GenBreed Publisher®

Molecular Plant Breeding 2024, Vol.15, No.5, 295-307
http://genbreedpublisher.com/index.php/mpb

Feature Review **Open Access**

Tailor-Made Rice: Using Haplotype Analysis to Design High-Yielding Varieties

Juan Li [1,2,3], Hui Zhang [1,3], Qian Zhu [1,2,3], Yanbo Xia [4], Zilin Duan [5], Jiancheng Wen [1,2], Lijuan Chen [1,2,3] ✉
1 Rice Research Institute, Yunnan Agricultural University, Kunming, 650201, Yunnan, China
2 The Key Laboratory for Crop Production and Smart Agriculture of Yunnan Province, Yunnan Agricultural University, Kunming, 650201, Yunnan, China
3 College of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming 650201, Yunnan, China
4 Seed Management Station of Yunnan Province, Kunming, 650031, Yunnan, China
5 Yunnan Yuanfang Agricultural Science and Technology Limited Company, Kunming, 650201, Yunnan, China
✉ Corresponding email: chenlijuan@hotmail.com

Molecular Plant Breeding, 2024, Vol.15, No.5 doi: 10.5376/mpb.2024.15.0028
Received: 10 Sep., 2024
Accepted: 11 Oct., 2024
Published: 22 Oct., 2024

Preferred citation for this article:
Li J., Zhang H., Zhu Q., Xia Y.B., Duan Z.L., Wen J.C., and Chen L.J., 2024, Tailor-made rice: using haplotype analysis to design high-yielding varieties, Molecular Plant Breeding, 15(5): 295-307 (doi: 10.5376/mpb.2024.15.0028)

Abstract This study explores the potential of haplotype analysis in the development of high-yielding rice varieties tailored for specific agronomic needs. Through detailed examination of genetic markers and haplotypes associated with key traits such as grain yield, quality, and resilience to environmental stresses, we demonstrate the efficacy of utilizing advanced genomic tools in rice breeding. Leveraging genome-wide association studies (GWAS) and haplotype-pheno analysis, we then identify specific haplotypes that contribute to enhanced agronomic traits, offering significant improvements in both yield and stability under varying climatic conditions. The integration of next-generation sequencing and machine learning in haplotype analysis has further refined the selection process, enabling the precise development of rice varieties that are not only productive but also suited to diverse environmental challenges and consumer preferences. We also discuss the broader implications of haplotype-based breeding (HBB) techniques, including their role in promoting sustainable agricultural practices and enhancing food security globally. These findings underline the transformative potential of haplotype analysis in rice genomics, paving the way for future innovations in crop improvement.

Keywords Haplotype analysis; Rice breeding; Genomic tools; High-yielding varieties; Sustainable agriculture

1 Introduction

Rice serves as a staple food for more than half of the world's population, making its yield and quality essential for global food security. Traditional breeding methods have been used extensively to enhance rice varieties. While these methods have achieved some successes, they often struggle to meet the increasing demands for higher yields and better quality. These conventional methods typically rely on phenotypic selection, which can be time-consuming and imprecise due to the complex nature of genetic traits affecting yield and quality. Moreover, the genetic gains from traditional breeding are generally incremental and insufficient to match the needs of a rising global population and changing environmental conditions, such as climate change and the necessity for sustainable agricultural practices. This slow pace of progress highlights the limitations of traditional methods in addressing urgent agricultural challenges (Abbai et al., 2019; Varshney et al., 2021a; Wang et al., 2023).

Haplotype is the combination of a series of genetic mutations that coexist on a single chromosom, contains a complete set of genetic information that is fundamental to the description of an individual genome and an essential aspect of genomic research. Haplotype analysis is a genetic research method for analyzing specific combinations of a set of continuous genetic markers (e.g., SNP) at a locus. These markers can come from different loci on the same chromosome and form haplotypes that are used to describe genetic variation between individuals (Li et al., 2018). Haplotype analysis offers a promising alternative to traditional breeding methods by enabling more precise and efficient selection of desirable genetic traits. This approach involves the identification and utilization of superior haplotype combinations of key genes that govern important traits such as grain yield and quality. Recent studies have demonstrated the potential of haplotype analysis to uncover significant genetic variations and superior haplotypes within large rice genome panels, paving the way for the development of tailor-made rice varieties with enhanced genetic gains (Abbai et al., 2019; Wei et al., 2024). By integrating genomics technologies with crop physiology, haplotype analysis can facilitate high-throughput identification of the genetic architecture of adaptive traits, making it a powerful tool for precision breeding (Varshney et al., 2021a).

This study aims to explore the application of haplotype analysis in creating high-yielding rice varieties. We focus on pinpointing optimal haplotype combinations within key genes that affect grain yield and quality, assessing their impact on global rice production, and examining the viability of using this approach to develop new, customized rice varieties that satisfy future nutritional and food demands. We seek to enhance precision breeding methods and respond to the critical need for more reliable and sustainable rice production.

2 Genetic Basis of Rice Yield

2.1 Understanding genetic factors influencing yield traits

Rice yield is a complex trait influenced by multiple genetic factors. The primary components determining rice yield include the number of grain number per panicle (GNP), thousand-grain weight (TGW), and effective panicle number per plant (ENP). These traits are controlled by numerous minor-effect loci, making it challenging to identify specific yield-related genes directly (Su et al., 2021). Quantitative trait loci (QTLs) play a significant role in controlling these yield components, and their effects can vary depending on the genetic background of the rice variety. For instance, editing QTLs using CRISPR/Cas9 has shown that the same QTL can have diverse effects on grain yield in different rice varieties (Shen et al., 2018). Additionally, genetic interactions, such as epistatic interactions, contribute significantly to yield-related traits, further complicating the genetic architecture of rice yield (Li et al., 2022; Wei et al., 2024).

2.2 Role of genetic diversity in yield optimization

Genetic diversity is crucial for optimizing rice yield. Studies have shown significant genetic variation among different rice genotypes for various yield-related traits, such as plant height, effective panicle number per plant, flag leaf area, grain number and grain size (Nath and Kole, 2021; Cao et al., 2023; Liu et al., 2023; Wang et al., 2023). This diversity allows for the selection of superior alleles that can enhance yield under different environmental conditions. For example, haplotype analysis of key genes governing grain yield and quality traits has revealed superior haplotypes that can be used to develop tailor-made rice varieties with enhanced genetic gains (Abbai et al., 2019). Moreover, understanding the genetic basis of correlations among growth duration and yield traits can help in developing crop varieties that efficiently use light and energy resources, thereby producing high yields (Li et al., 2018).

2.3 Overview of haplotype analysis and its relevance to genetics

Haplotype analysis is a powerful tool for understanding the genetic basis of complex traits like rice yield. It involves examining combinations of alleles at multiple loci to identify superior haplotypes that contribute to desirable traits. In rice, haplotype analysis has been used to identify superior haplotypes for genes influencing grain yield and quality traits (Abbai et al., 2019; Sivabharathi et al, 2024). This approach allows for the development of next-generation rice varieties with tailored genetic profiles that meet specific breeding goals. For instance, haplotype analysis of 120 functionally characterized genes revealed significant variations in the 3K rice genome panel, providing insights into the genetic basis of yield and quality traits. By leveraging haplotype-based breeding, it is possible to combine superior alleles from different genes to create high-yielding rice varieties that are better suited to future food and nutritional demands (Abbai et al., 2019; Liu et al., 2023).

