Discover Plants

Research

Superior haplotypes to enhance grain filling in rice (*Oryza sativa* L.) identified through GWAS

V. Jaldhani[1] · K. Suman[1] · D. Sanjeeva Rao[1] · Tony Travis[2] · I. Subhakara Rao[1] · Santosha Rathod[1] · S. R. Voleti[1] · D. Subrahmanyam[1] · P. Raghuveer Rao[1] · Kalyani M. Barbadikar[1] · S. K. Mangrauthia[1] · R. M. Sundaram[1] · Adam H. Price[2] · C. N. Neeraja[1]

Received: 31 July 2024 / Accepted: 9 December 2024
Published online: 27 December 2024

Abstract

Background Key traits for determining rice grain yield include the total number of spikelets per panicle (STOT), the total number of grains per panicle (GTOT), and the percentage of grain filling per panicle (GFTOT). Molecular genetic basis of grain yield is complex and regulates several interconnected genes. Identification of candidate genes and their favourable haplotypes associated with the spikelets, grains and grain filling would accelerate rice grain yield enhancement projects. The interface between genetic and environmental factors often influences the grain filling, affecting overall grain yield. Therefore, we executed a Genome-Wide Association Study (GWAS) to investigate the candidate genes associated with grain-filling traits and to explore genotype × environment (G × E) interactions influencing grain filling traits in rice—STOT, GTOT and GFTOT in the subset of Rice Diversity Panel 1 (RDP1).

Methods and results A set of 188 RDP1 accessions were evaluated across two environments (wet seasons—2021 and 2022) following the standard package of practices to raise healthy crop. Phenotyping analysis identified promising genotypes for three traits of study—STOT and GTOT (Priano Guaira and Tainan-Iku No. 512) and GFTOT (IR8 and Biser-1). GWAS was carried out using 5.2 M SNP dataset in the Parallel Identification of QTL's using EMMAX (PIQUE) pipeline. Five QTLs, one for STOT (*qSTOT6.6*), one for GTOT (*qGTOT4.2*) and three for GFTOT (*qGFTOT3.4, qGFTOT3.5* and *qGFTOT12.1*), were detected. Further, superior haplotypes for eight candidate genes (*LOC_Os6g38850* and *LOC_OS6g39050* for STOT*; LOC_OS4g11040* for GTOT; *LOC_OS3g62720*, *LOC_OS3g62750*, *LOC_OS3g62820, LOC_OS12g03450 and LOC_OS12g03470* for GFTOT) were identified.

Conclusions Our results provide key genetic information for the enhancement of grain filling traits in rice. The identified candidate genes and superior haplotypes can possibly be utilized in marker-assisted selection to improve rice grain yield through gene-pyramid breeding/haplotype-assisted breeding.

Keywords Rice · Grain filling · GWAS · Candidate genes · Haplotypes

Abbreviations

STOT	Spikelets total
GTOT	Grains total

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s44372-024-00069-0.

✉ C. N. Neeraja, cnneeraja@gmail.com | [1]Indian Council of Agricultural Research (ICAR) - Indian Institute of Rice Research (IIRR), Hyderabad, India. [2]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK.

GFTOT Grain filling percentage (%) total
GWAS Genome wide association study
RDP1 Rice diversity panel1
GGE Genotype and Genotype by Environment interaction effects
AMMI Additive main effect and multiplicative interactions
PIQUE Parallel identification of QTL's using EMMAX

1 Introduction

Rice is a staple cereal and feeds over 50% of the world's population, mostly living in Asia. Rice grain production has realized a three-fold increase during the era of green revolution. The ever-growing global population demands further yield improvement for rice. Hence, the prime objective of the rice breeding programs is to improve grain yield within the available input resources [1]. Grain yield is a complex trait influenced by panicle number per plant, spikelet number, grain number and seed setting rate (grain-filling percentage) in the panicle [2]. Grain filling is a critical stage of the rice crop that exerts a significant influence over both grain yield and grain quality attributes [3]. Grain filling begins in multiple branches simultaneously within a rice panicle and grains progress through various developmental stages prior to reaching their characteristic grain weight at physiological maturity [4]. Unfilled or partially-filled spikelets in rice panicles represents poor grain filling which is a major concern to be explored to maximize the yield potential of new plant types, hybrids and super rice [5]. On the genetic basis of grain yield improvement in rice, numerous quantitative trait loci (QTLs)/genes were reported. With the available literature, the genetic competency for enhanced grain filling in rice appears to be possible through the biotechnological interventions such as mapping/association mapping studies [6, 7].

Genome-Wide Association Study (GWAS) is an approach to identify genomic regions and candidate genes for complex traits in rice. This strategy identifies QTLs/genomic regions by examining marker-trait associations, leveraging linkage disequilibrium between various molecular markers and functional polymorphisms in a diverse germplasm pool [7, 8]. Using GWAS, several important agronomic characters including panicle traits were reported in many large sets of rice germplasm panels [8–10]. Based on GWAS, 17 marker trait associations were reported for number of spikelets per panicle [11]. Zhong et al. conducted a GWAS using 421 homozygous rice accessions, revealing 21 QTLs closely linked to total grain number per panicle [12]. Through GWAS and map based cloning, several candidate genes have been identified for yield components rather than specific mechanisms regulating grain filling [13]. However, the utility of reported candidate genes for grain yield traits viz., Gn1a, DEP1, IPA1/WFP, GIF1, GNP1, is limited in rice breeding programs [9]. Owing to the genetic complexity of grain filling traits, most of the identified QTLs found to be minor and explaining < 10% of phenotypic variance [14]. The associated genomic regions with grain filling traits harbour different kinds of candidate genes ranging from metabolic to regulatory nature [8, 15]. Thus, haplotype based selection for pooling/pyramiding of favourable alleles of the candidate genes identified for grain filling traits through GWAS appears to be a promising strategy [16]. The RDP1 represents a collection of 421 rice varieties from 79 countries [17]. This panel was genotyped initially using 44,100 SNPs [18], later with 700,000 SNPs [19] and recently with 5,231,435 SNPs [20]. The RDP1 has been analysed for more than 30 characters pertaining to agronomic, panicle, seed traits and disease reactions [21–24]. Using RDP1 and 5.2 M SNP dataset, a candidate gene *OsMTP8.1* for grain manganese concentration was identified by GWAS [25]. Twenty-three candidate genes related to panicle traits were identified in a set of 406 RDP1 accessions and a novel candidate gene (*Os01g0140100*) associated with total spikelets number per panicle was mapped [12]. Using *japonica* accessions of RDP1 as parents, biparental and recombinant inbred lines (RIL) mapping populations were generated to develop markers for panicle and grain related traits [26]. The primary objective of our study was to identify potential stable genotypes and candidate genes associated with the grain filling traits across environments. Given the scope of the current study, our focus was on identifying the stable candidate genes based on GWAS analysis and haplotype based associations across environments. In the present study, by deploying GWAS of 188 RDP1 accessions with the 5.2 M SNP data, five potential QTLs and eight candidate genes with superior haplotypes associated with three key yield traits of study viz., total spikelets (STOT), total grains (GTOT) and total grain filling% (GFTOT) were identified which can be utilized for breeding towards rice grain yield improvement. In addition, genotype × environment interactions were also studied for identifying the stable genotypes and QTLs for grain filling traits.

2 Materials and methods

2.1 Plant material

A subset of 188 accessions from the Rice Diversity Panel1 (RDP1) comprising *indica* (40), temperate *japonica* (45), tropical *japonica* (44), *aus* (31), aromatic (8) and admixed (20) [17, 18] was assessed for three main components of yield (Supplementary Table 1). Due to low accession numbers from each of six subpopulations (< 50), the subpopulations were not analysed separately. The RDP1 has publicly available 5.2 million SNPs with missing data < 5%, minor allele frequency (MAF) > 1% imputed with other SNP data sets of rice [20].

2.2 Field screening and sample collection

The 188 genotypes were evaluated in the Indian Council of Agricultural Research (ICAR)—Indian Institute of Rice Research (IIRR), Hyderabad, India (17.53°N and 78.27°E) for two consecutive wet seasons (2021 and 2022) as two environments (E1 and E2). The field evaluation followed a randomised complete block design with two replications following standard crop management and protection practices. Each genotype was planted in five rows with a spacing of 20 × 15 cm and five uniform plants in the centre row were tagged. Five panicles were harvested from each tagged plant, so that the grain filling of each genotype is truly represented. Thus, 25 panicles per genotype per one replication and 50 panicles per one environment were collected. The data of five plants from each replicate is considered as technical replicates. The panicles (from each plant) were harvested at physiological maturity and air dried.

2.3 Phenotyping

Three main traits of yield viz., spikelets total (STOT), grains total (GTOT), grain filling percentage total (GFTOT) per panicle were measured. The number of spikelets and grains were counted with a seed counter (Contador 2, PFEUFFER, Germany). The grain filling percentage was calculated based on the ratio of total number of grains (filled spikelets) to the total number of spikelets.

