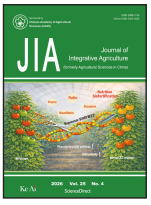
Journal of Integrative Agriculture 2026, 25(4): 1359–1372

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Multi-trait genome-wide association studies reveal novel pleiotropic loci associated with yield and yield-related traits in rice

Chunhai Liu[1*], Chao Wu[1*], Zheming Yuan[1], Bingchuan Tian[2], Peiyi Yu[2], Deze Xu[3], Xingfei Zheng[3], Lanzhi Li[1#]

[1] *Hunan Engineering & Technology Research Center for Agricultural Big Data Analysis & Decision-making/College of Plant Protection, Hunan Agricultural University, Changsha 410128, China*

[2] *Huazhi Biotechnology Co. Ltd., Changsha 410125, China*

[3] *Hubei Key Laboratory of Food Crop Germplasm and Genetic Improvement/Food Crop Institute, Hubei Academy of Agricultural Sciences, Wuhan 430064, China*

Highlights

- Multi-trait genome-wide association studies (GWAS) identified 44 pleiotropic quantitative trait loci (QTLs), including 29 novel loci undetected by single-trait methods.
- Key pleiotropic QTLs with common effects facilitate the simultaneous improvement of yield and yield-related traits, leading to the identification of 13 candidate genes.
- Pyramiding superior alleles of *GS3*, *GL3.1*, *OsCIPK17*, and *GNP12* yields optimal combinations for molecular breeding.

Abstract

Rice yield is a complex trait affected by many related traits. Traditional single-trait genome-wide association studies (GWAS) have limitations when studying complex traits, as they cannot account for the genetic relationships among multiple traits. Multi-trait GWAS can consider the relationships among multiple traits and identify pleiotropic loci, so it is more suitable for complex traits such as rice yield than single-trait GWAS. In this study, we conducted a multi-trait GWAS on 11 two-trait combinations of yield and yield-related traits with 575 hybrid rice varieties across two environments. All these yield-related traits showed significant genetic correlations with yield (YD), including filled grains per panicle (FGPP), 1,000-grain weight (KGW), tillers per plant (TP), primary branch number (PB), secondary branch number (SB), and main panicle length (MPL). In total, we identified 44 pleiotropic quantitative trait loci (pQTLs), including 29 new pQTLs not found in a single-trait GWAS. We then screened 23 pQTLs showing common effects in two traits as key pQTLs. These key pQTLs were subsequently analyzed by haplotype analysis, which identified 13 pleiotropic candidate genes. Finally, we identified two optimal yield-enhancing allele combinations by pyramiding the superior haplotypes: *GS3-GL3.1-OsCIPK17* for the YD-KGW combination and *GNP12* for the YD-FGPP and YD-SB combinations. This study provides pleiotropic candidate genes and allele combinations that exhibit superior differences in both yield and yield-related traits, offering valuable information for future high-yielding rice breeding.

Keywords: rice, yield, multi-trait GWAS, genetic correlation, pleiotropic SNP

1. Introduction

More than half of the global population relies on rice as a primary staple food. The world's population is anticipated to reach nearly 10 billion by 2050 (Jacquemin *et al.* 2013). To meet this substantial demand, a significant enhancement in rice yield will be necessary. However, rice yield is a complex trait related to many other agronomic traits and is susceptible to environmental conditions (Xing and Zhang 2010). Breeders often encounter tradeoff issues among yield traits when attempting to enhance rice productivity, such as increasing grains per panicle but reducing tillers (Song *et al.* 2022), or enhancing grain number but decreasing grain

Received 28 March 2024; Received in revised form 23 May 2024; Accepted 7 June 2024; Available online 19 July 2024
Chunhai Liu, E-mail: chunhai.liu@stu.hunau.edu.cn; Chao Wu, E-mail: 1481014337@qq.com; #Correspondence Lanzhi Li, Tel: +86-731-84618163, E-mail: lancy0829@163.com
* These authors contributed equally to this study.

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2024.07.026

weight (Li Y *et al.* 2023). Due to such tradeoffs, only a small fraction of cloned genes and alleles have been successfully used for breeding. Therefore, the identification of genes and alleles that can improve both yield and yield-related traits without encountering the tradeoff problem is essential for breeding high-yielding rice.

Yield is one of the most important and complex traits in rice. It is a quantitative trait controlled by multiple genes and heavily influenced by environmental factors (Zhang *et al.* 2017; Tu *et al.* 2022). In rice, yield is directly determined by three traits: filled grains per panicle (FGPP), 1,000-grain weight (KGW), and tillers per plant (TP). It is also influenced by several indirect traits, including primary branch number (PB), secondary branch number (SB), main panicle length (MPL), plant height (PH), seed-setting rate (SS), and growth period (GP) (Sakamoto and Matsuoka 2008; Huang *et al.* 2013). Understanding the relationships between yield and various yield-related traits will improve the efficiency of genetic selection in rice breeding programs. Correlation, path analysis and heritability estimation of agronomic traits are common and crucial analytical methods in the study of yield traits (Akinwale *et al.* 2011; Adjah *et al.* 2020; Nithya *et al.* 2020). Li *et al.* (2019) conducted a comprehensive yield and yield-related trait correlation analysis on 7,686 rice varieties and found that enhancing yield should prioritize increasing FGPP, KGW, PH, SS, and GP while reducing panicle number per unit area and seed length/width ratio in both *indica* inbreds and *indica* hybrid ecotypes. Path analyses conducted by Oladosu *et al.* (2018) showed that three direct traits had a positive impact on rice yield. When accounting for environmental factors, agronomic traits with high heritability and genetic independence or a positive correlation with rice yield can be more effective in assisting grain yield selection (Akinwale *et al.* 2011).

Genome-wide association study (GWAS) is a powerful tool for identifying genes associated with agronomic traits in crop species, and it plays an indispensable role in achieving efficient crop improvement (Yano *et al.* 2016; Xu *et al.* 2024). Although GWAS has been widely used in breeding for rice yield improvement, most studies have focused on single-trait analyses. For complex traits such as rice yield, which is affected by multiple traits, single-trait GWAS has limitations that can prevent the detection of some important pleiotropic SNPs (pSNPs), thus limiting deeper explanations of the genetic basis of complex traits. Multi-trait GWAS, which considers correlations among different traits, has shown greater power in detecting pleiotropic associations and demonstrated superiority in understanding complex traits (Bone *et al.* 2021). Currently, there are many multi-trait GWAS methods applicable to plants, such as mvLMM (Zhou and Stephens 2014), TATES (van der Sluis *et al.* 2013), PC-GWAS (Klei *et al.* 2008), and MTMM (Korte *et al.* 2012). These methods effectively control the false positives caused by population structure and efficiently detect pleiotropic loci. However, many of these methods focus on improving detection power and speed, while often overlooking the interpretability of pleiotropic loci. Methods like mvLMM, TATES, and PC-GWAS all encounter this issue. The MTMM method focuses more on dissecting the effects of pleiotropic loci by decomposing the effects into common and interactive effects, but it can analyze only two traits. When studying yield traits, the MTMM method can identify pleiotropic loci that show common effects in yield and yield-related traits, which is ideal for haplotype pyramiding and thus more effective in aiding molecular marker breeding. Therefore, we selected the MTMM method for the multi-trait GWAS analyses in this study. Akohoue *et al.* (2022) conducted a multi-trait GWAS using MTMM on 401 wheat varieties, and found a promising pQTL on chromosome 7B that enhances Fusarium head blight resistance while reducing plant height and anther retention. By including multiple correlated traits, multi-trait GWAS can effectively capture the genetic relationships between different traits to reveal a broader understanding of the underlying genetic mechanisms.

Quantitative trait locus (QTL) mapping is a highly effective approach for studying the genetic mechanisms of yield in rice (McCough and Doerge 1995). With QTL mapping, the roles of specific yield-associated loci can be described and the locations of genes controlling yield can be identified (Ren *et al.* 2021). To date, 131, 223, and 239 QTLs for FGPP (TO: 0000447), KGW (TO: 0000382), and TP (TO: 0000152) were identified respectively (http://www.gramene.org/), and they are densely distributed across the 12 chromosomes. While only a limited number of QTLs were found to be shared among multiple traits, these pleiotropic QTLs are important for the ultimate identification of the corresponding new pleiotropic genes. Li *et al.* (2018) investigated a diverse core collection of 266 cultivated rice accessions in two environments and identified 15 pQTLs and 47 candidate genes between heading date and yield traits. Rashid *et al.* (2022) detected 33 pQTLs associated with lodging resistance and further identified a candidate gene, *OsFBA2*, by integrating qRT-PCR and haplotype analysis, which provided a theoretical foundation for unraveling lodging resistance.

Marker-assisted selection (MAS) is a breeding technique that involves the identification and selection of superior alleles associated with a QTL, which facilitates the targeted improvement of desired traits and enhances the efficiency and precision of the breeding process (Dormatey *et al.* 2020). Liu *et al.* (2018) reported that rice grain yield QTL *qLGY3* encodes the MADS-domain transcription factor *OsMADS1*, which functions as a crucial downstream effector of G-protein βγ dimers. They demonstrated the effectiveness of integrating the *OsMADS1-lgy3* allele with high-yielding *dep1-1* and *gs3* alleles as a viable approach to concurrently enhance rice productivity and end-use quality. Wu *et al.* (2022) integrated two genes, *GE* (giant embryo) and *OsALDH7* (aldehyde dehydrogenase, golden-like endosperm), into colored rice through MAS, resulting in the development of functional rice with enhanced yield characteristics. MAS significantly accelerates breeding by pyramiding desirable alleles, which enables enhanced genetic diversity conservation, and makes it an essential tool for boosting agricultural sustainability and productivity.

In this study, we crossed 115 male inbred varieties with five male-sterile lines to generate 575 hybrid test crosses. Then, we conducted correlation and heritability analyses on eight yield traits of the progeny and selected two-trait combinations

of yield and yield-related traits for multi-trait GWAS to detect the pleiotropic loci. Next, we screened the pQTLs that showed the common effect in two traits for superior haplotype analysis and mined the candidate genes. This study provides valuable pQTLs and identifies the superior haplotypes of their candidate genes to understand the genetic basis of yield and assist in molecular breeding.

2. Materials and methods

2.1. Plant materials and field management

Under the diallel mating design, 115 *indica* rice accessions (Appendix A) as male parents were crossed with five male sterile varieties as female parents to produce 575 hybrid test crosses. Seeds of the 115 male parental varieties were sourced from the China National Crop GenBank (https://www.cgris.net/cgris_english.html), including restorer varieties of 29 three-line wild-deficient hybrid rice and 86 microcore germplasm accessions. The female parents consisted of five sterile lines, comprising one three-line sterile line and four two-line sterile lines. During 2011 and 2012, 120 parents were planted in Ezhou, Hubei Province, and Lingshui, Hainan Province, China. The resulting 575 F_1 hybrid rice plants were then planted in 2013, with separate plantings at the rice base of Huazhong Agricultural University (Environment I, Env I) and Hubei Academy of Agricultural Sciences (Env II), China. The field experiment was designed as a randomized block group, replicated twice, with 10 plants planted in one row per field at a planting density of 16.7 cm×26.7 cm. Standard agricultural procedures were used for field management, including irrigation, fertilizer application, and insect control.