3 Haplotype Analysis Techniques

3.1 Methods for identifying and analyzing haplotypes in rice

Haplotype analysis in rice involves the identification and characterization of specific combinations of alleles at multiple loci that are inherited together. Traditional methods rely on short-read sequencing technologies, which require computational assembly of shorter haplotype fragments due to the limited length of reads. This process can be challenging due to high variability and repeat sequences in the genome (Garg, 2021). Recent advancements have introduced long-read sequencing technologies, such as those provided by Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio), which allow for more direct and accurate haplotype phasing (Maestri et al., 2020; Fruzangohar et al., 2022). Additionally, targeted capture sequencing has been employed to analyze specific genes associated with important agronomic traits, such as grain size and weight, in rice. This method has successfully identified superior haplotypes and haplotype combinations that can be used in breeding programs (Liu et al., 2023).

3.2 Recent advances in genomic sequencing technologies

The field of genomic sequencing has seen significant advancements with the development of third-generation sequencing technologies. These technologies, such as ONT and PacBio, produce long reads that span larger genomic regions, facilitating direct haplotype phasing and reducing the reliance on statistical methods. Long-read sequencing has proven particularly useful in clinical settings for variant calling and phasing, enabling the analysis of complex genomic regions with high accuracy (Maestri et al., 2020). Additionally, linked-read sequencing techniques, such as haplotagging, have been developed to retain haplotype information while sequencing large populations, providing a cost-effective and efficient approach for whole-genome haplotyping (Meier et al., 2020; Kolesnikov et al., 2024). These advancements have significantly improved the resolution and accuracy of haplotype analysis, making it more accessible for various applications, including rice breeding.

3.3 Integration of bioinformatics tools in haplotype analysis

The integration of bioinformatics tools is crucial for the accurate assembly and analysis of haplotypes from sequencing data. Several algorithms and software tools have been developed to address the challenges of haplotype reconstruction (Sivabharathi et al, 2024). For instance, established statistical model of Quantitative Trait mapping to Haplotype (lcQTH) using low depth sequencing data for mapping population haplotype traceability, significantly reduce the cost of single sample genotyping, and the effective number significantly increased, marker density in the genome distribution, improve the efficiency of haplotype excavation, provides a new method for molecular design breeding. HaploMaker is a reference-based haplotype assembly algorithm that uses paired-end short reads and longer PacBio reads to phase long haplotypes with high accuracy (Fruzangohar et al., 2022). Another tool, Ranbow, is designed for polyploid genomes and integrates various types of small variants to reconstruct haplotypes efficiently (Moeinzadeh et al, 2020). Additionally, PERHAPS is a novel approach that directly calls haplotypes from short-read, paired-end sequencing data, demonstrating high concordance rates with traditional methods (Huang et al., 2020). These bioinformatics tools enhance the ability to analyze complex genomic data, facilitating the identification of superior haplotypes for breeding applications in rice.

4 Case Studies: Haplotype Analysis in Action

4.1 Examples of successful haplotype analysis in rice

Haplotype analysis has been instrumental in identifying superior genetic combinations that enhance grain yield and quality in rice. A comprehensive study analyzed 120 functionally characterized genes across the 3K rice genome panel, focusing on 87 genes related to grain yield and 33 genes associated with grain quality. This analysis revealed significant variations and identified superior haplotypes such as *SD1*-H8, *MOC1*-H9, and *IPA1*-H14, among others. These haplotypes were associated with improved traits like early flowering, medium duration flowering, and bold grains, demonstrating the potential for developing tailor-made rice varieties with enhanced genetic gains (Figure 1) (Abbai et al., 2019).

The candidate gene-based association analysis targeting 42-grain size- and genes in a panel of 180 rice accessions was conducted based on targeted capture sequencing. Among the 42 genes, 69 SNPs/Indels were detected to be associated with grain length (GL), rain width (GW), ratio of grain length-width (L/W) and thousand grain weight (TGW). Superior haplotypes and haplotype combinations for the seven genes were also identified. Furthermore, used the haplotype-specific marker panel for the superior haplotype combination *TGW*-HC4, it was analyzed in Minghui63 and some widely used restorer lines derived from Minghui63, as well as our breeding varieties and lines. The results showed that the *TGW*-HC4 is indeed an excellent haplotype combination of TGW, and it could be utilized in rice breeding. Haplotype-based breeding (HBB) provides selection targets for genomics-assisted breeding and contributes to the future development of rice varieties with high yield potential and high quality through this strategy (Figure 2) (Liu et al., 2023).

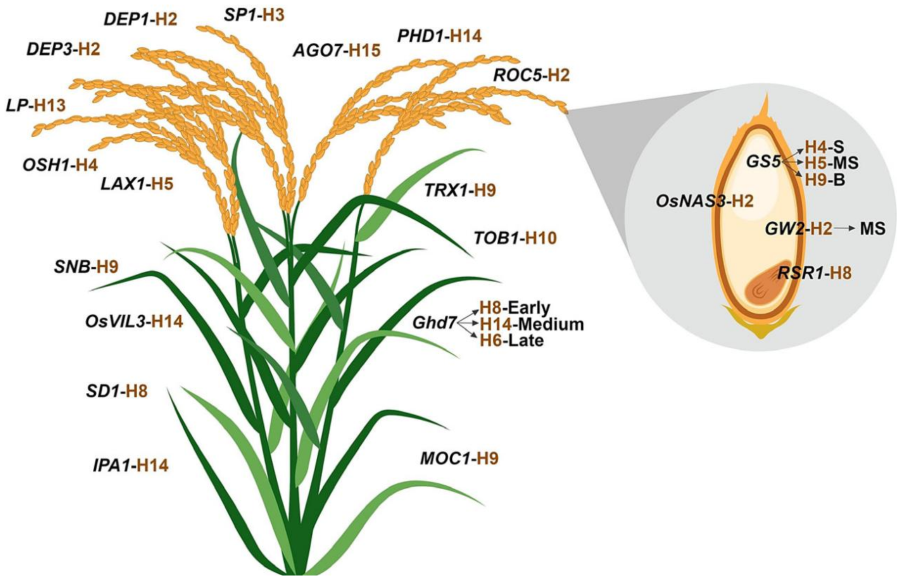
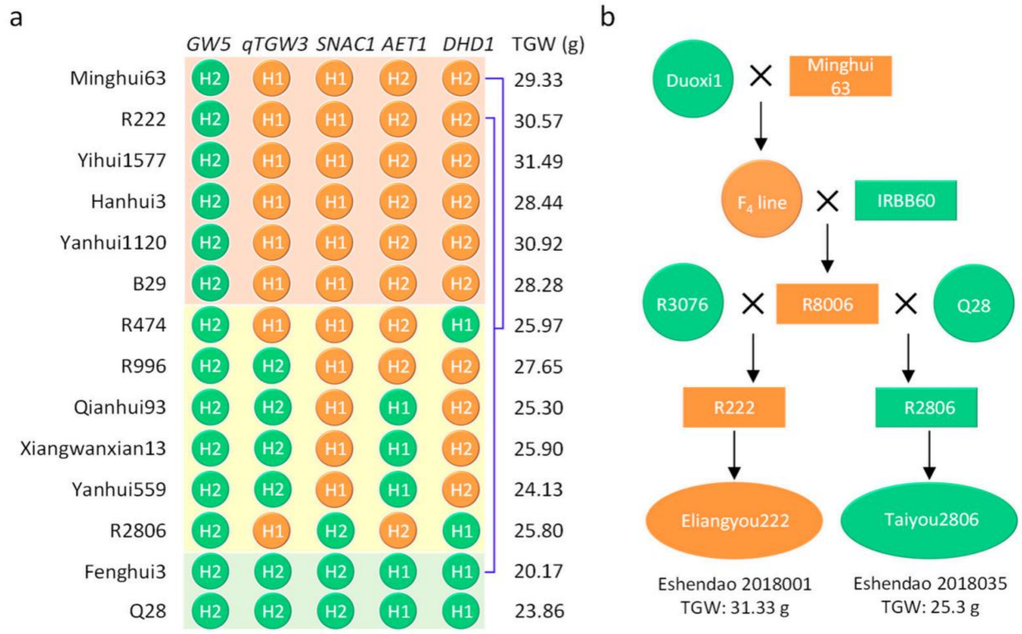


