Euphytica (2024) 220:154
https://doi.org/10.1007/s10681-024-03410-6

RESEARCH

Refining the major-effect QTL and candidate genes associated with grain number per panicle by QTL-seq in rice (*Oryza sativa* L.)

Gunasekaran Ariharasutharsan · Adhimoolam Karthikeyan · Seshadri Geetha · Ramasamy Saraswathi · Muthurajan Raveendran · Karuppasamy Krishna-Surendar · Latha-Devi Ananda-Lekshmi · Amudha Kailappan · Ramalingam Suresh · Natarajan Devasena

Received: 6 July 2024 / Accepted: 2 September 2024

Abstract Rice grain yield is a major focus of rice breeding, and with grain number per panicle being a major trait that largely determines overall grain yield. Despite its importance, the genetic architecture and underlying mechanisms governing grain number per panicle are not well understood. In this study, we adopted a whole-genome resequencing-based QTL-seq analysis to trace genomic regions related with grain number per panicle using a mapping population derived from a cross between CB12132 (High grain number) and IET28835 (Low grain number). This approach revealed five candidate genomic regions: *qGNPP1.1* (10.40 Mb to 12.76 Mb), *qGNPP1.2* (24.61 Mb to 25.33 Mb), *qGNPP1.3* (26.57 Mb to 27.26 Mb), *qGNPP4*.1 (27.70 Mb to 31.34 Mb), and *qGNPP5.1* (2.12 Mb to 5.50 Mb) on chromosomes 1, 4, and 5, respectively. Further, we searched for possible candidate genes using a

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10681-024-03410-6.

G. Ariharasutharsan · S. Geetha (✉) · R. Saraswathi (✉) · K. Krishna-Surendar · L.-D. Ananda-Lekshmi · A. Kailappan · R. Suresh · N. Devasena
Department of Rice, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641003, Tamil Nadu, India
e-mail: geethagovind1@gmail.com

R. Saraswathi
e-mail: sarasrice2004@yahoo.co.in

A. Karthikeyan
Subtropical Horticulture Research Institute, Jeju National University, Jeju, South Korea

S. Geetha
Department of Pulses, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641003, Tamil Nadu, India

R. Saraswathi
Department of Plant Genetic Resources, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641003, Tamil Nadu, India

M. Raveendran
Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore 641003, Tamil Nadu, India

M. Raveendran
Directorate of Research, Tamil Nadu Agricultural University, Coimbatore 641003, Tamil Nadu, India

A. Kailappan
Agricultural Research Station, Tamil Nadu Agricultural University, Bhavanisagar 638451, Tamil Nadu, India

N. Devasena
Division of Genetics and Plant Breeding, School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Coimbatore 641114, Tamil Nadu, India

Published online: 12 September 2024

comprehensive approach that included the analysis of gene sequences, functional annotation, and expression patterns. A total of 23 candidate genes, including most possible genes *Os01g0292900 (SPL1)*, *Os01g0622000 (OsCUGT1)*, *Os01g0655300 (SDG705)*, *Os04g0615000 (NAL1)*, *Os04g0559800 (SMG2)* and *Os05g0155200 (ERS2)*, were identified across the five candidate genomic regions. Collectively, our study results shed light on the genetic mechanisms underlying grain number per panicle in rice and will be helpful for improving grain yield in future rice breeding programs.

Keywords Candidate genes · Grain number per panicle · QTL seq · Whole genome resequencing

Introduction

Rice is a major food crop globally, and its consumption is projected to rise significantly by 2025, from 503.5 million tonnes to 800–900 million tonnes (Resilience 2017). The introduction of hybrid rice and semi-dwarf cultivars has already led to a substantial increase in rice production. However, despite reaching desired production levels in recent decades, rice yield growth has plateaued (Xu et al. 2015; Zhu et al. 2017). Therefore, it is necessary to intensify efforts to boost the yield potential of rice as the world's rapid population growth continues to pose a serious threat to food security. One of the key objectives of rice breeding remains the development of high-yielding improved cultivars. Grain yield is a complex quantitative trait, influenced by three major traits—panicle number, grain number per panicle, and grain weight—as well as genotype-environment (GE) interactions (Oladosu et al. 2017; Wang et al. 2020). Grain number per panicle is a major trait linked to grain yield, and understanding its genetic basis is useful for the development of higher-yielding rice cultivars. The genetic basis of grain number per panicle has been extensively studied in rice, and several major quantitative trait loci (QTL) or genes associated with grain numbers per panicle have been identified, including *GN1a* (Ashikari et al. 2005), *NOG1* (Huo et al. 2017), *LOG* (Kurakawa et al. 2007), *LP/EP3* (Li et al. 2011a), *GNP1* (Wu et al. 2016), *qGN4-1* (Singh et al. 2018), *GNP 4* (Zhang et al. 2011), *NAL1* (Fujita et al. 2013), *GSN1* (Guo et al. 2018), *APO 1* (Ikeda-Kawakatsu et al. 2009), *GHD 7* (Xue et al. 2008), *PAY1* (Zhao et al. 2015), *OsSPL14* (Miura et al. 2010), *DEP 1* (Huang et al. 2009), *TAW1* (Yoshida et al. 2013), *SP 1* (Li et al. 2009) and *RCN 1* (Nakagawa et al. 2002) in across all 12 chromosomes.

Despite extensive research, the mechanism underlying grain number trait formation remains unclear. Traditional QTL mapping studies require a large breeding population and numerous markers, which can be time-consuming and labor-intensive (Wang et al. 2019; Weng et al. 2021). In contrast, bulked segregant analysis (BSA) (Michelmore et al. 1991) offers a simple and cost-efficient approach to rapidly identify the polymorphic markers associated with traits of interest. The availability of the whole-genome sequence of rice, combined with advances in second and third-generation sequencing, enables the introduction of innovative genomics-driven breeding strategies (Yano et al. 2016). Furthermore, QTL-seq has emerged as a useful method for detecting QTL when it is integrated with traditional BSA and sequencing technological advances (Takagi et al. 2013). Numerous studies have demonstrated that QTL-seq, which involves tracing QTL from whole-genome resequencing of two DNA bulks of progeny with extreme phenotypes, offers a faster and more cost-effective approach compared to traditional QTL mapping (Jia et al. 2023; Pujol et al. 2019b; Vogel et al. 2021; Yang et al. 2021; Yuan et al. 2015). QTL-seq has been extensively used for the detection of QTL for many traits, including grain elongation (Arikit et al. 2019), plant height (Zhang et al. 2021a), panicle grain number (Ma et al. 2022), brown plant hopper resistance (Wang et al. 2022), and salt tolerance (Gao et al. 2023) in rice.

