Journal of Applied Biosciences and Agri-Food Technology 2026, Vol 1(1), 10-19
ISSN:

Review Article

Applications of Whole Genome Sequencing in Agricultural and Food Biotechnology: A Review

Fadilah Sugiharto*

Department of Food Science and Technology, Faculty of Agricultural Technology, Universitas Brawijaya, Malang 65145, Indonesia

Article Info

Article History
Received: 24 December 2025
Revised: 29 December 2025
Accepted: 29 December 2025
Available online:02 February 2026

*Email (Author Corresponding) : fadilahnurrs@gmail.com

ABSTRACT

This review highlights recent advances and applications of Whole Genome Sequencing (WGS) in agricultural biotechnology. Utilizing de novo assembly, reference-based resequencing, and emerging approaches such as pangenome analysis, WGS has been extensively applied to major horticultural and vegetable crops, including banana, citrus, apple, pear, and other economically important species. The reviewed studies reveal that reference-based resequencing remains the dominant approach for crops with available high-quality genomes, while de novo sequencing and long-read platforms are preferred for genetically complex or previously uncharacterized species. Importantly, WGS has enabled the identification of genes and genomic variants associated with nutritional quality, disease resistance, post-harvest traits, and shelf life, thereby accelerating precision breeding and trait improvement. Overall, WGS serves as a cornerstone technology in food biotechnology by supporting genome-informed crop enhancement, improving food quality and safety, and enabling sustainable innovation in modern agriculture.

Keywords: Whole Genome Sequencing, Agricultural Biotechnology, Plant Breeding

Introduction

The rapid advancement of DNA sequencing technologies has profoundly transformed agricultural and food biotechnology by enabling high-resolution exploration of plant genomes. From the early development of Sanger sequencing to the emergence of next-generation and third-generation sequencing platforms, the ability to decode entire genomes has become faster, more accurate, and increasingly affordable (Xue & Cui, 2025). Among these innovations, Whole Genome Sequencing (WGS) has emerged as a central tool for elucidating genome-wide variation, uncovering functional genes, and understanding the molecular basis of agronomically important traits in crops (Kumar et al., 2024a).

In plant and food biotechnology, WGS provides comprehensive insights into both coding and non-coding regions of the genome, allowing researchers to investigate genetic mechanisms underlying yield improvement, disease resistance, stress tolerance, and quality-related traits. Unlike targeted sequencing or marker-based approaches, WGS enables the detection of a wide range of genetic variations, including single nucleotide polymorphisms, insertions and deletions, structural variants, and copy number variations (Krasilnikova et al., 2025; Yoon et al., 2025). These capabilities have positioned WGS as a powerful platform for accelerating crop improvement programs, supporting molecular breeding strategies, and enhancing food quality and safety.

The application of WGS in plant sciences has expanded alongside the availability of diverse sequencing strategies, including de novo genome assembly, reference-based

resequencing, low-pass WGS, and pangenome analysis. Each approach offers distinct advantages depending on genome complexity, ploidy level, and research objectives. For instance, de novo sequencing is particularly valuable for wild relatives and underexplored species, while reference-based approaches dominate studies of major crops with established genome assemblies (Admas et al., 2025; Farooq et al., 2024). Recent advances in long-read sequencing and chromatin conformation capture technologies have further improved genome contiguity and accuracy, enabling more precise characterization of complex plant genomes (Liu & Conesa, 2025).

Despite the growing body of literature on WGS applications, existing reviews often focus on specific crop species, sequencing platforms, or technical workflows, with limited integration of methodological comparisons and food biotechnology outcomes (Brlek et al., 2024). Moreover, relatively few studies critically evaluate why certain WGS approaches are preferred for particular crops or how genome complexity and ploidy influence methodological choice (Kumar et al., 2024b). This gap underscores the need for a comprehensive synthesis that not only summarizes WGS applications but also interprets their strategic relevance for agricultural and food biotechnology.

The recent updates of this review lie in its integrative and comparative perspective on WGS applications across major horticultural and vegetable crops, emphasizing the rationale behind the selection of different sequencing strategies and their implications for food biotechnology. This review compares de novo sequencing, resequencing, and emerging pangenome approaches, highlighting their advantages, limitations, and suitability for diverse crop genomes.