Figure 1 The tailored rice with superior haplotypes for grain yield and quality (Adopted from Abbai et al., 2019)

Image caption: The findings of this study could be employed to develop a designer rice genotype comprising superior haplotype combinations of the target genes such as *MOC1*-H9 & *IPA1*-H14 for higher tiller number, *Ghd7*-H8 & *TOB1*-H10 for early, *Ghd7*-H14 & *OsVIL3*-H14 for medium duration, *Ghd7*-H6, *SNB*-H9 & *TRX1*-H9 for late flowering, *DEP3*-H2, *DEP1*-H2 & *SP1*-H3 for long panicles, *SD1*-H8 for semi-dwarf nature, *LAX1*-H5, *OSH1*-H4 & *LP*-H13 resulting in increased panicle branching, *PHD1*-H14, *AGO7*-H15 & *ROC5*-H2 for high yield, along with *GS5*-H4 for slender, *GS5*-H5 & *GW2*-H2 for medium slender, *GS5*-H9 for bold grains, *RSR1*-H8 for intermediate amylose content and *OsNAS3*-H2 for increased Fe and Zn concentration in grains. B, Bold grain; MS, Medium slender type grain; S, Slender type grain (Adopted from Abbai et al., 2019)



	GW5	qTGW3	SNAC1	AET1	DHD1	TGW (g)
Minghui63	H2	H1	H1	H2	H2	29.33
R222	H2	H1	H1	H2	H2	30.57
Yihui1577	H2	H1	H1	H2	H2	31.49
Hanhui3	H2	H1	H1	H2	H2	28.44
Yanhui1120	H2	H1	H1	H2	H2	30.92
B29	H2	H1	H1	H2	H2	28.28
R474	H2	H1	H1	H2	H1	25.97
R996	H2	H2	H1	H2	H2	27.65
Qianhui93	H2	H2	H1	H1	H2	25.30
Xiangwanxian13	H2	H2	H1	H1	H2	25.90
Yanhui559	H2	H2	H1	H1	H2	24.13
R2806	H2	H1	H2	H2	H1	25.80
Fenghui3	H2	H2	H2	H1	H1	20.17
Q28	H2	H2	H2	H1	H1	23.86

Figure 2 Haplotypes and pedigree of the varieties and lines derived from Minghui63 (Adopted from Liu et al., 2023)

Image caption: (a) Haplotype and haplotype combinations of the five genes related to TGW in the varieties derived from Minghui63. The orange circle indicates the superior haplotype of the gene, the green circle indicates the non-superior haplotype of the gene, the pink square indicates the varieties containing HC4, the yellow square indicates the varieties containing the superior haplotype of two or three of the five genes, the light green square indicates the varieties not containing the superior haplotype for the five genes, and the blue line indicates the varieties derived from Minghui63. (b) Pedigree of the two registered varieties derived from Minghui63. Orange indicates the varieties or lines that contain the superior haplotype combination HC4, while green indicates the varieties or lines that do not contain HC4 (Adopted from Liu et al., 2023)

Referencing the characteristics of haplotype blocks within the wheat genome, the detailed characterization of haplotype blocks across the genome provides insights into their distribution and the genes they contain on different chromosomes. This comprehensive mapping helps to understand the genetic architecture influencing various phenotypic traits in rice, thereby enriching the understanding of haplotype analysis applications in rice genetic improvement (Figure 3) (Brinton et al., 2020).

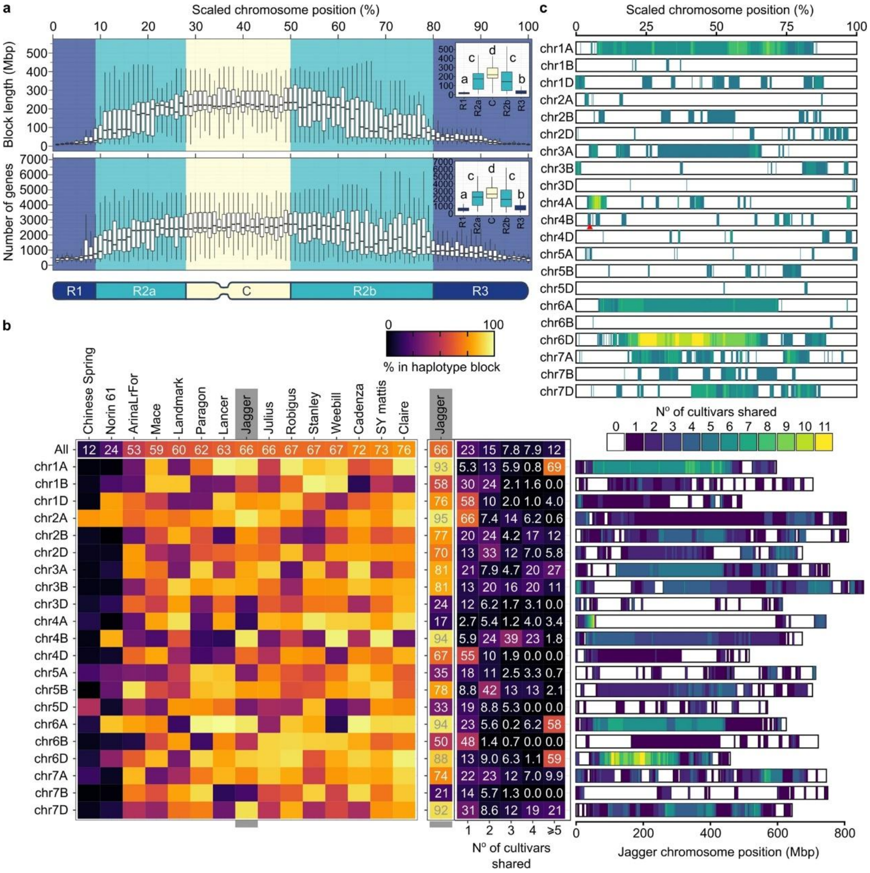


Figure 3 Genome-wide characterization of haplotype blocks (Adopted from Brinton et al., 2020)

