Rice Genomics and Genetics 2025, Vol.16, No.6, 304-311
http://cropscipublisher.com/index.php/rgg

Review Article Open Access

Integration of Long-read Sequencing for Accurate Detection of Structural Variants in *Oryza* Genomes

Zufan Chen, Yuandong Hong, Jianquan Li
Hier Rice Research Center, Hainan Institute of Tropical Agricultural Resources, Sanya, 572025, Hainan, China
Corresponding email: jianquan.li@hitar.org
Rice Genomics and Genetics, 2025, Vol.16, No.6 doi: 10.5376/rgg.2025.16.0027
Received: 20 Sep., 2025
Accepted: 31 Oct., 2025
Published: 21 Nov., 2025
Copyright © 2025 Chen et al., This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Preferred citation for this article:
Chen Z.F., Huang Y.D., and Li J.Q., 2025, Integration of long-read sequencing for accurate detection of structural variants in *Oryza* genomes, Rice Genomics and Genetics, 16(6): 304-311 (doi: 10.5376/rgg.2025.16.0027)

Abstract Structural Variants (SVs), including insertions, deletions, inversions and translocations, are the key factors shaping the genomic structure and phenotypic diversity of rice (Oryza genus). However, due to the limitations of short-read sequencing technology in terms of read length and alignment accuracy, traditional methods often have insufficient accuracy when identifying complex or large-scale SVS. This study constructed an integrated framework using long-read sequencing technology, mainly including PacBio HiFi and Oxford Nanopore platforms, to achieve high-resolution detection and annotation of SV in multiple rice genomes, and systematically analyzed the biological significance of SV in regulating agronomic traits such as yield, stress resistance and adaptability. In the comparative analysis among japonica rice, indica rice and wild rice varieties, a large number of structural variation differences were found in the study, and insertion and deletion sites related to key phenotypic traits were identified. The expression analysis and PCR verification of the candidate SV loci further revealed their potential role in gene regulation. This study has laid the foundation for integrating SV data into a multi-omics breeding platform, which is expected to accelerate the precise improvement of rice varieties in the context of modern agriculture.
Keywords Rice genome; Long-read sequencing; Structural variation; Precise genome assembly; Molecular breeding

1 Introduction

Rice (*Oryza sativa*), as one of the staple foods, feeds the majority of the world's population. Its adaptability, yield performance and various agronomic traits are largely attributed to the complex genetic diversity behind it. However, this diversity did not form overnight - it is not only the accumulation of natural evolution but also deeply influenced by long-term artificial domestication. For breeders, this is precisely an inexhaustible "genetic resource bank", especially in terms of enhancing stress resistance and yield traits. But things are not always that simple. Large-scale genomic analyses in recent years have identified a large number of variations, including SNPS and structural variations (SVs), which directly shape the phenotypic characteristics of different subpopulations of rice (Zhao et al., 2018; Wang et al., 2023). Among them, structural variations such as insertions, deletions, inversions and duplications often have a wide range of influence and deep regulatory levels, which can cause substantial intervention in gene expression and subsequently lead to significant differences in traits.

However, many past studies have encountered numerous bottlenecks in identifying SV due to the limitations of sequencing technology, especially those relying on short-read platforms. The rice genome has many repetitions and a complex structure, and it is very easy to miss key information using traditional methods. It was not until long-read sequencing technology gradually matured that this predicament found a breakthrough - now, researchers can more clearly "see" large variations in the genomic structure and more accurately assess the potential impact of these variations on the phenotype. Therefore, this study attempts to use long-read data to conduct a more systematic identification and annotation of SV in the rice genome, hoping to clarify the molecular mechanism of how SV affects key traits and provide a reference framework and data support for future rice molecular breeding (Fuentes et al., 2019; Qin et al., 2021; Zheng et al., 2023).

This study introduces the application of long-read sequencing technology in the detection of structural variations (SV), analyzes the impact of SV on gene expression and agronomic traits, explores its significance for rice

breeding, constructs high-quality SV maps, verifies candidate variations related to important traits, and demonstrates the application value of long-read sequencing data in overcoming the limitations of previous methods. This study significantly promotes the development of rice genomics by providing more comprehensive information on genomic structural diversity, facilitating functional genomics research, and supporting the cultivation of improved rice varieties through molecular breeding strategies.