Results and Discussions

Whole Genome Sequencing

The advancement of DNA sequencing technology has brought significant progress to the fields of genetics and biotechnology. It began with the Sanger sequencing method developed in 1977, which served as the fundamental technique for DNA sequencing for several decades. However, this method was limited by its relatively short read length (≤ 1000 bases), making it less efficient for analyzing large and complex genomes. To overcome this limitation, the *shotgun sequencing* approach was introduced in 1979, which involves randomly fragmenting long DNA segments into smaller pieces and sequencing them using the Sanger method (Kwong et al., 2015). Although this approach marked an early step toward WGS, it remained complex as it required genomic mapping to reassemble the fragmented sequences.

Subsequent developments led to the emergence of Next Generation Sequencing (NGS), which enables the parallel sequencing of millions of DNA fragments using platforms such as Illumina, Ion Torrent, and SOLiD (Zhao & F.A. Grant, 2011). This technology offers higher efficiency, lower cost, and faster analysis compared to the Sanger method. Later, Third Generation Sequencing (long-read sequencing) technologies such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) were developed, enabling direct reading of long DNA fragments without amplification. These methods are particularly effective for resolving complex genomic structures, detecting large variants, and analyzing difficult-to-sequence regions (Logsdon et al., 2020).

WGS is a technique used to determine the complete nucleotide sequence (adenine, thymine, cytosine, and guanine) of an organism's entire genome by aligning the results against a known reference genome (Dewey et al., 2014). WGS provides a comprehensive overview of DNA, covering both coding and non-coding regions, thereby allowing a complete understanding of genetic variation within a species (Lu et al., 2025). Through WGS, researchers can identify

various types of genetic variation, including Single Nucleotide Polymorphisms (SNPs), Insertions and Deletions (InDels), Structural Variations (SVs), and Copy Number Variations (CNVs) (Lu et al., 2025).

According to Illumina (2025), Whole Genome Sequencing can be classified into two main approaches based on genome size, namely small genome sequencing and large genome sequencing. Small genome sequencing is typically applied to microorganisms such as bacteria and viruses with genome sizes of less than or equal to five megabases. The resulting sequences are compared to available reference genomes and are widely used in food testing, infectious disease surveillance, molecular epidemiology, and environmental metagenomics. Its advantages include comprehensive genome analysis of single microbial cultures, the ability to sequence thousands of isolates in parallel, and the discovery of new biomarkers through in-depth genetic information.

In contrast, large genome sequencing is applied to more complex organisms such as plants, animals, and humans with genome sizes of more than five megabases. This approach is commonly used for studying genetic diseases, population-level genetic variation, plant and animal breeding selection, and gene expression analysis. It provides high-resolution genomic information by integrating data from short and long DNA fragments to achieve complete genome characterization and identify rare alleles and clinically relevant genetic variants (Satam et al., 2023). Overall, the development of Whole Genome Sequencing technology has opened new opportunities in various fields such as biotechnology, agriculture, and health due to its ability to provide a deep understanding of the structure, function, and variation of genomes.

De novo Whole Genome Sequencing

The *de novo* whole genome sequencing approach is a method of constructing a genome without relying on any pre-existing reference genome. This approach is particularly applied when genome sequences from closely related species are not yet available. In this method, the sequencing reads generated from shotgun sequencing are compared with one another to identify overlapping regions, which are then assembled into longer contiguous sequences known as contigs (Ng & Kirkness, 2010).

Unlike reference-based assembly, which uses an existing reference genome, *de novo* assembly is more complex because it requires reconstructing the entire genome from scratch. The complexity of this process depends on the length and number of DNA fragments sequenced as well as the sequencing technology employed. Consequently, *de novo* assembly demands high computational power and stringent data quality standards to ensure accurate and complete genome representation. The development of *de novo* sequencing began with Sanger sequencing technology, which was known for its high accuracy but limited data throughput. With the advancement of sequencing methods, Next Generation Sequencing (NGS) emerged, enabling the generation of millions of short reads at a lower cost, albeit with shorter read lengths. Subsequently, Third Generation Sequencing technologies such as Pacific Biosciences (PacBio) and Oxford Nanopore were introduced, producing long reads in the kilobase range that are highly useful for resolving complex genomic structures.

