation and histone modifications. Epigenetic changes have been linked to several illnesses, including cancer [88]. Understanding the causes of disease and identifying potential treatment targets is possible using epigenetic sequencing. Epigenetic alterations are essential for cellular differentiation and development. Understanding these mechanisms is facilitated using epigenetic sequencing. Epigenetic sequencing enables comprehensive profiling of epigenetic alterations throughout the whole genome, providing an integrated perspective of their distribution. Using methods such as bisulfite sequencing, it is possible to identify specific methylated cytosines with single-base resolution. Extensive bioinformatics analysis is needed to reliably identify altered areas when processing and analyzing epigenetic sequencing data. Understanding the functional importance of DNA methylation requires differentiation between its occurrence in different sequence contexts (CpG and CpH). Methods such as single-cell epigenome sequencing, which allows for the study of epigenetic heterogeneity at the single-cell level, and epigenetic sequencing continue to advance. Integrating epigenetic data with other omics data enables a more thorough understanding of gene regulatory networks. In summary, epigenetic sequencing methods have revolutionized the ability to investigate the dynamic epigenetic alterations that affect cellular function and gene expression. These methods offer important new insights into growth, illness, and the complex regulatory systems that control the function of the genome [89].

Long-read isoform sequencing

Full-length RNA isoforms are the focus of sequencing methods such as PacBio Iso-Seq. The understanding of alternative splicing and isoform diversity depends on this information. Long-read isoform sequencing, a specialized NGS method, aims to precisely characterize the entire range of isoforms that genes can create. The complexity of alternative splicing and the diversity of gene isoforms are frequently difficult to capture using conventional short-read sequencing techniques [90]. Using long-read isoform sequencing techniques, full-length transcripts and isoforms can be identified by directly sequencing larger RNA or DNA fragments.

Key techniques and concepts

For long-read isoform sequencing, TGS technologies from ONT and PacBio can be used. In contrast to traditional short-read platforms, these platforms generate substantially longer readings. TGS technologies use single-molecule sequencing, enabling the direct sequencing of individual RNA or DNA molecules without amplification (Table 9) [91]. This minimized the biases introduced by amplification processes. As part of long-read isoform sequencing, full-length cDNA sequencing involves ordering full-length cDNA molecules, which are copies of RNA transcripts that work with each other. The full transcript structure, including alternative splicing events, can be better understood using this method. Long-read isoform sequencing makes it possible for researchers to thoroughly examine alternative splicing patterns, including complicated isoform variants that were previously difficult to identify [92]. Gene isoform diversity is a method that makes it possible to find new and uncommon isoforms, which helps us to better understand the diversity of gene isoforms. Long-read isoform sequencing, which captures full-length transcripts and aids researchers in discovering novel transcript variants and fusion genes, highlights the intricacy of the transcriptome. Long-read isoform sequencing delivers exact whole RNA molecular sequences that enable researchers to precisely identify transcript structures and isoforms. The detection of complicated isoforms is especially helpful for locating complex isoforms that contain numerous alternative splicing events or structural changes. Short-read platforms have lower error rates than TGS platforms. Although advances in bioinformatics and sequencing chemistry are reducing errors, accurate isoform reconstruction remains a challenge. Current long-read sequencing methods have lower throughput than other short-read platforms, limiting the extent of transcriptome-wide research. Long-read isoform sequencing methods are evolving and are more precise, efficient, and cost-effective. The integration of long-read data with other sequencing data, such as short-read RNA-seq and ChIP-seq data, can provide a more thorough understanding of the activities of isoforms and gene regulatory networks. In summary, long-read isoform sequencing provides a useful method for studying a complex variety of gene isoforms and alternative splicing events. This technology contributes to our understanding of cellular functions and disease mechanisms by offering crucial insights into transcriptome complexity and gene regulation [93].

CRISPR-Cas9-based sequencing

Researchers can target specific genomic regions of interest before sequencing using methods such as CRISPR-Cas9-based enrichment sequencing, which improves sequencing productivity and lowers costs. The CRISPR-Cas9 genome editing technique targets and sequences specific genomic regions of interest using a novel method known as CRISPR-Cas9-based sequencing. Using this method, specific genomic areas, genes, or even individual alleles can be targeted for sequencing and analysis, providing researchers with high-resolution insights into genetic variants, mutations, and structural changes [93,94].

Key techniques and concepts

CRISPR-Cas9 targeting: Guide RNAs (gRNAs) that specifically target the desired genomic area can be created using CRISPR-Cas9. At the target location, the Cas9 enzyme introduces double-strand breaks under the control of gRNA. After Cas9-mediated cleavage, PCR is used to amplify the genomic DNA and enrich the area containing the target location. It is possible to create specialized primers that amplify the areas surrounding the target spot. The amplified DNA fragments are subsequently transformed into a sequencing library by the addition of the necessary adaptor sequences. Using the created library, sequence data specific to the targeted genomic region are generated by NGS. CRISPR-Cas9-based sequencing can be used to identify genetic alterations, including single-nucleotide polymorphisms (SNPs), insertions, deletions, and structural modifications within the targeted area. Allele-specific analysis facilitates the investigation of genetic disorders, population genetics, and allele-specific gene expression by allowing researchers to evaluate individual alleles or haplotypes [95]. CRISPR-Cas9-based sequencing has been used to identify mutations that cause disease, somatic mutations, and cancer-related mutations. Targeted sequencing enables precise and specific sequencing of genomic regions of interest, lowering sequencing costs and complicating data interpretation. Data on genetic variants within a particular genomic region are provided at a high resolution via CRISPR-Cas9-based sequencing. Functional analysis is a method that examines the effects of genetic changes and mutations on gene expression and protein function to investigate the resulting functional effects. Despite efforts to create highly specific gRNAs, Cas9 can still have off-target effects, which may result in inadvertent cleavage at nearby genomic locations. Preferential amplification of specific DNA fragments may occur during PCR amplification during construction. The precision and effectiveness of CRISPR-Cas9-based sequencing are improving. Improvements in gRNA design and specificity will help lessen off-target effects and increase the accuracy of targeted sequencing as CRISPR-Cas9 technology develops. In summary, CRISPR-Cas9-based sequencing provides an effective technique for targeted sequencing of particular genomic regions, enabling in-depth research on genetic variants, mutations, and structural alterations. Numerous fields, including genomics, genetics, and personalized medicine, can benefit from this method. Precision medicine has been transformed by CRISPR-based genome editing owing to developments in prime editing (template-guided insertions or deletions) and base editing (single-nucleotide modifications without double-strand breaks). Delivery strategies range from nonviral techniques (lipid nanoparticles and electroporation) for transient expression to viral vectors (adeno-associated virus and lentivirus) for stable integration. In hereditary illnesses, such as sickle cell disease and β-thalassemia, recent clinical trials have yielded encouraging results, with patients seeing long-lasting therapeutic effects. Precision medicine applications have been advanced by the combination of CRISPR and NGS, which improves off-target analysis, real-time genomic monitoring, and editing validation [96,97].