2 Types of Structural Variants and Their Biological Significance in Rice Genomes

2.1 Common types of SVs: insertions, deletions, inversions, translocations, and duplications

The genome of rice is not static. It often undergoes some invisible but influential changes, such as insertions, deletions, inversions, translocations, and duplications. These structural variations (SVS) do not necessarily have earth-shattering effects. Sometimes it is just the insertion of a few bases, and sometimes it is the rearrangement of millions of bases. An inversion of approximately 4.3 Mb between indica rice and japonica rice is a typical example. Once such variations involve regulatory regions or multiple genes, they often lead to differences in traits. There are also some moving components, such as SINE and mariner. Although they occur less frequently, they are also involved in the formation of SV - but a low frequency may also mean that they are not so "safe" (Kou et al., 2020; Qin et al., 2020).

2.2 Impact of SVs on rice adaptation, stress resistance, and yield-related traits

Not all SVS leave traces, but some do change the "character" of rice. Especially those SVS enriched in regions related to stress responses seem to be specifically speaking to the environment. Some variations altered the length of the grains, some were associated with plant type and seed setting rate, and others played a role in the expression differences among different subspecies (Figure 1). In other words, some SVS have been helpful in helping rice adapt to extreme environments; But it is not without cost - some variations brought about by domestication, although they have increased yields, may have lost some resistance. However, then again, it is precisely these variations that provide the material for improving traits (Fuentes et al., 2019; Zheng et al., 2023; Dan et al., 2025).

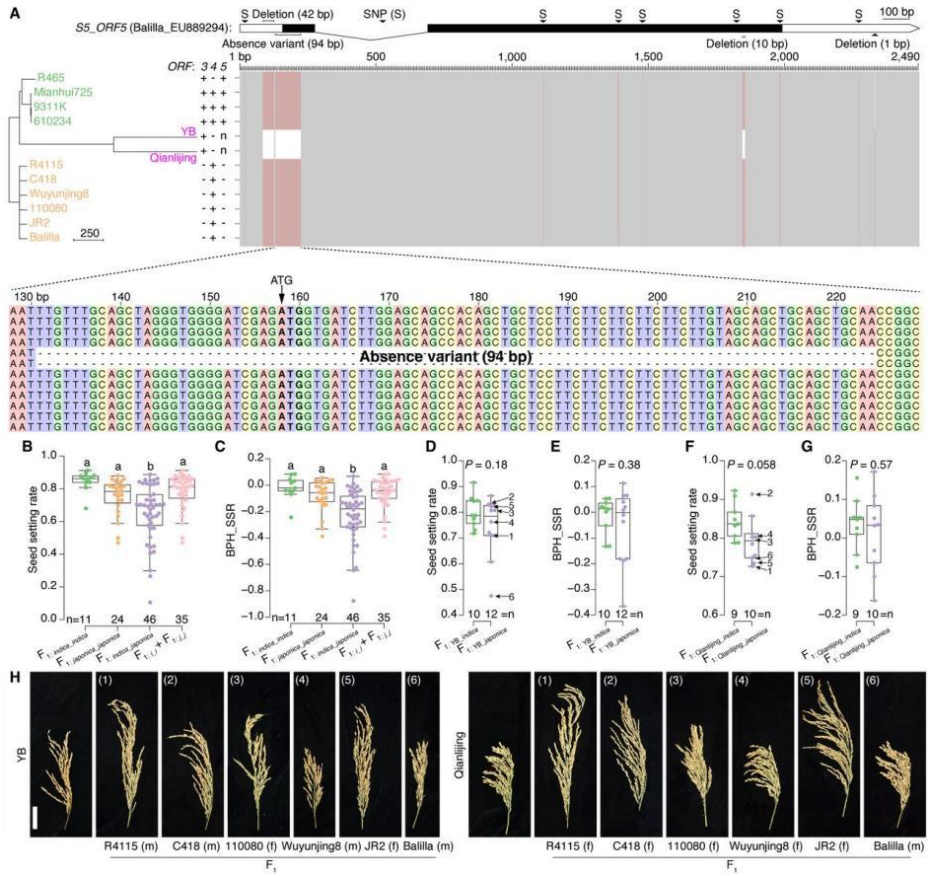

Figure 1 Contribution of AVs in *S5-ORF5* to heterosis for seed setting rate (Adopted from Dan et al., 2025)

2.3 Current progress in SV–phenotype association studies

In recent years, the maturation of long-read technology and the introduction of the pan-genome have made the research on SV more specific and reliable. Mutations that were hard to figure out in the past can now be traced in tens of thousands of rice samples. Some important traits, such as flowering time, flood resistance and grain weight, have begun to be associated with specific structural variations (Huan et al., 2025). Not only have the locations of the mutations been identified, but some studies have also targeted the "suspect" genes that might cause phenotypic differences. These achievements not only enrich our understanding of the evolution and domestication process of rice, but also provide a sense of direction for practical breeding: there is data support for which SVS are worth retaining and which can be screened out (Qin et al., 2021).