In our previous study, we investigated the genetic variation of grain number per panicle and its related traits using a diverse set of rice genotypes and identified high and low grain number per panicle genotypes of CB12132 (423.00 ± 12.56) and IET 28835 (108.00 ± 11.45) (Gunasekaran et al. 2023). To further elucidate the genetic and molecular basis of grain number per panicle, this study employed QTL-seq, a combination of BSA and whole genome resequencing. Our objective is to identify genomic regions associated with grain number per panicle and pinpoint potential candidate genes that can useful for rice breeding programs aimed at improving grain number per panicle in rice.

Materials and methods

Plant materials and generation of segregating population

In the present study, we used an F_2 segregating population composed of 256 individuals, derived from a cross between two *Oryza sativa* L. ssp. indica genotypes: CB12132 (High grain number) and IET28835 (Low grain number). To prepare the F_2 segregating population, IET28835 plants were crossed with CB12132 in the *summer* of 2020, and F_1 plants were obtained. After confirming the true F_1s, the plants were selfed to produce the F_2 population in the *summer* of 2021. All field experiments were conducted at the Paddy Breeding Station, Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, India. Each plant had a plant-plant distance of 20 cm × 20 cm, respectively. The 2012 edition of the TNAU-Crop Production Guide was followed for agronomic practices and plant protection techniques to ensure healthy crop growth.

Trait evaluation and construction of bulks

The parents and the F_2 segregating population were subjected to trait evaluation and used to construct bulks. Traits such as flag leaf area (FLA), panicle length (PL), number of spikelets per panicle (NOSPP), number of grains per panicle (NOGPP), number of primary branches (NOPB), number of secondary branches (NOSB), number of secondary branches per primary branch (NOSBPB), number of spikelets in primary branches (NOSIPB), and number of spikelets in secondary branches (NOSISB) were measured agreeing to IRRI's standard evaluation system (SES) (IRRI, 2014). In F_2 plants, the first-formed primary panicle was tagged in each plant, and at physiological maturity, it was collected and used to count the number of grains per panicle. From a total of 256 F_2 plants of CB12132 × IET28835 cross, 15 plants each with a high grain number per panicle and low grain number per panicle were selected to construct two bulks, referred to as the high grain number per panicle-bulk (HGNPP-B) and low grain number per panicle-bulk (LGNPP-B), respectively. F_2 individuals from the two groups were selected, and DNA was isolated by the CTAB method (Murray and Thompson 1980). The extracted DNA of the 15 individual plants in each group was mixed at equal molar concentrations and pooled together to construct the bulks.

Library preparation and whole-genome resequencing

We prepared four genomic libraries, including two parents (CB12132 and IET28835) and two extreme bulks (HGNPP-B and LGNPP-B) using Nextera XT DNA library preparation kit (Illumina, Inc., San Diego, CA, USA). The libraries had a size of 350b and were sequenced on the Illumina NovaSeq 6000 (Illumina, Inc., San Diego, CA, USA) with a pair-end read length of 150 bp. The raw sequence data has been deposited in the NCBI Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra) under the following Bio project and SRA accession numbers; PRJNA1094905, and SRX24115668- SRX24115671.

Data analysis via the QTL-seq pipeline

We performed quality control and preprocessing on the raw reads from four samples (CB12132, IET28835, HGNPP-B, and LGNPP-B) using Trimmomatic v0.39 (Bolger et al. 2014). The clean reads were then mapped to the reference genome, Nipponbare (Oryza_sativa_IRGSP-1.0, accessed on 10.02.2024) using Burrows-Wheeler Aligner (BWA) software (Version 0.7.17) (Li and Durbin 2009). To ensure that only uniquely mapped reads retained. We converted the SAM file to BAM format, sorted and indexed it using samtools (version 1.17), and removed duplicates using Picard Tools (version 3.00) (https://broadinstitute.github.io/picard/). The processed BAM files were used as input for QTLseq analysis using the pipeline reported by Sugihara et al. (2022) and Takagi et al. (2013). We calculated the SNP index for each SNP position in the HGNPP-B and LGNPP-B samples based on the count of reads harboring SNPs similar or dissimilar to the reads of CB12132. If all reads in the HGNPP-B was identical to CB12132, the SNP index was 0, and if all reads were different, the SNP index was 1. The ΔSNP index of each SNP was estimated according to the formula: ΔSNP index = SNP index of LGNPP-B – SNP index of HGNPP-B. The SNP index distribution among the 12 chromosomes was analysed using the sliding window method by setting 1 Mb of window size and 50 kb of increment steps. Genomic regions with an average ΔSNP index value greater than the surrounding region with a 95%

or 99% confidence level were considered effective. Additionally, Genome Analysis Toolkit (GATK), Haplotype Caller was used to predict the SNP and InDel variation (McKenna et al. 2010). SnpSift and SnpEff (version 5.1) were used for the filtering (parameters: QUAL > = 30 and MQ > = 30 and DP > = 10) and annotation of SNPs and InDels (Cingolani et al. 2012).

Search for candidate genes

We employed a three-stage approach to identify the possible candidate genes. First, we collected genes within the specific genomic region. Next, we pinpointed gene positions, sorted genes with nonsynonymous SNPs and InDels, and annotated them using the Rice Annotation Project Database (https://ricexpro.dna.affrc.go.jp accessed on 09.03.2024). We excluded genes labeled as '(retro) transposon', 'hypothetical', or 'unknown' from further analysis. Eventually, expression pattern of selected genes in flag leaf and inflorescence developmental stages were analysed using the RiceXpro database (https://ricexpro.dna.affrc.go.jp accessed on 09.03.2024). By integrating these analyses with literature knowledge, we prioritized the possible genes associated with grain numbers per panicle in each candidate genomic region.

Quantitative real-time PCR (qRT-PCR) analysis

Fresh, young inflorescence (0.5 cm length at the branch primordium stage) were collected from various plants of CB12132 and IET28835. A 0.5 g sample was used for total RNA isolation, which was performed using the TRIzol® reagent kit (Invitrogen, Carlsbad, CA, USA) following to the user guidelines. We used the iScript cDNA synthesis kit (Bio-Rad Laboratories, Inc., Hercules, CA, USA) to convert DNA-free RNA from high and low grain number parents into high-quality cDNA, following the product guidelines. High-quality cDNA was used as the template for qRT-PCR analysis, and reactions were performed on the Bio-Rad CFX 96 (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Each 20 µL PCR reaction contained 1 µL of cDNA, 1 µL of each forward and reverse primers, 10 µL AccuPower ® 2X GreenstarTM SYBR Green Master mix, and 6 µL dd H_2O. We used a $2^{-\Delta\Delta Ct}$ method to analyse the relative expression of the target genes, using *OsActin1* (Han et al. 2024) as the internal control. In addition, all the samples were amplified at three biological replications, with three technical replications at each biological replication. The primer sequences were listed in **supplementary Table 1**.