Metagenomic sequencing

Metagenomics involves the study of complex microbial communities by directly sequencing DNA from environmental materials. This reveals information regarding microbial interactions, biodiversity, and functional potential. "Metagenomic sequencing" is a potent method that directly extracts genetic material from ambient materials, eliminating the need for laboratory growth. This approach provides insights into the functional potential and diversity of microbial communities across many settings, including soil, water, and the human body [98].

Key techniques and concepts

Environmental samples, such as those from soil, water, air, the human body, animal microbiomes, and diverse ecosystems, are first collected before metagenomic sequencing can begin. The genetic material is collected from the complete microbial population by obtaining genomic DNA from the collected samples. Using adapters, the isolated DNA is broken apart and transformed into sequencing libraries. DNA fragments from these collections reveals the genetic diversity of the microbial community. The resulting libraries are subjected to NGS, which results in an important amount of DNA sequence data representing the community's collective genomes of microorganisms. Metagenomic sequencing reveals the taxonomic variety of microorganisms in a sample, enabling the identification of numerous species and strains present. By examining the existence of specific genes and pathways, this approach offers insights into the metabolic and functional potential of microbial communities [99]. The functions of microorganisms in many ecosystems, including nutrient cycling and pollution degradation, can be examined using metagenomic data. Microbial communities that live inside and outside the human body are identified via metagenomic sequencing. This information can be used to understand how the human microbiome affects health and diseases. Without the need to culture particular bacteria, metagenomic sequencing can capture the entire genetic diversity of microbial communities. This method enables the investigation of functional genes and pathways, revealing details about the potential functions of microorganisms in their surroundings [100]. Conventional approaches typically do no lead to the discovery of novel microbes; however, metagenomic data frequently reveal novel microbial species and genes. To assemble and annotate sequences, predict gene functions, and interpret functional potentials, metagenomic data analysis requires sophisticated bioinformatics tools. Inaccurate estimates of community composition can result from the cross-contamination of samples and contamination from laboratory chemicals. With advancements in sequencing technologies and analytical techniques, metagenomic sequencing is progressing. Efforts are being made to standardize procedures and create improved tools for the study of metagenomic data. As the discipline develops, metagenomics will become increasingly important to understand microbial ecology, human health, and environmental connections. In summary, metagenomic sequencing allows for a thorough examination of microbial communities and their genetic diversity, advancing our knowledge of ecosystems, human health, and interactions between microbes and their surroundings [101].

Single-cell sequencing

Profiling at the individual-cell level is made possible by single-cell RNA sequencing (scRNA-seq) and single-cell DNA sequencing, which reveal the cellular dynamics and heterogeneity within tissues and populations. Single-cell sequencing, a cutting-edge technique, enables sequencing and examination of the genetic material of individual cells. Unlike traditional sequencing methods, which require a large amount of starting material, single-cell sequencing allows for the study of genetic diversity at the single-cell level [102].

Key techniques and concepts

Cell isolation: The first step in single-cell sequencing is separation of distinct cells from a heterogeneous population. Techniques such as microfluidics, laser capture microdissection, and fluorescence-activated cell sorting (FACS) can be used to isolate cells effectively. After a cell is isolated, its DNA is amplified using several procedures, including whole-genome amplification (WGA). This step is essential for the collection of sufficient DNA for sequencing. The amplified DNA goes through a library preparation process in which DNA fragments are given adapters. These adapters enable the sequencing of DNA fragments in NGS systems. The generated libraries are sequenced using NGS technology. The genetic makeup of each individual cell is revealed by sequencing data. The sequencing data are examined using specialized bioinformatics tools. Genomic data from cells can be utilized to investigate gene expression and spot mutations and deduce cellular heterogeneity [103]. Single-cell sequencing is used to investigate genetic alterations that occur throughout embryonic development, enabling researchers to monitor cell fate and differentiation. This method offers insights into tumor heterogeneity and clonal evolution, assisting in the identification of uncommon cell subpopulations that may promote cancer development. Single-cell sequencing has been used to elucidate neural circuits, understand the variety of cell types in the brain, and examine gene expression patterns in specific neurons. Researchers have used single-cell sequencing to examine immune cell diversity, immunological responses, and the dynamics of immune cell populations. By identifying variances between individual cells, single-cell sequencing enables analysis of cellular variety within a population. Single-cell sequencing can reveal uncommon cell varieties or states that may be concealed by mass sequencing techniques. Single-cell sequencing offers a dynamic perspective on cellular changes during the course of disease development and evolution. Data accuracy may be affected by amplification bias and the technical noise induced during DNA amplification. Complex datasets produced by single-cell sequencing require specialized bioinformatics tools and computational techniques for appropriate interpretation. Single-cell sequencing techniques are becoming increasingly sensitive and effective as technology develops [104]. New techniques can analyze multiple layers of biological information simultaneously, including gene expression, epigenetic changes, and protein abundance. These advancements will provide a better understanding of cellular diversity and function. In summary, single-cell sequencing enables the investigation of genetic data at the single-cell level and provides information on cellular variety, growth, and disease. This method paves the way for greater comprehension of cellular complexity and has the potential to revolutionize many areas of biology [103].

