Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 06, Issue: 04 (JULY-AUG, 2026)
Available online at http://www.agriarticles.com
©Agri Articles, ISSN: 2582-9882

Next-Generation Sequencing for Metagenome Accessibility

*Vaishnavi R. Gadge[1], Riya Patel[2] and G. B. Patil[1]
[1]Department of Agricultural Biotechnology, Anand Agriculture University, Anand, Gujarat, India
[2]B. A. College of Agriculture, Anand Agriculture University, Anand, Gujarat, India
*Corresponding Author's email: vaishnavigadge1998@outlook.com

Microorganisms are ubiquitous in the environment, playing important roles in crop development, human health, and the environment; some beneficial microbes are not able to grow in the laboratory. The development of metagenomics with Illumina sequencing has given researchers directions to expand the unknown microbial universe by opening the way for the remarkable, high-quality, and accessible analysis of DNA without prior cultivation from environmental samples. This article illustrates how Illumina's Sequencing-by-Synthesis (SBS) technology has made metagenomics widely available to investigators on a large scale and explores its major applications in agriculture, environmental monitoring, medicine, and biotechnology. It also highlights the current challenges and future opportunities.

Keywords: Illumina sequencing, NGS, Metagenomics, Microbiome diversity

Introduction

Metagenomics is the study of genetic material recovered directly from the environment. It has been a game-changing technology that allows genomic scientists to study the DNA of whole microbial communities straight from environmental samples such as soil, water, and plant roots. In contrast to emphasising only the limited number of microbes that can be cultured, researchers can now examine the intense heterogeneity of the microbial universe in its natural environment. Next-generation sequencing (NGS), massively parallel second-generation sequencing technology that allows ultra-high throughput, scalability and speed, is used to determine the entire sequence order of nucleotides in a target genome. For latest Illumina and Ion torrent are originated from NGS platform to give 99.99% accuracy and the highest throughput for short reads; Ion Torrent, which detects hydrogen ions during base incorporation via ion semiconductor sequencing, hence delivering fast run times and lower instrument cost; PacBio SMRT (single molecule real time sequencing), which generates long reads for resolving complete genomes; Oxford Nanopore, which sequences DNA by passing it through protein nanopores for ultra-long, real-time reads; and prior systems such as 454 Pyrosequencing and SOLiD (sequencing by oligonucleotide ligation and deletion), which laid the foundation for high-throughput sequencing. Metagenomics has been able to become more cost-effective, precise, and faster, thanks to Illumina sequencing, a technology that has also been responsible for transformational advances. By producing millions of DNA sequences in a single run, Illumina has helped scientists worldwide uncover previously unknown microbes on a huge scale and is helping researchers to solve problems that were previously inaccessible by improving crop yields and soil health, monitoring infectious illnesses, and preserving the environment. This paper explains how this exceptional technology has altered metagenomics, uncovered the unknown microbiological domain, and enhanced the genomic study it allowed to a drastically more inclusive scientific community.

Illumina Sequencing

Illumina sequencing it worked based on sequencing-by-Synthesis (SBS). It sequences millions of DNA fragments simultaneously by adding one fluorescently labelled nucleotide (A, T, G, or C) at a time and detecting each base through imaging. This repeated process produces highly accurate DNA sequences quickly and cost-effectively, making SBS the foundation of Illumina platforms for applications such as whole-genome sequencing, microbiome research, 16S rRNA analysis, and shotgun metagenomics. The Workflow of Illumina Sequencing-by-Synthesis is illustrated in **Figure 1**.