3 Long-read Sequencing Technologies and Their Applications in Plant Genomics

3.1 Comparison of major platforms: PacBio HiFi and Oxford Nanopore sequencing

In recent years, research on plant genomes has made rapid progress, and one of the key driving forces is long-read sequencing technology. Currently, the two mainstream platforms are PacBio's HiFi and Oxford Nanopore (ONT). They each have their own focuses, and it's impossible to simply say which one is better. HiFi excels in precision. It can generate circular consistent read segments with extremely low error rates, making it particularly suitable for handling areas that are densely repeated. ONT, on the other hand, has a maximum read length, often reaching hundreds of kilobytes, making it particularly suitable for detecting large structural variations. Of course, ONT had a relatively high error rate in its early days, which was a major shortcoming of it. However, the improvement has been quite obvious in recent years. How to choose a platform? Usually, it still depends on the specific research goals: if ultra-high precision is pursued, PacBio is a conventional option; If structural variations are of concern or the budget is limited, the flexibility of ONT is also very attractive (Murigneux et al., 2020; Mascher et al., 2021; Pucker et al., 2022).

3.2 Read quality, coverage, and cost considerations of long-read data

Whether the reading paragraphs are long or not is one thing, but whether they are accurate and sufficient is another. The advantage of HiFi is that it can provide an accuracy rate of over 99% within a length of 15 kb to 20 kb. Therefore, even when the coverage does not need to be too high, the assembled result can still be very stable. Generally speaking, a coverage of 20 times is sufficient. Although ONT reads longer, the accuracy of the original data is slightly lower. Therefore, to achieve comparable accuracy, more depth needs to be added, and it also has to rely on later data processing to "remedy". As for the cost, ONT is relatively more flexible and suitable for experimental designs of all scales. Although PacBio is a bit more expensive, it is accurate enough to read it once and saves a lot of subsequent troubles. Nowadays, some methods for library preparation and DNA extraction are becoming increasingly mature. No matter which platform it is, the requirements for samples are actually much lower than before (Li et al., 2017; Russo et al., 2022; Kang et al., 2023).

3.3 Advantages of long-read over short-read sequencing in genome assembly and SV detection

The short-read long-read technique was once the main force, but its limitations have been exposed in polyploid crops or plants with a large number of repetitive regions. For a complex genome like that of rice, if the read segments are too short, it is prone to "skipping the needle" - it may seem like a deletion, but it is actually a rearrangement, or it may not be visible at all. In contrast, long-read long data can read large segments at once, and even structural variations can be captured more completely, such as insertions, deletions, inversions, and even translocations. In this way, the assembly is more coherent and the SV detection is more accurate. Therefore, many pan-genome studies have begun to rely on long-read data to construct high-quality references and even promote crop improvement. Those variations that were originally omitted by the short read segments thus have the opportunity to show up (Mascher et al., 2021; Pucker et al., 2022).

4 SV Detection Pipelines and Algorithms Based on Long-read Data

4.1 Data preprocessing: filtering, alignment, and error correction

Before conducting structural variation (SV) detection, there is actually one more "fundamental" thing that needs to

be clarified first, and that is data preprocessing. Don't underestimate this step. If the quality of the read segment is not up to standard, no matter how good the algorithm is later on, it will be in vain. Generally speaking, researchers will first filter out low-quality read segments and then use tools like Minimap2 to align them onto the reference genome. This type of tool is fast and accurate in matching, so it is commonly used. The problem is that although long-read sequencing has a significant advantage in reading length, mismatches are also prone to occur, and error correction algorithms have to be used as a fallback - some error correction modules are embedded in the comparison tool, while others run separately. In short, it is all to minimize false positives as much as possible. After all, in the context of the inherent error rate of long-read platforms, if the early data is not handled carefully, the SV detected later will not be very reliable (Helal et al., 2024; Liu et al., 2024; Meleshko et al., 2025).