Digital sequencing

Absolute quantification of DNA or RNA molecules is provided by digital sequencing platforms, such as digital droplet PCR

(ddPCR), which helps with accurate copy number variation and gene expression measurement. By partitioning individual DNA molecules into different reaction chambers or compartments, a novel DNA sequencing technique known as digital sequencing enables the counting or detection of individual DNA molecules. This approach has benefits in terms of precision, sensitivity, and capacity to identify uncommon mutations [105].

Key techniques and concepts

Single-molecule partitioning: Digital sequencing frequently divides individual DNA molecules into distinct compartments using microfluidic- or emulsion-based techniques. DNA molecules are amplified inside each compartment using methods such as PCR. Individual DNA molecules may be given fluorescent labels as part of digital sequencing techniques so that their identification is possible. The contents of each compartment are examined after amplification and labeled. Each compartment is checked for the presence or absence of labeled DNA molecules before being noted. The data collected are discrete counts of DNA molecules and are digital in character. To determine the sequence information and any mutations present in the original DNA sample, bioinformatic techniques are used to analyze the digital data produced by counting the number of DNA molecules in each compartment. Digital sequencing is useful for identifying genetic changes linked to diseases because it is highly sensitive and capable of identifying uncommon mutations or variants in a sample. Rare allele detection can detect uncommon alleles or low-frequency variants in a population that conventional sequencing techniques might overlook [106]. Digital sequencing can provide quantitative data on the prevalence of specific DNA molecules in a sample. Even at rates as low as 0.01%, digital sequencing can detect extremely rare mutations or variations. The ability to count individual DNA molecules reduces the probability of errors occurring during amplification. Accurate quantification of DNA molecules is possible owing to the digital characteristics of the data. The number of partitions or compartments that can be processed concurrently using digital sequencing techniques may be constrained, which will affect the overall throughput of the method. Workflow complexity can be increased by the use of microfluidic devices or emulsion-based methods. As technology advances, efforts are underway to increase the throughput and scalability of digital sequencing techniques. To increase accuracy and sensitivity, digital sequencing is being evaluated using other methods, such as SMRT sequencing and amplification-free technologies. In summary, digital sequencing is a powerful technique that uses single-molecule partitioning to enable the sensitive and accurate identification of genetic mutations and variants. This method has potential uses in illness research and diagnostics, including mutation detection, rare allele analysis, and quantitative assessment [105,106].

Conclusion

The field of gene sequencing has made great strides, progressing from traditional methods to state-of-the-art technologies. Conventional Sanger sequencing initially provided a fundamental understanding of DNA sequences; however, NGS revolutionized the field by making it possible to analyze genetic material quickly, affordably, and in large quantities. SAGE reveals gene expression patterns by integration with NGS. With the introduction of single-molecule sequencing, long-read capabilities that allow direct DNA analysis without amplification and TGS provide real-time sequencing based on nanopores. Innovative techniques, such as synthetic long-read sequencing, spatial transcriptomics, epigenetic sequencing, long-read isoform sequencing, CRISPR-Cas9-based sequencing, metagenomic sequencing, and digital sequencing, have transformed genomics research by offering deep insights into genome architecture, gene expression, epigenetic alterations, microbial ecosystems, and single-cell variability.

Personalized medicine has been transformed by the use of NGS in clinical diagnostics, particularly in the identification of uncommon genetic diseases. For example, WGS has proven crucial in detecting harmful variations in diseases such as cystic fibrosis and spinal muscular atrophy (SMA), allowing for targeted treatments such as gene replacement therapy (e.g., Zolgensma for SMA). A number of NGS-based diagnostics have received U.S. Food and Drug Administration (FDA) approval, such as FoundationOne CDx for oncology-targeted therapy guidance and Illumina TruSight Oncology 500 for thorough cancer profiling. As regulatory bodies increasingly acknowledge NGS as a common tool in precision medicine, the FDA clearance timeline has increased. However, payment remains a major problem because insurance companies frequently cover only certain purposes, which makes it expensive for many people to undergo NGS-based testing.

Prenatal genomic screening also raises ethical questions because whole exome or genome sequencing may identify variations with unclear relevance, necessitating difficult pregnancy care choices. Ethical discussions are more difficult because of the possibility of inadvertent findings and the danger of genetic discrimination. Regulatory agencies, physicians, and legislators must work together to overcome these obstacles and maintain the accessibility, morality, and therapeutic impact of NGS-driven precision medicine.

Summary

The ongoing development of gene sequencing technologies has made it possible to study the complexity of genetic information at a level never before possible. These developments have sped up scientific discoveries while deepening our understanding of the complexity of biological systems. Each new method adds a special element to the complex picture of life, from the study of gene expression patterns to the clarification of epigenetic changes and from the investigation of microbial communities to the deciphering of individual cell genomes. These technologies have the potential to transform our understanding of biology and genetics as they develop and become more sophisticated, opening the door for advances in personalized treatment, disease studies, and environmental studies. Ultimately, these many gene sequencing techniques enable researchers to decipher the mysteries hidden within DNA, leading to a greater understanding of the astounding variety and complexity of life on Earth.