2.2. Phenotype evaluation

After the materials matured, three plants with uniform growth were selected from the middle eight plants of each plot, and the agronomic traits of two replications were examined. Then, the average phenotypic values for each trait in the two repetitions were calculated for the GWAS analysis. The eight agronomic traits were yield per plant (YD, g), filled grains per panicle (FGPP, no.), 1,000-grain weight (KGW, g), tillers per plant (TP, no.), primary branch number (PB, no.), secondary branch number (SB, no.), main panicle length (MPL, cm), and plant height (PH, cm).

2.3. Resequencing and genotyping

Young leaves were collected and genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method and used to construct sequencing libraries following the Illumina protocol. The genomes of 120 parental varieties were sequenced on the Illumina HiSeq 2500 platform, achieving an average coverage of 11X. The raw reads were filtered to remove those with an N ratio of >10%, more than 50% of the Q value (<5) ratio, and average values of <15. For paired-end read mapping, BWA software (Li and Durbin 2009) was employed with default parameters, using Nipponbare as the reference genome (IRGSP-1.0, http://rapdb.dna.affrc.go.jp). SNP calling was conducted using the GATK UnifiedGenotyper set for diploids with default filtering parameters, which yielded a total of 7,734,465 raw SNPs from the 120 parents. After quality control, 1,894,012 SNPs met the criteria of a missing rate of <20% and a minor allele frequency (MAF) of >5%. The missing genotype was then imputed using the software NPUTE (Roberts *et al.* 2007) (version 4.0). The SNP information of the testcross hybrids was deduced from the SNP information of these five female parents and the 115 male parents. For subsequent GWAS analysis, all the SNP data were finally transformed into HapMap format.

2.4. Heritability estimation and correlation analysis

The narrow-sense heritability of each trait and the genetic correlation coefficients among traits were calculated using the software GCTA (Yang *et al.* 2011) (version 1.26.0). The Pearson correlation coefficient was used as the phenotypic correlation coefficient between the traits. The significance threshold used for correlation analysis was P-value<0.05.

2.5. Multi-trait genome-wide association study

A fully parameterized multi-trait mixed model (MTMM) implemented in R scripts was used for the multi-trait GWAS (Korte *et al.* 2012). Two-trait combinations with significant phenotypic and genetic correlations (P-values<0.05) were used as phenotypic data input for the multi-trait GWAS analysis. The kinship matrix was calculated for the genotypes of 575 hybrids using GCTA (Yang *et al.* 2011) (version 1.26.0). In MTMM, three tests are used to detect the differential effects of associated loci for each two-trait combination: (i) the common effect test identifies loci with the common effects across two traits; (ii) the interactve effect test is used to detect loci that have an effect on one trait or locus but the opposite effects on the other trait; and (iii) the full test is used to detect loci with both common and differing effects (Korte *et al.* 2012). In this study, we used these three tests to analyze the differential effects of pleiotropic loci.

For two traits, the single marker model can be simply described as:

$$y=\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}=s_1\mu_1+s_2\mu_2+x\times\beta+(x\times s_i)\alpha+\mathrm{v}$$

where y_i is the ith trait observation; s_i is a vector of 1 for all values belonging to the ith trait and 0 otherwise; x is the marker; and v–N(0, cov(y)) is a random variable capturing both the error and genetic random effects. The full model was evaluated against a null model where β=0 and α=0. This test identifies loci with both common and differing effects in one model but suffers in power from the extra degree of freedom (see more details in the original article (Korte *et al.* 2012)). The Manhattan plot shows the significant loci with a threshold of P-value<1e–6. The QQ plot includes GWAS results for the two single traits as well as results of the three tests for the two-trait combinations.

2.6. Identification of pSNPs, pQTLs and candidate genes

The Benjamin-Hochberg method (Benjamini and Hochberg 1995) was used to control for false-positive errors (FDR<0.05) in the P-value of the multi-trait GWAS results, and the threshold for selecting significant SNP was set at P-value<1e–6. Significant loci detected by the multi-trait GWAS were

considered pleiotropic SNPs (pSNPs). In this study, the pSNPs in multiple traits were classified into two types of effects, common effects and interactive effects. Common effects refer to pSNPs that show the same effect across multiple traits, whereas interactive effects are those that exhibit trait-specific effects or the opposite effects in different traits. Here, we defined the interval of significantly associated pSNP±LD decay distance as a pQTL. Based on previous studies in our laboratory (Li L *et al.* 2023), we chose approximately 200 kb as the LD decay distance for the 575 rice varieties. To minimize pQTL redundancy, we combined overlapping pQTL regions into one pQTL. Candidate genes associated with agronomic traits were identified by screening the cloned genes within each pQTL region, using information from the China Rice Data Center (https://www.ricedata.cn/gene/).

2.7. Haplotype analysis of pleiotropic QTLs

In this study, pQTLs that showed the common effect in two traits were denoted as key pQTLs. These key pQTLs were selected for further haplotype analysis. Haplotypes were constructed using significant pSNPs within pQTL. Information on the gene structure map, including coding and non-coding regions, was obtained from the Rice Annotation Project (https://rapdb.dna.affrc.go.jp). To ensure a reliable haplotype analysis and reduce the influence of extreme phenotypic values, we chose haplotypes corresponding to hybrid rice plants from at least 10 breeding lines.

3. Results

3.1. Relationships among yield traits in two environments

The phenotypic values for the eight rice agronomic traits were continuously normally distributed with rich variance (Appendix B), indicating they were suitable for Pearson correlation and GWAS analysis. The correlation analysis among these eight traits in two environments and the trait narrow-sense heritabilities are shown in Fig. 1.

YD showed significant positive phenotypic correlations with all other agronomic traits except PH in both environments. Regarding genetic correlations, YD was significantly positively correlated with all traits except PH and KGW in Env I, and PH in Env II. The results indicated that most of the agronomic traits measured were positively correlated with yield. All three yield direct traits (KGW, FGPP, and TP) displayed strong positive correlations in terms of both phenotypic and genetic correlations with YD, except for KGW in Env I. The different correlations of KGW with yield in the two environments indicated that the KGW trait is highly influenced by the environment. In Env II, KGW was moderately correlated with yield with a phenotypic correlation of 0.33 and a genetic correlation of 0.53. Notably, there was a strong positive phenotypic correlation between TP and YD (Env I, 0.7; Env II, 0.6). In addition, YD displayed a moderate positive correlation with FGPP (Env I, 0.59; Env II, 0.54). The genetic correlations of TP and FGPP with YD also exhibited strong positive associations, with (Env I, 0.91; Env II, 0.75) for TP and (Env I, 0.98; Env II, 0.86) for FGPP. These relationships suggested that FGPP and TP significantly affect YD.

Most of the yield indirect traits, except PH, showed moderately positive phenotypic and genetic correlations with YD. In addition, some of the indirect traits were strongly linked to the direct traits. Similar to studies reported in the literature (Xu *et al.* 2023; Zhou *et al.* 2023), FGPP had strong positive correlations with PB and SB, and the results of the genetic correlations were PB (Env I, 0.83; Env II, 0.66) and SB (Env I, 0.98; Env II, 0.94). TP had almost no phenotypic

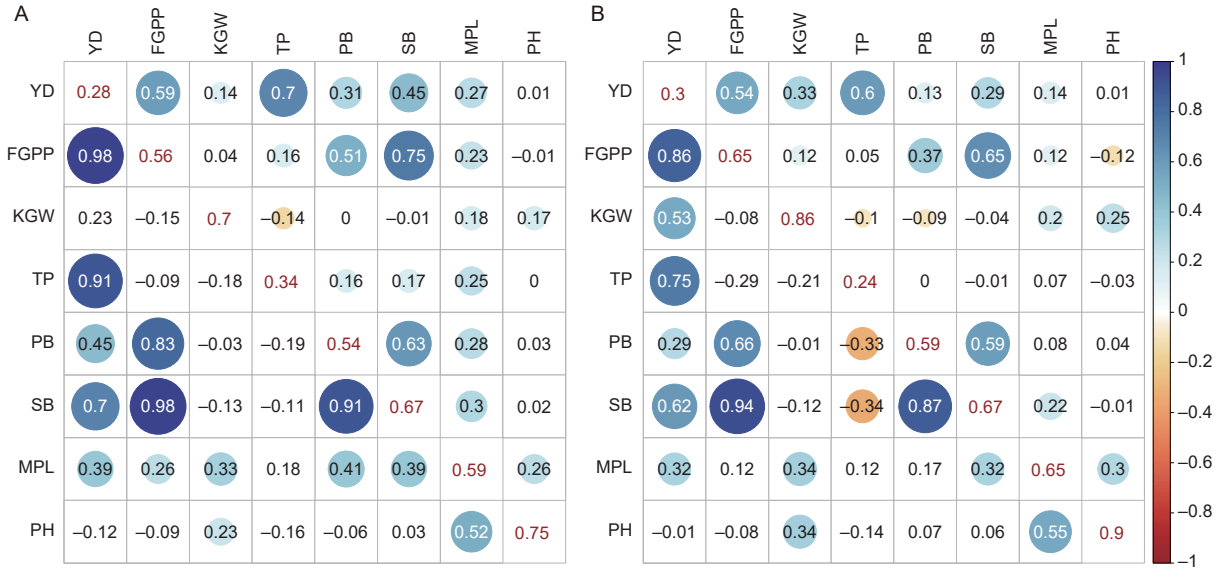


Fig. 1 Phenotypic and genetic correlations, and heritability of traits in two environments. A and B, corplots for Environment I and Environment II, respectively. YD, yield per plant; FGPP, filled grains per panicle; KGW, 1,000-grain weight; TP, tillers per plant; PB, primary branch number; SB, secondary branch number; MPL, main panicle length; PH, plant height. Upper triangle is the phenotypic correlation coefficient using Pearson's correlation; lower triangle is the genetic correlation coefficient; diagonal line is the narrow-sense heritability of the trait. Circles marked in the graph represent significant correlations with *P*-value<0.05.

correlation with any other trait, and only weak genetic correlations with PB and SB in Env II (PB, –0.33; SB, –0.34). These relationships suggest the possibility of improving TP without affecting other traits.

The narrow-sense heritability results indicated a relatively low heritability for YD in both environments (Env I, 0.28; Env II, 0.3), suggesting that direct selection of YD is difficult. The heritability of TP was also low in both environments, at 0.34 and 0.24, respectively. However, the heritability of the other traits was significantly higher, with h^2>0.54. Based on the analysis above, FGPP was chosen as the most suitable trait for improving yield through indirect selection. We then selected traits with significant phenotypic and genetic correlations with YD (*P*-values<0.05) and combined these traits with YD as a set of two-trait combinations. Ultimately, 11 two-trait combinations were identified: FGPP-YD, TP-YD, PB-YD, SB-YD, and MPL-YD in Env I and FGPP-YD, KGW-YD, TP-YD, PB-YD, SB-YD, and MPL-YD in Env II. These selected combinations were used in the subsequent multi-trait GWAS analysis.