4.2 Overview and comparison of common SV detection tools (e.g., Sniffles2, CuteSV, SVIM)

Nowadays, with so many tools available, I actually don't know which one to choose. Sniffles2 is a detection tool that many people default to using. It has good accuracy and is particularly good at handling complex structures. CuteSV is more sensitive, has better scalability, and can handle large samples without any problem. SVIM is relatively moderate, aiming to strike a balance between sensitivity and accuracy. The core ideas of these tools are mostly based on extracting mutation clues from comparison results. However, "newcomers" like SVHunter have begun to introduce deep learning models, such as Transformer, to identify complex mutation patterns(Naing et al., 2025), which is quite trendy. Ultimately, however, no tool can be "all-powerful" under all test conditions. The effect still depends on the data. For instance, genomic background, coverage depth, SV type, and so on all affect the final performance. So the more common practice nowadays is to use multiple tools in combination, with no one fully trusting and no one being absent (Jiang et al., 2020; Helal et al., 2024; Gao et al., 2025).

4.3 Integrated strategies and standardized workflows for accurate SV calling

If one tool can't handle it, many people eventually adopt the strategy of using multiple ones at once. Integrated processes usually do not only rely on comparison tools but also incorporate assembly tools, such as Sniffles2 and CuteSV, in combination with other methods. In this way, when dealing with large insertions or complex rearrangements, both the detection range and depth can be expanded. Some key nodes will also be designed in the process, such as read segment consistency checks, breakpoint cluster analysis, as well as screening criteria for comparison quality and read segment depth. Although this hybrid strategy is cumbersome, it can enhance the stability and explanatory power of the results. It is particularly useful for plant materials with complex genomic structures, numerous repetitive sequences and uneven coverage. Not only have the sensitivity and specificity been improved, but subsequent molecular breeding analyses can also be based on more solid data (He et al., 2025; Peng and Chong, 2025).

5 Case Study: SV Detection and Functional Annotation among Rice Varieties

5.1 Differences in SVs between japonica and indica rice and their influence on key traits

The differences between japonica rice and indica rice are not merely about the minor issues of "plant type" or "grain size" in terms of phenotype. More deeply, they lie in the differentiation of structural variations at the genomic level. For instance, there is approximately 4.3 Mb of inversions between indica rice and japonica rice. Coupled with large areas of insertions, deletions and repetitive sequences, it is already sufficient to draw a clear boundary between the two genomes. These SVS are not "harmless differences"; some directly affect stress responses, plant type regulation, and even yield-related traits. Group-level analysis also reveals that compared with wild rice, cultivated rice has accumulated more "new" SVS, which is inseparable from the selective preferences in the human domestication process. That is to say, many of the phenotypic differences between indica rice and japonica rice that we see today have actually been foreshadowed in genomic structure long ago (Qin et al., 2020).

5.2 Association of specific insertions/deletions with drought resistance, disease resistance, etc.

Not all SVS are useful, but those that fall into key areas are hard to go unnoticed. Some insertions or deletions occur precisely in regulatory regions, such as near promoters, affecting the expression of genes related to drought

resistance or disease resistance. Through long-read sequencing combined with pan-genome data for genome-wide association analysis (GWAS), it has been possible to clearly link these SVS with traits. Some SVS are also concentrated in pathways involved in abiotic stress or pathogen defense, which further indicates that they do not occur by chance. Looking for new alleles that can be put to use? Analyzing these SVS together with phenotypic information and environmental responses is a good entry point (Fuentes et al., 2019; Zheng et al., 2023).

5.3 Identification of beneficial SVs in wild rice using long-read sequencing

Domestication has brought about adaptability, but also the loss of genetic diversity. For instance, in wild rice, nearly half of the SV cannot be found in modern cultivated rice at all - behind these unique variations, there may be potential factors related to disease resistance, drought tolerance and even yield advantages. In the past, we couldn't see so finely, but now with long-read sequencing, these "invisible" SVS have become both detectable and annotable. More importantly, these SVS not only increase the diversity of genomic structure, but some of them have also been verified to be related to beneficial phenotypes. As long as high-quality assembly data can be combined with these variations to screen out the batch that can be introduced into cultivated rice, new candidate resources can be provided for breeding projects in a solid manner (Figure 2) (Kou et al., 2020; Qin et al., 2021).