Promising Future Prospects

The precision medicine revolution has led to the prediction that gene sequencing will play a crucial role in the era of personalized medicine. Clinical professionals can customize treatments on the basis of a patient's genetic profile with better insights into specific genetic variants, resulting in more effective and focused therapies. As methods advance, researchers will put more effort

into understanding how genomes operate. To better understand the intricate interactions between genes and their regulatory networks, it is important to understand the functions of ncRNAs, regulatory components, and epigenetic alterations. Single-cell sequencing will facilitate the understanding of cellular heterogeneity and the intricacies of tissues and diseases at the cellular level. Understanding unusual cell types, developmental processes, and disease progression can be greatly improved using this strategy. Advancements in long-read sequencing technologies will allow for more thorough analyses of repetitive sequences, structural variants, and complicated genomic areas. Accurate genome assembly and knowledge of the genetic causes of many diseases will benefit from this approach. Integrating data from omics, such as transcriptomics, proteomics, and metabolomics, will provide a comprehensive picture of how biological processes work. This multidimensional strategy will facilitate the understanding of intricate biological mechanisms [107,108]. Metagenomic sequencing will continue to facilitate the elucidation of the diversity of microorganisms found in many habitats and assist in environmental monitoring, bioremediation, and microbial ecosystem knowledge**.** Handling and analyzing the enormous volumes of data produced by high-throughput sequencing will require the use of advanced bioinformatics tools and artificial intelligence (AI) [109]. NGS data processing and analysis rely heavily on bioinformatics, with assemblers such as Canu, Flye, and Shasta serving as benchmarks for long-read sequencing [110]. Flye is tailored for Oxford Nanopore data, providing faster assembly with middling accuracy, whereas Canu is computationally intensive but extremely accurate for PacBio readings. Shasta, which combines speed and precision for assembling the genomes of bacteria and humans at extremely fast speeds, requires surroundings with considerable memory. Because of the increased scalability brought about by the move to cloud-based platforms, such as AWS Batch and Google Cloud Life Sciences, researchers can now analyze enormous datasets effectively without being constrained by local hardware. Furthermore, by encapsulating dependencies, containerization platforms, such as Docker and Singularity, improve reproducibility and guarantee consistent bioinformatics operations across various computing environments. These developments have simplified the interpretation of NGS data and increased the accessibility and reproducibility of large-scale sequencing initiatives. Algorithms for machine learning will facilitate pattern recognition, disease risk assessment, and the identification of prospective therapeutic targets. Advances in sequencing, gene synthesis, and editing technologies are expected to drive the field of synthetic biology. This includes the development of designer organisms for a range of uses, from pharmaceuticals to biofuels. As sequencing becomes more available and thorough, privacy, data sharing, and ethical issues become more crucial. Within the next 10 years, there will likely be a push for a $100 genome propelled by developments in AI-powered base calling and nanopore sequencing. Because quantum computing can solve complicated genomic patterns relatively quickly, it has the potential to revolutionize sequence alignment and variation detection. Blockchain technology can facilitate patient-controlled access, secure genetic data sharing, and prevent data breaches in precision medicine. However, global standardization remains difficult because different regulatory bodies have different sequencing techniques, which affects the interoperability and repeatability of data. For genetic medicine to be scalable and fair, these obstacles must be eliminated [111,112].

Acknowledgments

We thank the Luqman College of Pharmacy and the Novel Global Community Educational Foundation, Australia, for their support in this article. During the preparation of this work, the authors used Quill Bot and Bio render to paraphrase and generate figures. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.
Funding: No funding was received for this work.
Competing interests: The authors declare that they have no competing interests.

References

1. Levy SE, Boone BE. Next-generation sequencing strategies. *Cold Spring Harb Perspect Med*. 2019;9(7):a025791.
2. Mardis ER. A decade's perspective on DNA sequencing technology. *Nature*. 2011;470(7333):198–203.
3. Buermans HPJ, den Dunnen JT. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta Mol Basis Dis*. 2014;1842(10):1932–1941.
4. Shokralla S, Spall JL, Gibson JF, Hajibabaei M. Next-generation sequencing technologies for environmental DNA research. *Mol Ecol*. 2012;21(8):1794–1805.
5. Gong S, Kus L, Heintz N. Rapid bacterial artificial chromosome modification for large-scale mouse transgenesis. *Nat Protoc*. 2010;5(10):1678–1696.
6. Cviková K, Cattonaro F, Alaux M, Stein N, Mayer KFX, Doležel J, Bartoš J. High-throughput physical map anchoring via BAC-pool sequencing. *BMC Plant Biol*. 2015;15(1):99.
7. Shizuya H, Kouros-Mehr H. The development and applications of the bacterial artificial chromosome cloning system. *Keio J Med*. 2001;50(1):26–30.
8. Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH. Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. *J Agric Genomics*. 2000;5.
9. Lijavetzky D, Muzzi G, Wicker T, Keller B, Wing R, Dubcovsky J. Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. *Genome*. 1999;42(6):1176–1182.
10. Balloux F, Brønstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, Wang H, Eldholm V. From theory to practice: Translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol*. 2018;26(12):1035–1048.
11. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Kam KM, Perez E, et al. Pulsenet international: Vision for the implementation of whole genome sequencing (WGS) for global foodborne disease surveillance. *Eur Secur*. 2017;22(23):30544.
12. Zhang J, Chiodini R, Badr A, Zhang G. The impact of next-generation sequencing on genomics. *J Genet Genomics*. 2011;38(3):95–109.
13. Chen-Harris H, Borucki MK, Torres C, Slezak TR, Allen JE. Ultra-deep mutant spectrum profiling: Improving sequencing accuracy using overlapping read pairs. *BMC Genomics*. 2013;14(1):96.
14. Imelfort M, Duran C, Batley J, Edwards D. Discovering genetic polymorphisms in next-generation sequencing data: Review article. *Plant Biotechnol J*. 2009;7(4):312–317.
15. Edwards D, Batley J. Plant genome sequencing: Applications for crop improvement. *Plant Biotechnol J*. 2010;8(1):2–9.