3.2. Detection of pleiotropic loci by muti-trait GWAS

In this study, we conducted multi-trait GWAS with the full-test model, which can simultaneously detect common effects, interactive effects, and effects on individual traits. In total, 6,608 pSNPs (Env I, 4,477; Env II, 2,131) were identified, which were merged into 109 pQTLs (Env I, 51; Env II, 58) (Fig. 2; Appendices C and D). In addition, we identified 3,652 common effect SNPs (Env I, 1,466; Env II, 2,186) and 100 interactive effect SNPs (Env I, 71; Env II, 29). Notably, a significant proportion of these pSNPs showed clustering patterns. For example, one QTL interval (*qFGPP4.2*) of FGPP-YD contained 895 pSNPs. This QTL region is associated with the *LMPA* gene, an important gene that regulates panicle development (Hu *et al.* 2022). The 109 pQTLs were collectively associated with a total of 139 cloned genes (Appendix C). In Env I, the FGPP-YD, TP-YD, PB-YD, SB-YD, and MPL-YD trait combinations identified 1,651, 29, 865, 1,910, and 22 pSNPs, respectively (Fig. 2-A), corresponding to 10, 4, 17, 15, and 5 pQTLs, along with 15, 9, 20, 21, and 3 cloned genes. In Env II, the FGPP-YD, KGW-YD, TP-YD, PB-YD, SB-YD, and MPL-YD trait combinations identified 11, 646, 1,152, 1,176, and 145 pSNPs, respectively (Fig. 2-B), corresponding to 4, 16, 1, 12, 7, and 18 pQTLs, along with 11, 16, 0, 12, 7, and 25 cloned genes (Appendix D). Among the 109 pQTLs, certain cloned genes were repeatedly annotated to multiple trait combinations (Appendices C and E). For example, the pQTL associated with the *LMPA* gene, which showed clustering patterns, was repeated in four trait combinations: FGPP-YD, TP-YD, PB-YD, and SB-YD (Env I,

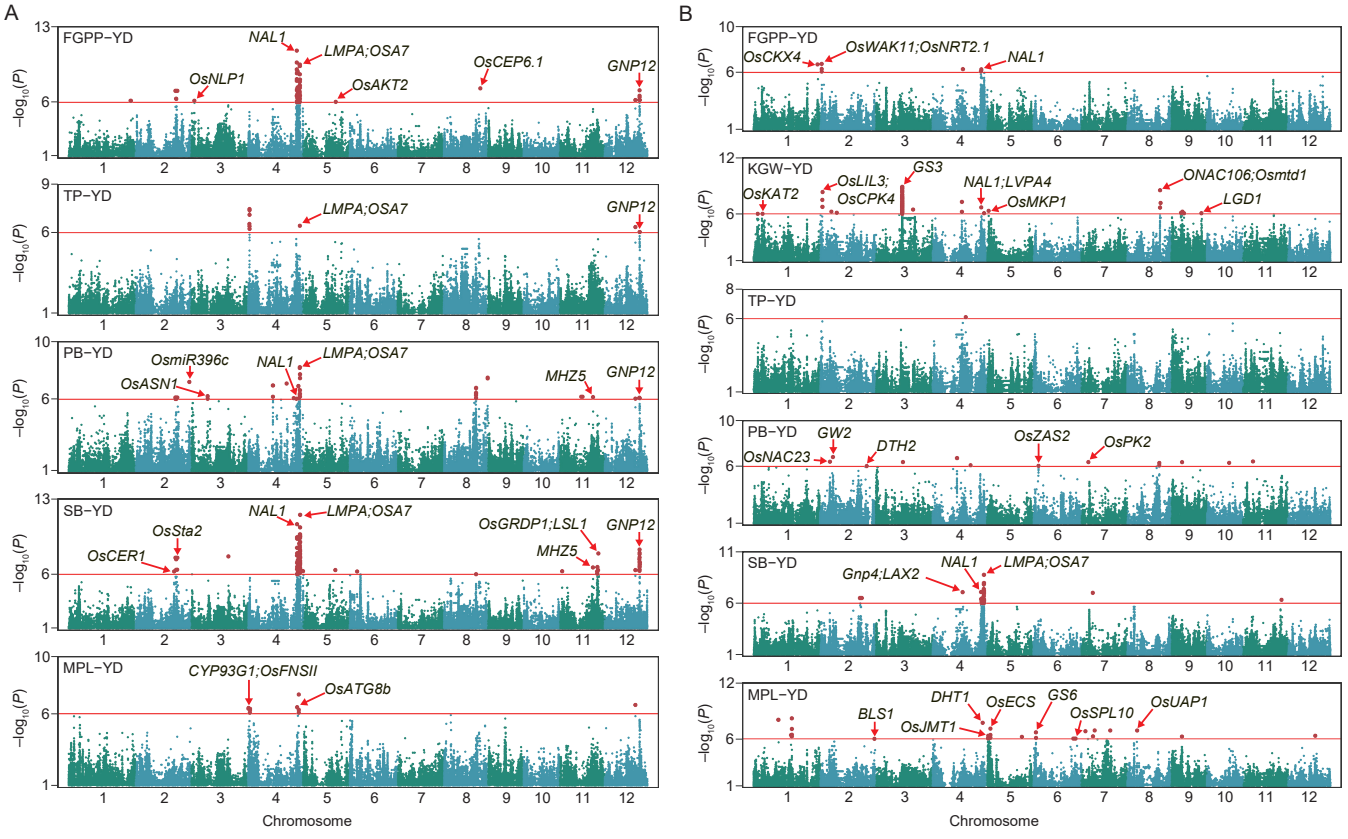


Fig. 2 Multi-trait genome-wide association studies (GWAS) results for each two-trait combination in Environment I (Env I) and Env II. A, Manhattan plot of the multi-trait GWAS results for five combinations in Env I. B, six combinations in Env II. YD, yield per plant; FGPP, filled grains per panicle; TP, tillers per plant; KGW, 1,000-grain weight; PB, primary branch number; SB, secondary branch number; MPL, main panicle length. Red lines show the genome-wide significance threshold of *P*-value=1e–6. Cloned genes reported to influence yield or yield-related traits are labeled in black at the nearest associated SNPs.

qFGPP4.2, *qTP4.2*, *qPB4.6*, and *qSB4.2*; Env II, *qSB4.6*). Similarly, the pQTL associated with the *NAL1* gene was detected in FGPP-YD, PB-YD, SB-YD, and KGW-YD (Env I, *qFGPP4.1*, *qPB4.4*, and *qSB4.1*; Env II, *qFGPP4.4*, *qSB4.5*, and *qKGW4.2*). The pQTL associated with *GNP12* was also detected in four trait combinations: FGPP-YD, TP-YD, PB-YD, and SB-YD in Env I (*qFGPP12.2*, *qTP12.2*, *qPB12.2*, *qSB12.2*). The quantile-quantile plot shows that the multi-trait GWAS results are more in line with the expected *P*-values than the single-trait GWAS results, indicating that the MTMM full-test method effectively controls the false positives (Appendices F and G).

Many cloned genes related to important yield traits were annotated in the pQTLs detected by multi-trait GWAS (Appendix C). For example, *NAL1* is a crucial gene that regulates the development of vascular bundles and positively affects panicle and grain yield (Qi *et al.* 2008; Huang *et al.* 2018). In the FGPP-YD trait combination, *OsCEP6.1* was detected on chromosome 8 (*qFGPP8*). Sui *et al.* (2016) showed that overexpression of the *OsCEP6.1* gene results in the alteration of multiple traits in rice, including the number of grains per spike, the length of grains, the width of grains, and the KGW. In the KGW-YD trait combination, the *qKGW3.1* pQTL annotated the *GS3* gene, the primary QTL governing grain weight in rice. Fan *et al.* (2006) demonstrated that the *GS3*-encoded protein negatively regulates the weight of grains. In the PB-YD trait combination, the *GW2* genes on chromosome 2 were annotated in Env II. *GW2* is a crucial gene controlling grain width and grain weight in rice. The allele *gw2.1* of *GW2* can regulate grain size and grain weight, which in turn increases rice yield (Huang *et al.* 2022).

Most significant SNPs identified in the single-trait GWAS, as well as the common and interactive effect loci of multiple traits were detected by the full-test model in the MTMM GWAS analysis (Fig. 3-A). Using the FGPP-YD results in Env I as an example, the full-test results overlapped with most of the single-trait GWAS significant loci (1,203 loci) and pleiotropic loci, but there were still 21 single-trait loci that were not detected by the full test, and the full test did not detect seven common effect loci (Fig. 3-A). In addition, multi-trait GWAS identified 457 more SNPs compared to the single-trait GWAS (Fig. 3-B). We used the newly identified pleiotropic SNPs from the multi-trait GWAS in 11 trait combinations for subsequent analyses.

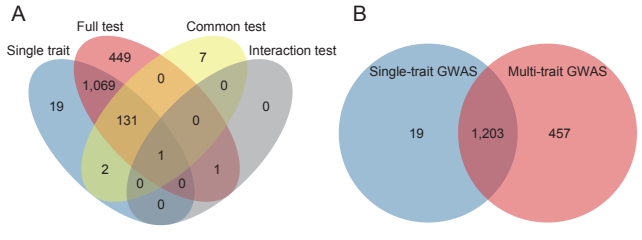


Fig. 3 Venn diagrams summarizing the genome-wide association studies (GWAS) of filled grains per panicle (FGPP)-yield per plant (YD) in Environment I (Env I). A, single-trait GWAS identified a total of 1,222 loci, of which 1,201 were shared with the full-test results, 134 with the common test results, and 1 with the interactive test results. The full test detected 450 new loci, while the common test detected seven new loci. B, the multi-trait mixed model (MTMM) method applied to multi-trait GWAS and single-trait GWAS shared 1,203 loci and was able to detect 457 new loci.

3.3. Screening of important common effect pQTLs and identification of candidate genes

In total, 322 novel loci with common or interactive effects, which were not detected in the single-trait GWAS, were merged into 44 pQTLs. We then functionally annotated these pQTLs into 84 genes, including 60 cloned genes (Appendix H). Among the 44 pQTLs, 15 pQTLs (common effect, 11; interactive effect, 4) were repeatedly detected in the single-trait GWAS, while the remaining 29 pQTLs (common effect, 12; interactive effect, 17) were newly discovered and not detected in the single-trait GWAS. The set of 15 pQTLs, although repeatedly detected in single traits, consisted of pSNPs that were newly identified. We then performed further screening of these 44 pQTLs. Compared to interactive loci with differential effects on two traits, the common loci with effects on yield and yield-related traits in the same direction are helpful for rice breeding. Thus, we selected 23 common effects pQTLs as key pQTLs for haplotype analyses (Figs. 4–6; Appendices I–S). Among these key pQTLs, 11 were associated with known cloned genes that affect yield or yield-related traits. The other 12 served as candidate pQTLs of unknown function. We identified 13 candidate genes within these 12 candidate pQTLs, which included eight newly detected pQTLs and four pQTLs that were also detected in the single-trait analyses. The haplotype results for these 13 candidate genes showed significant phenotypic differences in both traits, which is highly beneficial for molecular breeding (Appendix I).