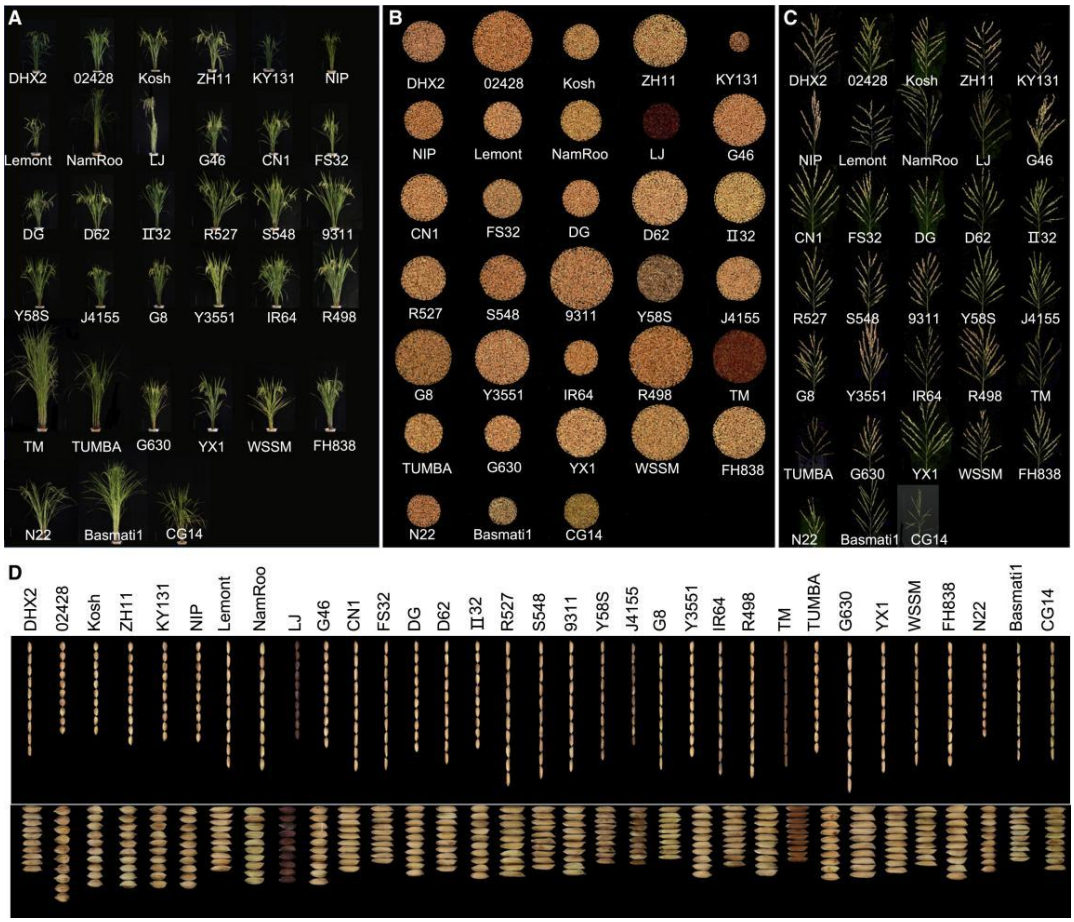

Figure 2 Diverse agronomic phenotypes for genetically diverse rice accessions (Adopted from Qin et al., 2021)

6 Functional Validation of SVs and Applications in Molecular Breeding

6.1 PCR/qPCR-based validation and expression profiling of candidate SV loci

Functional verification of structural variations (SVS) is often not directly obtained from the laboratory but starts with PCR or qPCR. These detection methods are mainly used to confirm whether the SV identified through long-read sequencing truly exists, and at the same time, they can precisely locate the position of the mutation breakpoint. On this basis, researchers further classified different germplasm materials of rice to see if the allele frequencies of these variations were representative. It is worth noting that some SVS do not directly alter the protein sequence but occur near the promoter. Such positional changes may also alter gene expression by

influencing the binding of transcription factors. Cases like this, which are further confirmed through expression profiling analysis, are not uncommon in plants. In conclusion, to determine whether a certain SV is worth considering in breeding, these steps are basically unavoidable (Li et al., 2025; Shi et al., 2025).