2.4 Statistical analysis

Descriptive statistics was performed using *Statistix 8.1* (Analytical software, 2003). The mean data of panicles for each genotype across two replications for each environment was considered for density plots using R studio packages—*ggplot2* and *reshape2* [27, 28]; *sjlabelled, sjPlot* and *ggcorrplot* [29]; *ggplot2, gridExtra, ggpmisc, ggpubr* [28–30]. Distribution of data for the three traits of study was checked with Shapiro–Wilk and Kolmogorov–Smirnov tests.

The G × E interaction effects were analysed using both the AMMI (Additive Main effects and Multiplicative Interaction) and GGE (Genotype and Genotype by Environment) biplots. The AMMI analysis was performed by first using an ANOVA to partition the total variation into genotype, environment, and genotype-by-environment interaction components. Then, the genotype × environment interaction was decomposed using singular value decomposition (SVD) to extract the interaction principal components (IPCA). The model used for the AMMI analysis was as follows;

$$Y_{ij} = \mu + g_i + e_j + \sum_{n=1}^{N} \gamma_n \beta_{in} \alpha_{ij} + \rho_{ij} + \varepsilon_{ij}$$

where, Y_ij represents the trait mean of the ith genotype in the jth environment, μ is the grand mean, g_i is the ith genotypic effect, e_j is the jth environmental effect, γ_n is eigenvalue of the nth IPCA, β_in is eigenvector for ith genotype for PC n, α_ij is eigenvector for jth environment for PC n, ρ_ij isAMMI residue, ε_ijis error associated with ith genotype in jth environment. Stability was interpreted using the magnitude of the IPC scores, with smaller scores indicating more stable genotypes across environments.

For the GGE biplots, the Sites Regression (SREG) linear-bilinear model was employed and it is depicted as follows;

$$Y_{ij} - \mu_j = \sum_{k=1}^{t} \gamma_k \beta_{in} \alpha_{ij} + e_{ij}$$

The SREG model emphasizes both the genotype and genotype-environment interaction (G + GE) by adjusting for the environment means, allowing clearer interpretation of genotypic performance and stability.

The stability values in GGE biplot analysis was calculated as the square root of the sum of squared scores for the principal components (PCs).

$$Stability\ Values = \sqrt{PC_1^2 + PC_2^2 + \ldots + PC_i^2}$$

PC_i PC_i represents the score of the i^{th} principal component. The more stable genotypes have lower stability values, indicating minimal G × E interaction effects and consistent performance across environments. The biplots were used to generate visualizations like the "mean vs. stability" plot, the "which-won-where" plot, and rankings of genotypes across environments. These methods were implemented using R packages such as *agricolae* [31], *GGE* [32] and *stability* [32–34]. These two analysis were implemented in R software [35]. Further, more details about the AMMI and GGE stability analysis can be found in Gollob [36], Gauch [37], Zobel et al. [38].

2.5 Genome-Wide Association Study (GWAS)

GWAS was performed in 188 RDP1 accessions for the three traits using 5.2 M SNP dataset [20] in the Parallel Identification of QTL's using EMMAX (PIQUE) pipeline (https://github.com/tony- travis/PIQUE). The pipeline used for the GWAS analysis in the present study included the implementation of principal component analysis (PCA), STRUCTURE analysis, kinship matrix, and MLM model by default [25]. GWA mapping was based on mixed model approach and EMMA (Efficient Mixed Model Analysis) as described in Norton et al. [39]. Based on the informative SNP data of RDP1, population structure was already assessed [40] and relatedness (K matrix) among the genotypes was also estimated by pairwise identity-by-state (IBS). From the earlier studies of RDP, the significance threshold for the association between SNP and traits was set at $P < 0.0001$ [39, 41]. The issue of false discovery rate (FDR) was also addressed using the Benjamini–Hochberg method for estimation of P values [42]. The pipeline identifies the SNP to be significantly associated only when the $P < 0.0001$ and a 10% FDR. The mean data of three traits per panicle of each genotype across two environments was considered for GWAS.

2.5.1 QTL identification

Significant SNP and trait associations were considered if $P < 0.0001$ as deployed in the previous study of RDP1 [39, 41] and false discovery rate (FDR) of 10% as represented in chromosome wise Manhattan plots. Using the CLUMP function in PLINK version 1.9, trait wise QTLs have been identified based on peak SNPs with $P < 0.0001$. For QTL identification, peak SNPs with $P < 0.0001$ harbouring ≥ 10 SNPs were only considered in the present study (Minimum 10 or more SNPs with $P < 0.0001$ within 250 kb). These peak SNPs were grouped into same locus, if the QTLs were within average linkage disequilibrium decay value of 250 kb and two peak SNPs with an $r^2 \geq 0.3$ [18, 25, 43]. The genomic positions of the QTLs in the present study were checked for their co-localization with the genomic regions of the reported QTLs for yield related traits in rice.

2.5.2 Analysis of candidate genes at the identified QTLs

LD heat map was constructed based on squared Pearson's correlation coefficient (r^2) using R packages- "*LD heatmap*" and "*genetics*" [44, 45] to narrow down the candidate genes in the identified QTL region [43]. The list of candidate genes with their function in each QTL region consisting of peak SNPs from the LD heat map was extracted using the Rice Genome Annotation Project (RGAP 7.0). Annotated genes with transposon, retrotransposon, unknown and hypothetical function were not included in the analysis.

2.5.3 Haplotype analysis

The genes in the peak SNP regions of QTLs for STOT, GTOT and GFTOT with relatively higher expression during reproductive stage and in reproductive tissues as reported in Rice Expression Profile Database (*RiceXpro*) or Rice Genome Annotation Project (RGAP 7.0) were considered for haplotype analysis. The sequences of target candidate genes were extracted using PLINK version 1.9. The non-synonymous SNPs in the coding region were extracted from the snp-seek.irri.org database for haplotype analysis. Association of haplotype groups and phenotype of STOT, GTOT and GFTOT was

Table 1 Phenotype variation (average) observed for grain filling traits in 188 accessions of RDP1 used in the study

Traits	Range	Average ± SD	CV
Spikelets total (STOT)	66.2–315.6	166.7 ± 37.2	22.3
Grains total (GTOT)	52.2–223.0	116.41 ± 31.1	26.7
Grain filling % total (GFTOT)	37.2–96	69.865 ± 10.0	14.2
Panicle length (PL)	16.3–32.5	23.8 ± 2.0	8.6

SD standard deviation, *CV* coefficient of variation

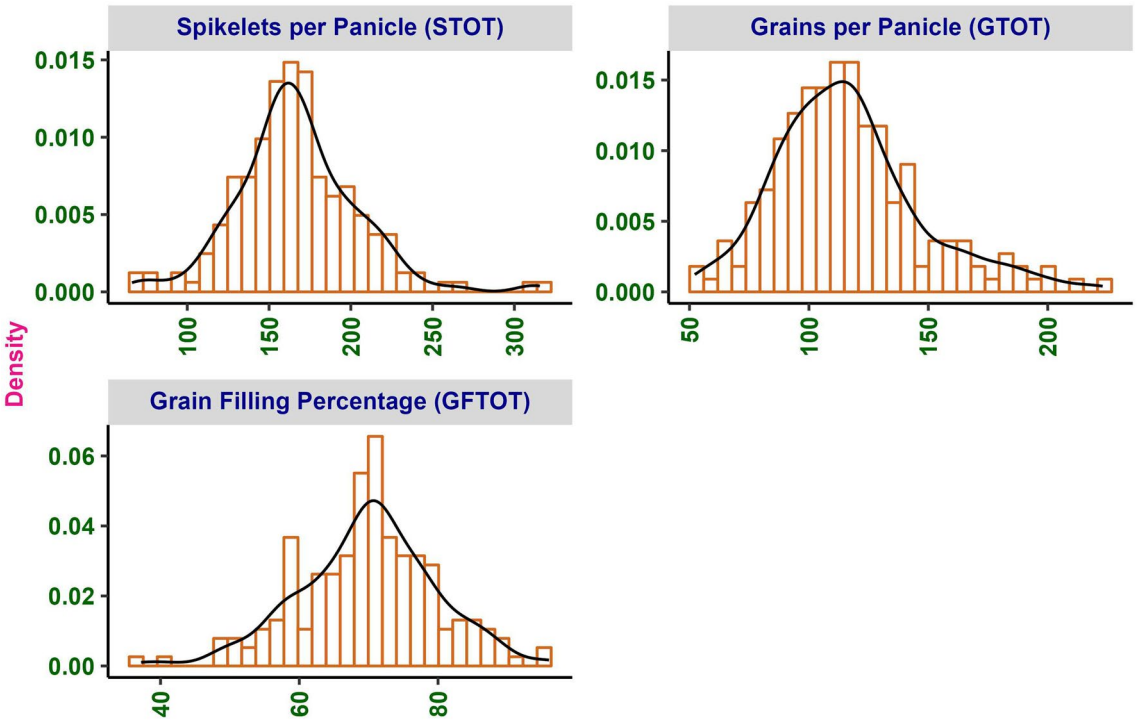


Fig. 1 Density histogram for grain filling traits in RDP1 (with mean of two environments)

analysed through analysis of variance (ANOVA) using Minitab statistical software version 21.4.3 (https://www.minitab.com/). The network analysis of the identified candidate genes were acquired from KnetMiner (https://knetminer.com). The identified candidate genes were analysed for their expression in *RiceXpro* [46] and TIGR [Rice Genome Annotation Project database (http://rice.uga.edu/cgi-bin/gbrowse/rice/#search, accessed on 14 October 2024)], representation in metabolic pathways in Gramene (https://plantreactome.gramene.org/) and co-expression in RiceFREND, a platform for retrieving coexpressed gene networks in rice (https://ricefrend.dna.affrc.go.jp/) [47].