In the FGPP-YD trait combination, three pQTLs (*iqFGPP11.1*, *iqFGPP11.2*, and *iqFGPP12*) were detected. Only one cloned gene (*GNP12*) and two candidate genes (*Os11g0572800* and *Os11g0537400*) were associated (Appendix I). Within the *iqFGPP12* interval, we identified *GNP12* (Figs. 4-A and 5-A), which encompasses six newly identified common effect pSNPs. The *GNP12* gene encodes the RGH1A protein, which regulates rice yield according to panicle length, grain number per panicle, and grain length (Pan *et al.* 2022). The gene *Os11g0572800* in *iqFGPP11.1* is a good example that was selected as a candidate gene due to its significant common effect (Figs. 4-B and 5-B). Phenotypic distribution analysis of the significant locus chr11_21492452 in *iqFGPP11.1* showed a strong positive effect on both FGPP and YD (Fig. 4-C). Similarly, in Env II, the gene *Os11g0537400* in *iqFGPP11.2* was also selected as a candidate gene (Fig. 5-C). Notably, *GNP12* was repeatedly detected in both the single-trait GWAS and multi-trait GWAS in the full-test and common results.

In the KGW-YD trait combination, we screened a total of seven pQTLs (*iqKGW2*, *iqKGW3.1*, *iqKGW3.2*, *iqKGW5*, *iqKGW8*, *iqKGW9.1*, and *iqKGW9.2*) in Env II (Appendix I). A total of 14 cloned genes and two candidate genes were associated. We detected *GS3* (Fan *et al.* 2006), *GL3.1* (Qi *et al.* 2012), and *OsCIPK17* (Kanwar *et al.* 2014) in *iqKGW3.1*, *iqKGW3.2* and *iqKGW5*, respectively, which are important genes affecting KGW and YD in rice (Fig. 5-D–F).

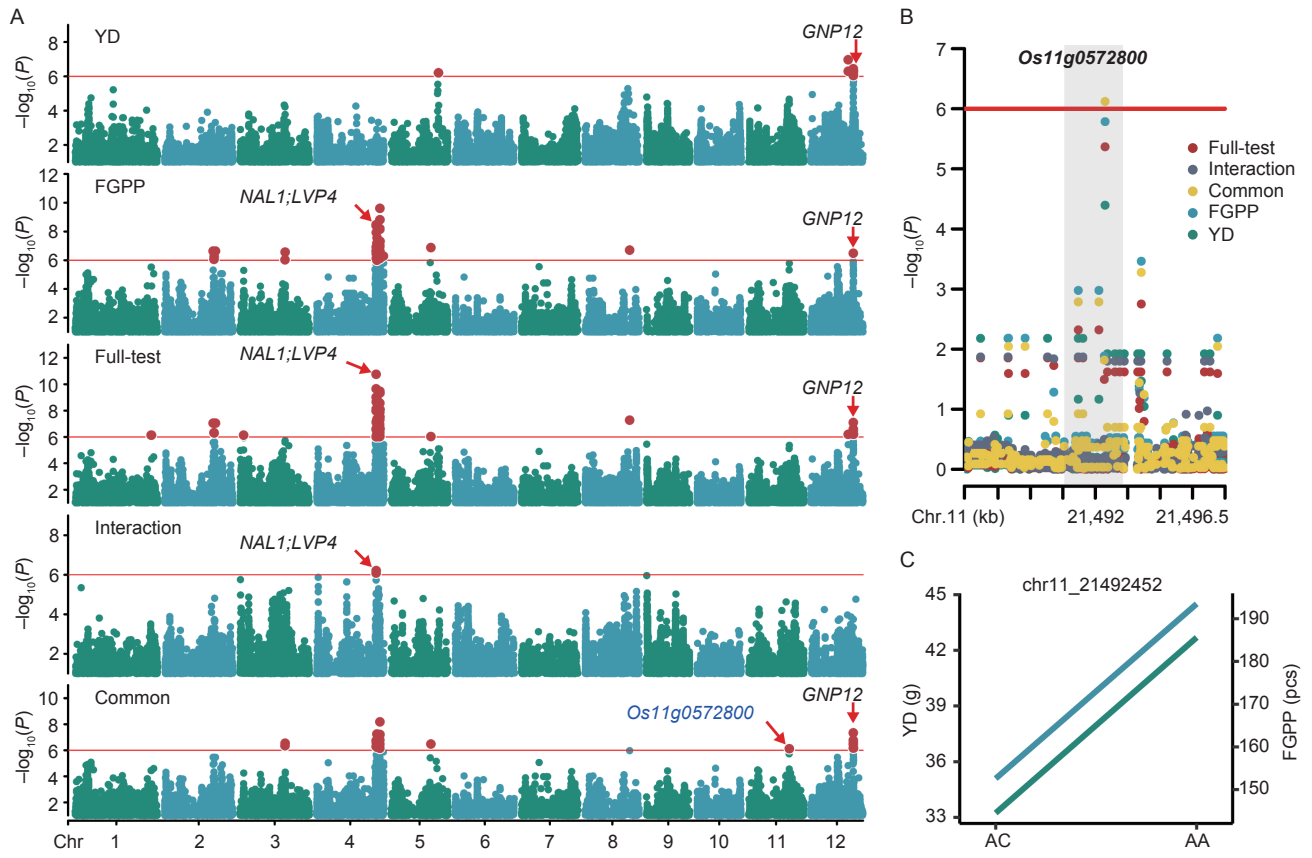


Fig. 4 Detection of candidate genes by multi-trait genome-wide association studies (GWAS). A, Manhattan plots of the multi-trait mixed model (MTMM) results for filled grains per panicle (FGPP)-yield per plant (YD) in Environment I (Env I). Reported cloned genes associated with pleiotropic loci are labeled in black; unreported candidate genes are labeled in bright blue. B, enlarged view of the *Os11g0572800* region on chromosome 11. Gray shading denotes the *Os11g0572800* gene region. C, estimated phenotypic effect of the peak SNP of chr11_21492452 in the *Os11g0572800* region, shown in light blue (FGPP) and green (YD).

Moreover, *OsLIL3* (Fig. 5-G) and *ONAC106* (Fig. 5-H) were detected in *iqKGW2* and *iqKGW8*. In addition, *Os09g0281700* (Fig. 5-I) and *Os09g0292900* (Fig. 5-J) were annotated as candidate genes on chromosome 9.

In the TP-YD trait combination, we screened three pQTLs (*iqTP5.2*, *iqTP8*, and *iqTP10*). A total of two cloned genes and two candidate genes were associated (Appendix I). We detected *OsCKX11* (Fig. 5-K) in the pQTL *iqTP8*. Note that *OsCKX11* functions antagonistically between cytokinins and ABA in leaf senescence in rice, and the mutant *osckx11* presented with significantly increased FGPP, TP, and PB compared with the wild type (Zhang *et al.* 2021). In *iqTP5.2* and *iqTP10*, we identified *Os05g0498300* (Fig. 5-L) and *Os10g0456800* (Fig. 5-M) as candidate genes, respectively. *Os02g0804900* is identical to *DCA1*, which acts as a transcriptional co-activator of DST and contributes to drought and salt tolerance in rice (Cui *et al.* 2015).

In the PB-YD trait combination, we screened four pQTLs (*iqPB1*, *iqPB2*, *iqPB9*, and *iqPB11*). Only one pQTL *iqPB2*-associated clone gene and three novel candidate genes were identified (Fig. 6-A–D; Appendix I). *OsPRR1*, identified in the *iqPB2* interval, inhibits *OsCCA1* expression and mediates tillering and spike development in rice (Wang *et al.* 2020). We found *Os01g0363900* (crown root defect 1), *Os09g0292900* (kelch repeat-containing F-box protein), and *Os11g0572200* (expressed protein) to be candidate genes in *iqPB1*, *iqPB9* and *iqPB11*, respectively.

In the SB-YD trait combination, we screened four pQTLs (*iqSB3.2*, *iqSB7*, *iqSB11*, and *iqSB12*), where two pQTLs are associated with cloned genes and two with candidate genes (Fig. 6-E–H; Appendix I). The pQTL *iqSB7* identified the *OsA2* gene (Fig. 6-E). Sperandio *et al.* (2020) found that the cell membrane proton ATPase *OsA2* regulates aboveground growth and seed yield in rice by affecting nitrate uptake. The *iqSB12* associated *GNP12* (Pan *et al.* 2022) and *OsCSLD4* appeared repeatedly (Fig. 6-F). *OsCSLD4* regulates not only rice growth under normal conditions but also salt tolerance, which helps plants to resist biomass and yield losses under salinity stress (Liu *et al.* 2022). The pQTL *iqSB3.2* covers a wide interval on chromosome 3, ranging from 6.12 to 6.65 Mb, which consists of 175 pSNPs and is associated with three genes (*Os03g0220100*, *Os03g0221700* and *Os03g0223400*). We selected the gene closest to the interval, namely *Os03g0220100* (Fig. 6-G), as the candidate gene. In addition, *Os11g0572800* (Fig. 6-H) was identified as a candidate gene in the *iqSB11* interval.

In the MPL-YD trait combination, we only screened two pQTLs (*iqMPL1.2* and *iqMPL5*) in Env II (Appendix I). In the *iqMPL5* intervals, we identified the *OsJMT1* gene (Fig. 6-I).