Image caption: a: Length (upper) and gene number (middle) of haplotype blocks sampled at 500 Kbp intervals across all 21 chromosomes (positions scaled to % of maximum chromosome length). Boxplots show distributions of 1% bins. Insets show boxplots for chromosomal compartments (bottom) based on recombination (R1 + R3 > R2a + R2b > C). Pairwise Wilcox test with Benjamini–Hochberg multiple testing correction was used for statistical analysis. b: Leftmost heatmap shows % of the genome (top row) or chromosomes contained within haplotype blocks for all cultivars. Jagger values are highlighted (middle), with a breakdown of how many other cultivars share these Jagger haplotype blocks. Rightmost panel shows the physical positions of Jagger haplotype blocks. c: Summary of highly conserved haplotype blocks across all cultivars (shared with ≥ 5 cultivars; positions scaled to % of maximum chromosome length per cultivar). Values based on haplotype blocks called using 5-Mbp bins. *RHT-B1* is indicated by a red arrowhead (Adopted from Brinton et al., 2020)

4.2 Comparative yield outcomes between haplotype-analyzed varieties and conventional breeds

The yield outcomes of haplotype-analyzed rice varieties have shown significant improvements compared to conventional breeds. For instance, the superior haplotypes identified in the 3K rice genome panel, such as *DEP1*-H2 and *GS5*-H4, have been linked to higher grain yield and better grain quality. These haplotypes were found to have higher frequencies in the population, indicating their advantageous nature. The study reported that varieties with these superior haplotypes exhibited better performance in terms of yield and quality traits, suggesting that HBB can lead to the development of high-yielding rice varieties that outperform traditional breeds. Figure 4 provides a detailed visualization of the haplotype distribution and their phenotypic effects, particularly highlighting the significant impact of haplotypes like H8 on early flowering and H6 on delayed flowering, further supporting the phenotypic outcomes discussed here (Abbai et al., 2019).

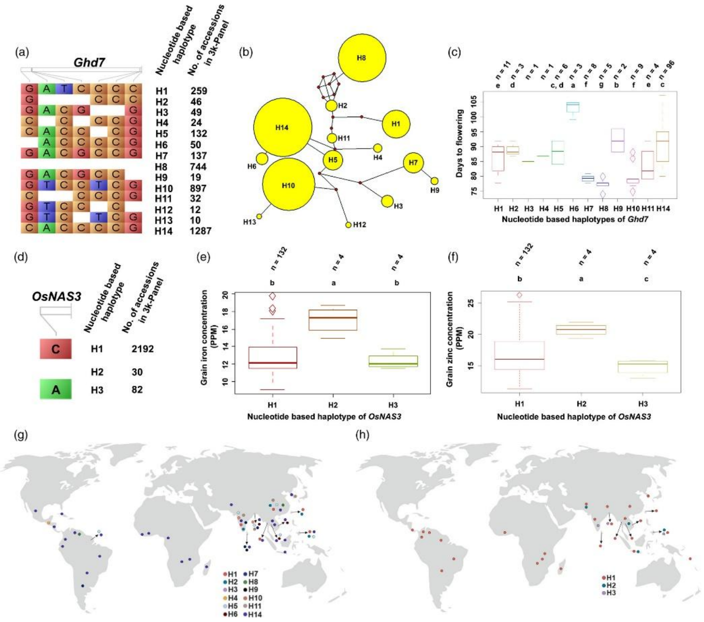


Figure 4 Haplotype analysis of *Ghd7* and *OsNAS3* across the 3K RG panel (Adopted from Abbai et al., 2019)
Image caption: (a) *Ghd7*, a key gene associated with heading date has about 14 haplotypes in the 3K RG panel with wide phenotypic variations. (b) *Ghd7*-H8 was the most diverse one based on SNP and (c) interestingly was the earliest to flower, while *Ghd7*-H6 took greater than 100 days to flower. (d) *OsNAS3* that influences grain Fe and Zn concentration has three haplotypes in the 3K RG panel with significant phenotypic variations in the subset. *OsNAS3*-H2 had the highest grain (e) Fe and (f) Zn profile. The geographical distribution of various haplotypes of (g) *Ghd7* and (h) *OsNAS3* (Adopted from Abbai et al., 2019)

Furthermore, the researchers discovered the elite haplotype *EPN4*-1 of the effective panicle number per plant (*EPN4*), the elite haplotype *GNP1*-6 of the grain number per panicle (*GNP*). Pyramiding *of EPN4*-1 and *GNP1*-6 could be a preferred approach to obtain high yield without affecting other yield-related traits, which has application value in high-yield rice breeding (Figure 5) (Wang et al., 2023).

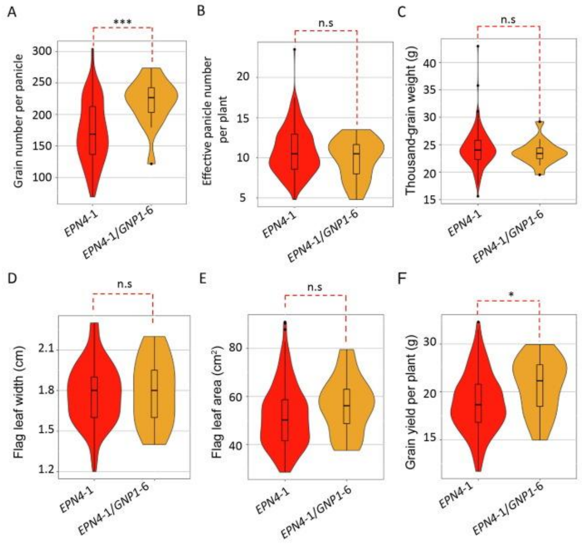


Figure 5 Comparison of yield and yield-related traits between accessions carrying an *EPN4*-1/*GNP1*-6 combination and accessions carrying the *EPN4*-1 haplotype (Adopted from Wang et al., 2023)

4.3 Lessons learned from practical applications

The practical applications of haplotype analysis in rice breeding have provided several valuable lessons. Firstly, the identification of superior haplotypes requires a comprehensive understanding of the genetic basis of key traits. The study on the 3K rice genome panel highlighted the importance of integrating meta-expression analysis and co-expression networks to gain insights at the systems level. Secondly, the success of HBB depends on the accurate phenotyping of genetic variations. The identification of haplotypes like *SD1*-H8 and *MOC1*-H9 underscores the need for precise phenotypic data to correlate genetic variations with desirable traits. Lastly, the deployment of superior haplotypes in breeding programs can significantly enhance the genetic gains in rice, paving the way for the development of next-generation tailor-made rice varieties that meet future food and nutritional demands (Abbai et al., 2019).

5 Breeding High-Yielding Rice Varieties

5.1 Strategies for incorporating haplotype analysis into breeding programs

Haplotype analysis has emerged as a powerful tool in modern rice breeding programs, enabling the identification and utilization of superior genetic combinations to enhance yield and other desirable traits. By leveraging genome-wide association studies (GWAS) and marker-trait associations (MTAs), researchers can pinpoint specific haplotypes associated with high-yield traits. For instance, the identification of superior haplotypes for grain size and weight-related genes has been shown to significantly impact yield (Liu et al., 2023). Additionally, the integration of HBB with traditional breeding methods, such as marker-assisted selection (MAS) and genomic selection (GS), can accelerate the development of high-yielding rice varieties (Thudi et al., 2020; Singh et al., 2022b; Sivabharathi et al., 2024).