16. Luo C, Tsementzi D, Kyrpides N, Read T, Konstantinidis KT. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PLOS ONE*. 2012;7(2):e30087.
17. Benaglio P, Rivolta C. Ultra high throughput sequencing in human DNA variation detection: A comparative study on the NDUFA3-PRPF31 region. *PLOS ONE*. 2010;5(9):e13071.
18. Forgetta V, Leveque G, Dias J, Grove D, Lyons R, Genik S, Wright C, Singh S, Peterson N, Zianni M, et al. Sequencing of the Dutch elm disease fungus genome using the Roche/454 GS-FLX titanium system in a comparison of multiple genomics core facilities. *J Biomol Tech*. 2013;24(1):39–49.
19. Dames S, Durtschi J, Geiersbach K, Stephens J, Voelkerding KV. Comparison of the illumina genome analyzer and roche 454 GS FLX for resequencing of hypertrophic cardiomyopathy-associated genes. *J Biomol Tech*. 2010;21(2):73–80.
20. Soares AR, Pereira PM, Santos MAS. Next-generation sequencing of miRNAs with Roche 454 GS-FLX technology: Steps for a successful application. *Methods Mol Biol*. 2012;822:189–204.
21. Quail MA, Swerdlow H, Turner DJ. Improved protocols for the illumina genome analyzer sequencing system. *Curr Protoc Hum Genet*. 2009.
22. Kircher M, Stenzel U, Kelso J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol*. 2009;10(8):R83.
23. Ondov BD, Varadarajan A, Passalacqua KD, Bergman NH. Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic applications. *Bioinformatics*. 2008;24(23):2776–2777.
24. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, Swerdlow H, Turner DJ. A large genome center's improvements to the Illumina sequencing system. *Nat Methods*. 2008;5(12):1005–1010.
25. Okou DT, Locke AE, Steinberg KM, Hagen K, Athri P, Shetty AC, Patel V, Zwick ME. Combining microarray-based genomic selection (MGS) with the illumina genome analyzer platform to sequence diploid target regions. *Ann Hum Genet*. 2009;73(5):502–513.
26. Ondov BD, Cochran C, Landers M, Meredith GD, Dudas M, Bergman NH. An alignment algorithm for bisulfite sequencing using the Applied Biosystems SOLiD System. *Bioinformatics*. 2010;26(15):1901–1902.
27. Mahmood SA, Al-Gburi NM. Detection of listeria monocytogenes and atypical hemolytic listeria innocua from mastitis cows' milk in Baghdad Province. *J Med Pharm Chem Res*. 2025;7(8):1774–1786.
28. Ossiboff RJ. Molecular diagnostic techniques. In: *Schalm's veterinary hematology, seventh edition*. Hoboken (NJ): Wiley; 2020. p. 1331–1336.
29. Gibriel AA, Adel O. Advances in ligase chain reaction and ligation-based amplifications for genotyping assays: Detection and applications. *Mutat Res Rev Mutat Res*. 2017;773:66–90.
30. McKernan KJ, Peckham HE, Costa GL, McLaughlin SF, Fu Y, Tsung EF, Clouser CR, Duncan C, Ichikawa JK, Lee CC, et al. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Res*. 2009;19(9):1527–1541.
31. Bormann Chung CAB, Boyd VL, McKernan KJ, Fu Y, Monighetti C, Peckham HE, Barker M. Whole methylome analysis by ultra-deep sequencing using two-base encoding. *PLOS ONE*. 2010;5(2):e9320.
32. Homer N, Merriman B, Nelson SF. Local alignment of two-base encoded DNA sequence. *BMC Bioinformatics*. 2009;10:175.
33. Huang W, Li L, Myers JR, Marth GT. ART: A next-generation sequencing read simulator. *Bioinformatics*. 2012;28(4):593–594.
34. Wiedmann M, Wilson WI, Czajka J, Luo J, Barany F, Batt CA. Ligase chain reaction (LCR)—Overview and applications. *Genome Res*. 1994;3(4):S51–S64.
35. Pfeffer M, Wiedmann M, Batt CA. Applications of DNA amplification techniques in veterinary diagnostics. *Vet Res Commun*. 1995;19(5):375–407.
36. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: Using high-throughput sequencing to discover protein-DNA interactions. *Methods*. 2009;48(3): 240–248.
37. Omar A, Alwaseef MA, Mosa MFI, Adbelgawad SS, Hussein NR, Ajlan MA, Elmorsy WA, Abdulwehab MM, Elshehabi AAM, Seoudi MA, et al. Impact of oxidative stress and TPMT genetic variants on the pharmacodynamics of mercaptopurine in pediatric acute lymphoblastic leukemia. *J Med Pharmaceut Chem Res*. 2025;7(11):2422–2434.
38. Zhu JY, Sun Y, Wang ZY. Genome-wide identification of transcription factor-binding sites in plants using chromatin immunoprecipitation followed by microarray (ChIP-chip) or sequencing (ChIP-seq). *Methods Mol Biol*. 2012;876:173–188.
39. Tao Z, Shen L, Liu C, Liu L, Yan Y, Yu H. Genome-wide identification of SOC1 and SVP targets during the floral transition in Arabidopsis. *Plant J*. 2012;70(4):549–561.
40. Hino S, Sato T, Nakao M. Chromatin immunoprecipitation sequencing (ChIP-seq) for detecting histone modifications and modifiers. *Methods Mol Biol*. 2023;2577:55–64.
41. O'Geen H, Echipare L, Farnham PJ. Using ChIP-seq technology to generate high-resolution profiles of histone modifications. *Methods Mol Biol*. 2011;791:265–286.
42. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008;26(10):1135–1145.
43. Allahverdi A, Pesch E, Pinedo M, Werner F. Scheduling in manufacturing systems: New trends and perspectives. *Int J Prod Res*. 2018;56(19):6333–6335.
44. Ng SB, Nickerson DA, Bamshad MJ, Shendure J. Massively parallel sequencing and rare disease. *Hum Mol Genet*. 2010;19(R2):R119–R214.
45. Tucker T, Marra M, Friedman JM. Massively parallel sequencing: The next big thing in genetic medicine. *Am J Hum Genet*. 2009;85(2):142–154.
46. Hu TC. Parallel sequencing and assembly line problems. *Oper Res*. 1961;9(6):841–848.
47. Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, Jovanovich SB, Nelson JR, Schloss JA, Schwartz DC, et al. The challenges of sequencing by synthesis. *Nat Biotechnol*. 2009;27(11):1013–1023.
48. Brockman W, Alvarez P, Young S, Garber M, Giannoukos G, Lee WL, Russ C, Lander ES, Nusbaum C, Jaffe DB. Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Res*. 2008;18(5):763–770.
49. Chen CY. DNA polymerases drive DNA sequencing-by-synthesis technologies: Both past and present. *Front Microbiol*. 2014;5:305.
50. Pomraning KR, Smith KM, Bredeweg EL, Connolly LR, Phatale PA, Freitag M. Library preparation and data analysis