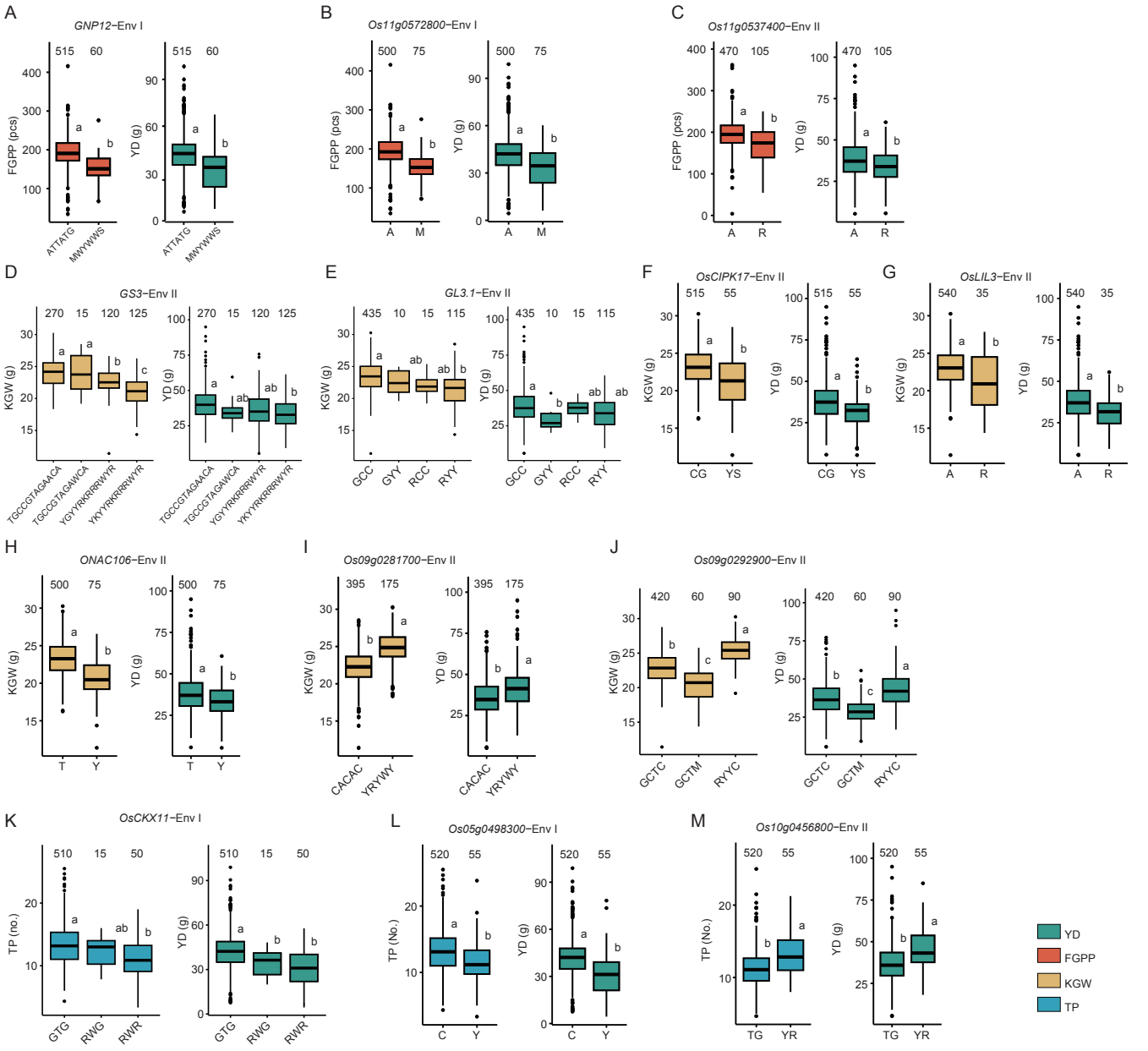


Fig. 5 Haplotype boxplots for two-trait combinations of yield and yield direct traits. A–C, for the filled grains per panicle (FGPP)-yield per plant (YD) trait combination, haplotype box plots of *GNP12* (A), *Os11g0572800* (B), and *Os11g0537400* (C). D–J, for the 1,000-grain weight (KGW)-YD combination, haplotype box plot of *GS3* (D), *GL3.1* (E), *OsCIPK17* (F), *OsLIL3* (G), *ONAC106* (H), *Os09g0281700* (I), and *Os09g0292900* (J). K–M, for the tillers per plant (TP)-YD trait combination, haplotype box plots of *OsGLP1* (K), *Os05g0498300* (L), and *Os10g0456800* (M). Env, environment. Letters in each subplot suggest the significance by ANOVA (for >three haplotypes) or the *t*-test (for two haplotypes) at $P<0.05$. The value on each box is the sample size of each haplotype. The *x*-axis coordinates are genotypes corresponding to significant SNPs in the QTL interval.

Overexpression of *OsJMT1* results in reduced plant height, panicle length, and grain yield in rice, while not affecting the number of panicles per plant (Qi *et al.* 2016). The *Os01g0385400* (Fig. 6-J) gene was annotated as a candidate gene in the *iqMPL1.2* interval, and *Os01g0385400* is identical to *SLAC7,* which is a light-responsive gene crucial for maintaining the stability of chloroplasts in rice (Fan *et al.* 2015).

3.4. Identification of optimal allele combinations by pyramiding superior haplotypes

Haplotype-based breeding, a recent and promising approach for developing tailor-made crop varieties, involves identifying superior haplotypes and incorporating them into breeding programs (Sinha *et al.* 2020). Among the above 23 key pQTLs, we screened two groups of pQTLs for optimal allele combination identification using two different methods. Method I involved selecting pQTLs associated with several important cloned genes in the same trait. Method II involved selecting pQTLs associated with the same cloned genes in multiple traits. Using method I, we selected pQTLs associated with three important cloned genes (*GS3*, *GL3.1*, and *OsCIPK17*) in KGW, corresponding to pQTLs (*iqKGW3.1*,

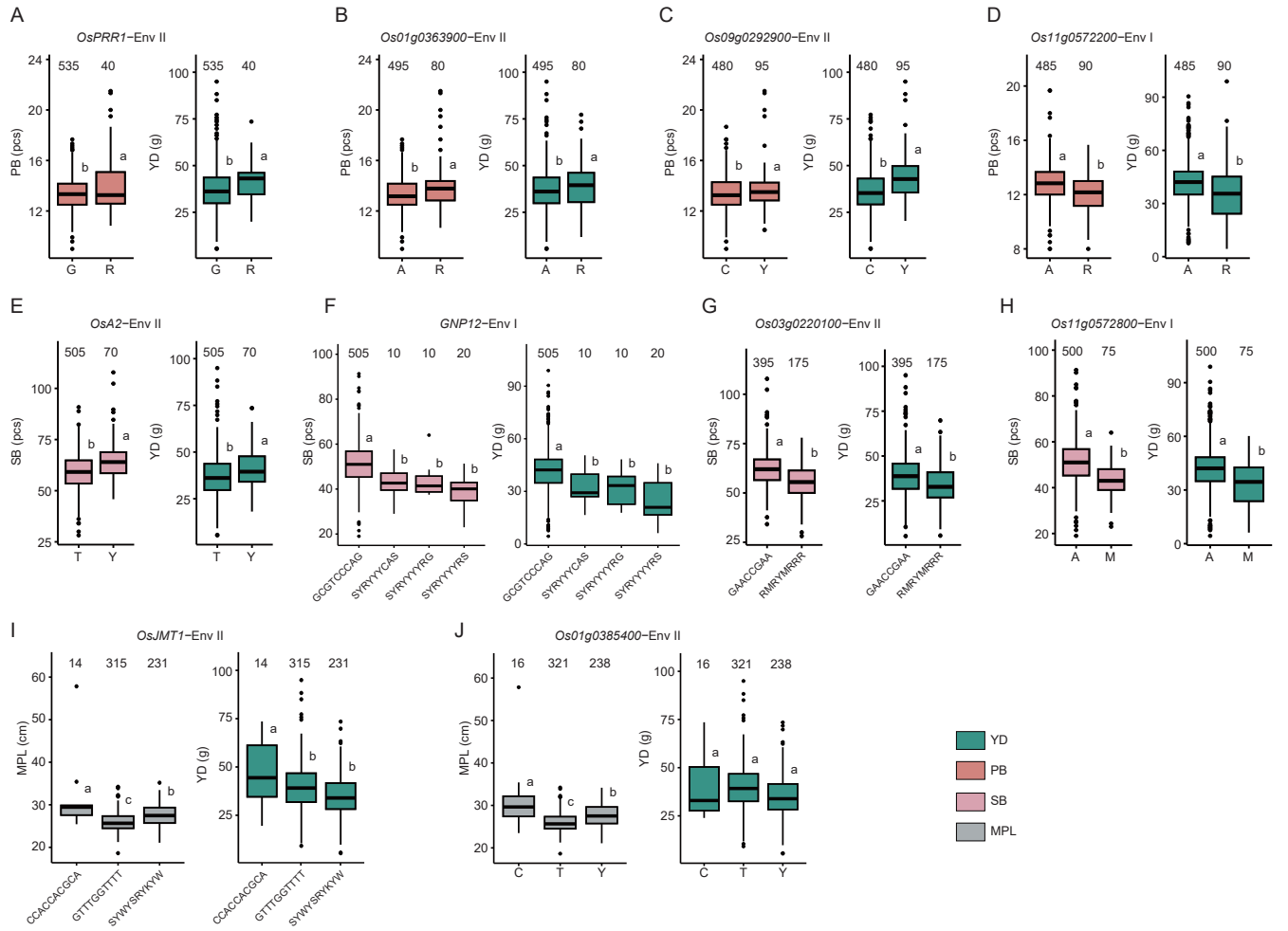


Fig. 6 Haplotype boxplots for two-trait combinations of yield and indirect traits of yield. A–D, for the primary branch number (PB)-yield per plant (YD) trait combination, haplotype box plots of *OsPRR1* (A), *Os01g0363900* (B), *Os09g0292900* (C), and *Os11g0572200* (D). E–H, for the secondary branch number (SB)-YD combination, haplotype box plots of *OsA2* (E), *GNP12* (F), *Os03g0220100* (G), and *Os11g0572800* (H). I–J, for the main panicle length (MPL)-YD trait combination, haplotype box plots of *OsJMT1* (I) and *Os01g0385400* (J). Letters suggest the significance by ANOVA (for >three haplotypes) or the *t*-test (for two haplotypes) at $P<0.01$. The value on each box is the number of individuals of each haplotype. The *x*-axis coordinates are the corresponding significant SNPs, with light blue boxes for yield-related traits and green boxes for yield.

iqKGW3.2, and *iqKGW5*). Using method II, we selected two pQTLs (*iqFGPP12* and *iqSB12*) for two different traits associated with the cloned gene *GNP12*. These two sets of pQTLs were then used for superior haplotype pyramiding to obtain the optimal allele combination.

For KGW, haplotype analysis revealed four haplotypes for *GS3*: Hap A (YKYYRKRRRWYR), Hap B (YGYYRKRRRWYR), Hap C (TGCCGTAGAACA), and Hap D (TGCCGTAGAWCA) (Fig. 7-A). *GL3.1* has two haplotypes: Hap A (GCC) and Hap B (RYY) (Fig. 7-B). *OsCIPK17* also has two haplotypes: Hap A (CG) and Hap B (YS) (Fig. 7-C). Regarding the cloned gene *GNP12*, which affects FGPP and SB and is associated with two pQTLs, *iqFGPP12* and *iqSB12*, we found that *iqFGPP12* has two haplotypes: Hap A (ATTATG) and Hap B (MWYWWS), while *iqSB12* has four haplotypes: Hap A (GCGTCCCAG), Hap B (SYRYYYCAS), Hap C (SYRYYYYRG), and Hap D (SYRYYYYRS) (Fig. 7-D).

In the KGW-YD trait combination, we found that the maximum values of both YD and KGW were present in the Hap D (*GS3*_Hap3-*GL3.1*_Hap1-*OsCIPK17*_Hap1) allele combination, and this combination also had the largest number of samples, 230 in total (Fig. 8-A). In contrast, Hap B (*GS3*_Hap1-*GL3.1*_Hap2-*OsCIPK17*_Hap2) showed the smallest values of YD and KGW. In addition, Hap C and Hap F–I also showed more significant phenotypic advantages.

For the *GNP12* gene, haplotype pyramiding analysis revealed that Hap A (*iqFGPP12*_Hap1-*iqSB12*_Hap1) was maximal for all three traits, FGPP, SB, and YD, and showed significant differences compared to the other allele combinations (Fig. 8-B). This optimal allele combination also had the largest number of samples with 505.