3 Results

3.1 Phenotyping of 188 genotypes from RDP1

Among the 188 accessions of RDP1 panel, based on the trait mean of two environments, STOT ranged from 66.2 (Chiem Chanh) to 315.6 (Tainan-Iku No.512) with a mean of 166.77; GTOT ranged from 52.2 (Chiem Chanh) to 223 (Priano Guaira) with a mean of 116.41 and GFTOT ranged from 37.2% (Nucleoryza) to 96.0% (Biser 1) with a mean of 69.87% (Table 1). The density histograms were represented in Fig. 1. By using Shapiro–Wilk and Kolmogorov–Smirnov tests, GFTOT found to normally distributed, STOT appeared to leptokurtic and GTOT observed to skew to the right. The descriptive statistics are given in Supplementary Table 2.

3.1.1 Stability analysis

Stability analysis of 188 accessions of RDP1 in two environments (E1 and E2) showed significant genotype × environment effects for the three traits of study (Supplementary Fig. 1). For STOT, R43 (DJ24) from the IPCA line was found to be a relatively stable genotype across environments. Genotype R121 (Dom Zard) and R10 (CA902/B/2/1) were found to be the best genotypes in E1, R166 (Desvauxii) and R110 (RTS14) in E2. Based on the average vs. stability, R43 (DJ24) was more stable. As per the Which Won Where/What graph, R23 (Tainan-Iku No.512) and R182 (Priano Guaira) won in E1; R96 (Canella De Ferro) won in E2 for GTOT, R138 from the IPCA line was found relatively stable genotype across the seasons. R93 (Kiang-Chou-Chiu) and R125 (Caucasica) were found to be the best genotypes in E1, R66 (NSFTV5) was found the best genotype in E2. Based on the mean vs. stability, R138 (Vavilovi) was more stable. As per the Which Won Where/What graph, R23 (Tainan-Iku No.512) won in E1; R96 (Canella De Ferro) won in E2; R182 (Priano Guaira) won in E1 and E2. For GFTOT, R175 (Kalubala Vee) from the IPCA line was found to be a relatively stable genotype and R162 (Chinese) farther from the IPCA line was found to be the specific adapter across the seasons. R22 (Leuang Hawn) was the best genotype in E1, R109 (Creole) and R150 (DK 12) were the best genotypes in E2. Based on the mean vs. stability, R175 (Kalubala Vee) and R162 (Chinese) were more stable. As per the Which Won Where/What graph, R4 (IR8) and R173 (CTG1516) won in E1; R4 (IR8) and R131 (Biser 1) won in E1 and E2.

3.1.2 Marker trait associations identified through GWAS

Based on high density 5.2 M SNP data and the phenotype data of 188 accessions, 17 QTLs for the three traits of study have been identified (Supplementary Table 3, Fig. 2). The genomic regions/QTLs were considered only when a minimum of 10 or more peak SNPs with P < *0.0001* were present (Supplementary Table 3). Manhattan plots of STOT, GTOT and GFTOT were represented in Fig. 2. The QTLs—*qSTOT6.6* (23.14–23.39 Mb), *qGTOT4.1* (5.97–6.05 Mb), *qGFTOT3.1* (35.57–35.58 Mb), *qGFTOT3.2* (35.50–35.89 Mb), and *qGFTOT12.1* (1.33–1.58 Mb) were found to be significantly associated. The pairwise LD correlations have highlighted the span of these QTL regions and based on the LD value of 250 kb, the regions were further analysed to locate the candidate genes for haplotype analysis. Based on the genomic regions narrowed down from LD

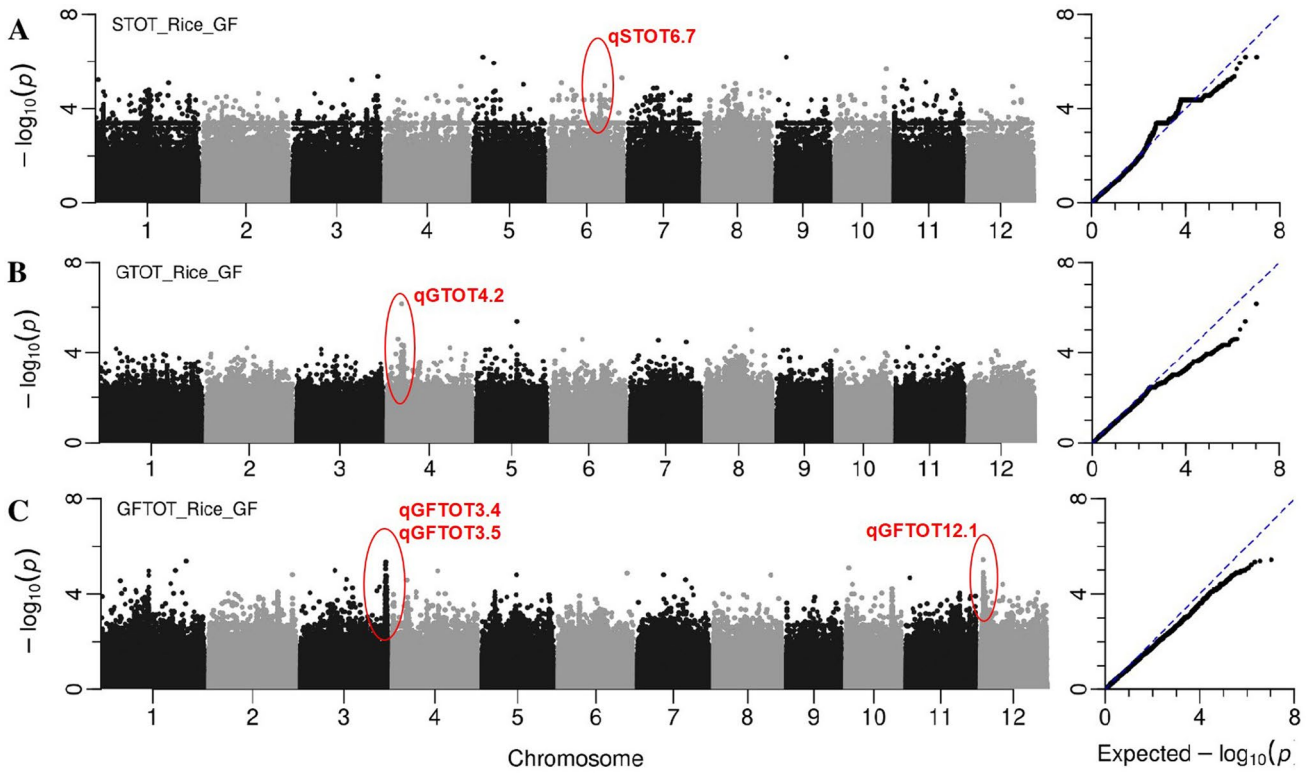


Fig. 2 Genome wide association Manhattan plots for **A** spikelets total (STOT), **B** grains total (GTOT) and **C** grain filling percentage total (GFTOT) with Quantile–quantile plots in RDP1

heat maps, candidate genes were identified from public database of rice (Rice Genome Annotation project –RGAP 7.0). The list of candidate genes in the genomic regions of identified QTLs viz., *qSTOT6.6* (23.14–23.39 Mb)—44 genes; *qGTOT4.1* (5.97–6.05 Mb)—39 genes; *qGFTOT3.1* (35.57–35.58 Mb) and *qGFTOT3.2* (35.50–35.89 Mb)—47 genes and *qGFTOT12.1* (1.33–1.58 Mb)—35 genes are provided in Supplementary Tables 4 and 5.

3.2 Candidate genes

The Spatio-temporal gene expression of the candidate genes in the four LD heat map regions (spanning 250 kb) consisting STOT (chromosome 6), GTOT (chromosome 4) and GFTOT (chromosomes 3 and 12) were explored in *RiceXpro* database. The candidate genes present in the LD heat map region for STOT, GTOT and GFTOT were given in Supplementary Tables 4 and 5. Based on their enhanced expression in the reproductive tissues such as pistil, ovary, anther, endosperm and embryo or during reproductive stage, two candidate genes (*LOC_Os06g38850* and *LOC_Os06g39050)* in *qSTOT6.6*; one candidate gene (*LOC_Os04g11040*) in *qGTOT4.1*; two candidate genes (*LOC_OS3g62720*, *LOC_OS3g62750*) in *qGFTOT3.1*; one candidate gene (*LOC_Os03g62820*) in *qGFTOT3.2* and two candidate genes (*LOC_OS12g03450* and *LOC_OS12g03470*) in *qGFTOT12.1* were considered for haplotype analysis. The Spatio-temporal expression of these genes extracted from RiceXpro and TIGR databases are presented in Supplementary Fig. 3.