6.2 Integration of SV data with QTL mapping and gene expression analysis

Whether SV itself can reflect the phenotype or not still needs to be seen more clearly in combination with QTL and transcriptome data. Sometimes, it is difficult to determine whether a structural variation is functionally related just by looking at it. However, if this SV happens to fall within the QTL region related to a certain agronomic trait, such as disease resistance or yield trait, then its "suspicion" is quite high. Furthermore, if through expression analysis it can be found that the expression of genes in this region varies among different materials, or even certain quantitative trait loci (eQTL) that match it can be identified, then the role of SV will be clear. This type of integration strategy has been maturely applied in rice, and many new allelic variations have been discovered precisely in this way - these might have been directly ignored in conventional SNP analysis (Qin et al., 2021; Sehrawat et al., 2022; Zhang et al., 2024).

6.3 Development of SV-based molecular markers and prospects for marker-assisted selection

In the selection of breeding markers, SNPS are no longer the sole leading role. Structural variations, due to their large scale and more direct impact on gene functions, have gradually become a new focus of attention in recent years. SV sites confirmed by PCR or sequencing can be directly developed as molecular markers, which are particularly suitable for tracking complex traits in marker-assisted selection (MAS). In contrast, SV often has a stronger explanatory power in many traits. Of course, the prerequisite is that these SVS should have standardized tag groups to facilitate efficient screening in large groups. At present, high-throughput sequencing and data analysis technologies have become relatively mature, and the detection and labeling design of SV are no longer as troublesome as they used to be. Further studies have also incorporated SV into genomic selection (GS) models, which have been proven to improve the accuracy of trait prediction - which is of great significance for enhancing breeding efficiency and crop adaptability (Panahi et al., 2024; Zong et al., 2024).

7 Conclusions and Future Perspectives

Research on the rice genome has made rapid progress in recent years, especially after the introduction of long-read sequencing. In the past, when using the short-read long-read technique, I always felt that there was something not detailed enough, especially when encountering repetitive sequences or structural variations (SVS), the parsing was often unclear. But now, with the successive establishment of unbroken reference genomes for multiple rice populations, many new sequences and unannotated genes that were previously unclear have also come to light. It can be said that the introduction of this technology has brought us one step closer to "seeing the whole picture", especially in revealing the genomic structural differences and evolutionary paths (SVs) between cultivated rice and wild rice.

However, one should not speak too confidently. This technology has indeed opened up a new window, but it is not without problems. For instance, some platforms themselves have a high error rate, and complex repetitive areas are prone to errors when pieced together. Even with data available, they may not be assembled accurately. Furthermore, to associate these structural variations with specific traits, relying solely on genomic data is not enough; multi-omics information such as transcriptomics and phenotypes is also required. However, in terms of functional verification, many genes have not yet been clearly defined for exactly what they do. Coupled with the fact that the regulatory mechanism itself is very complex, these factors have stuck the subsequent analysis. Therefore, in addition to collecting data, it is also necessary to promptly make up for the "shortcomings" in the standard processes and tools.

The path ahead is actually quite clear: on one hand, continue to enrich genomic details through long-read sequencing; on the other hand, focus on multi-omics integration. Elements like the transcriptome, epigenome, and proteome all need to be included in order to truly explain the functions of SV and other genetic variations thoroughly. Marker development cannot be halted either, especially those related to important agronomic traits. If

these structural variations can be applied to genomic selection or precision breeding, the speed of rice improvement is sure to increase. Ultimately, only by integrating high-quality genomic resources, intelligent analysis methods and gene editing tools into one can there be hope to maximize natural variations and ensure that rice is grown steadily and well.

Acknowledgments

We are grateful to Dr. Z. Hu for his assistance with the serious reading and helpful discussions during the course of this work.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Dan Z., Chen Y., and Huang W., 2025, Structural variations contribute to subspeciation and heterosis in rice, bioRxiv, 2024: 11.
https://doi.org/10.1101/2024.11.03.621459

Fuentes R., Chebotarov D., Duitama J., Smith S., De La Hoz J., Mohiyuddin M., Wing R., McNally K., Tatarinova T., Grigoriev A., Mauleon R., and Alexandrov N., 2019, Structural variants in 3000 rice genomes, Genome Research, 29: 870-880.
https://doi.org/10.1101/gr.241240.118

Gao R., Hu H., Jiang Z., Cao S., Wang G., Zhao Y., and Jiang T., 2025, SVHunter: long-read-based structural variation detection through the transformer model, Briefings in Bioinformatics, 26(3): 819-833.
https://doi.org/10.1093/bib/bbaf203