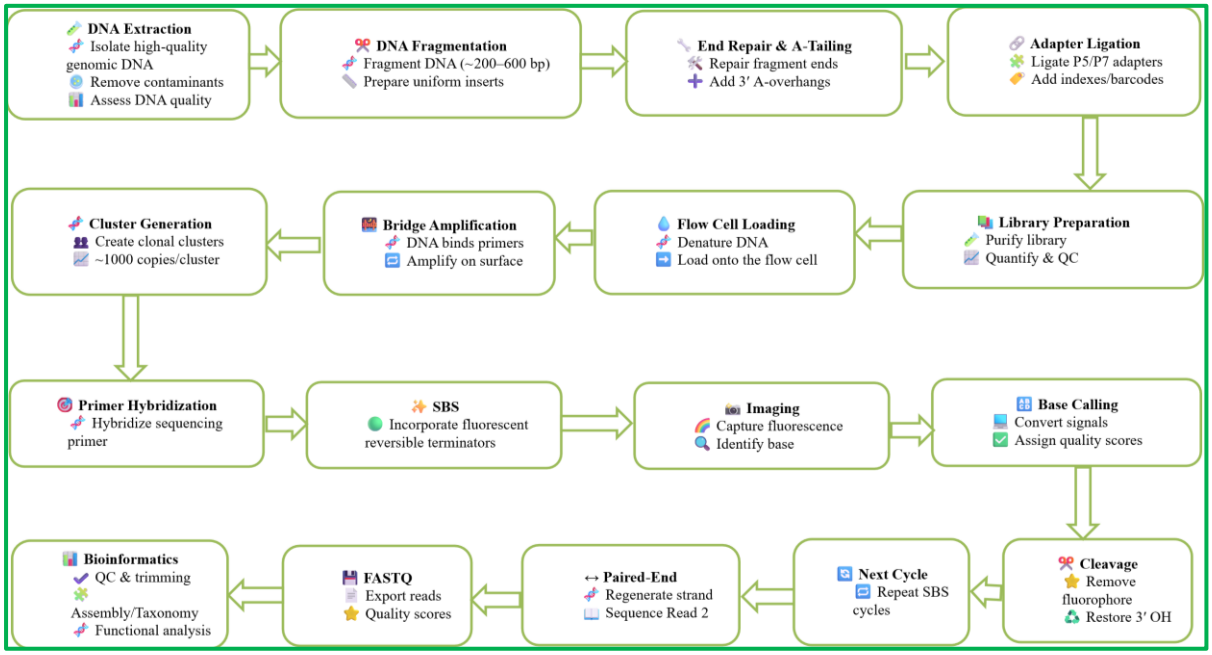

Figure 1. Workflow of Illumina Sequencing-by-Synthesis (SBS).

The Workflow of Illumina sequencing in short. First, DNA is extracted from a sample such as soil, water, plant tissue, or a clinical specimen. The DNA is then broken into small fragments, and short adapter sequences are attached to their ends to prepare them for sequencing. These fragments are loaded onto a flow cell, where they are copied many times through bridge amplification, creating clusters of identical DNA molecules. During Sequencing-by-Synthesis (SBS), fluorescently labelled nucleotides are added one base at a time, and each incorporated nucleotide is recorded by high-resolution imaging. The instrument repeats this process until the complete DNA sequence is generated. Finally, advanced computer software converts the sequencing signals into DNA sequences, which are analysed to identify microorganisms, genes, and their biological functions. For many studies, both ends of each DNA fragment are sequenced (paired-end sequencing), providing more accurate and reliable results.

Why Illumina Became the Gold Standard for Metagenomics

Illumina sequencing has become the gold standard for metagenomics because it synthesize outstanding sequencing accuracy, rapid data generation, excellent cost-effectiveness, and reproducibility in a single platform. Compared with first-generation sequencing technology, it can analyse millions of DNA fragments simultaneously, generating large volumes of high-quality data within a short time. This allows scientists to correctly find microorganisms, study complicated groups of microbes, and get trustworthy results for different scientific uses. Ongoing advances in sequencing technology have lowered costs, made lab work easier, and improved bioinformatics tools, allowing metagenomic research to be done by universities, hospitals, agricultural research centres, and environmental labs around the world. Because of this, Illumina sequencing keeps helping make discoveries in areas like agriculture, environmental tracking, biotechnology, and human health by making it easier than ever for researchers to study microbes in depth.