packages for rapid genome sequencing. *Methods Mol Biol*. 2012;944:1–22.
51. Wu WW, Phue JN, Lee CT, Lin C, Xu L, Wang R, Zhang Y, Shen RF. Robust sub-nanomolar library preparation for high throughput next generation sequencing. *BMC Genomics*. 2018;19(1):326.
52. Chen SJ, Rai CI, Wang SC, Chen YC. Point-of-care testing for infectious diseases based on class 2 CRISPR/Cas technology. *Diagnostics*. 2023;13(13):2255.
53. Umar M, Tahir M, Shabbir CA. A biochemical analysis of detrimental effects of COVID-19 severity on multiple organ systems. *TSF J Biol*. 2023;1(2):28–45.
54. Schröter C, Günther R, Rhiel L, Becker S, Toleikis L, Doerner A, Becker J, Schönemann A, Nasu D, Neuteboom B, et al. A generic approach to engineer antibody pH-switches using combinatorial histidine scanning libraries and yeast display. *MAbs*. 2015;7(1):138–151.
55. Jiang D, Tang C, Zhang A. Cluster analysis for gene expression data: A survey. *IEEE Trans Knowl Data Eng*. 2004;16(11):1370–1386.
56. Trayhurn P. Northern blotting. *Proc Nutr Soc*. 1996;55(1B):583–589.
57. Várallyay É, Burgyán J, Havelda Z. MicroRNA detection by northern blotting using locked nucleic acid probes. *Nat Protoc*. 2008;3(2):190–196.
58. Mocharla H, Mocharia R, Hodes ME. Coupled reverse transcription-polymerase chain reaction (RT-PCR) as a sensitive and rapid method for isozyme genotyping. *Gene*. 1990;93(2):271–275.
59. Postollec F, Falentin H, Pavan S, Combrisson J, Sohier D. Recent advances in quantitative PCR (qPCR) applications in food microbiology. *Food Microbiol*. 2011;28(5):848–861.
60. Bumgarner R. Overview of dna microarrays: Types, applications, and their future. *Curr Protoc Mol Biol*. 2013.
61. Pabinger S, Rödiger S, Kriegner A, Vierlinger K, Weinhäusel A. A survey of tools for the analysis of quantitative PCR (qPCR) data. *Biomol Detect Quantif*. 2014;1(1):23–33.
62. Prado M, Ortea I, Vial S, Rivas J, Calo-Mata P, Barros-Velázquez J. Advanced DNA- and protein-based methods for the detection and investigation of food allergens. *Crit Rev Food Sci Nutr*. 2016;56(15):2511–2542.
63. Oue N, Hamai Y, Mitani Y, Matsumura S, Oshimo Y, Aung PP, Kuraoka K, Nakayama H, Yasui W. Gene expression profile of gastric carcinoma: Identification of genes and tags potentially involved in invasion, metastasis, and carcinogenesis by serial analysis of gene expression. *Cancer Res*. 2004;64(7):2397–2405.
64. Yamamoto M, Wakatsuki T, Hada A, Ryo A. Use of serial analysis of gene expression (SAGE) technology. *J Immunol Methods*. 2001;250(1–2):45–66.
65. Zhang M, Liu YH, Chang CS, Zhi H, Wang S, Xu W, Smith CW, Zhang HB. Quantification of gene expression while taking into account RNA alternative splicing. *Genomics*. 2019;111(6):1517–1528.
66. Morioka MS, Kawaji H, Nishiyori-Sueki H, Murata M, Kojima-Ishiyama M, Carninci P, Itoh M. Cap analysis of gene expression (CAGE): A quantitative and genome-wide assay of transcription start sites. *Methods Mol Biol*. 2020;2120:277–301.
67. Singhi AD, Nikiforova MN, Chennat J, Papachristou GI, Khalid A, Rabinovitz M, das R, Sarkaria S, Ayasso MS, Wald AI, et al. Integrating next-generation sequencing to endoscopic retrograde cholangiopancreatography (ERCP)-obtained biliary specimens improves the detection and management of patients with malignant bile duct strictures. *Gut*. 2020;69(1):52–61.