4. Discussion

Yield is a complex trait controlled by polygenes and influenced by multiple traits. The intricate genetic structure, low heritability, and susceptibility to environmental factors hinder the direct selection of yield-controlling genes in rice breeding (Xing and Zhang 2010). Therefore, selecting the

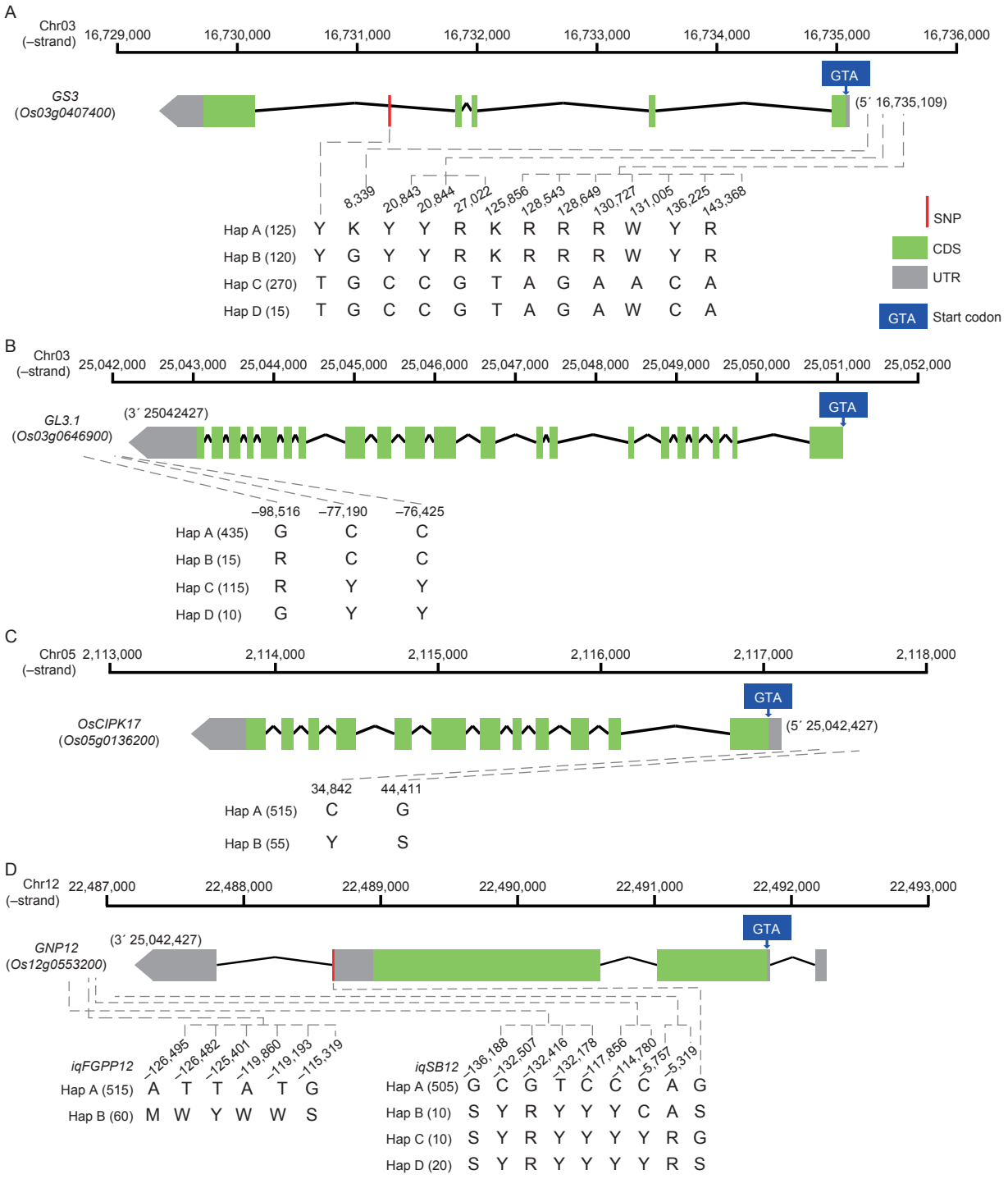
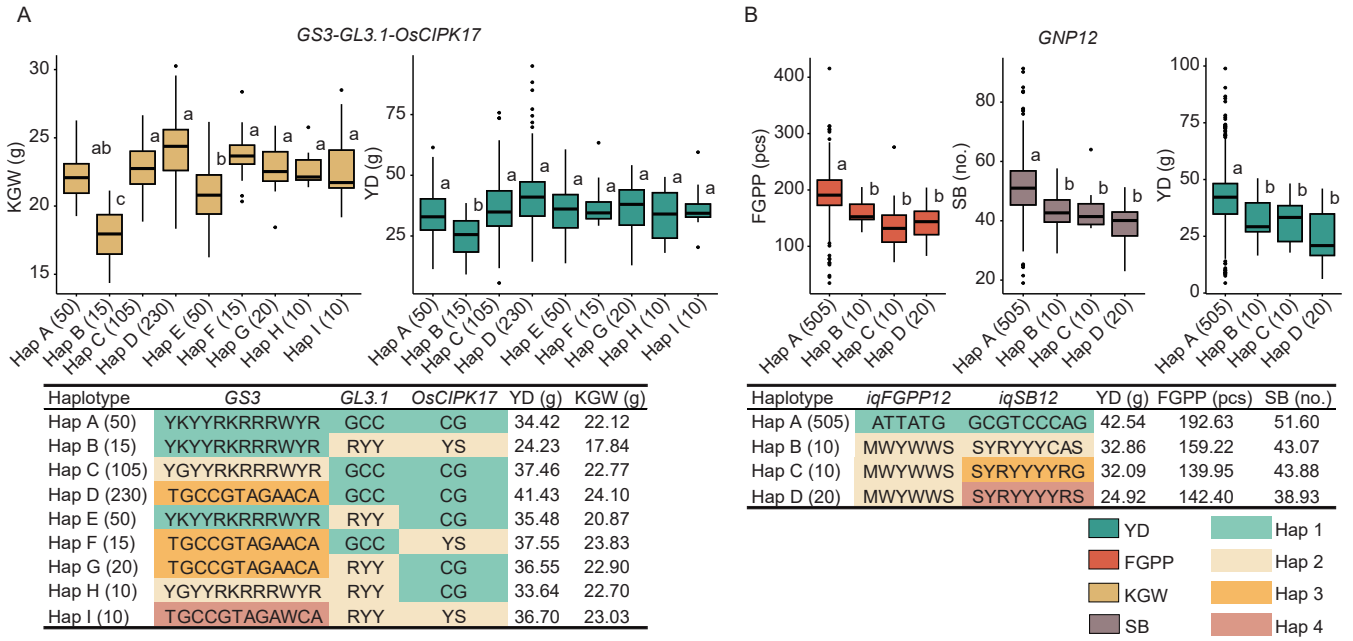


Fig. 7 Gene structure, significant SNP loci, and alleles. A, *GS3*, Hap A (YKYYRKRRRWYR), Hap B (YGYYRKRRRWYR), Hap C (TGCCGTAGAACA), and Hap D (TGCCGTAGAWCA). B, *GL3.1*, Hap A (GCC) and Hap B (RYY). C, *OsCIPK17*, Hap A (CG) and Hap B (YS). D, *iqFGPP12*, Hap A (ATTATG) and Hap B (MWYWWS). *iqSB12*, Hap A (GCGTCCCAG), Hap B (SYRYYYCAS), Hap C (SYRYYYYRG) and HapD (SYRYYYYRS). Blue arrow indicates the location of each start codon "GTA"; number in brackets represents the 5′ or 3′ end position of each cloned gene; number above the significant SNP is its position relative to the 5′ or 3′ end position of the cloned gene.

genes for yield-related traits has been suggested as an additional strategy for enhancing rice yield production, which has also been emphasized by national and international rice breeding programs (Su *et al.* 2021; He *et al.* 2024). This study used a multi-trait GWAS approach to include both yield and yield-related traits in the model, with the aim of identifying new pleiotropic loci and their superior alleles that can improve rice yield.

4.1. Exploring rice yield pleiotropy with multi-trait GWAS

Most GWAS to date have been conducted using the



Haplotype	*GS3*	*GL3.1*	*OsCIPK17*	YD (g)	KGW (g)
Hap A (50)	YKYYRKRRRWYR	GCC	CG	34.42	22.12
Hap B (15)	YKYYRKRRRWYR	RYY	YS	24.23	17.84
Hap C (105)	YGYYRKRRRWYR	GCC	CG	37.46	22.77
Hap D (230)	TGCCGTAGAACA	GCC	CG	41.43	24.10
Hap E (50)	YKYYRKRRRWYR	RYY	CG	35.48	20.87
Hap F (15)	TGCCGTAGAACA	GCC	YS	37.55	23.83
Hap G (20)	TGCCGTAGAACA	RYY	CG	36.55	22.90
Hap H (10)	YGYYRKRRRWYR	RYY	CG	33.64	22.70
Hap I (10)	TGCCGTAGAWCA	RYY	YS	36.70	23.03

Haplotype	*iqFGPP12*	*iqSB12*	YD (g)	FGPP (pcs)	SB (no.)
Hap A (505)	ATTATG	GCGTCCCAG	42.54	192.63	51.60
Hap B (10)	MWYWWS	SYRYYYCAS	32.86	159.22	43.07
Hap C (10)	MWYWWS	SYRYYYYRG	32.09	139.95	43.88
Hap D (20)	MWYWWS	SYRYYYYRS	24.92	142.40	38.93

Fig. 8 Boxplots and tables for the allele combinations. A, boxplots and table for the allele combination of three cloned genes (*GS3*, *GL3.1*, and *OsCIPK17*) in relation to 1,000-grain weight (KGW) and yield per plant (YD). B, boxplots and table for the allele combination of *GNP12* in relation to *iqFGPP12* and *iqSB12*. FGPP, filled grains per panicle; SB, secondary branch number. Hap A–I represent each of the different allele combinations, followed by a number in parentheses representing the number of samples in each combination. Hap 1–4 represent the different haplotypes in each gene or pQTL.

variations of a single-trait general linear model, like MLM (Yu *et al.* 2006), SUPER (Wang *et al.* 2014), FarmCPU (Liu *et al.* 2016), and others. The advantages of the single-trait GWAS model are model simplicity, calculation speed, and high accuracy. Given that most traits of interest are multifactorial, this clearly leads to model misspecification, ultimately leading to a proportion of false positives and loss of pleiotropic loci (Korte *et al.* 2012). Multi-trait analysis increases the statistical power compared to single-trait analysis, especially in the case of complex biological processes and correlated traits (Ruotsalainen *et al.* 2021). However, due to increased model complexity and reduced interpretability, the application of multi-trait GWAS has been relatively limited. The two most popular plant multi-trait GWAS methods are mvLMM (Zhou and Stephens 2014) and MTMM (Korte *et al.* 2012). These two methods are widely used in plant research (Jaiswal *et al.* 2016; Carlson *et al.* 2019), but they have different modeling focuses. The mvLMM method uses a multivariate linear mixed model, which can assess more than two correlated traits at the same time and has a faster computational speed. However, the mvLMM approach has shortcomings in interpreting the effects of detected pleiotropic loci, which is a problem for many other multi-trait methods such as TATES (van der Sluis *et al.* 2013) and PC-GWAS (Klei *et al.* 2008). In contrast, the MTMM method employs a fully parameterized multi-trait mixed model that is capable of accurately identifying the specific effects of loci in two traits. Although MTMM can only consider two traits at a time, it provides a clear explanation of the effects of pleiotropic loci, which is very useful for haplotype-based molecular breeding. In addition, there are few studies on multi-trait GWAS for rice yield. To effectively explore pleiotropic genes linked to yield in the population, this study employed the MTMM model to conduct multi-trait GWAS between yield and yield-related agronomic traits (Korte *et al.* 2012). Compared with other multi-trait GWAS methods, the MTMM method not only addresses population structure but also provides a more in-depth classification of pleiotropic loci. The MTMM method can identify the pleiotropy of SNPs associated with two traits, including the interactive effect, which is useful for one specific trait, and the common effect, which has a positive effect on both traits. By decomposing the effects of SNPs, we can gain a deep understanding of the pleiotropy of these loci. In this study, 243 novel common effect pSNPs were detected and merged into 23 pQTLs, as well as 89 interactive effect pSNPs, which were merged into 21 pQTLs (Appendix H). Newly discovered interactive effect loci have limited utility compared to common effect loci but may have potential value in specific contexts. For example, in some breeding situations, breeders may focus on improving one trait without affecting others. This approach has significant advantages in yield breeding by considering relatively less complex yield-related traits, such as FGPP or KGW, to improve rice yield traits. These pleiotropic loci have a greater advantage in breeding than single traits. The identified pSNPs, pQTLs, and genes collectively play a crucial role in elucidating the genetic mechanisms of rice yield.