Results

Phenotypic variation in F_2 population and construction of bulks

The high and low grain number per panicle parents, CB12132 and IET28835 were crossed and advanced to F_1 and F_2 generation. The evaluation of traits such as, FLA, PL, NOSPP, NOGPP, NOPB, NOSBPB, NOSIPB, and NOSISB showed significant differences between parents and its F_1 generation (Table 1). In the F_1 generation, the grain number per panicle had an intermediate value of 187.50 ± 12.56 compared to CB12132 (423.00 ± 12.56) and IET 28835 (108.00 ± 11.45). In the F_2 generation, 225 individual plants were evaluated, and the number of grains per panicle ranged from 89 to 420. The majority of plants had grain number per panicle ranging between 150 and 200. A total of 15 F_2 plants with grain number per panicle ranging from 355 to 420 were selected to generate the HGNPP-B, while 15 F_2 plants with grain number per panicle ranging from 89 to 125 were used to produce the LGNPP-B. The two parents and high and low-grain numbers F_2 individuals are presented in Fig. 1.

Whole-genome resequencing of parental lines and HGNPP and LGNPP bulks

Sequencing of CB12132, IET28835, HGNPP-B and LGNPP-B yielded 60.45, 56.40, 60.00 and 48.39 million reads, respectively (Table 2). After filtering, we obtained the following number of high-quality clean reads for each sample: 60.07 million for CB12132, 55.98 million for IET28835, 59.40 million for HGNPP-B, and 48.08 million for LGNPP-B. Alignment with the reference genome produced 8.98 Gb and 8.16 Gb of clean data (22.74X and 21.42X coverage) for CB12132 and IET28835, and 8.87 Gb, and 7.19 Gb clean data (21.53X and 18.36X coverage) for HGNPP-B and LGNPP-B. The number of genome-wide SNPs and InDels

Table 1 The phenotypic observation on grain number per panicle and its related traits

Sl.no	Traits	CB12132	IET28834	F_1
1	FLA	44.25 ± 2.64	26.45 ± 1.89	30.18 ± 1.75
2	PL	29.38 ± 1.64	22.75 ± 1.75	25.63 ± 1.95
3	NOSPP	437.50 ± 10.50	111.75 ± 9.85	195.00 ± 12.65
4	NOGPP	423.00 ± 12.56	108.00 ± 11.45	187.50 ± 12.56
5	NOPB	17.50 ± 1.32	9.33 ± 1.36	13.50 ± 1.65
6	NOSB	74.75 ± 3.85	17.83 ± 3.45	32.50 ± 3.84
7	NOSBPB	4.26 ± 0.74	1.91 ± 0.80	2.43 ± 0.95
8	NOSIPB	83.50 ± 3.24	48.17 ± 3.15	68.50 ± 3.29
9	NOSISB	354.00 ± 13.50	63.58 ± 12.50	126.50 ± 11.45

FLA, Flag leaf area; PL, Panicle length; NOSPP, Number of spikelets per panicle; NOGPP, Number of grains per panicle; NOPB, Number of primary branches; NOSB, number of secondary branches; NOSBPB, number of secondary branches per primary branch; NOSIPB, number of spikelets in primary branches; NOSISB, number of spikelets in secondary branches

identified in HGNPP-B and LGNPP-B after mapping cleaned reads onto the reference genome were 877,063 and 846,524, respectively. Using a criterion of read depth exceeding 15 in both HGNPP-B and LGNPP-B and a SNP index surpassing 0.30 in at least one, a total of 204,923 SNPs and 66,459 InDels commonly detected in both HGNPP-B and LGNPP-B were used for QTL seq analysis (Fig. 2).

Tracing candidate genomic region associated with grain number per panicle using QTLseq

QTL-seq analysis was done to trace the candidate genomic region associated with grain number per panicle. This analysis calculated the Δ (SNP index) values across the genome in 1 Mb window with a 50 kb increment. It revealed five candidate QTL regions associated with grain number per panicle on chromosomes 1, 4, and 5 (Table 3). Of these, three QTL regions on chromosome 1 designated as *qGNPP1.1*, *qGNPP1.2*, and *qGNPP1.3* spanned between 10.40 to 12.76 Mb (2.36 Mb), 24.61 to 25.33 Mb (0.72 Mb) and 26.57 to 27.26 Mb (0.68 Mb) with Δ SNP index values of 0.451, 0.459, and 0.233. The QTL on chromosome 4, *qGNPP4.1* (3.64 Mb) spanned between 27.70 to 31.34 Mb with Δ SNP index values of 0.471. Another QTL, *qGNPP5.1* (3.38 Mb) on chromosome 5, was identified in the region of 2.12 to 5.50 Mb with a Δ SNP index values of 0.452. The SNP index plot and candidate genomic regions across the rice genome are presented in Fig. 3 and 4.

Briefing the putative candidate genes associated with grain number per panicle

A total of 199, 87, 73, 508, and 359 genes were identified in the targeted candidate genomic regions *qGNPP1.1*, *qGNPP1.2*, *qGNPP1.3*, *qGNPP4.1* and *qGNPP5.1*, respectively. Further, we identified 23 candidate genes associated with grain number per panicle in five candidate genomic regions based on nonsynonymous SNPs/InDels, expression in inflorescence, and annotation details (Table 4 and Fig. 5). On chromosome 1, the *qGNPP1.1* region includes gene such as squamosa-promoter binding-like protein 2 (*Os01g0292900*), similar to kinase interactor 1 (*Os01g0310800*), R2R3-MYB transcription factor 6 (*Os01g0298400*) and polygalacturonase (*Os01g0296200*). The *qGNPP1.2* region contains genes like ATPase, AAA-type, core domain-containing protein (*Os01g0623500*), myosin tail 2 domain-containing protein (*Os01g0621700*), and glycosyltransferase (*Os01g0622000*), while the *qGNPP1.3* region contains the TRITHORAX-like protein (*Os01g0655300*). On chromosome 4, the *qGNPP4.1* region encompasses eight possible candidate genes, including the previously reported serine protease (*NAL1; Os04g0615000)* gene. Meanwhile chromosome 5's *qGNPP5.1* region includes six possible candidate genes such as serine carboxypeptidase (*Os05g0158500*), 4Fe-4S ferredoxin (*Os05g0157300*), S-Domain receptor-like kinase-36 (*Os05g0166600*), ethylene receptor-like (*Os05g0155200*), KIP1-like domain-containing protein (*Os05g0168800*) and similar to P-glycoprotein