In general, the main stages of *de novo* assembly using NGS data consist of three processes: contig assembly, scaffolding, and gap filling. During the contig assembly stage, sequencing reads are aligned based on sequence similarity to form long, continuous consensus sequences called contigs (Ekblom & Wolf, 2014). In the scaffolding stage, multiple contigs are connected using paired-end or mate-pair read data, which typically originate from large DNA fragments or fosmid inserts spanning several kilobases (Fuentes-Pardo & Ruzzante, 2017). The collection of ordered and connected contigs forms a structure known as a scaffold. Remaining gaps

between contigs are then filled using independent reads through a process known as gap filling to achieve a more complete genome assembly (Sohn & Nam, 2018). Scaffolding and gap filling are often performed iteratively to improve the overall assembly quality. With the rapid progress in sequencing technologies and bioinformatics tools, the *de novo* approach has become an essential method in modern genomics projects, particularly for characterizing new organisms or species that lack a complete reference genome.

Reference-based Whole Genome Sequencing

The reference-based whole genome sequencing approach, also known as mapping assembly or read mapping, is a genome assembly method that aligns sequencing reads to one or more pre-existing reference genomes. In this process, DNA fragments obtained from sequencing are aligned to the most appropriate positions on the reference genome, generating a consensus sequence that represents both the genetic similarities and differences between the sample and the reference genome (Deng et al., 2016). This method is computationally more efficient than the *de novo* assembly approach because it does not require reconstructing the genome from scratch (Lu et al., 2025). Consequently, reference-based assembly is widely used in population studies, species identification, and the detection of genetic variants in organisms for which reliable reference genomes are available.

Common sequencing platforms used in this approach include short-read technologies such as Illumina, which can generate paired-end reads of up to approximately 250 base pairs per read. Meanwhile, third-generation sequencing technologies such as PacBio and Oxford Nanopore Technologies (ONT) are often employed complementarily to produce long reads that help close sequencing gaps and improve read placement within repetitive regions (Worley et al., 2017). The advantages of the reference-based approach include rapid analysis and high accuracy in base positioning. However, the accuracy and reliability of the results are highly dependent on the quality, completeness, and suitability of the reference genome used.

This approach also allows direct comparison with type strains or verified strains. According to the European Food Safety Authority (EFSA, 2021), bacterial species identification can be validated using genomic similarity metrics such as digital DNA–DNA hybridization (dDDH) and average nucleotide identity (ANI), with conservative thresholds of dDDH $\geq$ 70% and ANI $\geq$ 95% to confirm species-level identity. In the case of yeasts and filamentous fungi, species confirmation is achieved through phylogenomic analysis using several conserved genes, such as AFToL markers including 18S rDNA and ITS, to generate phylogenetic relationships against available related genomes or by performing alignment to a complete reference genome from the same species.

Application of Whole Genome Sequencing

The advancement of WGS technology has revolutionized the fields of agriculture and plant biotechnology. Through this approach, the complete DNA composition of an organism can be mapped to understand its genetic structure, evolutionary relationships, and key genes involved in agronomic traits such as disease resistance, environmental stress tolerance, and yield quality. WGS also enables the analysis of genetic variation among cultivars, the construction of high-resolution genetic maps, and the development of marker-assisted breeding programs. Numerous fruit and crop species have been analyzed using WGS, providing valuable insights into the genetic mechanisms underlying desirable agronomic traits. Several applications of WGS in horticultural and fruit crops are summarized in Table 1.

Table 1. Application of Whole Genome Sequencing (WGS) in Major Fruit Crops

Scientific Name	Sequencing Method	Findings	References
Musa itinerans (Banana)	Reference-Based WGS	Identified disease resistance and cold-tolerance genes	(Wu et al., 2016)
Fusarium oxysporum f. sp. *cubense* (Banana pathogen)	WGS (comparative genomic analysis)	Identified virulence and SIX gene diversity	(Raman et al., 2021)
Malus domestica (Golden Delicious, GDDH13, HFTH1)	WGS (Illumina, PacBio, Hi-C, long-read)	Improved assemblies, trait associated polymorphisms	(Varnderzande et al., 2024)
Malus domestica (Golden Delicious)	WGS (Illumina, short-read, Hi-C)	Draft genome, 81% coverage, 17-chromosome mapping	(Peace et al., 2019)
*Vitis vinifera (*Pinot noir, PN40024)	WGS (Sanger, 454, Illumina)	First grape genome, 30.000 genes, SNP map	(Chagné, 2015)
Pyrus bretschneideri / Pyrus communis (Pear)	WGS (NGS, BAC-by-BAC, 454)	High synteny with apple, ~43.000 genes predicted	(Chagné, 2015)
Phoenix dactylifera (Date Palm)	Whole Genome Resequencing	Identified genetic diversity and domestication patterns	(Hazzouri et al., 2015)
Actinidia eriantha (Kiwifruit)	WGS	Identified structural variation and key genes for breeding	(Yao et al., 2022)
Citrus reticulata (Kinnow mandarin)	Whole Genome Resequencing	Identified key variants linked to fruit quality traits	(Jabeen et al., 2023)
Mangifera odorata, *M. altissima*, *M. indica* (Mango)	WGS	Identified genome-wide variants linked to fruit development and disease resistance	(Cortaga et al., 2022)