5.2 Development of tailor-made rice varieties with desired traits

The development of tailor-made rice varieties involves the strategic selection and combination of superior haplotypes to meet specific breeding goals. For example, the identification of haplotypes associated with low glycemic index (GI) and preferred grain quality traits has led to the creation of rice varieties that cater to health-conscious consumers (Selvaraj et al., 2021). Similarly, haplotypes linked to drought tolerance, nitrogen uptake efficiency and salt resistance have been utilized to develop resilient rice varieties capable of thriving under adverse environmental conditions (Sinha et al., 2020; Elangovan et al., 2023; Wu et al., 2024). The use of haplo-pheno analysis, which correlates haplotypes with phenotypic traits, further aids in the precise selection of desirable genetic combinations (Chen et al., 2023).

5.3 Challenges and solutions in molecular breeding

Despite the promising potential of HBB, several challenges must be addressed to fully realize its benefits. One major challenge is the complexity of genetic interactions and the need for comprehensive genomic data to accurately identify superior haplotypes (Li et al., 2022). Additionally, the integration of haplotype analysis into existing breeding programs requires significant investment in technology and expertise (Wang et al., 2021; Sivabharathi et al, 2024). To overcome these challenges, researchers are employing advanced genomic tools and bioinformatics approaches to streamline the identification and utilization of superior haplotypes (Thudi et al., 2020). Collaborative efforts and the sharing of genomic resources among research institutions can also facilitate the widespread adoption of haplotype-based breeding strategies (Faysal et al., 2022).

6 Impact of Tailor-Made Varieties on Rice Production

6.1 Enhancements in yield and plant health

The development and adoption of tailor-made rice varieties, particularly those designed through haplotype analysis, have shown significant improvements in yield and plant health. For instance, the adoption of improved rice varieties (IRVs) in Nigeria resulted in an increase of 452 kg of rice grains per hectare, highlighting the substantial yield benefits of these varieties (Bello et al., 2020). Similarly, the characterization of haplotypes associated with the *Gn1a* gene has been linked to high grain number formation, which directly contributes to increased grain yield in rice plants (Gouda et al., 2020). Additionally, the performance of doubled haploid elite rice germplasm in Zimbabwe demonstrated a yield advantage of 66% over local checks, further emphasizing the potential of these advanced breeding techniques to enhance rice productivity (Chitanda et al., 2022). The etile haplotype *OsGATA8*-H of *OsGATA8* was discovered, and excellent nitrogen efficiency breeding materials were created using gene editing and backcrossing breeding techniques, which showed that *OsGATA8*-H was an elite haplotype with high NUE (Wu et al., 2024). Under drought stress, the dominant haplotype *LEA12* $^{\mathrm{OR}}$ of the rice salt tolerance gene *LEA12*, regulates the expression of 9-is-epoxycarotenoid dioxygenase *OsNCED3*, a key enzyme in the ABA biosynthesis pathway, thus enhancing salt tolerance and providing a new way to improve salt tolerance and yield in rice.

6.2 Environmental and economic benefits of high-yielding varieties

High-yielding rice varieties not only improve productivity but also offer significant environmental and economic benefits. The adoption of these varieties can lead to more efficient use of land and resources, reducing the need for expansion into ecologically sensitive areas. For example, the development of low glycemic index rice with preferred grain quality can cater to specific dietary needs while maintaining high yield, thus addressing both health and environmental concerns (Selvaraj et al., 2021). Economically, the increased productivity from high-yielding varieties translates to higher incomes for farmers. In Yunnan, China, the adoption of improved rice varieties increased both rice income and total household income, demonstrating the economic viability of these varieties (Wang et al., 2020). Furthermore, the genetic diversity introduced through the International Rice Genebank has been shown to significantly boost rice productivity, thereby contributing to economic stability in farming communities (Villanueva et al., 2020).

6.3 Acceptance and adoption by the farming community

The acceptance and adoption of tailor-made rice varieties by the farming community are crucial for their success. Factors such as educational attainment, experience in rice farming, and access to extension services and credit have been identified as significant determinants of IRV adoption in Nigeria (Bello et al., 2020). Government programs and extension services also play a vital role in promoting the adoption of these varieties, as evidenced by the increased adoption rates in Yunnan due to government initiatives (Wang et al., 2020). Additionally, the development of varieties that are stable across different environmental conditions, such as those evaluated for heat tolerance, ensures that farmers can rely on these varieties under varying climatic conditions, thereby increasing their acceptance (Senguttuvel et al., 2021). The positive correlation between yield and traits such as grain width and number of grains per panicle further supports the adoption of these varieties, as they meet the farmers' expectations for high productivity (Ayyenar et al., 2022; Singh et al., 2022a).

7 Integration and Scalability of HBB

7.1 Implementing haplotype analysis on a large scale

Implementing haplotype analysis on a large scale involves leveraging advanced genomic technologies to identify and utilize genetic variations effectively. The process begins with the collection of extensive genomic data from diverse rice cultivars. For instance, Brinton et al. (2020) demonstrated the utility of genome assemblies from multiple wheat cultivars to identify haplotypes and their potential for trait improvement. Similarly, the identification of superior haplotypes for drought tolerance in pigeonpea highlights the importance of whole-genome re-sequencing data in HBB programs (Sinha et al., 2020). By integrating high-throughput sequencing and advanced computational tools, large-scale haplotype analysis can be systematically applied to rice breeding programs to enhance precision and efficiency.

7.2 Overcoming technical and logistical challenges

The technical and logistical challenges in HBB include accurate haplotype phasing, data management, and the integration of diverse datasets. Consensus strategies for haplotype phasing, such as combining multiple independent phasing estimates, have been shown to improve accuracy and reduce errors (Bkhetan et al., 2020). Additionally, the development of specialized software, like PolyOrigin for tetraploid species, facilitates haplotype reconstruction and enhances QTL detection power (Zheng et al., 2020). Addressing these challenges requires robust computational infrastructure, standardized protocols, and collaboration among researchers to ensure the seamless integration of haplotype data into breeding programs.

7.3 Strategies for enhancing farmer and industry adoption

To enhance the adoption of HBB by farmers and the industry, it is crucial to demonstrate the tangible benefits of this approach. Data-driven decentralized breeding, which combines genomics, farmers' knowledge, and environmental analysis, has shown promise in improving local adaptation and productive performance in challenging environments (Sousa et al., 2021). Additionally, the identification of superior haplotypes for key traits, such as grain size and weight in rice, provides clear targets for breeding programs and can be communicated effectively to stakeholders (Liu et al., 2023). Engaging farmers through participatory breeding programs and providing training on the use of genomic tools can further facilitate the adoption of HBB practices.