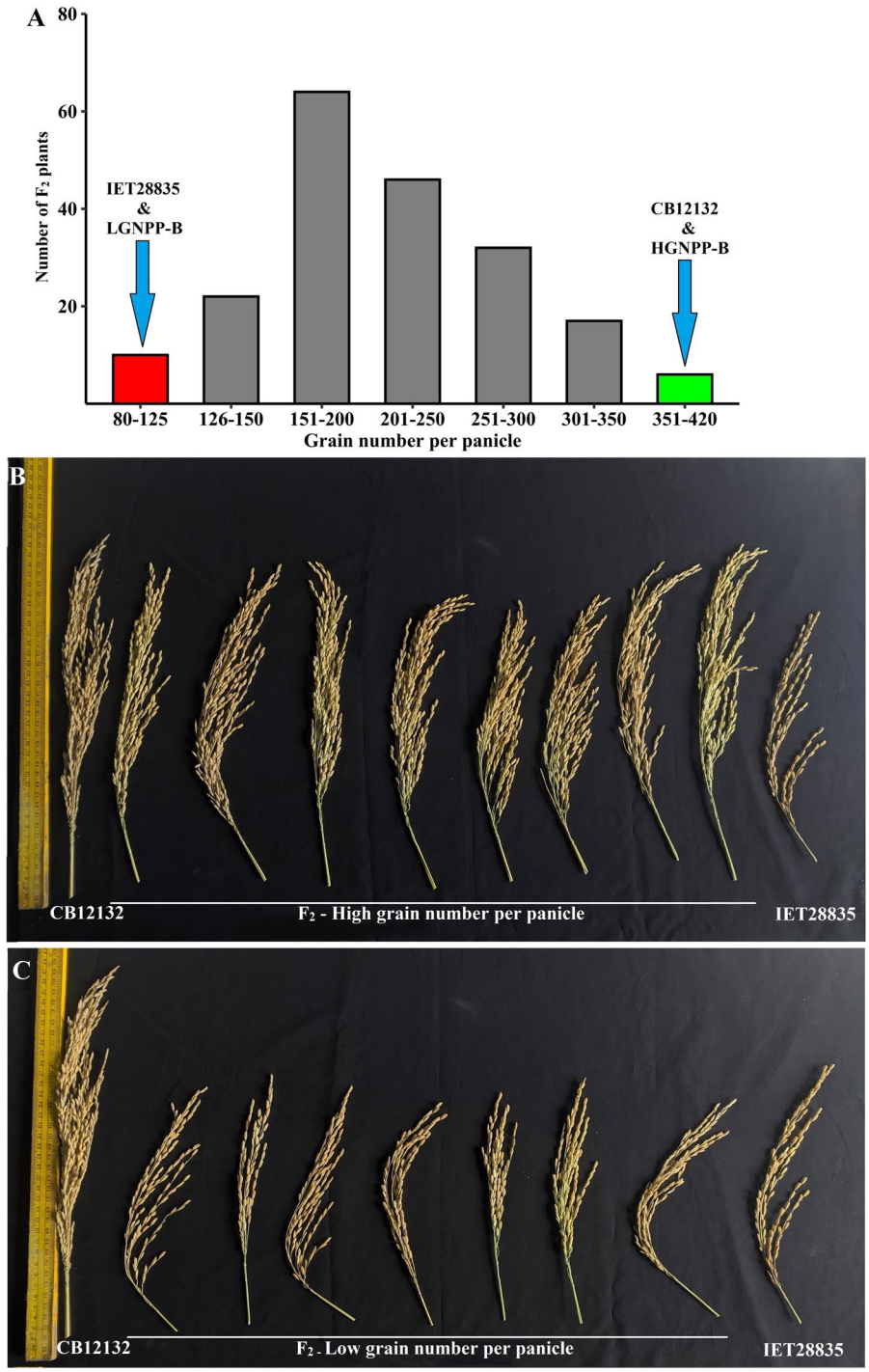

Fig. 1 Grain number per panicle distribution in the F_2 generation derived from the cross between CB12132 and IET28835. Parents and F_2 individuals **(A)**, F_2 individuals with a high grain number per panicle **(B)**, and F_2 individuals with a low grain number per panicle **(C)**. *Note*: HGNPP-B stands for high grain number per panicle-bulk, and LGNPP-B stands for low grain number per panicle-bulk

ABCB5 (*Os05g0137200*). These genes are potential contributors to grain number per panicle and useful for further investigation.

Expression profiling of candidate genes

A total of six candidate genes such as *Os01g0292900 (SPL1)*, *Os01g0622000 (OsCUGT1)*, *Os01g0655300*

Table 2 The overview of whole genome resequencing data of parents (CB12132 and IET28835), and HGNPP-B and LGNPP-B derived from F_2 individuals

Sl.no	Genotypes	Raw reads (Millions)	Cleaned reads (Million)	Cleaned bases (Gb)	Alignment (%)	Average depth (X)	Genome coverage (%)
1	CB12132	60.45	60.07	8.98	97.23	22.74	91.89
2	IET28835	56.40	55.98	8.16	98.27	21.42	92.29
3	HGNPP-B	60.00	59.40	8.87	93.60	21.53	94.17
4	LGNPP-B	48.39	48.08	7.19	98.30	18.36	94.18

HGNPP-B, High grain number per panicle-bulk and LGNPP-B: Low grain number per panicle-bulk;

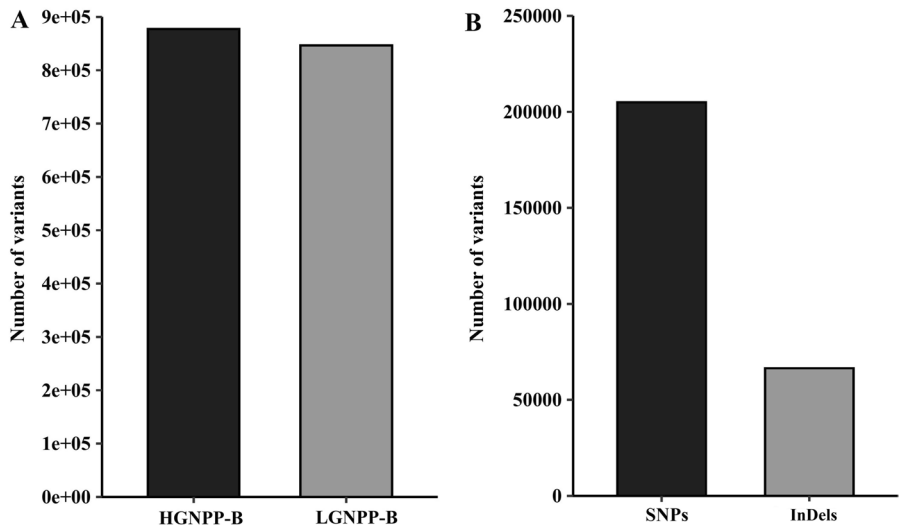

Fig. 2 A summary of SNP and InDel variants identified in the high and low bulks mapped to the reference genome **(A)**, and the common SNPs and InDels between the high and low grain number per panicle bulks **(B)**. *Note*: HGNPP-B stands for high grain number per panicle-bulk, and LGNPP-B stands for low grain number per panicle-bulk

Table 3 Details of candidate genomic regions associated with the grain number per panicle