4.2. Pyramiding breeding based on superior haplotypes of common effect pQTLs

Pyramiding superior haplotypes is an effective strategy in molecular design breeding (Lv *et al.* 2021). In the study of

Xiao *et al.* (2021) on maize, they identified an optimal allele combination by pyramiding the superior haplotypes of the *MADS69*, *ZCN8*, and *myb74* genes. They used this optimal allele combination to propose an effective molecular design with the goal of achieving a shorter plant height (PH) and earlier days to tasseling (DTT) without reducing spike weight (EW). Gene pleiotropy is another crucial consideration in molecular breeding, as the optimal allele of the target gene may have negative effects on other traits (He *et al.* 2024). By pyramiding superior haplotypes of common effect pQTLs, we obtained six superior allele combinations with three genes associated with KGW (*GS3*, *GL3.1*, and *OsCIPK17*) and one superior allele combination with the *GNP12* gene. These superior allele combinations exhibit significantly higher yields and associated yield-related traits compared to the other haplotypes. This indicates that integrating the common effect pQTL haplotypes through pyramiding will enhance the effectiveness of the MAS method in achieving trait improvement goals. The haplotype results of the validated cloned genes in this study suggest that similar types of identified candidate genes can affect rice yield and yield-related traits. These candidate genes of unknown function with common effects can facilitate molecular breeding for high yield in rice.

4.3. Enlightenment and prospects of rice yield multi-trait breeding

At present, scientific and technological workers in the field of genetic breeding have shown great interest in multi-trait breeding (Bone *et al.* 2021; Shirai *et al.* 2022). In a study by Schulthess *et al.* (2017), a multi-trait association study unveiled the genetic basis of wheat grain yield by identifying 345 significant markers linked to various yield-related traits, such as plant height, KGW, grains per ear, and ear weight. In this study, we investigated the genetic relationships between yield and yield-related traits using the MTMM multi-trait GWAS method and gained valuable insights. However, the GWAS statistical model is always accompanied by the presence of false positives when explaining complex traits. Therefore, understanding the complex interactions between yield and its related traits remains a challenge. Conducting multi-trait GWAS analyses involving three or more traits may complicate calculations and the interpretation of results. Meanwhile, our results suggest the existence of pleiotropic loci affecting multiple traits that are susceptible to environmental conditions. To obtain a comprehensive understanding of the genetic mechanisms underlying rice yield and its related traits, breeders should conduct more experiments and studies that account for environmental effects. This integrated approach will facilitate the development of stable, high-yielding hybrid rice varieties that are suitable for different growing conditions and refine the multi-trait breeding strategy.

5. Conclusion

In summary, we conducted an extensive multi-trait GWAS analysis on eight yield-related traits across 575 rice varieties under two different environments. A total of 44 pQTLs were identified, including 29 novel pQTLs that were not detected in the single-trait GWAS. In addition, we screened 23 key pQTLs exhibiting common effects and identified 13 candidate genes. Subsequently, we performed haplotype analyses of these pQTLs and selected three cloned genes (*GS3*, *GL3.1*, and *OsCIPK17*) associated with KGW, as well as the gene *GNP12*, which affects multiple traits, for superior haplotype pyramid analysis. The haplotype pyramid analysis identified specific optimal allele combinations influencing yield, which represent potential targets for breeding strategies. By shedding light on the genetic mechanisms governing rice yield and yield-related traits, this study provides valuable insights for guiding the development of targeted breeding approaches aimed at enhancing rice productivity.

Acknowledgements

This work was supported by the Science and Technology Innovation Program of Hunan Province, China (2023NK2001), the Hubei Key Laboratory of Food Crop Germplasm and Genetic Improvement, China (2022LZJJ08), and the Changsha Science and Technology Project, China (kq2402109). The authors thank Prof. Sibin Yu and Tongmin Mou from Huazhong Agricultural University, Prof. Shuangcheng Li from Sichuan Agricultural University, Prof. Shuzhu Tang from Yangzhou University, Prof. Guanghua He from Southwest University, and Prof. Huaxiong Qi from Hubei Academy of Agricultural Sciences for providing materials.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

Data availability

The raw DNA sequencing data of 120 rice inbred lines analyzed in the current study are available in the NCBI Sequence Read Archive under accession numbers SRP080763 and SRP080834.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2024.07.026

References

Adjah K L, Abe A, Adetimirin V O, Asante M D. 2020. Genetic variability, heritability and correlations for milling and grain appearance qualities in some accessions of rice (*Oryza sativa* L.). *Physiology and Molecular Biology of Plants*, **26**, 1309–1317.

Akinwale M G, Gregorio G, Nwilene F, Akinyele B O, Ogunbayo S A, Odiyi A C. 2011. Heritability and correlation coefficient analysis for yield and its components in rice (*Oryza sativa* L.). *African Journal of Plant Science*, **5**, 207–212.

Akohoue F, Koch S, Plieske J, Miedaner T. 2022. Separation of the effects of two reduced height (*Rht*) genes and genomic background to select for less Fusarium head blight of short-strawed winter wheat (*Triticum aestivum* L.) varieties. *Theoretical and Applied Genetics*, **135**, 4303–4326.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the*