8 Future Directions in Rice Genomics

8.1 Emerging trends in genomic research for crop improvement

The field of crop genomics is rapidly evolving, with significant advancements that promise to revolutionize rice breeding and production. Over the past two decades, the sequencing of numerous crop genomes, including rice, has laid a robust foundation for genomic research. The integration of genome-scale information across various biological scales is expected to enhance our understanding of crop biological processes, thereby facilitating the translation of laboratory findings to field applications (Purugganan and Jackson, 2021). Using18K-rice “hybrid” population, the method of rapid excavation of quantitative trait genes in RiceG2G rice was developed, which greatly improved the efficiency of genetic analysis of agronomic traits, opened up the key step from rapid screening of genetic loci to candidate genes, and systematically evaluated the additive and epistatic effects of rice QTL genes. This study provides comprehensive gene interaction information for quantitative traits for rice genetic research and will provide theoretical support for rice molecular design breeding (Wei et al., 2024). Genomics-assisted breeding (GAB) has emerged as a pivotal approach, leveraging modern genome resources to exploit allelic variation for germplasm enhancement and cultivar development. Future iterations of GAB (GAB 2.0) will focus on the targeted manipulation of allelic variation to create novel diversity, which is crucial for developing climate-smart and nutritionally superior rice cultivars (Varshney et al., 2021b). Additionally, the development of platinum standard reference genomes for the *Oryza* genus will enable the efficient identification and utilization of adaptive traits from wild rice relatives, further enhancing crop improvement efforts (Mussurova et al., 2020).

8.2 Potential for integrating AI and machine learning in haplotype analysis

The integration of artificial intelligence (AI) and machine learning (ML) in haplotype analysis holds immense potential for accelerating rice genomic research. Recent advancements in deep learning have demonstrated its effectiveness in modeling the flow of information from genomic DNA sequences to molecular phenotypes, as well as in identifying functional variants in natural populations (Wang et al., 2020). Machine learning approaches are also being utilized for genomic selection (GS), significantly reducing the need for resource-intensive phenotyping by predicting agronomically relevant traits from genotypic data (Tong and Nikoloski, 2020). These technologies can enhance the precision of genomic prediction models, incorporating environmental effects and genotype-by-environment interactions to improve the accuracy of trait predictions (Toda et al., 2020). The integration of high-throughput molecular phenotypic data with biological networks through AI and ML will further streamline the selection of elite genotypes, thereby shortening the breeding cycle and improving crop yields (Tong and Nikoloski, 2020).

8.3 Prospects for global collaboration in rice research

Global collaboration is essential for advancing rice genomics and addressing the challenges of food security. Collaborative efforts have already led to significant progress in rice functional genomics, with contributions from experts in major rice-producing countries (Xiong et al., 2020). The establishment of international research centers and consortia, such as the International Maize and Wheat Improvement Center (CIMMYT) and the National Key Laboratory of Crop Genetic Improvement, has facilitated the sharing of knowledge and resources (Xu et al., 2021). Future collaborations should focus on the development and dissemination of genomic resources, such as the platinum standard reference genomes, to ensure that researchers worldwide can access and utilize these tools for crop improvement (Mussurova et al., 2020). Additionally, fostering partnerships between academic institutions, government agencies, and private sector organizations will be crucial for translating genomic research into practical applications that benefit farmers and consumers globally.

9 Concluding Remarks

Haplotype analysis has emerged as a pivotal tool in modern rice breeding, enabling the identification of superior genetic variants associated with desirable traits. By leveraging GWAS and haplotype-pheno analysis, researchers can pinpoint specific haplotypes that contribute to traits such as low glycemic index, drought tolerance, and enhanced grain size and weight. This approach allows for the precise selection of genetic combinations that can be utilized in breeding programs to develop high-yielding, resilient rice varieties tailored to specific environmental conditions and consumer preferences.

The benefits include increased yield and stability, even under adverse climatic conditions, which is crucial for ensuring food security in the face of climate change. Additionally, enhanced nutritional quality has been achieved, with the development of rice varieties that have higher zinc content and lower glycemic indexes, addressing malnutrition and health concerns. Furthermore, varieties developed through haplotype analysis tend to be more resource-efficient, requiring less water and labor, especially beneficial in dry direct-seeded rice (DDSR) systems. However, there are challenges, including the complex genetic architecture of many desirable traits such as grain yield and drought tolerance, which complicates the identification and utilization of superior haplotypes. Environmental interactions also pose a challenge as the expression of genetic traits can be influenced by environmental factors, making it difficult to achieve consistent results across different growing conditions. Additionally, the advanced genomic tools and techniques required for haplotype analysis can be resource-intensive and may limit their accessibility in resource-poor settings.

The future of tailor-made rice in sustainable agriculture appears promising, with several advancements on the horizon. The continued integration of advanced genomics, including next-generation sequencing (NGS) and machine learning in haplotype analysis, will enhance the precision and efficiency of breeding programs. This integration enables the development of rice varieties that are not only high-yielding but also resilient to biotic and abiotic stresses. As climate change increasingly impacts agriculture, developing rice varieties capable of withstanding extreme weather conditions becomes crucial. HBB will play a key role in identifying and combining traits that confer resilience to drought, heat, and flooding. Furthermore, future breeding efforts will likely focus on biofortification, aiming to enhance the nutritional content of rice to combat malnutrition by increasing the content of essential micronutrients such as zinc and iron. Additionally, the adoption of rice varieties tailored for specific growing conditions, such as DDSR, will promote sustainable agricultural practices by reducing the need for water and labor inputs, thus conserving resources and reducing the environmental footprint of rice cultivation.

Acknowledgments

We extend our sincere thanks to two anonymous peer reviewers for their feedback on the manuscript, whose constructive suggestions have greatly contributed to the improvement of our manuscript.

Funding

This work was supported by the grants from the Key and Major Science and Technology Projects of Yunnan (grant nos. 202202AE09002102) and the Major Science and Technology Projects in Yunnan Province (202402AE090026).

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

Abbai R., Singh V., Nachimuthu V., Sinha P., Selvaraj R., Vipparla A., Singh A., Singh U., Varshney R., and Kumar A., 2019, Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains, Plant Biotechnology Journal, 17: 1612-1622.
https://doi.org/10.1111/pbi.13087
PMid:30701663 PMCid:PMC6662101

Ayyenar B., Premnath A., Sudhakar D., and Muthurajan R., 2022, Allelic diversity of *OsGW5.1* regulating grain width in rice, Madras Agricultural Journal, 109: 92-98.

Bello L., Baiyegunhi L., and Danso-Abbeam G., 2020, Productivity impact of improved rice varieties' adoption: case of smallholder rice farmers in Nigeria, Economics of Innovation and New Technology, 30: 750-766.
https://doi.org/10.1080/10438599.2020.1776488

Bkhetan Z., Chana G., Ramamohanarao K., Verspoor K., and Goudey B., 2020, Evaluation of consensus strategies for haplotype phasing, Briefings in Bioinformatics, 22(4): bbaa280.
https://doi.org/10.1093/bib/bbaa280
PMid:33236761

Brinton J., Ramírez-González R., Simmonds J., Wingen L., Orford S., Griffiths S., Haberer G., Spannagl M., Walkowiak S., Pozniak C., and Uauy C., 2020, A haplotype-led approach to increase the precision of wheat breeding, Communications Biology, 3: 712.
https://doi.org/10.1038/s42003-020-01413-2
PMid:33239669 PMCid:PMC7689427

Cao L., Li T., Geng S., Zhang Y., Pan Y., Zhang X., Wang F., and Hao C., 2023, *TaSPL14-7A* is a conserved regulator controlling plant architecture and yield traits in common wheat (*Triticum aestivum* L.), Front. Plant Sci., 14: 1178624.
https://doi.org/10.3389/fpls.2023.1178624
PMid:37089636 PMCid:PMC10113487

Chen Z., Bu Q., Liu G., Wang M., Wang H., Liu H., Li X., Li H., Fang J., Liang Y., Teng Z., Kang S., Yu H., Cheng Z., Xue Y., Liang C., Tang J., Li J., and Chu C., 2023, Genomic decoding of breeding history to guide breeding-by-design in rice, National Science Review, 10(5): nwad029.
https://doi.org/10.1093/nsr/nwad029
PMid:37056426 PMCid:PMC10089590

Chitanda L., Mativavarira M., Manjeru P., Kang K., Nzuma J., Kamunhukamwe T., Hove T., and Madzingaidzo L., 2022, Performance of doubled haploid elite rice (*Oryza sativa* L.) germplasm for grain yield and associated traits in Harare, Zimbabwe, Advances in Agriculture, 1: 7393896.