3.3 STOT

LOC_Os6g38850 (23,057,309 bp to 23,057,584 bp) encoded an expressed protein with higher expression in ovary/inflorescence (*RiceXpro*) and has shown 18 non- synonymous SNPs. Out of four haplotype groups observed as haplotype A (n = 97, mean: 169.85), haplotype B (n = 26, mean: 168.89), haplotype C (n = 12, mean: 148.57) and haplotype D (n = 8, mean: 194.00), haplotype D was found to be superior for STOT compared to other groups. Another gene, *LOC_Os6g39050* (23,185,057 bp to 23,186,296 bp) which encodes syntaxin, has shown increased expression in the reproductive tissues, especially in embryo (*RiceXpro*). One non-synonymous SNP observed at 23,185,435 bp (C/A polymorphism) with amino acid substitution from Arg(R) to Ser(S) led to two haplotype groups (Fig. 3). Haplotype A (n = 160, mean: 172.53) is observed to be significantly associated with STOT than haplotype B (n = 14, mean: 153.34).

3.4 GTOT

LOC_Os4g11040 (5,989,515 bp to 5,990,234 bp) encoding an expressed protein has shown higher expression in roots at vegetative and reproductive stages with five non- synonymous SNPs. Two haplotypes A (n = 45, mean: 120.95) and B (n = 130, mean: 110.76) were identified with haplotype A being significantly superior for GTOT (Fig. 4).

3.5 GFTOT

LOC_OS03g62720 (35,494,058 bp to 35,497,784 bp) encoding exosome complex exonuclease has shown higher expression in the reproductive tissues, especially in pistil and seed (RGAP database). The gene harboured three non-synonymous SNPs and three haplotype groups were formed (Fig. 5). Haplotype A (n = 81, mean: 71.7%) was found to be a superior group than haplotypes B (n = 19, mean: 63.9%) and C (n = 69, mean: 71.3%). Another candidate gene *LOC_Os03g62750* (35,511,319 bp to 35,518,031 bp) which encodes inner membrane protein with increased expression in the inflorescence tissue (*RiceXpro*) also observed in the same QTL region as *LOC_OS03g62720*. Ten non-synonymous SNPs and three haplotype groups were identified in this gene. No significant difference was observed between haplotypes A (n = 51, mean: 70.9%) and B (n = 59, mean: 71.3%), while haplotype C (n = 15, mean: 63.5%), which was inferior for the value of GFTOT and was significantly different from haplotypes A and B. One more candidate gene *LOC_Os03g62820* in the same region (35,539,015 bp to 35,540,147 bp) encoding expressed protein has 13 non-synonymous SNPs with enhanced expression in anthers, young panicles and endosperm (RGAP database). Three haplotype groups were identified and haplotype C (n = 11, mean: 79.1%) is superior and significantly different from haplotype A (n = 65, mean: 71.2%) and haplotype B (n = 55, mean: 71%).

Two candidate genes, *LOC_OS12g03450* (1,346,249–1,348,258 bp) and *LOC_OS12g03470* (1,359,572–1,366,649 bp) on chromosome 12 with higher expression in reproductive tissues as per *RiceXpro* were identified for GFTOT. Five non-synonymous SNPs in *LOC_OS12g03450* and two haplotype groups as A (n = 140, mean: 71%) and B (n = 33, mean: 67.7%) were found. Haplotype A was significantly different from haplotype B. Four non-synonymous SNPs in *LOC_OS12g03470* led to three

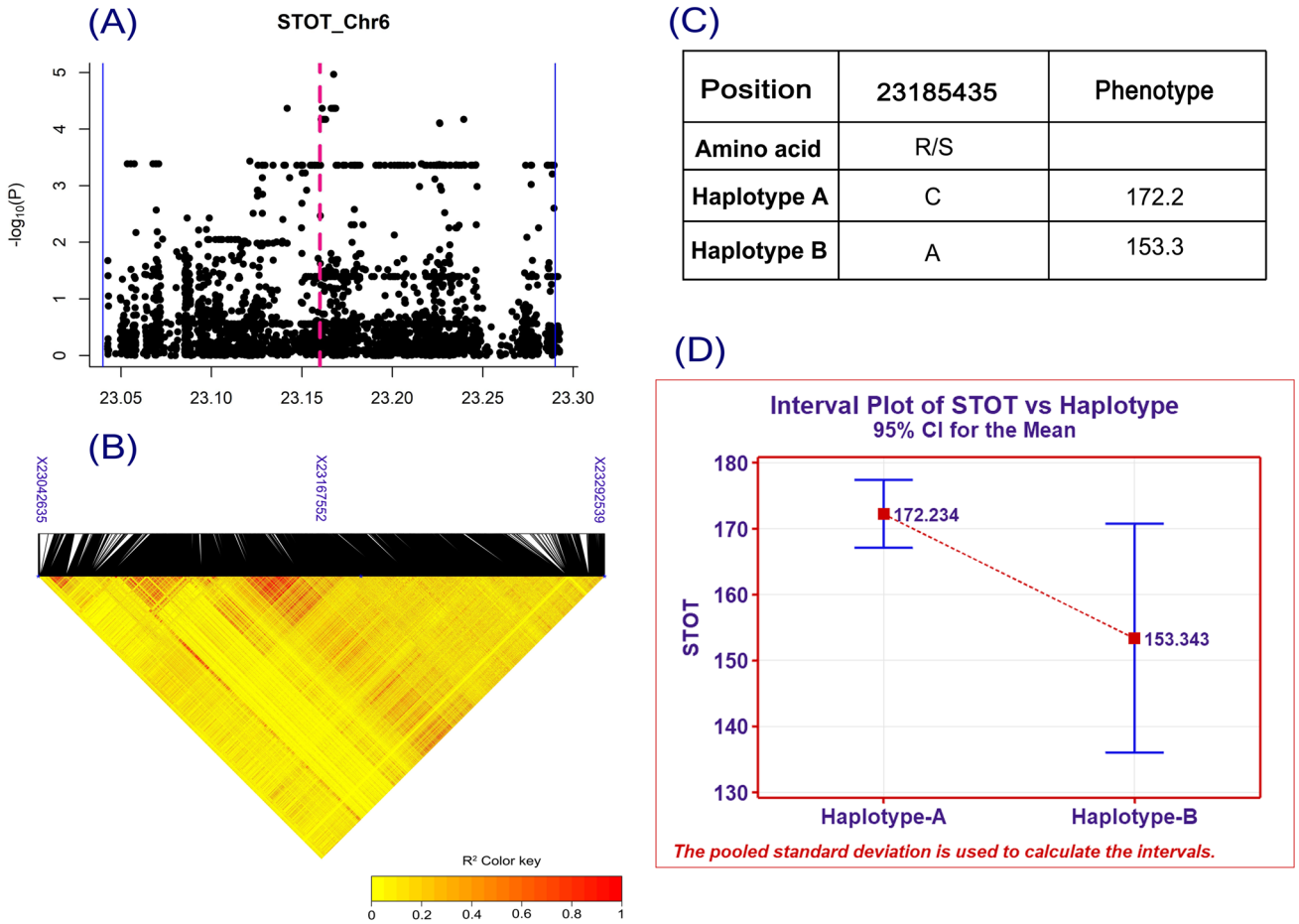
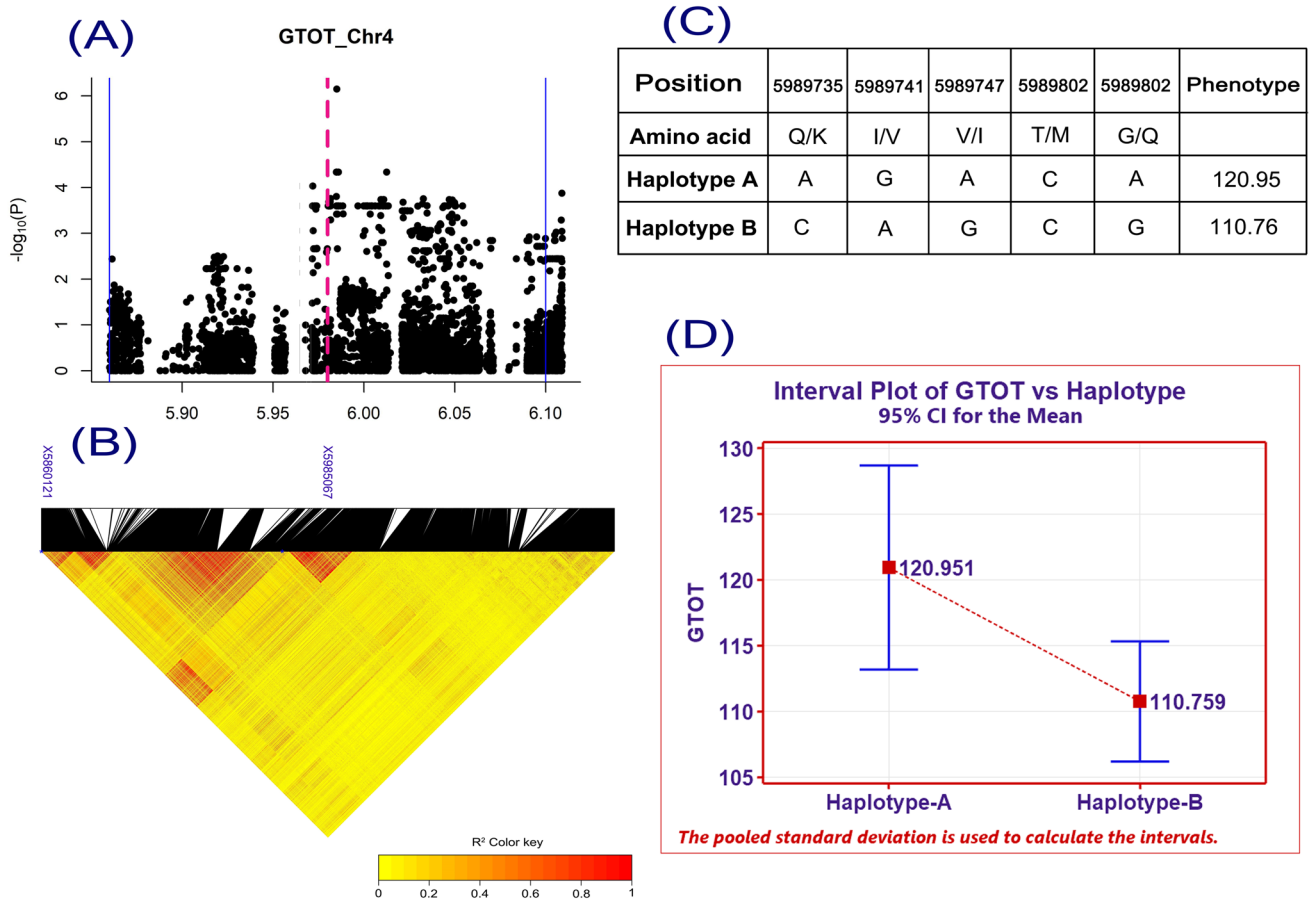