Sl.no	QTL	Chromosome	QTL region interval (Mb)	QTL size (Mb)	HGNPP-B SNP index	LGNPP-BSNP index	Δ SNP index	Total genes	Causal genes
1	*qGNPP1.1*	Chr 01	10.40–12.76	2.36	0.216	0.667	0.451	199	54
2	*qGNPP1.2*	Chr 01	24.61–25.33	0.72	0.367	0.828	0.459	87	12
3	*qGNPP1.3*	Chr 01	26.57–27.26	0.68	0.341	0.574	0.233	73	3
4	*qGNPP4.1*	Chr 04	27.70–31.34	3.64	0.305	0.776	0.471	508	156
5	*qGNPP5.1*	Chr 05	2.12–5.50	3.38	0.386	0.838	0.452	359	69

HGNPP-B, High grain number per panicle-bulk and LGNPP-bulk: Low grain number per panicle-bulk

(SDG705), *Os04g0615000 (NAL1)*, *Os04g0559800 (SMG2)* and *Os05g0155200 (ERS2*) were selected from the 23 identified genes and their expression pattern were analysed in the young inflorescence of the two parents, CB12132 and IET28835 using qRT-PCR analysis. The results showed that all genes, except for *Os04g0559800*, were expressed at significantly higher levels in CB12132 compare to IET28835. Figure 6 shows the expression levels of these genes in CB12132 and IET28835.

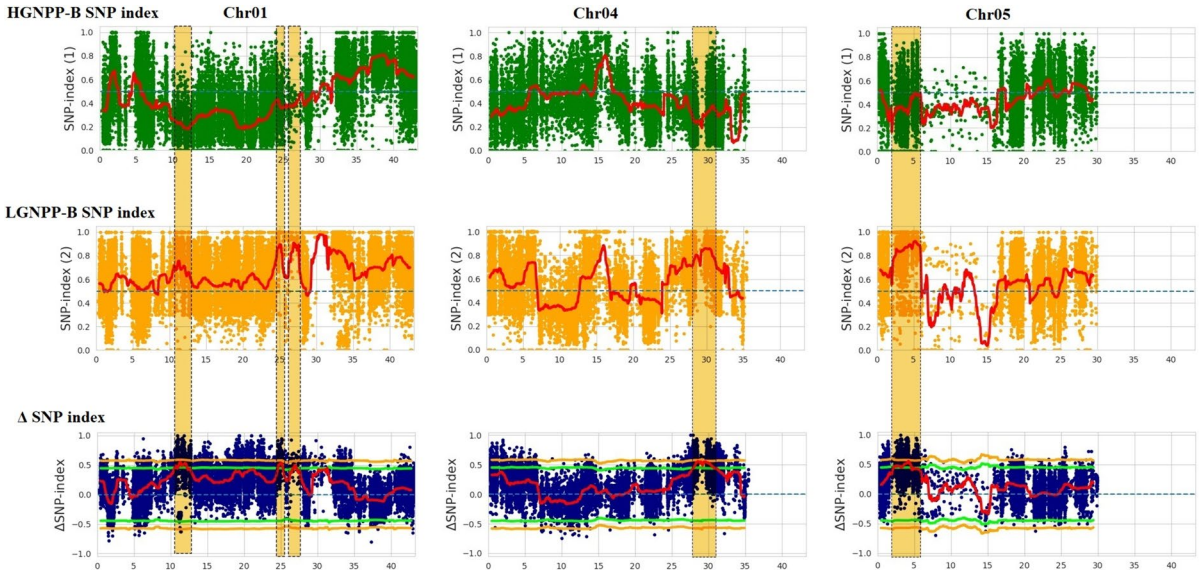

Fig. 3 SNP index plots for HGNPP-B (green dotted) **(A)**, LGNPP-B (orange dotted) **(B)**, and Δ SNP index plots for the two bulks **(C)**. Sliding window plots of the average SNP index, using a 1-Mb window size and 50-kb steps, are shown as red lines. The green and orange lines in the Δ SNP index plots represent the 95% and 99% confidence intervals, respectively. *Note:* HGNPP-B refers to the high grain number per panicle bulk, and LGNPP-B refers to the low grain number per panicle bulk

Discussion

The grain number per panicle is closely associated with grain yield and is an ideal trait targeted by rice breeders to improve the rice yield. In this study, we used a segregating F_2 population derived from the cross between CB12132 and IET28835, to investigate the genetic mechanism of this trait. The F_2 population showed a wide distribution for grain number per panicle, ranging from 89 to 420, indicating that the trait governed by additive gene action. This finding is consistent with the reports of Priyanka et al. (2019), Sekhar et al. (2021) and Wang et al. (2024). Furthermore, our study used whole-genome resequencing-based QTL-seq approach integrated with BSA to identify the genomic region regulating grain number per panicle in rice. We used a total of 15 individuals to construct the HGNPP-B and LGNPP-B, which is adequate for detecting the major loci associated with trait of interest, consistent with previous reports (Bommisetty et al. 2023; Huang et al. 2022; Kaur et al. 2022; Singh et al. 2022; Zhang et al. 2021b). The genome coverage for the HGNPP-B and LGNPP-B was 94.17% and 94.18%, respectively, which is within the ranges reported in previous studies: 96.72–96.95%, 93.70–94.39%, 93.51–94.05% and 93.37–95.53% (Gao et al. 2023; Luo et al. 2018; Ma et al. 2022). The average depth was above 20X obtained in both bulks, similar to findings reported in other studies (Nubankoh et al. 2020; Singh et al. 2022 and Takagi et al. 2013). We identified five candidate QTL regions associated with grain number per panicle using QTL-seq. These regions, *qGNPP1.1*-*qGNPP1.3*, *qGNPP4.1*, and *qGNPP5.1*, were located on chromosomes 1, 4, and 5. Notably, several of these QTLs coincide with or near previously reported QTLs for grain number per panicle. For instance, *qGNPP1.2* was positioned 1.5 Mb downstream from *LOG1* (Kurakawa et al. 2007), while *qGNPP1.3* was near *RGN1a* (Zhang et al. 2022b), with a distance of 1 Mb. Additionally, *qGNPP4.1* overlaps with *NAL1* (Fujita et al. 2013) and *qGN4.1*, and *qGNPP5.1* is linked to *qSPP5* (Luo et al. 2013). Interestingly, the *qGNPP1.1* region appears to be novel, with no reported genes related to grain number in this area.