68. Lin CH, Chen PL, Tai CH, Lin HI, Chen CS, Chen ML, Wu RM. A clinical and genetic study of early-onset and familial parkinsonism in Taiwan: An integrated approach combining gene dosage analysis and next-generation sequencing. *Mov Disord*. 2019;34(4):506–515.
69. Nguyen YTN, Duong-Dinh C, Vu-Quang H, Dinh LTL, Nguyen-Minh T, Nguyen ND, Nguyen AT. LC-ESI-QTOF-HRMS-based myxobacterial metabolite profiling for potential anti-breast cancer extracts. *J Med Chem Sci*. 2023;6(6):2767–2777.
70. Kuo MC, Liu SCH, Hsu YF, Wu RM. The role of noncoding RNAs in Parkinson's disease: Biomarkers and associations with pathogenic pathways. *J Biomed Sci*. 2021;28(1):78.
71. Sun YM, Chen YQ. Principles and innovative technologies for decrypting noncoding RNAs: From discovery and functional prediction to clinical application. *J Hematol Oncol*. 2020;13(1):109.
72. Hernández-Romero IA, Guerra-Calderas L, Salgado-Albarrán M, Maldonado-Huerta T, Soto-Reyes E. The regulatory roles of non-coding rnas in angiogenesis and neovascularization from an epigenetic perspective. Front. *Oncologia*. 2019;9:1091.
73. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, Rowan TN, Low WY, Zimin A, Couldrey C, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020;9(3):giaa021.
74. Yip M, Chen J, Zhi Y, Tran NT, Namkung S, Pastor E, Gao G, Tai PWL. Querying recombination junctions of replication-competent adeno-associated viruses in gene therapy vector preparations with single molecule, real-time sequencing. *Viruses*. 2023;15(6):1228.
75. Martin-Abadal M, Ruiz-Frau A, Hinz H, Gonzalez-Cid Y. Jellytoring: Real-time jellyfish monitoring based on deep learning object detection. *Sensors*. 2020;20(6):1708.
76. Muncan J, Tei K, Tsenkova R. Real-time monitoring of yogurt fermentation process by aquaphotomics near-infrared spectroscopy. *Sensors*. 2021;21(1):1–18.
77. Perez M, Aroh O, Sun Y, Lan Y, Juniper SK, Young CR, Angers B, Qian PY. Third-generation sequencing reveals the adaptive role of the epigenome in three deep-sea polychaetes. *Mol Biol Evol*. 2023;40(8):msad172.
78. Huang R, Liu Y, Xu J, Lin D, Mao A, Yang L, Zhong G, Wang H, Xu R, Chen Y, et al. Back-to-back comparison of TGS and next-generation sequencing in carrier screening of thalassemia. *Arch Pathol Lab Med*. 2024;148(7):797–804.
79. Liang Q, He J, Li Q, Zhou Y, Liu Y, Li Y, Tang L, Huang S, Li R, Zeng F, et al. Evaluating the clinical utility of a long-read sequencing-based approach in prenatal diagnosis of thalassemia. *Clin Chem*. 2023;69(3):239–250.
80. Kernohan KD, Boycott KM. The expanding diagnostic toolbox for rare genetic diseases. *Nat Rev Genet*. 2024;25(6):401–415.
81. Zhang F, Xue H, Dong X, Li M, Zheng X, Li Z, Xu J, Wang W, Wei C. Long-read sequencing of 111 rice genomes reveals significantly larger pan-genomes. *Genome Res*. 2022;32(5):853–863.
82. Fan Y, Andrusivová Ž, Wu Y, Chai C, Larsson L, He M, Luo L, Lundeberg J, Wang B. Expansion spatial transcriptomics. *Nat Methods*. 2023;20(8):1179–1182.
83. Song Q, Su J. DSTG: Deconvoluting spatial transcriptomics data through graph-based artificial intelligence. *Brief Bioinform*. 2021;22(5):bbaa414.