Illumina Sequencing Platforms

Illumina has held by appending to its sequencing options to help scientific projects that vary in size and how complex they are. The journey started with the Genome Analyser and Genome Analyser II models, which brought in high-throughput DNA sequencing and set the foundation for today's genomics. This was followed by MiSeq, a small platform that is commonly used for 16S rRNA sequencing, identifying microbes, and working on small genome projects because it can read longer sequences. The HiSeq series improved the ability to sequence entire genomes, study gene expression, and analyse microbial communities in a sample, while NextSeq offered a good mix of speed and amount of data processed, making it suitable for studies that are not too large. In recent times, the NovaSeq and the latest versions, NovaSeq X and NovaSeq X Plus, have greatly improved sequencing speed, data production, and cost savings. These systems are effective for big projects like sequencing entire genomes, studying microbes, advancing personalised medicine, researching plants and crops, and looking into environmental changes. These platforms all use the same precise Sequencing-by-Synthesis (SBS) technology, but they have different capacities for handling samples. The major characteristics of Illumina sequencing platforms are summarised in **Table 1**.

Table 1. Comparison of major Illumina sequencing platforms.

Platform	Year Introduced	Read Length Range	Typical Output per Run	Major Applications	References
Genome Analyzer	2006	35–75 bp	<1 Gb	Early whole-genome sequencing, proof-of-concept NGS	Bentley *et al.*, 2008
GAII/GAIIx	2008	36–150 bp	20–95 Gb	Genome sequencing, transcriptomics	Bentley *et al.*, 2008
HiSeq 2000/2500/3000/4000/X	2010	50–150 bp	120–1800 Gb	Whole-genome sequencing, RNA-seq, shotgun metagenomics	Glenn, 2011
MiSeq	2011	50–300 bp	up to ~15 Gb	16S rRNA sequencing, ITS sequencing, targeted sequencing, microbial identification	Caporaso *et al.*, 2012
NextSeq 500/550/1000/2000	2014	50–150 bp	100–540 Gb	Exome sequencing, transcriptomics, microbial genomics	Goodwin *et al.*, 2016
NovaSeq 6000	2017	50–150 bp	up to ~6 Tb)	Population genomics, large-scale microbiome studies, precision medicine	Logsdon *et al.*, 2020
NovaSeq X / X Plus	2022	50–150 bp	(up to ~16 Tb)	National genome projects, population metagenomics, large clinical genomics programs	Illumina, 2022
MiSeq i100 / MiSeq i100 Plus	2024	50–300 bp	(up to ~15 Gb)	16S rRNA sequencing, amplicon sequencing, microbial genomics, targeted panels, small genome sequencing	Illumina, 2024

Bioinformatics Pipeline for Illumina Sequencing Data

After sequencing unfiltered data are produced and then it transformed into FASTQ files through base calling. The reads are then quality-checked, trimmed, and filtered to remove adapters and low-quality sequences. High-quality reads are analysed to create OTUs or ASVs and are allocated taxonomy using reference databases such as SILVA or UNITE. Finally, diversity analysis, statistical analysis, and data visualisation are performed to characterise microbial communities and analysed their composition and potential functions. The sequence work-flow of sequencing data analysis was represented in tabular form in table.2

Table 2. Sequence workflow of sequencing data analysis

Step	Purpose	Common Software/Database
1. Base Calling	Convert fluorescence signals into nucleotide sequences	Illumina DRAGEN, bcl2fastq
2. FASTQ Generation	Generate FASTQ files containing sequences and quality scores	bcl2fastq
3.Quality Assessment	Evaluate sequence quality	FastQC, MultiQC
4. Adapter Trimming	Remove sequencing adapters and poor-quality bases	Cutadapt, Trimmomatic, fastp
5. Quality Filtering	Remove short or low-quality reads	QIIME2, DADA2
6. Read Processing	Merge paired-end reads and correct sequencing errors	DADA2, USEARCH, VSEARCH
7. Chimera Removal	Eliminate PCR-generated chimeric sequences	DADA2, UCHIME, VSEARCH
8.OTU/ASV Generation	Generate Operational Taxonomic Units or Amplicon Sequence Variants	DADA2, Deblur, USEARCH
9.Taxonomic Assignment	Assign taxonomy to bacterial or fungal sequences	SILVA, Greengenes, RDP, UNITE
10.Diversity Analysis	Assess microbial diversity within and between samples	QIIME2, Phyloseq, vegan (R)
11.Statistical Analysis	Compare microbial communities	R, QIIME2, MicrobiomeAnalyst
12.Functional Prediction	Predict microbial metabolic functions	PICRUSt2, Tax4Fun, FAPROTAX
13. Visualization	Create graphs and plots	R (ggplot2, Phyloseq), QIIME2
14. Interpretation	Draw biological conclusions	Literature and downstream analyses