He S., Song B., Tang Y., Qu X., Li X., Yang X., Bao Q., Fang L., Jiang J., Tang Z., and Yi G., 2025, Systematic benchmarking of tools for structural variation detection using short- and long-read sequencing data in pigs, iScience, 28(3): 111983.
https://doi.org/10.1016/j.isci.2025.111983

Helal A., Saad B., Saad M., Mosaad G., and Aboshanab K., 2024, Benchmarking long-read aligners and SV callers for structural variation detection in Oxford nanopore sequencing data, Scientific Reports, 14(1): 1-22.
https://doi.org/10.1038/s41598-024-56604-2

Huang D.S., Chen R.C., and Li J.Q., 2025, Dissecting complex traits in rice: insights from recent GWAS findings, Plant Gene and Trait, 16(2): 47-55.
https://doi.org/10.5376/pgt.2025.16.0006

Jiang T., Liu Y., Jiang Y., Li J., Gao Y., Cui Z., Liu Y., and Wang Y., 2020, Long-read-based human genomic structural variation detection with cuteSV, Genome Biology, 21(189): 1-17.
https://doi.org/10.1186/s13059-020-02107-y

Kang M., Chanderbali A., Lee S., Soltis D., Soltis P., and Kim S., 2023, High‐molecular‐weight DNA extraction for long‐read sequencing of plant genomes: an optimization of standard methods, Applications in Plant Sciences, 11(3): e11528.
https://doi.org/10.1002/aps3.11528

Kou Y., Liao Y., Toivainen T., Lv Y., Tian X., Emerson J., Gaut B., Zhou Y., and Purugganan M., 2020, Evolutionary genomics of structural variation in Asian rice (*Oryza sativa*) domestication, Molecular Biology and Evolution, 37: 3507-3524.
https://doi.org/10.1093/molbev/msaa185

Li C., Lin F., An D., Wang W., and Huang R., 2017, Genome sequencing and assembly by long reads in plants, Genes, 9(1): 1-11.
https://doi.org/10.3390/genes9010006

Li J., Yiming S., Zheng W., Wu W., Bi J., Li F., and Zhu M., 2025, Identification of candidate genes for reproductive traits in Xinjiang sheep breeds based on genomic structural variation, Frontiers in Veterinary Science, 12: 1551293.
https://doi.org/10.3389/fvets.2025.1551293

Liu Y., Luo C., Golding S., Ioffe J., and Zhou X., 2024, Tradeoffs in alignment and assembly-based methods for structural variant detection with long-read sequencing data, Nature Communications, 15(1): 2447.
https://doi.org/10.1038/s41467-024-46614-z

Mascher M., Wicker T., Jenkins J., Plott C., Lux T., Koh C., Ens J., Gundlach H., Boston L., Tulpová Z., Holden S., Hernández-Pinzón I., Scholz U., Mayer K., Spannagl M., Pozniak C., Sharpe A., Šimková H., Moscou M., Grimwood J., Schmutz J., and Stein N., 2021, Long-read sequence assembly: a technical evaluation in barley, The Plant Cell, 33: 1888-1906.
https://doi.org/10.1093/plcell/koab077

Meleshko D., Yang R., Maharjan S., Danko D., Korobeynikov A., and Hajirasouliha I., 2025, Blackbird: structural variant detection using synthetic and low-coverage long-reads, Bioinformatics Advances, 5(1): vbaf151.
https://doi.org/10.1093/bioadv/vbaf151

Murigneux V., Rai S., Furtado A., Bruxner T., Tian W., Ye Q., Wei H., Yang B., Harliwong I., Anderson E., Mao Q., Drmanac R., Wang O., Peters B., Xu M., Wu P., Topp B., Coin L., and Henry R., 2020, Comparison of long-read methods for sequencing and assembly of a plant genome, GigaScience, 9(12): giaa146.

Naing N.N.Z.N., Wang C.L., Zhou X.L., Zhang C., Li J.J., Li J., Zhu Q., Lee D.S., and Chen L.J., 2025, Molecular mechanisms of rice drought resistance genes and their prospects in breeding, Molecular Plant Breeding, 16(3): 165-179.
https://doi.org/10.5376/mpb.2025.16.0017

Panahi B., Jalaly H., and Hamid R., 2024, Using next-generation sequencing approach for discovery and characterization of plant molecular markers, Current Plant Biology, 40: 100412.
https://doi.org/10.1016/j.cpb.2024.100412