Fig. 3 Significant association of spikelets total (STOT) on chromosome 6 at 23.0–23.3 Mb. **A** Local Manhattan plot; the blue solid line indicates the candidate regions (250 kbp upstream and downstream of the peak SNP) and pink dash line represents the candidate gene (LOC_Os6g39050). **B** LD heat map for STOT underlying the QTL. **C** The non-synonymous SNPs with the amino acid variations in the candidate gene LOC_Os6g39050 significantly associated STOT. **D** Variation in the phenotype identified haplotypes associated with STOT

haplotype groups. Out of which, haplotype A (n = 37, mean: 72.8%) is superior to haplotype B (n = 28, mean: 66.1%) and C (n = 97, mean: 69.7%) (Fig. 6).

3.5.1 Pleotropic analysis

The network analysis by Knetminer revealed the candidate gene *LOC_OS6g39050* for STOT encoding syntaxin and found to be involved in pollen tube growth (Supplementary Fig. 2). Five candidate genes identified for GFTOT, viz., *LOC_OS3g62720* encoding exosome complex exonuclease was linked to days to flowering, seed germination and development; *LOC_OS3g62750* annotated as inner membrane protein was associated with root morphology, days to 50% flowering, leaf shape and heat tolerance. Another candidate gene for GFTOT, *LOC_OS12g03450* encoding an expressed protein, was associated with regulation of translational process, mitochondrial mRNA modifications, zinc ion binding and pentatricopeptide repeats (PPR) containing proteins and *LOC_OS12g03470* was associated with plant height and root morphology traits. Knetminer analysis was not available for candidate genes *LOC_OS4g11040* (GTOT), *LOC_Os6g38850* (STOT) and *LOC_OS3g62820* (GFTOT). Coexpression analysis has shown two of the candidate genes (LOC_Os12g03450 and LOC_Os12g03470) to be associated with carbohydrate metabolism (Supplementary Fig. 4).



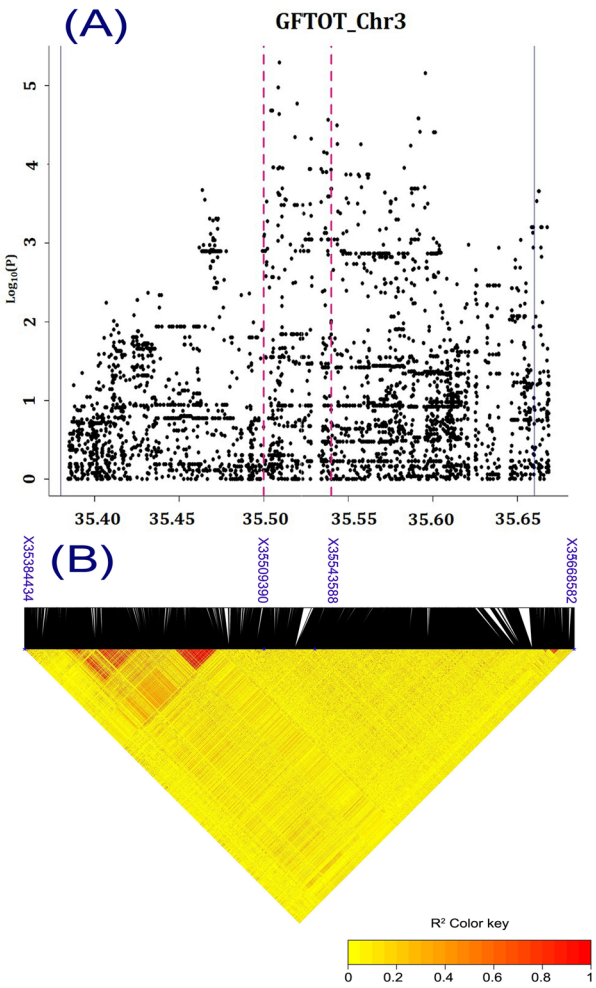
Position	5989735	5989741	5989747	5989802	5989802	Phenotype
Amino acid	Q/K	I/V	V/I	T/M	G/Q	
Haplotype A	A	G	A	C	A	120.95
Haplotype B	C	A	G	C	G	110.76

Fig. 4 Significant association for grains total (GTOT) on chromosome 4 at 5.8–6.1 Mb. **A** Local Manhattan plot; the blue solid line indicates the candidate regions (250 kbp upstream and downstream of the peak SNP) and pink dash line represents the candidate gene (LOC_Os04g11040). **B** LD heat map for GTOT underlying the QTL. **C** The non-synonymous SNPs with the amino acid variations in the candidate gene LOC_Os04g11040 for GTOT. **D** Variation in the phenotype identified haplotypes associated with GTOT

4 Discussion

Grain filling in rice is a critical factor influencing the targeted crop returns. Modern rice cultivars/hybrids holding abundant spikelets per panicle exhibited poor grain filling which limits their yield potential [48]. Wide genetic variation was reported for spikelets, grains and grain filling percentage per panicle in rice germplasm [15, 49–51]. Screening extensive germplasm collection, including landraces, presents a promising approach to identify potential donors for various panicle traits contributing yield. The genetic variation observed in rice germplasm can be utilized in future breeding programs to enhance rice productivity [52]. From the germplasm characterized in the present study, promising genotypes were identified with maximum trait values for the number of spikelets, grains and grain filling percentage indicating the occurrence of promising natural variability in rice and their possible utility in the breeding programs for enhancing the yield. The data distribution of the three traits of the study as evidenced from the graphical representation also indicates the presence of genotypes for higher spikelet number, grain number and grain filling percentage, though in lesser frequency. Similar range for total grain number were also reported from the earlier evaluation of RDP1 accessions [21]. Though genetic variability of STOT, GTOT and GFOT was earlier studied, reports on their stability across environments are limited in rice. Stable donors for grain filling traits can accelerate the breeding program for increasing yield in rice as demonstrated for other traits [53, 54]. Through stability analyses based on the mean value of two environments, we identified stable promising donors' viz., STOT and GTOT (Priano Guaira and Tainan-Iku No. 512) and GFTOT (IR8 and Biser-1) in the present study.