Applications of Illumina Sequencing in Metagenomics

Illumina sequencing has applicable in many divert field like medical, pharmaceutical, agricultural, biotechnology, but we focus on application in agriculture system it adapted the field of metagenomics by facilitating researchers to study the unkonown world of microorganisms in ways previously impossible. Today, it is commonly used in farming to help make the soil better and grow more crops, in environmental studies to watch over nature and check for pollution, in medicine to learn about the bacteria in our bodies and find diseases, and in food safety to spot bad germs and keep food safe to eat. The major applications of Illumina sequencing are summarised in **Table 3.**

Table 3. Sequence-based identification of microbiome in agriculture.

Application Area	Sample Type	Illumina Sequencing Approach	Representative Findings	Reference
Rhizosphere microbiome analysis	Rhizosphere soil	16S rRNA amplicon sequencing (MiSeq)	Revealed diverse bacterial communities involved in nutrient cycling and plant growth promotion	Bulgarelli *et al.* (2013)
Plant growth-promoting bacteria (PGPR)	Root-associated soil	16S rRNA sequencing	Identified *Bacillus*, *Pseudomonas*, and *Rhizobium* as dominant PGPR	Compant *et al.* (2019)
Nitrogen-fixing microbial communities	Rhizosphere soil	16S rRNA and shotgun metagenomics	Detected abundant nitrogen-fixing taxa associated with sustainable agriculture	Bahram *et al.* (2018)
Disease suppression	Agricultural soil	16S rRNA sequencing	Identified beneficial microbes suppressing soil-borne pathogens	Mendes *et al.* (2011)
Soil health monitoring	Agricultural soil	16S rRNA sequencing	Demonstrated that conservation practices improve microbial diversity	Hartmann *et al.* (2015)
Crop response to drought	Root and rhizosphere	16S rRNA sequencing	Identified drought-enriched bacterial taxa associated with stress tolerance	Naylor *et al.* (2017)

Application Area	Sample Type	Illumina Sequencing Approach	Representative Findings	Reference
Fertilizer management	Agricultural soil	Amplicon sequencing	Organic amendments increased microbial diversity and beneficial taxa	Lupatini *et al.* (2017)
Plant pathogen detection	Plant tissues and soil	Amplicon sequencing (16S/ITS)	Enabled rapid identification of plant pathogens before symptom development	Agler *et al.* (2016)
Functional metagenomics	Soil	Shotgun metagenomic sequencing (HiSeq/NovaSeq)	Identified genes involved in nitrogen cycling, phosphorus solubilization, and stress adaptation	Fierer *et al.* (2012)
Precision agriculture	Soil and roots	Shotgun metagenomics	Linked microbial community composition with crop yield and soil fertility	Trivedi *et al.* (2020)

Current Challenges in Illumina-Based Metagenomics

Granted with many advantages, Illumina-based metagenomics still faces several challenges. Short read lengths can make complicate genome assembly, while PCR bias and host DNA contamination may affect the accuracy of microbial analysis. In addition, the large volume of sequencing data requires advanced bioinformatics tools and substantial computing resources, and incomplete reference databases can limit the identification and functional interpretation of microorganisms. However, developing advances such as hybrid sequencing, improved library preparation methods, artificial intelligence, cloud-based bioinformatics, and expanding microbial databases are helping overcome these limitations, making metagenomic analysis more accurate, efficient, and accessible.