According to the Nipponbare reference genome sequence, the five genomic regions (*qGNPP1.1*, *qGNPP1.2*, *qGNPP1.3*, *qGNPP4.1* and *qGNPP5.1*) contain a total of 1226 genes. To identify possible candidate genes associated with grain number per

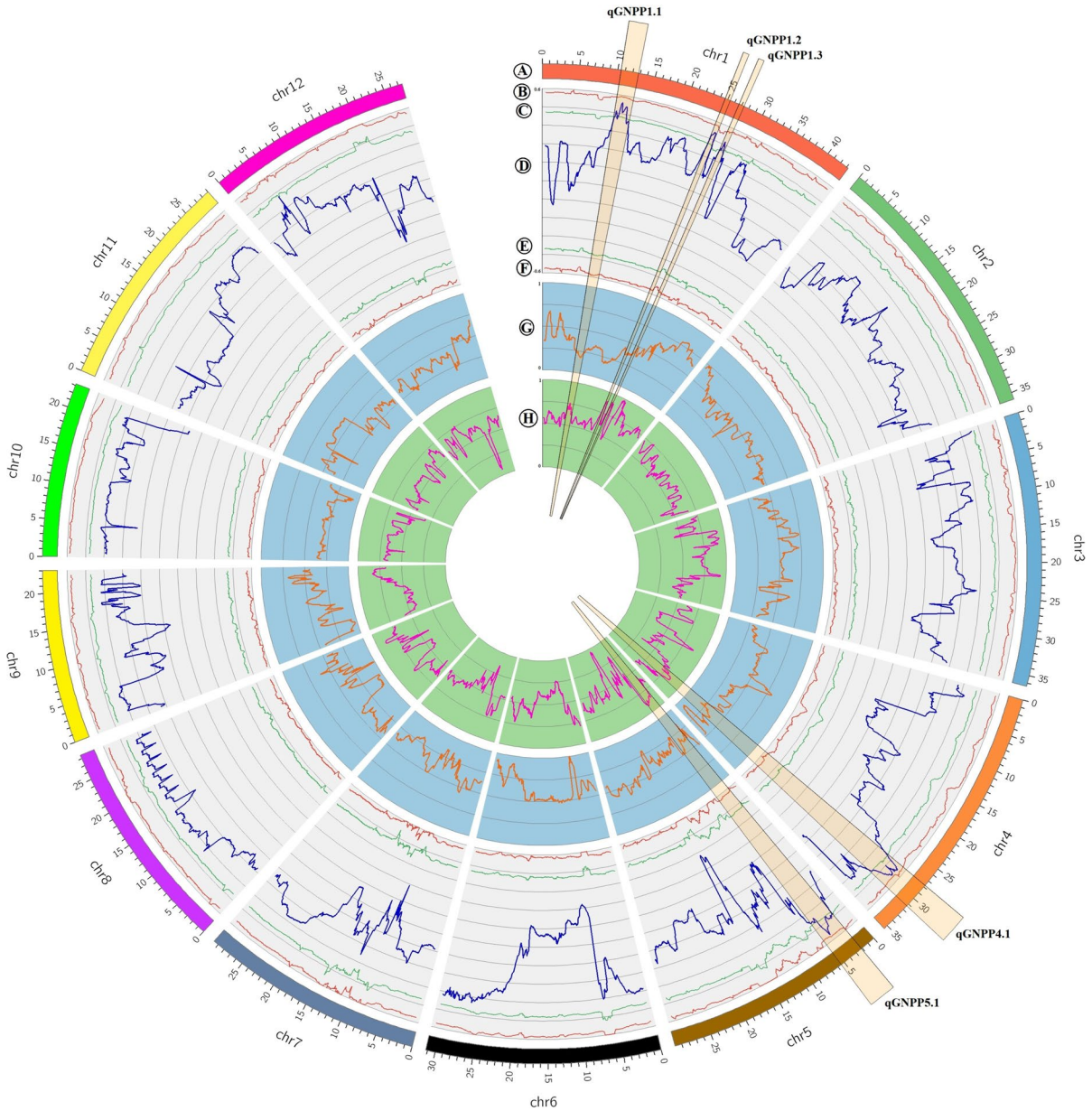

Fig. 4 Sliding window plots depict the SNP index for HGNPP-B and LGNPP-B, providing a comparison of the SNP index between the two bulks. The plots include: Pseudomolecules of the Nipponbare reference genome according to IRGSP 1.0 **(A)**, upper probability threshold at 99% confidence level ($P < 0.01$) **(B)** and 95% confidence level ($P < 0.05$) **(C)**, sliding window analysis of the ΔSNP index with a window size of 1 Mb and steps of 50 kb **(D)**, lower probability threshold at 95% confidence level ($P < 0.05$) **(E)** and 99% confidence level ($P < 0.01$) **(F)**, sliding window plots of average SNP index values in the HGNPP-B **(G)** and LGNPP-B **(H)** with a 1 Mb window and 50 kb steps. The identified candidate genomic regions for grain number per panicle are highlighted in the circos plot. *Note:* HGNPP-B stands for high grain number per panicle-bulk, and LGNPP-B stands for low grain number per panicle-bulk

panicle, we filtered these genes based on nonsynonymous SNPs/InDels, expression in inflorescence, and annotation. We first narrowed down the list to 294 genes with nonsynonymous SNPs/InDels and then selected 23 possible genes across the five regions using the expression and annotation data. These 23

Table 4 Summary of possible candidate genes identified in the genomic region associated with grain number per panicle

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP1.1	Chr 01	*Os01g0292900*	Similar to Squamosa-promoter binding-like protein 2	10,653,714	C	G	Missense_variant	0.861
				10,654,586	G	A	Missense_variant	0.875
				10,655,208	A	C	Missense_variant	0.819
qGNPP1.1	Chr 01	*Os01g0310800*	Similar to Pto kinase interactor 1	11,678,041	G	A	Missense_variant	0.840
				11,678,805	T	G	Missense_variant	0.770
				11,678,934	G	T	Missense_variant	0.796
qGNPP1.1	Chr 01	*Os01g0298400*	R2R3-MYB transcription factor 6	10,935,015	TCGGCGG	T	Disruptive_inframe_insertion	0.496
				10,935,677	A	C	Missense_variant	0.617
qGNPP1.1	Chr 01	*Os01g0296200*	Photo-sensitive leaf rolling 1	10,830,126	G	A	Missense_variant	0.667
qGNPP1.2	Chr 01	*Os01g0623500*	ATPase, AAA-type, core domain containing protein	24,896,190	G	T	Missense_variant	0.733
				24,897,157	C	A	Missense_variant	0.790
				24,863,549	G	GGCCTTC	Disruptive_inframe_insertion	0.856
				24,896,536	CTTG	C	Conservative_inframe_deletion	0.494
				24,863,725	TCTC	T	Conservative_inframe_deletion	0.284
qGNPP1.2	Chr 01	*Os01g0621700*	Myosin tail 2 domain containing protein	24,785,497	T	C	Missense_variant	0.654
				24,785,694	T	C	Missense_variant	0.312
qGNPP1.2	Chr 01	*Os01g0622000*	Glycosyltransferase	24,786,776	C	T	Missense_variant	0.702
				24,787,278	G	T	Missense_variant	0.685
qGNPP1.3	Chr 01	*Os01g0655300*	Trithorax-like protein	26,592,888	TGACGA TGACGA TGACGA TGACGA GGACGA G	TGACGA TGACGA TGACGA TGACGA TGACGA GGACGA G	Disruptive_inframe_insertion	0.586

Table 4 (continued)