Elangovan D., Pandey R., Sharma S., Balamurugan B., Anand N., Das A., Kumar T., Ellur R., Kalia S., and Rane J., 2023, Haplo-pheno association for *OsNRT1.1* paralog in rice reveals superior haplogroup with high nitrate uptake efficiency, bioRxiv, 12: 1-46.
https://doi.org/10.1101/2023.12.07.570681

Faysal A., Ali L., Azam M., Sarker U., Ercişli S., Golokhvast K., and Marc R., 2022, Genetic variability, character association, and path coefficient analysis in transplant Aman rice genotypes, Plants, 11(21): 2952.
https://doi.org/10.3390/plants11212952
PMid:36365406 PMCid:PMC9655179

Fruzangohar M., Timmins W., Kravchuk O., and Taylor J., 2022, HaploMaker: an improved algorithm for rapid haplotype assembly of genomic sequences, GigaScience, 11: giac038.
https://doi.org/10.1093/gigascience/giac038
PMid:35579550 PMCid:PMC9112781

Garg S., 2021, Computational methods for chromosome-scale haplotype reconstruction, Genome Biology, 22: 101.
https://doi.org/10.1186/s13059-021-02328-9
PMid:33845884 PMCid:PMC8040228

Gouda G., Gupta M., Donde R., Kumar J., Parida M., Mohapatra T., Dash S., Pradhan S., and Behera L., 2020, Characterization of haplotypes and single nucleotide polymorphisms associated with *Gn1a* for high grain number formation in rice plant, Genomics, 112(3): 2647-2657.
https://doi.org/10.1016/j.ygeno.2020.02.016
PMid:32087244

Huang J., Pallotti S., Zhou Q., Kleber M., Xin X., King D., and Napolioni V., 2020, PERHAPS: paired-end short reads-based HAPlotyping from next-generation sequencing data, Briefings in Bioinformatics, 22(4): bbaa320.

Kolesnikov A., Cook D., Nattestad M., Brambrink L., McNulty B., Gorzynski J., Goenka S., Ashley E., Jain M., Miga K., Paten B., Chang P., Carroll A., and Shafin K., 2024, Local read haplotagging enables accurate long-read small variant calling, Nature Communications, 15(1): 5907.
https://doi.org/10.1038/s41467-024-50079-5
PMid:39003259 PMCid:PMC11246426

Li F., Xie J., Zhu X., Wang X., Zhao Y., Ma X., Zhang Z., Rashid M., Zhang Z., Zhi L., Zhang S., Li J., Li Z., and Zhang H., 2018, Genetic basis underlying correlations among growth duration and yield traits revealed by GWAS in rice (*Oryza sativa* L.), Frontiers in Plant Science, 9: 650.
https://doi.org/10.3389/fpls.2018.00650
PMid:29872443 PMCid:PMC5972282

Li L., Wu X., Chen J., Wang S., Wan Y., Ji H., Wen Y., and Zhang J., 2022, Genetic dissection of epistatic interactions contributing yield-related agronomic traits in rice using the compressed mixed model, Plants, 11(19): 2504.

Li S., Zhang Y., Fan C., Chen Y., Deng C., and Hu Z., 2018, Advances in haplotype analysis technique, Chinese Journal of Biotechnology, 34(6): 852-861.

Liu G., Qiu D., Lu Y., Wu Y., Han X., Jiao Y., Wang T., Yang J., You A., Chen J., and Zhang Z., 2023, Identification of superior haplotypes and haplotype combinations for grain size- and weight-related genes for breeding applications in rice (*Oryza sativa* L.), Genes, 14(12): 2201.
https://doi.org/10.3390/genes14122201
PMid:38137023 PMCid:PMC10742856

Maestri S., Maturo M., Cosentino E., Marcolungo L., Iadarola B., Fortunati E., Rossato M., and Delledonne M., 2020, A long-read sequencing approach for direct haplotype phasing in clinical settings, International Journal of Molecular Sciences, 21(23): 9177.
https://doi.org/10.3390/ijms21239177
PMid:33271988 PMCid:PMC7731377

Meier J., Salazar P., Kučka M., Davies R., Dréau A., Aldás I., Power O., Nadeau N., Bridle J., Rolian C., Barton N., McMillan W., Jiggins C., and Chan Y., 2020, Haplotype tagging reveals parallel formation of hybrid races in two butterfly species, Proceedings of the National Academy of Sciences of the United States of America, 118(25): e2015005118.
https://doi.org/10.1073/pnas.2015005118
PMid:34155138 PMCid:PMC8237668

Moeinzadeh M., Yang J., Muzychenko E., Gallone G., Heller D., Reinert K., Haas S., and Vingron M., 2020, Ranbow: a fast and accurate method for polyploid haplotype reconstruction, PLoS Computational Biology, 16(5): e1007843.
https://doi.org/10.1371/journal.pcbi.1007843
PMid:32469863 PMCid:PMC7310859

Mussurova S., Al-Bader N., Zuccolo A., and Wing R., 2020, Potential of platinum standard reference genomes to exploit natural variation in the wild relatives of rice, Frontiers in Plant Science, 11: 579980.
https://doi.org/10.3389/fpls.2020.579980
PMid:33072154 PMCid:PMC7539145

Nath S., and Kole P., 2021, Genetic variability and yield analysis in rice, Electronic Journal of Plant Breeding, 12(1): 253-258.
https://doi.org/10.37992/2021.1201.039

Purugganan M., and Jackson S., 2021, Advancing crop genomics from lab to field, Nature Genetics, 53: 595-601.
https://doi.org/10.1038/s41588-021-00866-3
PMid:33958781

Selvaraj R., Singh A., Singh V., Abbai R., Habde S., Singh U., and Kumar A., 2021, Superior haplotypes towards development of low glycemic index rice with preferred grain and cooking quality, Scientific Reports, 11: 10082.
https://doi.org/10.1038/s41598-021-87964-8
PMid:33980871 PMCid:PMC8115083

Senguttuvel P., Sravanraju N., Jaldhani V., Divya B., Beulah P., Nagaraju P., Manasa Y., Prasad A., Brajendra P., Gireesh C., Anantha M., Suneetha K., Sundaram R., Madhav M., Tuti M., Subbarao L., Neeraja C., Bhadana V., Rao P., Voleti S., and Subrahmanyam D., 2021, Evaluation of genotype by environment interaction and adaptability in lowland irrigated rice hybrids for grain yield under high temperature, Scientific Reports, 11: 15825.
https://doi.org/10.1038/s41598-021-95264-4
PMid:34349182 PMCid:PMC8338964

Shen L., Wang C., Fu Y., Wang J., Liu Q., Zhang X., Yan C., Qian Q., and Wang K., 2018, QTL editing confers opposing yield performance in different rice varieties, Journal of Integrative Plant Biology, 60: 89-93.
https://doi.org/10.1111/jipb.12501
PMid:27628577

Singh A., Verma O., Singh A., and Choudhary A., 2022a, Association analysis for yield and its attributing components in rice (*Oryza sativa* L) under two environments, Journal of Agriculture Research and Technology, 1: 3-10.