GWAS for STOT, GTOT and GFOT in rice precisely characterizes the associated genomic regions to the level of candidate gene/s or favourable alleles of candidate gene/s in the present study. LD heatmaps were generated to further



(C)

Position	35496399	35497210	35497424	Phenotype
Amino acid	I/M	M/I	R/H	
Haplotype A	A	G	G	71.7
Haplotype B	A	A	G	63.9
Haplotype C	G	G	A	71.3

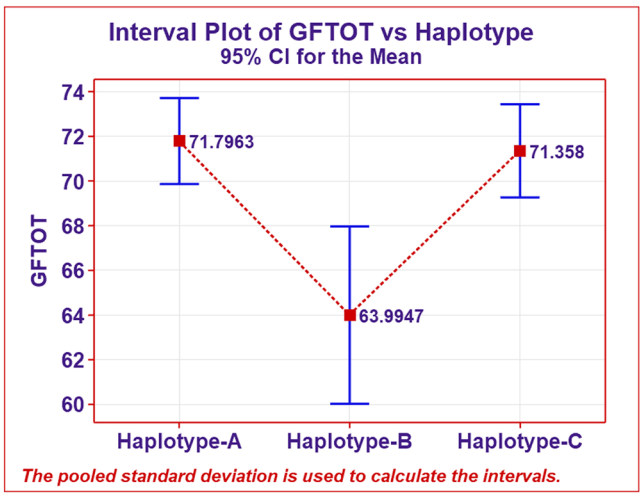
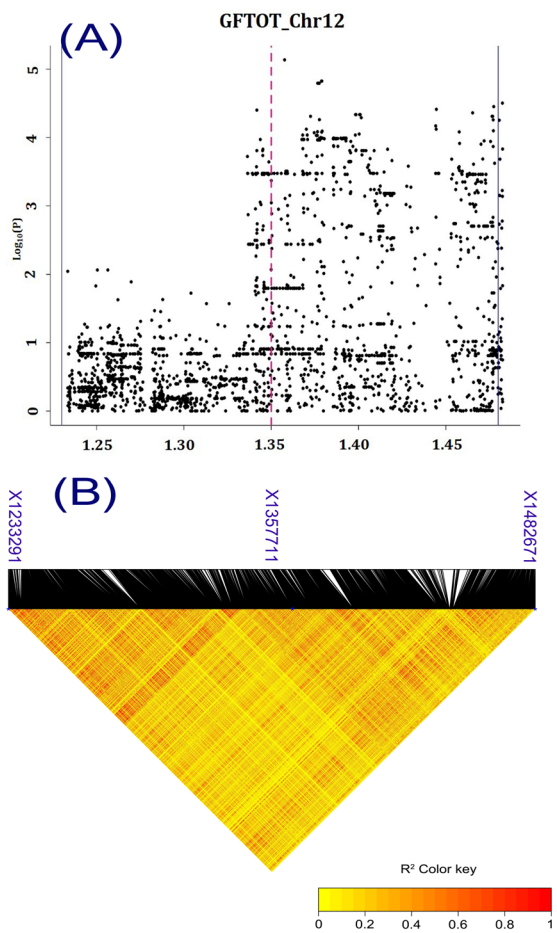


Fig. 5 Significant association for grains filling percentage total (GFTOT) on chromosome 3 at 35.4–36.5 Mb. **A** Local Manhattan plot; the blue solid line indicates the candidate regions (250 kbp upstream and downstream of the peak SNP) and pink dash line represents the candidate gene (LOC_Os03g62750). **B** LD heat map (bottom) for GFTOT underlying the QTL. **C** The non-synonymous SNPs with the amino acid variations in the candidate gene LOC_Os03g62750 for GFTOT. **D** Variation in the phenotype identified haplotypes associated with GFTOT

narrowing down of the genomic regions associated with grain filling traits. From LD heat maps, candidate genes were identified from public database of rice (Rice Genome Annotation Project -RGAP 7.0). The accurate identification of the candidate genes for the trait of interest through GWAS has been successfully demonstrated by previous studies in rice [9, 11, 12, 55–57]. A subset of *indica* panel from 3 K was subjected to GWAS and a locus with large effect association with upper secondary rachis branch (USRB) has been recently identified [15]. The RDP1 was screened for several agronomically important characters such as panicle related, seed and disease tolerance traits and genomic regions have been identified using 700,000 SNP data [21], spikelet number [11] and different panicle architecture traits [57] were reported using 5.2 M SNP data. No common SNP loci were identified in the study compared with Zhong et al. [12], where candidate genomic regions for six panicle related traits (total spikelets number per panicle, grain number per panicle, empty grain number per panicle, primary branch number, panicle length and panicle number per plant) identified in 406 individuals of RDP1 using 411,066 SNPs. This could be due to differences in the evaluated environments and subset of the individual genotypes subjected in GWAS analysis.

From the local Manhattan plots and LD heat map analysis of five genomic regions on four chromosomes**,** eight candidate genes were identified based on their expression in the reproductive tissues or during the reproductive stage from *RiceXpro* database. The association of haplotypes of eight candidate genes with the respective trait phenotype was confirmed through ANOVA in the present study. The identified QTLs in the present study were co-localized with the genomic regions of the reported QTLs for yield and yield-associated traits in rice (Supplementary Table 7). Close to the QTL region for total spikelets (STOT) on chromosome 6 (23.14–23.29 Mb) identified in the present study, a cloned gene



(C)

Position	1346396	1346855	1347129	1347402	Phenotype
Amino acid	R/C	R/C	A/T	A/V	
Haplotype A	C	C	C	G	71.3
Haplotype B	T	T	T	A	67.7

Position	1360516	1360755	1362336	1365682	Phenotype
Amino acid	L/P	A/V	S/P	S/R	
Haplotype A	C	C	C	A	72.8
Haplotype B	C	C	T	A	66.1
Haplotype C	T	C	T	A	69.7

(D)

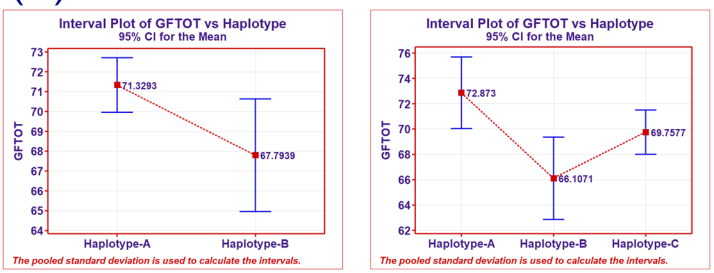


Fig. 6 Significant association for grains filling percentage total (GFTOT) on chromosome 12 at 1.23–1.48 Mb. **A** Local Manhattan plot; the blue solid line indicates the candidate regions (250 kbp upstream and downstream of the peak SNP) and pink dash line represents the candidate genes (LOC_Os12g03450 and LOC_Os12g03470). **B** LD heat map (bottom) for GFTOT underlying the QTL. **C** The non-synonymous SNPs with amino acid variations in the candidate genes—LOC_Os12g03450 (Upper Table) and LOC_Os12g03470 (Lower Table). **D** Variation in the phenotype showed haplotypes associated with GFTOT

SPED1 (*LOC_Os06g39650*) (23.54 Mb) that shortens the pedicels and secondary branches was reported [58]. Also, *MQTL6.6* for grain number per panicle and thousand grain weight, reported by Aloryi et al. [59] (22.29 Mb) is in the vicinity of the same QTL region for STOT. Similarly, for the QTL region for GTOT on chromosome 4 (5.97–6.05 Mb), a genomic region controlling to the grain weight QTL *qGW4.1* (6.57 Mb) [60].

Co-localization of QTLs associated with GFTOT in the regions of chromosome 3.4 (35.5–35.8 Mb) and 3.5 (35.5–35.89 Mb) and chromosome 12.1 (1.33–1.58 Mb) was also observed from several previously reported QTLs/genes. QTL *qSBN3a* at RM148-RM227 (35.83 Mb) reported in two reciprocal introgressive line (IL) populations population, was observed to near to the genomic region for GFTOT on chromosome 3 [61]. Significant SNPs for secondary branch number at 36,088,094 bp region [56] were reported ~ 380 kb away from our QTL region on chromosome 3 for GFTOT. A QTL, *qTGW3c* (35.37–35.77 Mb) on chromosome 3 identified in 56 k SNP array of 3 K genome project by Wang et al. [8] co-localized exactly within the QTL region of GFTOT. Gene *GNP1* that positively regulates the number of secondary rachis number and grain number per panicle on chromosome 3 is observed to be in close proximity to the identified QTL for GFTOT in the present study [62]. This co-localization of QTLs/genes and/or presence of reported QTL in the very close to the identified QTLs in the present study reiterates that these genomic regions can possibly be deployed for improving the grain filling traits in rice. Bioinformatics analyses of the eight identified candidate genes through public databases supported their role in grain filling traits of rice. Six of out of the eight identified candidate genes have known biological functions in rice, especially during reproductive tissues and reproductive associated with carbohydrate metabolism and hormone regulation. These known functions align with the phenotypic traits observed in our study through haplotype analyses, providing evidence that supports their involvement in grain filling. The coexpression analysis and pathway analysis using

Knetminer and RiceFREND underscored the possible role in the metabolic pathways associated grain filling traits. Interestingly, three candidate genes (LOC_Os03g62750, LOC_Os04g11040 and LOC_Os06g39050) have also shown synteny with other poaceae members' viz., sorghum, foxtail millet and maize. The candidate gene—*LOC_Os03g62720* which was found to be associated with GFTOT in the present study was reported to be associated with the degradation of unstable mRNAs containing AU-rich elements (AREs) within their 3′-untranslated regions and suggested to be crucial in the regulation for protein synthesis under high temperature stress in rice [63]. Another candidate gene—*LOC_Os03g62750* was also associated with GFTOT encodes albino like (OsALB3) proteins needed for the development of the chloroplast by accumulating of light harvesting chlorophyll binding proteins in rice [64]. Being complex trait, a very few candidate genes for grain filling traits have been functionally validated in rice despite the reports of hundreds of QTLs. The reported candidate genes for grain filling traits range from metabolic genes (e.g. carbohydrate metabolism) to regulatory genes (e.g. transcription factors). Thus, the identified new candidate genes for grain filling in the present study can probably be added to the repository of the rice yield genes and can be validated further through functional characterization to elucidate their role in molecular mechanisms of grain filling.