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP4.1	Chr 04	*Os04g0555000*	GRAS (GAI-RGA-SCR) plant-specific transcription factor	27,765,440	A	C	Missense_variant	0.274
				27,765,467	A	G	Missense_variant	0.192
				27,765,836	A	G	Missense_variant	0.354
qGNPP4.1	Chr 04	*Os04g0559800*	Mitogen-activated protein kinase (MAPK) 10	28,050,666	T	C	Missense_variant	0.617
				28,053,739	A	G	Missense_variant	0.623
				28,055,009	C	T	5_prime_UTR_premature_start_codon_gain_variant	0.573
qGNPP4.1	Chr 04	*Os04g0560600*	Calcium-dependent protein kinase	28,082,063	CGCCGCCG	C	Frameshift_variant	0.556
				28,082,071	AAGACGAGC	A	Frameshift_variant	0.574
				28,082,103	G	C	Missense_variant	0.429
qGNPP4.1	Chr 04	*Os04g0563900*	Similar to H0409D10.5 protein	28,265,430	TTGC	T	Conservative_inframe_insertion	0.606

Table 4 (continued)

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP4.1	Chr 04	*Os04g0591900*	F-box associated domain, type 3 domain containing protein	29,958,138	C	CG	Frameshift_variant	0.541
				29,958,141	T	C	Missense_variant	0.565
				29,958,204	T	TGATGCG	Conservative_inframe_deletion	0.527
				29,958,381	G	C	Missense_variant	0.701
				29,958,394	CGGCTTGTAACCCCAGCACAGCGTGTA	C	Frameshift_variant	0.611
				29,958,501	G	A	Missense_variant	0.689
				29,958,584	A	G	Missense_variant	0.729
				29,958,713	G	A	Missense_variant	0.375
				29,959,319	G	C	missense_variant	0.736
				29,959,581	TCGACGC	T	Conservative_inframe_insertion	0.573
qGNPP4.1	Chr 04	*Os04g0597400*	Nitrate transporter, Control of grain yield and nitrogen use efficiency	30,147,663	A	G	Missense_variant	0.652
qGNPP4.1	Chr 04	*Os04g0598300*	Probable transcription factor	30,185,147	G	C	Missense_variant	0.355
qGNPP4.1	Chr 04	*Os04g0615000*	Serine protease	31,214,019	C	T	Missense_variant	0.602
				31,214,045	G	A	Missense_variant	0.745

Table 4 (continued)

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP4.1	Chr 04	*Os04g0615700*	Argonaute (AGO) family protein	31,239,297	A	G	Missense_variant	0.750
				31,239,372	C	G	Missense_variant	0.690
				31,239,383	G	GGT	Frameshift_variant	0.452
				31,239,387	C	CT	Frameshift_variant	0.402
				31,239,485	A	G	Missense_variant	0.267
				31,239,694	A	AGCCCCC	Disruptive_inframe_insertion	0.350
				31,239,711	C	CGGCGCCGGCTCA	Disruptive_inframe_insertion	0.465
				31,239,747	A	T	Missense_variant	0.395
				31,239,748	G	T	Missense_variant	0.650
				31,239,770	A	T	Missense_variant	0.505
				31,240,150	A	T	Missense_variant	0.525
				31,240,160	A	G	Missense_variant	0.645
				31,240,217	C	A	Missense_variant	0.381
				31,242,782	T	G	Missense_variant	0.671
qGNPP5.1	Chr 05	*Os05g0137200*	Similar to P-glycoprotein ABCB5	2,171,489	C	A	Missense_variant	0.620
qGNPP5.1	Chr 05	*Os05g0155200*	Similar to ethylene receptor	3,234,718	C	T	Missense_variant	0.393
				3,235,129	G	A	Missense_variant	0.650

Table 4 (continued)

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP5.1	Chr 05	*Os05g0157300*	4Fe-4S ferredoxin, iron-sulpur binding domain containing protein	3,349,978	C	G	Missense_variant	0.740
				3,350,579	G	T	Missense_variant	0.705
				3,350,731	G	C	Missense_variant	0.560
				3,350,931	T	A	Missense_variant	0.650
				3,350,935	T	C	Missense_variant	0.785
				3,351,318	G	C	Missense_variant	0.690
				3,351,343	T	C	Missense_variant	0.840
				3,351,345	C	T	Missense_variant	0.880
				3,351,369	G	A	Missense_variant	0.716
				3,351,755	T	C	Missense_variant	0.745
				3,351,768	A	C	Missense_variant	0.635
				3,351,779	A	C	Missense_variant	0.294
				3,352,741	A	G	Missense_variant	0.520
				3,352,742	G	A	Missense_variant	0.630
				3,352,746	CG	C	Frameshift_variant	0.580
				3,352,748	CGGCGGCGG	C	Frameshift_variant	0.615
				3,352,751	G	T	Missense_variant	0.620
qGNPP5.1	Chr 05	*Os05g0158500*	Serine carboxypeptidase	3,439,806	G	T	Missense_variant	0.487
				3,442,782	T	TGG	Frameshift_variant	0.480
				3,442,785	TAG	T	Frameshift_variant	0.460
				3,443,576	TCGCCGC	T	Conservative_inframe_insertion	0.451

Table 4 (continued)

QTL	Chromosome	Gene ID	Description	Position (bp)	HGNPP-B (Allele)	LGNPP-B (Allele)	Variant effect	Δ SNP index
qGNPP5.1	Chr 05	*Os05g0166600*	S-domain receptor like kinase-36	3,970,975	C	T	Missense_variant	0.458
				3,971,451	G	A	Missense_variant	0.648
				3,971,866	A	G	Missense_variant	0.620
				3,971,872	G	C	Missense_variant	0.487
				3,971,988	A	G	Missense_variant	0.550
				3,972,200	T	G	Missense_variant	0.382
				3,972,984	G	A	Missense_variant	0.826
				3,973,300	G	C	Missense_variant	0.740
				3,973,348	T	G	Missense_variant	0.360
				3,973,363	C	T	Missense_variant	0.480
	Chr 05	*Os05g0168800*	KIP1-like domain containing protein	4,130,144	A	C	Missense_variant	0.665
				4,131,212	C	T	Missense_variant	0.673