Peng Z., and Chong Z., 2025, Abstract 5056: Full spectrum of somatic structural variations (SVs) detection in COLO829 with long-read sequencing, Cancer Research, 85(8_Suppl_1): 5056.
https://doi.org/10.1158/1538-7445.am2025-5056

Pucker B., Irisarri I., De Vries J., and Xu B., 2022, Plant genome sequence assembly in the era of long reads: progress challenges and future directions, Quantitative Plant Biology, 3: e5.
https://doi.org/10.1017/qpb.2021.18

Qin P., Lu H., Du H., Chen W., Wang H., Ou S., Chen Z., Li X., Li Y., Liao Y., Gao Q., Tu B., Yuan H., Ma B., Wang Y., Qiang Y., Fan S., Li W., Wang J., He M., Yin J., Jiang N., Chen X., Liang C., and Li S., 2020, A catalog of structural and gene copy number variations of cultivated rice, Social Science Research Network, 11: 218.
https://doi.org/10.2139/ssrn.3525548

Qin P., Lu H., Du H., Wang H., Chen W., Chen Z., He Q., Ou S., Zhang H., Li X., Li X., Li Y., Liao Y., Gao Q., Tu B., Yuan H., Ma B., Wang Y., Qian Y., Fan S., Li W., Wang J., He M., Yin J., Li T., Jiang N., Chen X., Liang C., and Li S., 2021, Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations, Cell, 184(13): 3542-3558.e16.
https://doi.org/10.1016/j.cell.2021.04.046

Russo A., Mayjonade B., Frei D., Potente G., Kellenberger R., Frachon L., Copetti D., Studer B., Frey J., Grossniklaus U., and Schlüter P., 2022, Low-input high-molecular-weight DNA extraction for long-read sequencing from plants of diverse families, Frontiers in Plant Science, 13: 883897.
https://doi.org/10.3389/fpls.2022.883897

Sehrawat S., Najafian K., and Jin L., 2022, Predicting phenotypes from novel genomic markers using deep learning, Bioinformatics Advances, 3(1): vbad028.
https://doi.org/10.1093/bioadv/vbad028

Shi L., Zhang P., Yu B., Cheng L., Liu S., Liu Q., Zhou Y., Xiang M., Zhao P., and Chen H., 2025, Genomic analysis of indel and SV reveals functional and adaptive signatures in hubei indigenous cattle breeds, Animals, 15(12): 1755.
https://doi.org/10.3390/ani15121755

Wang J., Yang W., Zhang S., Hu H., Yuan Y., Dong J., Chen L., Ma Y., Yang T., Zhou L., Chen J., Liu B., Li C., Edwards D., and Zhao J., 2023, A pangenome analysis pipeline provides insights into functional gene identification in rice, Genome Biology, 24: 282.
https://doi.org/10.1186/s13059-023-02861-9

Zhang Z., Viana J., Zhang B., Walden K., Paul H., Moose S., Morris G., Daum C., Barry K., Shakoor N., and Hudson M., 2024, Major impacts of widespread structural variation on sorghum, Genome Research, 34: 286-299.
https://doi.org/10.1101/gr.278396.123

Zhao Q., Feng Q., Lu H., Li Y., Wang A., Tian Q., Zhan Q., Lu Y., Zhang L., Huang T., Wang Y., Fan D., Zhao Y., Wang Z., Zhou C., Chen J., Zhu C., Li W., Weng Q., Xu Q., Wang Z., Wei X., Han B., and Huang X., 2018, Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice, Nature Genetics, 50: 278-284.
https://doi.org/10.1038/s41588-018-0041-z

Zheng X., Zhong L., Pang H., Wen S., Li F., Lou D., Ge J., Fan W., Wang T., Han Z., Qiao W., Pan X., Zhu Y., Wang J., Tang C., Wang X., Zhang J., Xu Z., Kim S., Kohli A., Ye G., Olsen K., Fang W., and Yang Q., 2023, Lost genome segments associate with trait diversity during rice domestication, BMC Biology, 21 (1): 20.
https://doi.org/10.1186/s12915-023-01512-6

Zong W., Zhao R., Wang X., Zhou C., Wang J., Chen C., Niu N., Zheng Y., Chen L., Liu X., Hou X., Zhao F., Wang L., Wang L., Song C., and Zhang L., 2024, Population genetic analysis based on the polymorphisms mediated by transposons in the genomes of pig, DNA Research, 31(2): dsae008.
https://doi.org/10.1093/dnares/dsae008

Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.