The superior haplotypes identified across two environments were also found to be stable when checked with individual environment. The identified superior haplotypes for grain filling traits are being tested for their performance in multi-location studies under All India Coordinated Research Project—Rice (AICRP-R) and the stable haplotype/s will be used in the rice breeding programs.

5 Conclusions

In the present study, a set of 188 Rice diversity Panel (RDP1) accessions were evaluated to identify the promising genotypes and marker trait associations for three grain filling traits (STOT, GTOT and GFTOT). Promising genotypes identified for STOT, GTOT viz., Priano Guaira, Tainan-Iku No.512 and IR8, Biser1 for GFTOT will be used in the breeding programs for yield enhancement. Stability analysis, from which-won-where view of GGE biplot, also revealed the above-mentioned four genotypes have won in different environments for the three studied traits. GWAS identified five QTLs on chromosomes 3, 4, 6 and 12 for STOT, GTOT and GFTOT. Bioinformatics analyses of the eight identified candidate genes through public databases supported their role in grain filling. Superior haplotypes can possibly be deployed in the haplotype assisted breeding for grain filling traits.

Acknowledgements The authors acknowledge the Director-Indian Institute of Rice Research (IIRR) support for conducting the experiment. We gratefully acknowledge Mrs Vani Badri for the extensive support in phenotyping of the grain filling traits. We also acknowledge South Asian Nitrogen Hub (SANH)-WP2.1b training for GWAS analyses. We would like to express our deepest gratitude to Dr Malathi for her contributions to this manuscript. Her dedication and expertise significantly advanced our work. We honour her memory with this manuscript.

Author contributions CNN conceptualized and designed the study. SRI, KS and VJ conducted the field experiments. SR, VJ and KS analysed the phenotype data. DS, SRD and VJ analysed the GWAS and clump pipeline. TT and AHP developed GWAS pipeline, revision of manuscript. CNN prepared and edited the manuscript. All authors commented on previous versions of the manuscript. All authors read and approved the submitted version of the manuscript.

Funding The authors acknowledge the Institute Rice Council, Indian Council of Agricultural Research (ICAR)-Indian Institute of Rice Research (IIRR) and ICAR-CRP "Biofortification in selected crops for nutritional security" for the financial support.

Data availability All data supporting the findings of this study are available within the manuscript and supplementary information. It is also available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate All the experiments carried out on plants were carried out in accordance with the guidelines of ICAR – Indian Institute of Rice Research.

Competing interests The authors declare no competing interests.

References

1. Mohapatra PK, Sahu BB. Diversity of panicle architecture and traits influencing grain filling. In: Mohapatra PK, Sahu BB, editors. Panicle Architecture of Rice and Its Relationship with Grain Filling. Cham, Switzerland: Springer International Publishing; 2022. p. 107–28.
2. Liu Z, Sun H, Zhang Y, et al. Mining the candidate genes of rice panicle traits via a genome-wide association study. Front Genet. 2023;14:1239550. https://doi.org/10.3389/FGENE.2023.1239550/BIBTEX.
3. Kato T, Morita R, Ootsuka S, et al. Evaluation of alleles at OsAGPS2, OsAGPL2, and OsSUT1 related to grain filling in rice in a common genetic background. Crop Sci. 2021. https://doi.org/10.1002/csc2.20429.
4. Durbha SR, Siromani N, Jaldhani V, et al. Dynamics of starch formation and gene expression during grain filling and its possible influence on grain quality. Sci Rep. 2024;141(14):1–20. https://doi.org/10.1038/s41598-024-57010-4.
5. Kato T. An approach to the "grain-filling problem" in rice through the improvement of its sink strength. J Crop Res. 2020;65:1–11. https://doi.org/10.18964/JCR.65.0_1.
6. Parida AK, Sekhar S, Panda BB, et al. Effect of panicle morphology on grain filling and rice yield: genetic control and molecular regulation. Front Genet. 2022;13:876198. https://doi.org/10.3389/FGENE.2022.876198/BIBTEX.
7. Zia MAB, Yousaf MF, Asim A, Naeem M. An overview of genome-wide association mapping studies in Poaceae species (model crops: wheat and rice). Mol Biol Rep. 2022;49:12077–90. https://doi.org/10.1007/S11033-022-08036-2/TABLES/3.
8. Wang Y, Wang J, Zhai L, et al. Identify QTLs and candidate genes underlying source-, sink-, and grain yield-related traits in rice by integrated analysis of bi-parental and natural populations. PLoS ONE. 2020. https://doi.org/10.1371/JOURNAL.PONE.0237774.
9. Ta KN, Khong NG, Ha TL, et al. A genome-wide association study using a Vietnamese landrace panel of rice (Oryza sativa) reveals new QTLs controlling panicle morphological traits. BMC Plant Biol. 2018;18:1–15. https://doi.org/10.1186/S12870-018-1504-1/FIGURES/6.
10. Crowell S, Korniliev P, Falcão A, et al. Genome-wide association and high-resolution phenotyping link Oryza sativa panicle traits to numerous trait-specific QTL clusters. Nat Commun. 2016;71(7):1–14. https://doi.org/10.1038/ncomms10527.
11. Rebolledo MC, Peña AL, Duitama J, et al. Combining image analysis, genome wide association studies and different field trials to reveal stable genetic regions related to panicle architecture and the number of spikelets per panicle in rice. Front Plant Sci. 2016;7:207616. https://doi.org/10.3389/FPLS.2016.01384/BIBTEX.
12. Zhong H, Liu S, Meng X, et al. Uncovering the genetic mechanisms regulating panicle architecture in rice with GPWAS and GWAS. BMC Genomics. 2021;22:1–13. https://doi.org/10.1186/S12864-021-07391-X/FIGURES/4.
13. Ashikari M, Sakakibara H, Lin S, et al. Cytokinin oxidase regulates rice grain production. Science. 2005;309:741–5. https://doi.org/10.1126/SCIENCE.1113373.
14. Xu X, Zhang M, Xu Q, et al. Quantitative trait loci identification and genetic diversity analysis of panicle structure and grain shape in rice. Plant Growth Regul. 2020;90:89–100. https://doi.org/10.1007/S10725-019-00549-4/TABLES/4.
15. Pasion EA, Badoni S, Misra G, et al. OsTPR boosts the superior grains through increase in upper secondary rachis branches without incurring a grain quality penalty. Plant Biotechnol J. 2021;19:1396–411. https://doi.org/10.1111/PBI.13560.
16. Sachdeva S, Singh R, Maurya A, et al. Multi-model genome-wide association studies for appearance quality in rice. Front Plant Sci. 2023;14:1304388. https://doi.org/10.3389/FPLS.2023.1304388/BIBTEX.
17. Eizenga GC, Ali ML, Bryant RJ, et al. Registration of the rice diversity panel 1 for genomewide association studies. J Plant Regist. 2014;8:109–16. https://doi.org/10.3198/jpr2013.03.0013crmp.
18. Zhao K, Tung CW, Eizenga GC, et al. Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat Commun. 2011. https://doi.org/10.1038/NCOMMS1467.
19. McCouch SR, Wright MH, Tung CW, et al. Open access resources for genome-wide association mapping in rice. Nat Commun. 2016;71(7):1–14. https://doi.org/10.1038/ncomms10532.
20. Wang DR, Agosto-Pérez FJ, Chebotarov D, et al. An imputation platform to enhance integration of rice genetic resources. Nat Commun. 2018. https://doi.org/10.1038/S41467-018-05538-1.
21. Eizenga GC, Jia MH, Jackson AK, et al. Validation of yield component traits identified by genome-wide association mapping in a tropical japonica × tropical japonica rice biparental mapping population. Plant Genome. 2019. https://doi.org/10.3835/PLANTGENOME2018.04.0021.
22. Wang D, Liu J, Li C, et al. Genome-wide association mapping of cold tolerance genes at the seedling stage in rice. Rice. 2016;9:1–10. https://doi.org/10.1186/S12284-016-0133-2/FIGURES/4.
23. Yang B, Chen M, Zhan C, et al. Identification of OsPK5 involved in rice glycolytic metabolism and GA/ABA balance for improving seed germination via genome-wide association study. J Exp Bot. 2022;73:3446–61. https://doi.org/10.1093/JXB/ERAC071.
24. Davidson H, Shrestha R, Cornulier T, et al. Spatial effects and GWA mapping of root colonization assessed in the interaction between the rice diversity panel 1 and an arbuscular mycorrhizal fungus. Front Plant Sci. 2019;10:451825. https://doi.org/10.3389/FPLS.2019.00633/BIBTEX.
25. Ruang-areerate P, Travis AJ, Pinson SRM, et al. Genome-wide association mapping for grain manganese in rice (Oryza sativa L.) using a multi-experiment approach. Heredity. 2021;126:505. https://doi.org/10.1038/S41437-020-00390-W.
26. Eizenga GC, Jackson AK, Edwards JD. Prototype for developing SNP markers from GWAS and biparental QTL for rice panicle and grain traits. Agric Environ Lett. 2021;6:e20047. https://doi.org/10.1002/AEL2.20047.
27. Wickham H. Reshaping data with the reshape package. J Stat Softw. 2007;21:1–20. https://doi.org/10.18637/JSS.V021.I12.
28. Wickham H. ggplot2: Elegant graphics for data analysis. Springer-Verlag. 2016. https://ggplot2.tidyverse.org.
29. Kassambara A. "ggplot2" Based Publication Ready Plots [R package ggpubr version 0.6.0]. 2023. https://doi.org/10.32614/CRAN.PACKAGE.GGPUBR.
30. Auguie B. Miscellaneous Functions for "Grid" Graphics [R package gridExtra version 2.3]. 2017. https://doi.org/10.32614/CRAN.PACKAGE.GRIDEXTRA.