Royal Statistical Society (Series B: Methodological), **57**, 289–300.
Bone W P, Siewert K M, Jha A, Klarin D, Damrauer S M, Chang K M, Tsao P S, Assimes T L, Ritchie M D, Voight B F. 2021. Multi-trait association studies discover pleiotropic loci between Alzheimer's disease and cardiometabolic traits. *Alzheimer's Research & Therapy*, **13**, 1–14.
Carlson M O, Montilla-Bascon G, Hoekenga O A, Tinker N A, Poland J, Baseggio M, Sorrells M E, Jannink J L, Gore M A, Yeats T H. 2019. Multivariate genome-wide association analyses reveal the genetic basis of seed fatty acid composition in oat (*Avena sativa* L.). *G3 — Genes*, *Genomes*, *Genetics*, **9**, 2963–2975.
Cui L G, Shan J X, Shi M, Gao J P, Lin H X. 2015. DCA1 acts as a transcriptional co-activator of DST and contributes to drought and salt tolerance in rice. *PLoS Genetics*, **11**, e1005617.
Dormatey R, Sun C, Ali K, Coulter J A, Bi Z, Bai J. 2020. Gene pyramiding for sustainable crop improvement against biotic and abiotic stresses. *Agronomy*, **10**, 1255.
Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q. 2006. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics*, **112**, 1164–1171.
Fan X, Wu J, Chen T, Tie W, Chen H, Zhou F, Lin Y. 2015. Loss-of-function mutation of rice SLAC7 decreases chloroplast stability and induces a photoprotection mechanism in rice. *Journal of Integrative Plant Biology*, **57**, 1063–1077.
He X, Li Z, Guo S, Zheng X, Liu C, Liu Z, Li Y, Yuan Z, Li L. 2024. Epistasis-aware genome-wide association studies provide insights into high-yield and high-quality rice efficient breeding. *Journal of Integrative Agriculture*, **23**, 2541–2556.
Hu P, Tan Y, Wen Y, Fang Y, Wang Y, Wu H, Wang J, Wu K, Chai B, Zhu L, Zhang G, Gao Z, Ren D, Zeng D, Shen L, Xue D, Qian Q, Hu J. 2022. LMPA regulates lesion mimic leaf and panicle development through ROS-induced PCD in rice. *Frontiers in Plant Science*, **13**, 875038.
Huang J, Chen Z, Lin J, Guan B, Chen J, Zhang Z, Chen F, Jiang L, Zheng J, Wang T, Chen H, Xie W, Huang S, Wang H, Huang Y, Huang R. 2022. *gw2.1*, a new allele of *GW2*, improves grain weight and grain yield in rice. *Plant Science*, **325**, 111495.
Huang R, Jiang L, Zheng J, Wang T, Wang H, Huang Y, Hong Z. 2013. Genetic bases of rice grain shape: So many genes, so little known. *Trends in Plant Science*, **18**, 218–226.
Huang Y, Zhao S, Fu Y, Sun H, Ma X, Tan L, Liu F, Sun X, Sun H, Gu P, Xie D, Sun C, Zhu Z. 2018. Variation in the regulatory region of *FZP* causes increases in secondary inflorescence branching and grain yield in rice domestication. *The Plant Journal*, **96**, 716–733.
Jacquemin J, Bhatia D, Singh K, Wing R A. 2013. The international oryza map alignment project: Development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. *Current Opinion in Plant Biology*, **16**, 147–156.
Jaiswal V, Gahlaut V, Meher P K, Mir R R, Jaiswal J P, Rao A R, Balyan H S, Gupta P K. 2016. Genome wide single locus single trait, multi-locus and multi-trait association mapping for some important agronomic traits in common wheat (*T. aestivum* L.). *PLoS ONE*, **11**, e0159343.
Kanwar P, Sanyal S K, Tokas I, Yadav A K, Pandey A, Kapoor S, Pandey G K. 2014. Comprehensive structural, interaction and expression analysis of CBL and CIPK complement during abiotic stresses and development in rice. *Cell Calcium*, **56**, 81–95.
Klei L, Luca D, Devlin B, Roeder K. 2008. Pleiotropy and principal components of heritability combine to increase power for association analysis. *Genetic Epidemiology*, **32**, 9–19.
Korte A, Vilhjálmsson B J, Segura V, Platt A, Long Q, Nordborg M. 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics*, **44**, 1066–1071.
Li F, Xie J, Zhu X, Wang X, Zhao Y, Ma X, Zhang Z, Rashid M A, Zhang Z, Zhi L, Zhang S, Li J, Li Z, 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.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
Li L, Zheng X, Wang J, Zhang X, He X, Xiong L, Song S, Su J, Diao Y, Yuan Z, Zhang Z, Hu Z. 2023. Joint analysis of phenotype-effect-generation identifies loci associated with grain quality traits in rice hybrids. *Nature Communications*, **14**, 3930.
Li R, Li M, Ashraf U, Liu S, Zhang J. 2019. Exploring the relationships between yield and yield-related traits for rice varieties released in China from 1978 to 2017. *Frontiers in Plant Science*, **10**, 430305.
Li Y, Wu S, Huang Y, Ma X, Tan L, Liu F, Lv Q, Zhu Z, Hu M, Fu Y, Zhang K, Gu P, Xie D, Sun H, Sun C. 2023. OsMADS17 simultaneously increases grain number and grain weight in rice. *Nature Communications*, **14**, 3098.
Liu Q, Han R, Wu K, Zhang J, Ye Y, Wang S, Chen J, Pan Y, Li Q, Xu X, Zhou J, Tao D, Wu Y, Fu X. 2018. G-protein βγ subunits determine grain size through interaction with MADS-domain transcription factors in rice. *Nature Communications*, **9**, 852.
Liu X, Huang M, Fan B, Buckler E S, Zhang Z. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genetics*, **12**, e1005767.
Liu Z, Hu Y, Du A, Yu L, Fu X, Wu C, Lu L, Liu Y, Wang S, Huang W, Tu S, Ma X, Li H. 2022. Cell wall matrix polysaccharides contribute to salt–alkali tolerance in rice. *International Journal of Molecular Sciences*, **23**, 15019.
Lv C, Lu W, Quan M, Xiao L, Li L, Zhou J, Li P, Zhang D, Du Q. 2021. Pyramiding superior haplotypes and epistatic alleles to accelerate wood quality and yield improvement in poplar breeding. *Industrial Crops and Products*, **171**, 113891.
McCough S R, Doerge R W. 1995. QTL mapping in rice. *Trends in Genetics*, **11**, 482–487.
Nithya N, Beena R, Stephen R, Abida P, Jayalekshmi V, Viji M, Manju R. 2020. Genetic variability, heritability, correlation coefficient and path analysis of morphophysiological and yield related traits of rice under drought stress. *Chemical Science Review and Letters*, **9**, 48–54.
Oladosu Y, Rafii M, Magaji U, Abdullah N, Miah G, Chukwu S C, Hussin G, Ramli A, Kareem I. 2018. Genotypic and phenotypic relationship among yield components in rice under tropical conditions. *BioMed Research International*, **2018**, 8936767.
Pan Y H, Chen L, Guo H F, Feng R, Lou Q J, Rashid M A R, Zhu X Y, Qing D J, Liang H F, Gao L J, Huang C C, Zhao Y, Deng G F. 2022. Systematic analysis of NB-ARC gene family in rice and functional characterization of GNP12. *Frontiers in Genetics*, **13**, 887217.
Qi J, Li J, Han X, Li R, Wu J, Yu H, Hu L, Xiao Y, Lu J, Lou Y. 2016. Jasmonic acid carboxyl methyltransferase regulates development and herbivory-induced defense response in rice. *Journal of Integrative Plant Biology*, **58**, 564–576.
Qi J, Qian Q, Bu Q, Li S, Chen Q, Sun J, Liang W, Zhou Y, Chu C, Li X, Ren F, Palme K, Zhao B, Chen J, Chen M, Li C. 2008. Mutation of the rice *Narrow leaf1* gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiology*, **147**, 1947–1959.
Qi P, Lin Y S, Song X J, Shen J B, Huang W, Shan J X, Zhu M Z, Jiang L, Gao J P, Lin H X. 2012. The novel quantitative trait locus GL3.1 controls rice grain size and yield by regulating Cyclin-T1;3. *Cell Research*, **22**, 1666–1680.
Rashid M A R, Zhao Y, Azeem F, Zhao Y, Ahmed H G M D, Atif R M, Pan Y, Zhu X, Liang Y, Zhang H, Li D, Zhang Z, Li Z. 2022. Unveiling the genetic architecture for lodging resistance in rice (*Oryza sativa* L.) by genome-wide association analyses. *Frontiers in Genetics*, **13**, 960007.
Ren Y, Chen D, Li W, Tao L, Yuan G, Cao Y, Li X, Deng Q, Wang S, Zheng A, Zhu J, Liu H, Wang L, Li P, Li S. 2021. Genome-wide pedigree analysis of elite rice Shuhui 527 reveals key regions for breeding. *Journal of Integrative Agriculture*, **20**, 35–45.
Roberts A, McMillan L, Wang W, Parker J, Rusyn I, Threadgill D. 2007. Inferring missing genotypes in large SNP panels using fast nearest-neighbor searches over sliding windows. *Bioinformatics*, **23**, i401–i407.
Ruotsalainen S E, Partanen J J, Cichonska A, Lin J, Benner C, Surakka I, Reeve M P, Palta P, Salmi M, Jalkanen S, Ahola-Olli A, Palotie A, Salomaa V, Daly M J, Pirinen M, Ripatti S, Koskela J. 2021. An expanded analysis framework for multivariate GWAS connects inflammatory biomarkers to functional variants and disease. *European Journal of Human Genetics*, **29**, 309–324.
Sakamoto T, Matsuoka M. 2008. Identifying and exploiting grain yield genes in rice. *Current Opinion in Plant Biology*, **11**, 209–214.
Schulthess A W, Reif J C, Ling J, Plieske J, Kollers S, Ebmeyer E, Korzun V, Argillier O, Stiewe G, Ganal M W, Röder M S, Jiang Y. 2017. The roles of pleiotropy and close linkage as revealed by

association mapping of yield and correlated traits of wheat (*Triticum aestivum* L.). *Journal of Experimental Botany*, **68**, 4089–4101.
Shirai Y, Nakanishi Y, Suzuki A, Konaka H, Nishikawa R, Sonehara K, Namba S, Tanaka H, Masuda T, Yaga M, Satoh S, Izumi M, Mizuno Y, Jo T, Maeda Y, Nii T, Oguro-Igashira E, Biobank Japan Project, Morisaki T, *et al.* 2022. Multi-trait and cross-population genome-wide association studies across autoimmune and allergic diseases identify shared and distinct genetic component. *Annals of the Rheumatic Diseases*, **81**, 1301–1312.
Sinha P, Singh V K, Saxena R K, Khan A W, Abbai R, Chitikineni A, Desai A, Molla J, Upadhyaya H D, Kumar A, Varshney R K. 2020. Superior haplotypes for haplotype-based breeding for drought tolerance in pigeonpea (*Cajanus cajan* L.). *Plant Biotechnology Journal*, **18**, 2482–2490.
Van Der Sluis S, Posthuma D, Dolan C V. 2013. TATES: Efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genetics*, **9**, e1003235.
Song X, Meng X, Guo H, Cheng Q, Jing Y, Chen M, Liu G, Wang B, Wang Y, Li J, Yu H. 2022. Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nature Biotechnology*, **40**, 1403–1411.
Sperandio M V L, Santos L A, Tavares O C H, Fernandes M S, De Lima M D F, De Souza S R. 2020. Silencing the *Oryza sativa* plasma membrane H^+-ATPase isoform *OsA2* affects grain yield and shoot growth and decreases nitrogen concentration. *Journal of Plant Physiology*, **251**, 153220.
Su J, Xu K, Li Z, Hu Y, Hu Z, Zheng X, Song S, Tang Z, Li L. 2021. Genome-wide association study and Mendelian randomization analysis provide insights for improving rice yield potential. *Scientific Reports*, **11**, 6894.
Sui Z, Wang T, Li H, Zhang M, Li Y, Xu R, Xing G, Ni Z, Xin M. 2016. Overexpression of peptide-encoding *OsCEP6.1* results in pleiotropic effects on growth in rice (*O. sativa*). *Frontiers in Plant Science*, **7**, 228.
Tu D, Jiang Y, Zhang L, Cai M, Li C, Cao C. 2022. Effect of various combinations of temperature during different phenological periods on *indica* rice yield and quality in the Yangtze River Basin in China. *Journal of Integrative Agriculture*, **21**, 2900–2909.
Wang F, Han T, Song Q, Ye W, Song X, Chu J, Li J, Chen Z J. 2020. The rice circadian clock regulates tiller growth and panicle development through strigolactone signaling and sugar sensing. *Plant Cell*, **32**, 3124–3138.
Wang Q, Tian F, Pan Y, Buckler E S, Zhang Z. 2014. A SUPER powerful method for genome wide association study. *PLoS ONE*, **9**, e107684.
Wu Y P, Chang Y C, Kuo H I, Lin B N, Wang S M, Tseng Y C. 2022. The development of two high-yield and high-quality functional rice cultivars using marker-assisted selection and conventional breeding methods. *International Journal of Molecular Sciences*, **23**, 4678.
Xiao Y, Jiang S, Cheng Q, Wang X, Yan J, Zhang R, Qiao F, Ma C, Luo J, Li W, Liu H, Yang W, Song W, Meng Y, Warburton M L, Zhao J, Wang X, Yan J. 2021. The genetic mechanism of heterosis utilization in maize improvement. *Genome Biology*, **22**, 148.
Xing Y, Zhang Q. 2010. Genetic and molecular bases of rice yield. *Annual Review of Plant Biology*, **61**, 421–442.
Xu H, Toikumo S, Crist R C, Glogowska K, Jinwala Z, Deak J D, Justice A C, Gelernter J, Johnson E C, Kranzler H R, Kember R L. 2023. Identifying genetic loci and phenomic associations of substance use traits: A multi-trait analysis of GWAS (MTAG) study. *Addiction*, **118**, 1942–1952.
Xu P, Li H, Li H, Zhao G, Dai S, Cui X, Liu Z, Shi L, Wang X. 2024. Genome-wide and candidate gene association studies identify *BnPAP17* as conferring the utilization of organic phosphorus in oilseed rape. *Journal of Integrative Agriculture*, **23**, 1134–1149.
Yang J, Lee S H, Goddard M E, Visscher P M. 2011. GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, **88**, 76–82.
Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P C, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K, Matsuoka M. 2016. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nature Genetics*, **48**, 927–934.
Yu J, Pressoir G, Briggs W H, Vroh B I, Yamasaki M, Doebley J F, McMullen M D, Gaut B S, Nielsen D M, Holland J B, Kresovich S, Buckler E S. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, **38**, 203–208.
Zhang L, Yu H, Ma B, Liu G, Wang J, Wang J, Gao R, Li J, Liu J, Xu J, Zhang Y, Li Q, Huang X, Xu J, Li J, Qian Q, Han B, He Z, Li J. 2017. A natural tandem array alleviates epigenetic repression of IPA1 and leads to superior yielding rice. *Nature Communications*, **8**, 14789.
Zhang W, Peng K, Cui F, Wang D, Zhao J, Zhang Y, Yu N, Wang Y, Zeng D, Wang Y, Cheng Z, Zhang K. 2021. Cytokinin oxidase/dehydrogenase OsCKX11 coordinates source and sink relationship in rice by simultaneous regulation of leaf senescence and grain number. *Plant Biotechnology Journal*, **19**, 335–350.
Zhou W, Yan L, Fu Z, Guo H, Zhang W, Liu W, Ye Y, Long P. 2023. Increasing planting density and reducing N application improves yield and grain filling at two sowing dates in double-cropping rice systems. *Plants*, **12**, 2298.
Zhou X, Stephens M. 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, **11**, 407–409.

Executive Editor-in-Chief Xueyong Zhang
Managing Editor Ning Wang