Singh G., Kaur N., Khanna R., Kaur R., Gudi S., Kaur R., Sidhu N., Vikal Y., and Mangat G., 2022b, 2Gs and plant architecture: breaking grain yield ceiling through breeding approaches for next wave of revolution in rice (*Oryza sativa* L.), Critical Reviews in Biotechnology, 44(1): 139-162.
https://doi.org/10.1080/07388551.2022.2112648
PMid:36176065

Sinha P., Singh V., Saxena R., Khan A., Abbai R., Chitikineni A., Desai A., Molla J., Upadhyaya H., Kumar A., and Varshney R., 2020, Superior haplotypes for haplotype-based breeding for drought tolerance in pigeonpea (*Cajanus cajan* L.), Plant Biotechnology Journal, 18: 2482-2490.

Sivabharathi R., Rajagopalan V., Suresh R., Sudha M., Karthikeyan G., Jayakanthan M., and Raveendran M., 2024, Haplotype-based breeding: a new insight in crop improvement, Plant Science, 346: 112129.
https://doi.org/10.1016/j.plantsci.2024.112129
PMid:38763472

Sousa K., Etten J., Poland J., Fadda C., Jannink J., Kidane Y., Lakew B., Mengistu D., Pè M., Solberg S., and Dell'Acqua M., 2021, Data-driven decentralized breeding increases prediction accuracy in a challenging crop production environment, Communications Biology, 4: 944.

Su J., Xu K., Li Z., Hu Y., Hu Z., Zheng X., Song S., Tang Z., and Li L., 2021, Genome-wide association study and Mendelian randomization analysis provide insights for improving rice yield potential, Scientific Reports, 11: 6894.
https://doi.org/10.1038/s41598-021-86389-7
PMid:33767346 PMCid:PMC7994632

Thudi M., Palakurthi R., Schnable J., Chitikineni A., Dreisigacker S., Mace E., Srivastava R., Satyavathi C., Odeny D., Tiwari V., Lam H., Hong Y., Singh V., Li G., Xu Y., Chen X., Kaila S., Nguyen H., Sivasankar S., Jackson S., Close T., Shubo W., and Varshney R., 2020, Genomic resources in plant breeding for sustainable agriculture, Journal of Plant Physiology, 257: 153351.
https://doi.org/10.1016/j.jplph.2020.153351
PMid:33412425 PMCid:PMC7903322

Toda Y., Wakatsuki H., Aoike T., Kajiya-Kanegae H., Yamasaki M., Yoshioka T., Ebana K., Hayashi T., Nakagawa H., Hasegawa T., and Iwata H., 2020, Predicting biomass of rice with intermediate traits: modeling method combining crop growth models and genomic prediction models, PLoS One, 15(6): e0233951.
https://doi.org/10.1371/journal.pone.0233951
PMid:32559220 PMCid:PMC7304626

Tong H., and Nikoloski Z., 2020, Machine learning approaches for crop improvement: leveraging phenotypic and genotypic big data, Journal of Plant Physiology, 257: 153354.
https://doi.org/10.1016/j.jplph.2020.153354
PMid:33385619

Varshney R., Barmukh R., Roorkiwal M., Qi Y., Kholová J., Tuberosa R., Reynolds M., Tardieu F., and Siddique K., 2021a, Breeding custom-designed crops for improved drought adaptation, Advanced Genetics, 2(3): e202100017.
https://doi.org/10.1002/ggn2.202100017
PMid:36620433 PMCid:PMC9744523

Varshney R., Bohra A., Yu J., Graner A., Zhang Q., and Sorrells M., 2021b, Designing future crops: genomics-assisted breeding comes of age, Trends in Plant Science, 26(6): 631-649.
https://doi.org/10.1016/j.tplants.2021.03.010
PMid:33893045

Villanueva D., Smale M., Jamora N., Capilit G., and Hamilton R., 2020, The contribution of the International Rice Genebank to varietal improvement and crop productivity in Eastern India, Food Security, 12: 929-943.

Wang Y., Wang X., Zhai L., Zafar S., Shen C., Zhu S., Chen K., Wang Y., and Xu J., 2023, A novel Effective Panicle Number per Plant 4 haplotype enhances grain yield by coordinating panicle number and grain number in rice, The Crop Journal, 12(1): 202-212.
https://doi.org/10.1016/j.cj.2023.11.003

Wang H., Cimen E., Singh N., and Buckler E., 2020, Deep learning for plant genomics and crop improvement, Current Opinion in Plant Biology, 54: 34-41.
https://doi.org/10.1016/j.pbi.2019.12.010
PMid:31986354

Wang H., Pandey S., and Feng L., 2020, Econometric analyses of adoption and household-level impacts of improved rice varieties in the uplands of Yunnan, China, Sustainability, 12(17): 6873.
https://doi.org/10.3390/su12176873

Wang X., Jing Z., He C., Liu Q., Jia H., Qi J., and Zhang H., 2021, Breeding rice varieties provides an effective approach to improve productivity and yield sensitivity to climate resources, European Journal of Agronomy, 124: 126239.
https://doi.org/10.1016/j.eja.2021.126239

Wei X., Chen M., Zhang Q., Gong J., Liu J., Yong K., Wang Q., Fan J., Chen S., Hua H., Luo Z., Zhao X., Wang X., Li Wei., Cong J., Yu X., Wang Z., Huang R., Chen J., Zhou X., Qiu J., Xu P., Murray J., Wang H., Xu Y., Xu C., Xu G., Yang J., Han B., and Huang X., 2024, Genomic investigation of 18 421 lines reveals the genetic architecture of rice, Science, 385: 6704.
https://doi.org/10.1126/science.adm8762
PMid:38963845

Wu W., Dong X., Chen G., Lin Z., Chi W., Tang W., Yu J., Wang S., Jiang X., Liu X., Wu Y., Wang C., Cheng X., Zhang W., Xuan W., Terzaghi W., Ronald P., Wang H., Wang C., and Wan J., 2024, The elite haplotype *OsGATA8*-H coordinates nitrogen uptake and productive tiller formation in rice, Nature Genetics, 56: 1516-1526.
https://doi.org/10.1038/s41588-024-01795-7
PMid:38872029 PMCid:PMC11250373

Xiong L., Uga Y., and Li Y., 2020, Rice functional genomics: theories and practical applications, Molecular Breeding, 40: 72.
https://doi.org/10.1007/s11032-020-01150-8

Xu J., Xing Y., Xu Y., and Wan J., 2021, Breeding by design for future rice: genes and genome technologies, Crop Journal, 9: 491-496.
https://doi.org/10.1016/j.cj.2021.05.001

Zheng C., Amadeu R., Muñoz P., and Endelman J., 2020, Haplotype reconstruction in connected tetraploid F_1 populations, Genetics, 219(2): iyab106.
https://doi.org/10.1093/genetics/iyab106
PMid:34849879 PMCid:PMC8633103