Future Perspectives

The future of Illumina sequencing reclines in faster, cheaper, and higher-throughput sequencing, making metagenomics more accessible worldwide. The integration of multi-omics, artificial intelligence (AI), and hybrid sequencing with long-read technologies will improve genome assembly, functional analysis, and data interpretation. These advances will support precision agriculture, precision medicine, environmental monitoring, and cloud-based bioinformatics, enabling more sustainable farming, personalized healthcare, better ecosystem management, and easier access to large-scale metagenomic analysis**.**

Conclusions

Next generation sequencing is a foundation of omics study. Illumina sequencing has become the first priority of researchers for metagenomics by providing a reliable, accurate, and cost-effective way to explore the hidden diversity of microbial communities. Its ability to generate high-quality sequencing data has transformed research across agriculture, environmental science, medicine, biotechnology, and industry, enabling scientists to better understand microbial diversity, functional genes, and ecological interactions. Although challenges such as short read lengths and the need for advanced bioinformatics remain, continuous improvements in sequencing technology, computational tools, and multi-omics approaches are steadily overcoming these limitations. As these innovations continue, Illumina sequencing is expected to remain a driving force in microbiome research, supporting sustainable agriculture, environmental conservation, biotechnology, and precision medicine.

References

1. Agler, M. T., Ruhe, J., Kroll, S., Morhenn, C., Kim, S. T., Weigel, D., & Kemen, E. M. (2016). Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLoS Biology, 14(1), e1002352. https://doi.org/10.1371/journal.pbio.1002352

2. Aird, D., Ross, M. G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D. B., Nusbaum, C., & Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biology, 12*(2), R18.
3. Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R., Mundra, S., Olsson, P. A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., & Bork, P. (2018). Structure and function of the global topsoil microbiome. Nature, 560(7717), 233–237.
4. Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, *456*(7218), 53–59.
5. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114–2120.
6. Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L., & Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. Annual Review of Plant Biology, 64, 807–838.
7. Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., & Knight, R. (2012). *Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms*. The ISME Journal, 6(8), 1621–1624.
8. Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. Microbiome, 7(1), 29.
9. Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., Owens, S., Gilbert, J. A., Wall, D. H., & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences, 109(52), 21390–21395.
10. Glenn, T. C. (2011). Field guide to next-generation DNA sequencers. *Molecular Ecology Resources, 11*(5), 759–769.
11. Glenn, Travis C. (2011). Field guide to next-generation DNA sequencers. *Molecular Ecology Resources*, *11*(5), 759–769.
12. Goodwin, Sara, McPherson, John D., & McCombie, W. Richard. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351.
13. Hartmann, M., Frey, B., Mayer, J., Mader, P., & Widmer, F. (2015). Distinct soil microbial diversity under long-term organic and conventional farming. Proceedings of the National Academy of Sciences, 112(40), 12307–12312.
14. Illumina. (2022). *NovaSeq X series system guide*. Illumina, Inc.
15. Illumina. (2024). *MiSeq i100 series specifications*. Illumina, Inc.
16. Jain, Miten, Olsen, Hugh E., Paten, Benedict, & Akeson, Mark. (2016). The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biology*, *17*, 239.
17. Logsdon, G. A., Vollger, M. R., & Eichler, E. E. (2020). Long-read human genome sequencing and its applications. *Nature Reviews Genetics*, *21*(10), 597–614.
18. Lupatini, M., Korthals, G. W., de Hollander, M., Janssens, T. K. S., & Kuramae, E. E. (2017). Soil microbiome is more heterogeneous in organic than in conventional farming systems. *Frontiers in Microbiology*, 7, 2064.
19. Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J. H. M., Piceno, Y. M., DeSantis, T. Z., Andersen, G. L., Bakker, P. A. H. M., & Raaijmakers, J. M. (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science*, 332(6033), 1097–1100.

20. Naylor, D., DeGraaf, S., Purdom, E., & Coleman-Derr, D. (2017). Drought and host selection influence bacterial community dynamics in the grass root microbiome. *The ISME Journal*, 11(12), 2691–2704.
21. Streit, W. R., & Daniel, R. (2009). *Metagenomics: A treasure trove for the discovery of novel biocatalysts*. Current Opinion in Biotechnology, 20(3), 285–290.
22. Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: From community assembly to plant health. *Nature Reviews Microbiology*, 18(11), 607–621.