The application of WGS in various fruit crops demonstrates the crucial role of this technology in accelerating the discovery of functional genes and the construction of reference genome maps. Through WGS, research can be conducted with greater precision to elucidate the genetic mechanisms underlying key agronomic traits, thereby supporting the development of superior cultivars that are more adaptive, productive, and resistant to diseases.

The studies summarized in **Table 1** demonstrate that different WGS approaches are selected based on specific research objectives, genome complexity, and available genomic resources. Reference-based resequencing is the most frequently applied approach in fruit and vegetable crops such as banana, citrus, apple, and date palm, primarily because high-quality reference genomes are available for these species. WGS method is cost-effective within less than one week analysis, computationally efficient, and well suited for detecting single

nucleotide polymorphisms (SNPs) and small insertions/deletions across large populations, making it ideal for diversity analysis, domestication studies, and marker-assisted breeding (Chavhan et al., 2024; Dickinson et al., 2024).

In contrast, de novo genome sequencing is preferred for wild relatives or species lacking reliable reference genomes, such as *Musa itinerans* and *Actinidia eriantha*. Although de novo assembly is more expensive and computationally demanding, especially for large or repetitive plant genomes, it provides an unbiased and complete representation of genome structure (Raghavan et al., 2022; Thomas et al., 2022). This approach is particularly advantageous for identifying novel genes, structural variations, and lineage-specific genomic regions that are often missed by reference-based methods.

Whole-genome resequencing and low-pass WGS approaches offer a balance between cost and resolution and are increasingly used for population-scale studies where the primary objective is variant discovery rather than full genome reconstruction (Lv et al., 2021; Zheng et al., 2022). These approaches are especially suitable for diploid crops with relatively low genome complexity. However, their resolution may be limited for highly repetitive or polyploid vegetable species.

More recently, pangenome analysis has emerged as a powerful extension of WGS, enabling the comparison of multiple genomes within a species to capture core and dispensable genes. This approach is particularly relevant for crops with high genetic diversity, where single reference genomes fail to represent the full spectrum of functional variation (Hammond et al., 2020; Sherman & Salzberg, 2020). Despite higher costs and analytical complexity, pangenome studies provide critical insights into trait diversity, adaptation, and domestication.

Prospect of Whole Genome Sequencing in Food Biotechnology

The application of whole genome sequencing (WGS) is not limited to the mapping of plant genomes or their pathogens, but also holds vast potential in the development of modern food biotechnology. Through comprehensive genetic information analysis, WGS enables an in-depth understanding of the molecular basis of key traits in food crops, such as disease resistance, environmental stress tolerance, enhanced nutritional content, and improved flavor and texture quality.

In addition, WGS plays a vital role in food safety due to its ability to identify sources of microbial contamination and detect foodborne pathogens with high accuracy. This technology allows for rapid and precise traceability of pathogen origins, thereby accelerating responses to outbreaks of foodborne diseases (Oniciuc et al., 2018). The approach has also been used to construct phylogenetic trees of various foodborne pathogens, helping to elucidate their evolutionary relationships and patterns of dissemination across food environments and production chains.

Nevertheless, the application of WGS in food biotechnology still faces several challenges. The analysis of massive genomic datasets requires advanced bioinformatics expertise and high computational capacity. Moreover, there is currently no standardized international protocol for WGS methods, pipelines, and result interpretation. The lack of integration between genomic data, epidemiological data, and food traceability systems also limits the broader implementation of this technology (Dong et al., 2025). On the other hand, the relatively high operational and maintenance costs of WGS equipment remain a barrier for many laboratories, particularly in developing countries (Deng et al., 2016).Overall, despite existing technical and logistical challenges, the prospects for applying WGS in food biotechnology are highly promising. This technology has great potential to enhance food safety, accelerate genome-

based product innovation, and support more efficient, safe, and sustainable food production systems in the future.