31. Mendiburu FD. agricolae: Statistical procedures for agricultural research (Version 1.3–7) [R package]. CRAN. 2023. https://CRAN.R-project.org/package=agricolae.
32. Dumble S. GGE Biplots with "ggplot2" [R package GGEBiplots version 0.1.3]. 2022. https://doi.org/10.32614/CRAN.PACKAGE.GGEBIPLOTS.
33. Wright K, Laffont JL. GGE: Genotype Plus Genotype-by-Environment Biplots. R Packag version 14. 2018.
34. Yaseen M, Eskridge KM, Murtaza G. stability: Stability Analysis of Genotype by Environment Interaction (GEI). R Packag version 050. 2018
35. R Core Team. R: A Language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. 2018.
36. Gollob HF. A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika. 1968;33:73–115. https://doi.org/10.1007/BF02289676.
37. Gauch Jr HG. Model selection and validation for yield trials with interaction. Biometrics. 1988;44(3):705–15. https://doi.org/10.2307/2531585.
38. Zobel RW, Wright MJ, Gauch HG. Statistical analysis of a yield trial. Agron J. 1988;80:388–93. https://doi.org/10.2134/agronj1988.00021962008000030002x.
39. Norton GJ, Douglas A, Lahner B, et al. Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (Oryza sativa L.) grown at four international field sites. PLoS ONE. 2014. https://doi.org/10.1371/JOURNAL.PONE.0089685.
40. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nat Rev Genet. 2010;11:459–63. https://doi.org/10.1038/NRG2813.
41. Famoso AN, Zhao K, Clark RT, et al. Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping. PLoS Genet. 2011. https://doi.org/10.1371/JOURNAL.PGEN.1002221.
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995;57:289–300.
43. Anandan A, Panda S, Sabarinathan S, et al. Superior haplotypes for early root vigor traits in rice under dry direct seeded low nitrogen condition through genome wide association mapping. Front Plant Sci. 2022;13:911775. https://doi.org/10.3389/FPLS.2022.911775/BIBTEX.
44. Shin JH, Blay S, McNeney B, Graham J. LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. J Stat Softw. 2006;16:1–9. https://doi.org/10.18637/JSS.V016.C03.
45. Warnes G, Gorjanc G, Leisch F et al Genetics: population genetics. cran.uvigo.es
46. Sato Y, Takehisa H, Kamatsuki K, et al. RiceXPro Version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids Res. 2013;41:D1206–13. https://doi.org/10.1093/NAR/GKS1125.
47. Sato Y, Namiki N, Takehisa H, et al. RiceFREND: a platform for retrieving coexpressed gene networks in rice. Nucleic Acids Res. 2013;41:D1214–21. https://doi.org/10.1093/NAR/GKS1122.
48. Yang J, Zhang J. Grain-filling problem in 'super' rice. J Exp Bot. 2010;61:1–5. https://doi.org/10.1093/JXB/ERP348.
49. Kato T. Variation and association of the traits related to grain filling in several extra-heavy panicle type rice under different environments. Plant Prod Sci. 2010;13:185–92. https://doi.org/10.1626/PPS.13.185.
50. Panda D, Sahu N, Behera PK, Lenka K. Genetic variability of panicle architecture in indigenous rice landraces of Koraput region of Eastern Ghats of India for crop improvement. Physiol Mol Biol Plants. 2020;26:1961. https://doi.org/10.1007/S12298-020-00871-6.
51. Rao IS, Srikanth B, Kishore VH, et al. Indel polymorphism in sugar translocation and transport genes associated with grain filling of rice (Oryza sativa L.). Mol Breed. 2011;28:683–91. https://doi.org/10.1007/s11032-011-9618-2.
52. Vishwakarma C, Karwa S, Kapoor RT, et al. Exploring component traits of biomass and yield in Indian rice mini-core germplasm. Plant Physiol Reports. 2024;29:226–38. https://doi.org/10.1007/S40502-024-00791-7/FIGURES/6.
53. Suman K, Neeraja CN, Madhubabu P, et al. Identification of promising RILs for high grain zinc through genotype × environment analysis and stable grain zinc QTL using SSRs and SNPs in rice (Oryza sativa L.). Front Plant Sci. 2021. https://doi.org/10.3389/fpls.2021.587482.
54. Babu PM, Neeraja CN, Rathod S, et al. Stable SNP allele associations with high grain zinc content in polished rice (Oryza sativa L.) identified based on ddRAD sequencing. Front Genet. 2020;11:763. https://doi.org/10.3389/FGENE.2020.00763/BIBTEX.
55. Bai X, Zhao H, Huang Y, et al. Genome-wide association analysis reveals different genetic control in panicle architecture between and rice. Plant Genome. 2016. https://doi.org/10.3835/PLANTGENOME2015.11.0115.
56. Bai S, Hong J, Li L, et al. Dissection of the genetic basis of rice panicle architecture using a genome-wide association study. Rice. 2021;14:1–12. https://doi.org/10.1186/S12284-021-00520-W/FIGURES/5.
57. Thapa R, Tabien RE, Septiningsih EM. Genome-wide association study to identify chromosomal regions related to panicle architecture in rice (Oryza sativa). Genet Resour Crop Evol. 2021;68:2849–65. https://doi.org/10.1007/S10722-021-01159-8.
58. Jiang G, Xiang Y, Zhao J, et al. Regulation of inflorescence branch development in rice through a novel pathway involving the pentatricopeptide repeat protein sped1-D. Genetics. 2014;197:1395–407. https://doi.org/10.1534/GENETICS.114.163931/-/DC1/FILES2.XLS.
59. Aloryi KD, Okpala NE, Amo A, et al. A meta-quantitative trait loci analysis identified consensus genomic regions and candidate genes associated with grain yield in rice. Front Plant Sci. 2022;13:1035851. https://doi.org/10.3389/FPLS.2022.1035851/BIBTEX.
60. Mallikarjuna Swamy BP, Kaladhar K, Reddy G, et al. Mapping and introgression of QTL for yield and related traits in two backcross populations derived from Oryza sativa cv. Swarna and two accessions of O. nivara. J Genet. 2014;93:643–54. https://doi.org/10.1007/S12041-014-0420-X.
61. Mei HW, Xu JL, Li ZK, et al. QTLs influencing panicle size detected in two reciprocal introgressive line (IL) populations in rice (Oryza sativa L.). Theor Appl Genet. 2006;112:648–56. https://doi.org/10.1007/S00122-005-0167-0.
62. Yin C, Zhu Y, Li X, Lin Y. Molecular and genetic aspects of grain number determination in rice (Oryza sativa L.). Int J Mol Sci. 2021;22:1–19. https://doi.org/10.3390/IJMS22020728.
63. Onaga G, Wydra K, Koopmann B, et al. High temperature effects on Pi54 conferred resistance to Magnaporthe oryzae in two genetic backgrounds of Oryza sativa. J Plant Physiol. 2017;212:80–93. https://doi.org/10.1016/J.JPLPH.2017.02.004.
64. Zhang C, Mao X, Feng X, et al. OsALB3 is required for chloroplast development by promoting the accumulation of light-harvesting chlorophyll-binding proteins in rice. Plants. 2023;12:4003. https://doi.org/10.3390/PLANTS12234003/S1.