84. Sun ED, Ma R, Navarro Negredo P, Brunet A, Zou J. TISSUE: Uncertainty-calibrated prediction of single-cell spatial transcriptomics improves downstream analyses. *Nat Methods*. 2024;21(3):444–454.
85. Zhang L, Lu Q, Chang C. Epigenetics in health and disease. *Adv Exp Med Biol*. 2020;1253:3–55.
86. Ebrahimi V, Soleimanian A, Ebrahimi T, Azargun R, Yazdani P, Eyvazi S, Tarhriz V. Epigenetic modifications in gastric cancer: Focus on DNA methylation. *Gene*. 2020;742:144577.
87. Salgotra RK, Gupta M. Exploring the role of epigenetics in cereal and leguminous crops exposed to abiotic stress. In: *Epigenetics in plants of agronomic importance: Fundamentals and applications: Transcriptional regulation and chromatin remodelling in plants: Second edition*. London: Springer Nature Link; 2019. p. 149–170.
88. Gazaem MAH, Ismail NH, Sadiran SH, Mutalib NA, Alsayadi MMS, Alkhudhayri DA, Mohamad SAS. Secondary metabolites of endophytic fungi from malaysian local herbs as potential antimicrobial and anticancer agents. *J Med Pharm Chem Res*. 2025;7(7):1435–1452.
89. King DE, Sparling AC, Joglekar R, Meyer JN, Murphy SK. Direct comparisons of bisulfite pyrosequencing versus targeted bisulfite sequencing. *MicroPubl Biol*. 2021.
90. Yao S, Liang F, Gill RA, Huang J, Cheng X, Liu Y, Tong C, Liu S. A global survey of the transcriptome of allopolyploid Brassica napus based on single-molecule long-read isoform sequencing and Illumina-based RNA sequencing data. *Plant J*. 2020;103(2):843–857.
91. Il LJ, Ramekar R, Kim JM, Hung NN, Seo JS, Kim JB, Choi IK, Park KC, Kwon SJ. Unraveling the complexity of faba bean (Vicia faba L.) transcriptome to reveal cold-stress-responsive genes using long-read isoform sequencing technology. *Sci Rep*. 2021;11(1):21094.
92. De Paoli-Iseppi R, Gleeson J, Clark MB. Isoform age-splice isoform profiling using long-read technologies. *Front Mol Biosci*. 2021;8:711733.
93. Mohsin M, Li Y, Zhang X, Wang Y, Huang Z, Yin G, Zhang Z. Development of CRISPR-CAS9 based RNA drugs against Eimeria tenella infection. *Genomics*. 2021;113(6):4126–4135.
94. Fatek AS, Esaily RMH, Madkhali FMH, Otayf YMA, Hakami MAA, Frahan RM, Ageeli NAI, Abbadi IM, Hseelh AEA, Almasoudi KH, et al. CRISPR-based pathology diagnostics: Game-changer in precision. *J Med Chem Sci*. 2024;7:1808–1823.
95. Tran MT, Doan DTH, Kim J, Song YJ, Sung YW, Das S, Kim E-J, Son GH, Kim SH, Vu VT, et al. CRISPR/Cas9-based precise excision of SlHyPRP1 domain(s) to obtain salt stress-tolerant tomato. *Plant Cell Rep*. 2021;40(6):999–1011.
96. Kamiya Y, Abe F, Mikami M, Endo M, Kawaura K. A rapid method for detection of mutations induced by crispr/cas9-based genome editing in common wheat. *Plant Biotechnol*. 2020;37(2):247–251.
97. Hus K, Betekhtin A, Pinski A, Rojek-Jelonek M, Grzebelus E, Nibau C, Gao M, Jaeger KE, Jenkins G, Doonan JH, et al. A CRISPR/Cas9-based mutagenesis protocol for Brachypodium distachyon and its allopolyploid relative, Brachypodium hybridum. *Front Plant Sci*. 2020;11:614.
98. Hong HL, Flurin L, Thoendel MJ, Wolf MJ, Abdel MP, Greenwood-Quaintance KE, Patel R. Targeted versus shotgun metagenomic sequencing-based detection of microorganisms in sonicate fluid for periprosthetic joint infection diagnosis. *Clin Infect Dis*. 2023;76(3):e1456–e1462.
99. Chen H, Liang Y, Wang R, Wu Y, Zhang X, Huang H, Yu X, Hong M, Yang J, Liao K, et al. Metagenomic next-generation sequencing for the diagnosis of Pneumocystis jirovecii Pneumonia in critically pediatric patients. *Ann Clin Microbiol Antimicrob*. 2023;22(1):6.
100. Marter P, Huang S, Brinkmann H, Pradella S, Jarek M, Rohde M, Bunk B, Petersen J. Filling the gaps in the cyanobacterial tree of life—Metagenome analysis of Stigonema ocellatum DSM 106950, Chlorogloea purpurea SAG 13.99 and Gomphosphaeria aponina DSM 107014. *Genes*. 2021;12(3):389.
101. Minor NR, Ramuta MD, Stauss MR, Harwood OE, Brakefield SF, Alberts A, Vuyk WC, Bobholz MJ, Rosinski JR, Wolf S, et al. Metagenomic sequencing detects human respiratory and enteric viruses in air samples collected from congregate settings. *Sci Rep*. 2023;13(1):21398.
102. King HW, Orban N, Riches JC, Clear AJ, Warnes G, Teichmann SA, James LK. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol*. 2021;6(56):eabe6291.
103. Salma M, Andrieu-Soler C, Deleuze V, Soler E. High-throughput methods for the analysis of transcription factors and chromatin modifications: Low input, single cell and spatial genomic technologies. *Blood Cells Mol Dis*. 2023;101, 102745.
104. Gao H, Di J, Yin M, He T, Wu D, Chen Z, Li S, He L, Rong L. Identification of chondrocyte subpopulations in osteoarthritis using single-cell sequencing analysis. *Gene*. 2023;852:147063.
105. Mizuno K, Sumiyoshi T, Okegawa T, Terada N, Ishitoya S, Miyazaki Y, Kojima T, Katayama H, Fujimoto N, Hatakeyama S, et al. Clinical impact of detecting low-frequency variants in cell-free DNA on treatment of castration-resistant prostate cancer. *Clin Cancer Res*. 2021;27(22):6164–6173.
106. Dikau R. The application of a digital relief model to landform analysis in geomorphology. In: *Three dimensional applications in GIS*. Boca Raton (FL): CRC Press; 1989. p. 51–77.
107. Wagner DE, Klein AM. Lineage tracing meets single-cell omics: Opportunities and challenges. *Nat Rev Genet*. 2020;21(7):410–427.
108. Ji Y, Lotfollahi M, Wolf FA, Theis FJ. Machine learning for perturbational single-cell omics. *Cell Syst*. 2021;12(6): 522–537.
109. LFA T, Alharbi SM, Albadrani NM, Alhouri LARA, Bajahzar EM, Alanazy MJ, Hassani MH, Hakami AYM, Alsaleh SAI, Yahya EAY, et al. Advancing diagnostic pathology with artificial intelligence: Innovations, applications, and future directions. *J Med Chem Sci*. 2024;7:1899–1915.
110. Shinde , Deshmukh AD, More UK. Neonatal sickle cell disease assessment in rural Indian community: Demand for point of care testing (POCT). *J Med Chem Sci*. 2023;7(1):1–8.
111. Ali PA, Hanif F, Nettour H, Rehman M. Strategies and techniques of drug discovery from natural products. *Glob Drug Des Dev Rev*. 2017;II(I):34–43.
112. Pur DR, Krance SH, Pucchio A, Miranda RN, Felfeli T. Current uses of artificial intelligence in the analysis of biofluid markers involved in corneal and ocular surface diseases: A systematic review. *Eye*. 2023;37(10):2007–2019.