Conclusion

Whole Genome Sequencing represents a major breakthrough in agricultural and food biotechnology, enabling comprehensive exploration of genetic variation, functional gene discovery, and genome-wide trait analysis. Through de novo assembly, reference-based resequencing, and advanced comparative genomics approaches, WGS has significantly accelerated crop improvement programs and enhanced food safety through precise pathogen detection and traceability.

Looking forward, the future of WGS in food biotechnology lies in its integration with genome editing technologies such as CRISPR/Cas systems, which will enable targeted modification of genes identified through sequencing studies. In addition, the incorporation of artificial intelligence and machine learning–based analytical tools will improve the interpretation of complex genomic datasets and facilitate predictive breeding strategies. Coupling WGS with high-throughput and advanced phenotyping platforms will further strengthen genotype–phenotype associations, enabling more efficient development of crops with improved nutritional quality, extended shelf life, and enhanced post-harvest performance. Together, these integrated approaches position WGS as a foundational technology for the next generation of sustainable and precision-driven food biotechnology.

Acknowledgement

Thank you to Belvania Nabila Putri for giving moral supports to write this manuscript.

Conflict of Interest

No conflict of interest.

References

Admas, T., Jiao, S., Pan, R., & Zhang, W. (2025). Pan-omics insights into abiotic stress responses: bridging functional genomics and precision crop breeding. *Functional & Integrative Genomics*, *25*(1), 128.

Brlek, P., Bulić, L., Bračić, M., Projić, P., Škaro, V., Shah, N., Shah, P., & Primorac, D. (2024). Implementing whole genome sequencing (WGS) in clinical practice: advantages, challenges, and future perspectives. *Cells*, *13*(6), 504.

Chagné, D. (2015). Whole Genome Sequencing of Fruit Tree Species. In *Advances in Botanical Research* (Vol. 74). Elsevier Ltd. https://doi.org/10.1016/bs.abr.2015.04.004

Chavhan, R. L., Hinge, V. R., Wankhade, D. J., Deshmukh, A. S., Mahajan, N., & Kadam, U. S. (2024). Bioinformatics for molecular breeding and enhanced crop performance: applications and perspectives. *Bioinformatics for Plant Research and Crop Breeding*, 21–74.

Cortaga, C. Q., Lachica, J. A. P., Lantican, D. V., & Ocampo, E. T. M. (2022). Genome-wide SNP and InDel analysis of three Philippine mango species inferred from whole-genome sequencing. *Journal of Genetic Engineering and Biotechnology*, *20*(1), 46. https://doi.org/10.1186/s43141-022-00326-3

Deng, X., Bakker, H. C. Den, & Hendriksen, R. S. (2016). *Genomic Epidemiology : Powered Surveillance and Outbreak Investigation of Foodborne Bacterial Pathogens*. 353–374. https://doi.org/10.1146/annurev-food-041715-033259

Dewey, F. E., Grove, M. E., Pan, C., Goldstein, B. A., Bernstein, J. A., Chaib, H., Merker, J. D., Goldfeder, R. L., Enns, G. M., David, S. P., Pakdaman, N., Ormond, K. E., Caleshu, C., Kingham, K., Klein, T. E., Whirl-Carrillo, M., Sakamoto, K., Wheeler, M. T., Butte, A. J., … Quertermous, T. (2014). Clinical interpretation and implications of whole-genome sequencing. *Jama*, *311*(10), 1035–1044. https://doi.org/10.1001/jama.2014.1717

Dickinson, M. C., Wirth, S. E., Baker, D., Kidney, A. M., Mitchell, K. K., Nazarian, E. J., Shudt, M., Thompson, L. M., Gubbala Venkata, S. L., & Musser, K. A. (2024). Implementation of a high-throughput whole genome sequencing approach with the goal of maximizing efficiency and cost effectiveness to improve public health. *Microbiology Spectrum*, *12*(4), e03885-23.

Dong, K., Song, D., Li, S., Wang, X., Dai, L., Pei, X., Yang, X., & Jiang, Y. (2025). Significance of Whole-Genome Sequencing for the Traceability of Foodborne Pathogens: During the Processing of Meat and Dairy Products. *Foods*, *14*(8), 1–22. https://doi.org/10.3390/foods14081410

Ekblom, R., & Wolf, J. B. W. (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary Applications*, *7*(9), 1026–1042. https://doi.org/10.1111/eva.12178

Farooq, A., Khan, U. M., Riaz, A., Maqbool, R., Ali, Z., & Sajjad, M. (2024). Genomes, Pangenomes, and Super Pangenomes of Cereals: How to Translate the Generated Information into Genetic Gains. In *Omics and System Biology Approaches for Delivering Better Cereals* (pp. 49–70). CRC Press.

Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, *26*(20), 5369–5406. https://doi.org/10.1111/mec.14264

Gilchrist, C. A., Turner, S. D., Riley, M. F., Petri, W. A., & Hewlett, E. L. (2015). Whole-genome sequencing in outbreak analysis. *Clinical Microbiology Reviews*, *28*(3), 541–563. https://doi.org/10.1128/CMR.00075-13

Hammond, J. A., Gordon, E. A., Socarras, K. M., Chang Mell, J., & Ehrlich, G. D. (2020). Beyond the pan-genome: current perspectives on the functional and practical outcomes of the distributed genome hypothesis. *Biochemical Society Transactions*, *48*(6), 2437–2455.

Hazzouri, K. M., Flowers, J. M., Visser, H. J., Khierallah, H. S. M., Rosas, U., Pham, G. M., Meyer, R. S., Johansen, C. K., Fresquez, Z. A., Masmoudi, K., Haider, N., El Kadri, N., Idaghdour, Y., Malek, J. A., Thirkhill, D., Markhand, G. S., Krueger, R. R., Zaid, A., & Purugganan, M. D. (2015). Whole genome re-sequencing of date palms yields insights into diversification of a fruit tree crop. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms9824

Jabeen, S., Saif, R., Haq, R., Hayat, A., & Naz, S. (2023). Whole-genome sequencing and variant discovery of Citrus reticulata “Kinnow” from Pakistan. *Functional and Integrative Genomics*, *23*(3), 1–13. https://doi.org/10.1007/s10142-023-01153-6

Krasilnikova, L. A., Sater, M. R. A., & Anahtar, M. N. (2025). Application-Specific Considerations for Clinical and Epidemiologic Sequencing. *Manual of Molecular Microbiology: Fundamentals and Applications*, 335–351.

Kumar, R., Das, S. P., Choudhury, B. U., Kumar, A., Prakash, N. R., Verma, R., Chakraborti, M., Devi, A. G., Bhattacharjee, B., & Das, R. (2024a). Advances in genomic tools for plant breeding: harnessing DNA molecular markers, genomic selection, and genome editing. *Biological Research*, *57*(1), 80.

Kumar, R., Das, S. P., Choudhury, B. U., Kumar, A., Prakash, N. R., Verma, R., Chakraborti, M., Devi, A. G., Bhattacharjee, B., & Das, R. (2024b). Advances in genomic tools for

plant breeding: harnessing DNA molecular markers, genomic selection, and genome editing. *Biological Research*, *57*(1), 80.

Kwong, J. C., Mccallum, N., Sintchenko, V., & Howden, B. P. (2015). Whole genome sequencing in clinical and public health microbiology. *Pathology*, *47*(3), 199–210. https://doi.org/10.1097/PAT.0000000000000235

Liu, T., & Conesa, A. (2025). Profiling the epigenome using long-read sequencing. *Nature Genetics*, *57*(1), 27–41.

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. https://doi.org/10.1038/s41576-020-0236-x

Lu, Y., Li, M., Gao, Z., Ma, H., Chong, Y., Hong, J., Wu, J., Wu, D., Xi, D., & Deng, W. (2025). Advances in Whole Genome Sequencing: Methods, Tools, and Applications in Population Genomics. *International Journal of Molecular Sciences*, *26*(1), 1–18. https://doi.org/10.3390/ijms26010372

Lv, F., Liu, Y., Zhang, Y., Liu, X., Zhihua, P., Yu, J., Chen, W., & He, J. (2021). *Accurately estimating the homologous recombination deficiency based on low-pass whole-genome sequencing and target sequencing.* Wolters Kluwer Health.

Ng, P. C., & Kirkness, E. F. (2010). Chapter 12: Whole Genome Sequencing. *Methods in Molecular Biology*, *628*, 215–226. https://doi.org/10.1007/978-1-60327-367-1

Oniciuc, E. A., Likotrafiti, E., Alvarez-Molina, A., Prieto, M., Santos, J. A., & Alvarez-Ordóñez, A. (2018). The present and future of whole genome sequencing (WGS) and whole metagenome sequencing (WMS) for surveillance of antimicrobial resistant microorganisms and antimicrobial resistance genes across the food chain. *Genes*, *9*(5), 1–28. https://doi.org/10.3390/genes9050268

Peace, C. P., Bianco, L., Troggio, M., van de Weg, E., Howard, N. P., Cornille, A., Durel, C. E., Myles, S., Migicovsky, Z., Schaffer, R. J., Costes, E., Fazio, G., Yamane, H., van Nocker, S., Gottschalk, C., Costa, F., Chagné, D., Zhang, X., Patocchi, A., … Vanderzande, S. (2019). Apple whole genome sequences: recent advances and new prospects. *Horticulture Research*, *6*(1). https://doi.org/10.1038/s41438-019-0141-7

Raghavan, V., Kraft, L., Mesny, F., & Rigerte, L. (2022). A simple guide to de novo transcriptome assembly and annotation. *Briefings in Bioinformatics*, *23*(2), bbab563.

Raman, T., Raj, E. E., Muthukathan, G., & Loganathan, M. (2021). Comparative Whole-Genome Sequence Analyses of Fusarium (AAA) Bananas in India, with a Special Emphasis on Pathogenicity Mechanisms. *Journal of Fungi*, *7*(717), 1–19.

Satam, H., Joshi, K., Mangrolia, U., Waghoo, S., Zaidi, G., Rawool, S., Thakare, R. P., Banday, S., Mishra, A. K., Das, G., & Malonia, S. K. (2023). Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology*, *12*(7). https://doi.org/10.3390/biology12070997

Sherman, R. M., & Salzberg, S. L. (2020). Pan-genomics in the human genome era. *Nature Reviews Genetics*, *21*(4), 243–254.

Sohn, J. Il, & Nam, J. W. (2018). The present and future of de novo whole-genome assembly. *Briefings in Bioinformatics*, *19*(1), 23–40. https://doi.org/10.1093/bib/bbw096

Thomas, M., O'Boyle, N. M., Bender, A., & De Graaf, C. (2022). Re-evaluating sample efficiency in de novo molecule generation. *ArXiv Preprint ArXiv:2212.01385*.

Varnderzande, S., Peace, C., & de Weg, E. van. (2024). Whole genome sequence improvement with pedigree information and reference genotypic profiles, demonstrated in outcrossing apple. *BioR*, *xiv*. https://doi.org/https://doi.org/10.1101/2024.08.08.607141 ;

Worley, K. C., Richards, S., & Rogers, J. (2017). The value of new genome references. *Experimental Cell Research*, *358*(2), 433–438. https://doi.org/10.1016/j.yexcr.2016.12.014

Wu, W., Yang, Y. L., He, W. M., Rouard, M., Li, W. M., Xu, M., Roux, N., & Ge, X. J. (2016). Whole genome sequencing of a banana wild relative Musa itinerans provides insights into lineage-specific diversification of the Musa genus. *Scientific Reports*, *6*(August), 1–11. https://doi.org/10.1038/srep31586

Xue, A., & Cui, Y. (2025). The research progress on crop genomics and genome-wide association studies: A review. *Advances in Resources Research*, *5*(1), 123–145.

Yao, X., Wang, S., Wang, Z., Li, D., Jiang, Q., Zhang, Q., Gao, L., Zhong, C., Huang, H., & Liu, Y. (2022). The genome sequencing and comparative analysis of a wild kiwifruit Actinidia eriantha. *Molecular Horticulture*, *2*(1), 1–14. https://doi.org/10.1186/s43897-022-00034-z

Yoon, C. W., Kang, Y., Chung, J.-W., Bulos, M., Hwang, T.-Y., & Yu, J.-K. (2025). Harnessing genomic technologies for modern soybean breeding. *Journal of Crop Science and Biotechnology*, 1–12.

Zhao, J., & F.A. Grant, S. (2011). Advances in Whole Genome Sequencing Technology. *Current Pharmaceutical Biotechnology*, *12*(2), 293–305. https://doi.org/10.2174/138920111794295729

Zheng, Y., Zhu, B., Tan, J., Guan, Y., Consortium, C. G. S. V., Morton, C. C., & Lu, G. (2022). Experience of low-pass whole-genome sequencing-based copy number variant analysis: a survey of chinese tertiary hospitals. *Diagnostics*, *12*(5), 1098.