HGNPP-B, High grain number per panicle-bulk and LGNPP-B: Low grain number per panicle-bulk

genes show promise for further investigation into their role in regulating grain number per panicle. In the *qGNPP1.1* genomic region (Chr 1; 10.40 Mb to 12.76 Mb), we identified four potential genes: *Os01g0292900* (*SPL1*), *Os01g0310800* (*WAK4*), *Os01g0298400* (*2R_MYB6*), and *Os01g0296200* (*PSL*). Notably, *SPL1* belongs to a plant transcription factor family that regulates growth, development, and grain yield. *SPL* genes increase grain number per panicle by promoting panicle branching and enhance grain length and plant yield by regulating cell size (Jiao et al. 2010; Miura et al. 2010). We found three missense SNPs in *SPL1*, which may contribute to higher grain number per panicle. Additionally, *WAK4*, involved in cell expansion (Delteil et al. 2016), may also play a role in regulating panicle development, as loss of *WAK* function has been linked to decreased florets in rice panicles (Zhang et al. 2022b). In the *qGNPP1.2* genomic region (Chr 1; 24.61 Mb to 25.33 Mb), we identified three potential candidate genes: *Os01g0623500*, *Os01g0621700*, and *Os01g0622000* (*OsCUGT1*). Notably, *OsCUGT1* is a Glycosyltransferase involved in regulating physiological and stress responses in rice. It was previously identified as a cold stress-responsive gene, and its mutant showed a sterile inflorescence phenotype with altered cytokinin metabolism and phenylpropanoid biosynthesis (Zhao et al. 2023). *Os01g0655300* (*SDG705*) is the most possible candidate gene identified in the *qGNPP1.3* genomic region (Chr 1; 26.57 Mb to 27.26 Mb). *SDG705* belongs to the SET domain family, which is involved in histone modification and various biological processes, including flowering regulation. Enhanced expression of *SDG* genes promotes florigen genes and *Ehd3*, critical promoters of rice flowering. Interestingly, overexpression of a

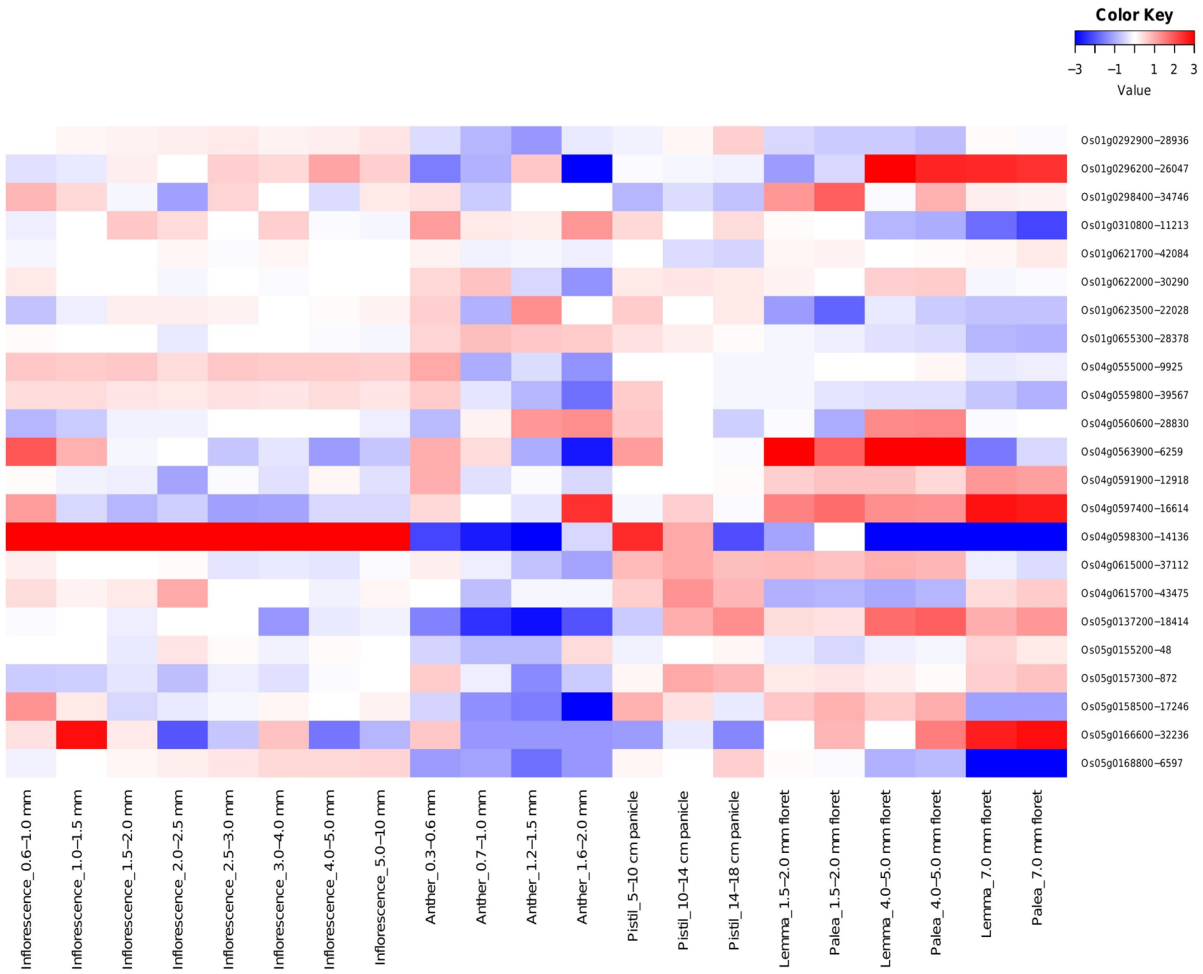

Fig. 5 The expression patterns of 23 possible candidate genes across various reproductive stages and tissues, as obtained from the RiceXPro database (accessed on 09.03.2024)

similar gene, *GhD7*, delayed flowering and increased panicle size, leading to enhanced grain yield in rice (Xue et al. 2008).

In the *qGNPP4.1* genomic region (Chr 4; 27.70 Mb to 31.34 Mb), we identified nine potential candidate genes, including *Os04g0615700* (*OsAGO2*), *Os04g0563900* (*OsRLCK157*), *Os04g0597400* (*OsNPF7.6*), *Os04g0559800* (*SMG2*), *Os04g0598300* (*RFL*), *Os04g0560600* (*CDPK12*), *Os04g0615000* (*NAL1*), *Os04g0555000* (*OsHAM3*) and *Os04g0591900* (*OsFbox231*). Notably, *NAL1* (*Os04g0615000*) encodes a serine protease, previously reported to regulate grain number per panicle, with two missense SNPs. *NAL1* has been extensively studied for its pleiotropic effects on various traits, including grain number per panicle, leaf width, and photosynthetic efficiency (Fujita et al. 2013; Takai et al. 2013; Xu et al. 2015; Yano et al. 2016; Zhang et al. 2014). Other notable candidate genes in this region include: *OsNPF7.6* (*Os04g0597400*), involved in nitrate transportation and uptake, with potential for improving nitrogen utilization efficiency and yield (Zhang et al. 2022a). *SMG2* (*Os04g0559800*), interacting with signaling cascades to modulate grain size (Xu et al. 2018). *OsAGO2* (*Os04g0615700*), regulating anther development by controlling ROS levels and tapetal PCD through DNA methylation-mediated gene expression (Zheng et al. 2019). In the *qGNPP5.1* genomic region (Chr 5; 2.12 Mb to 5.50 Mb), we identified six potential candidate

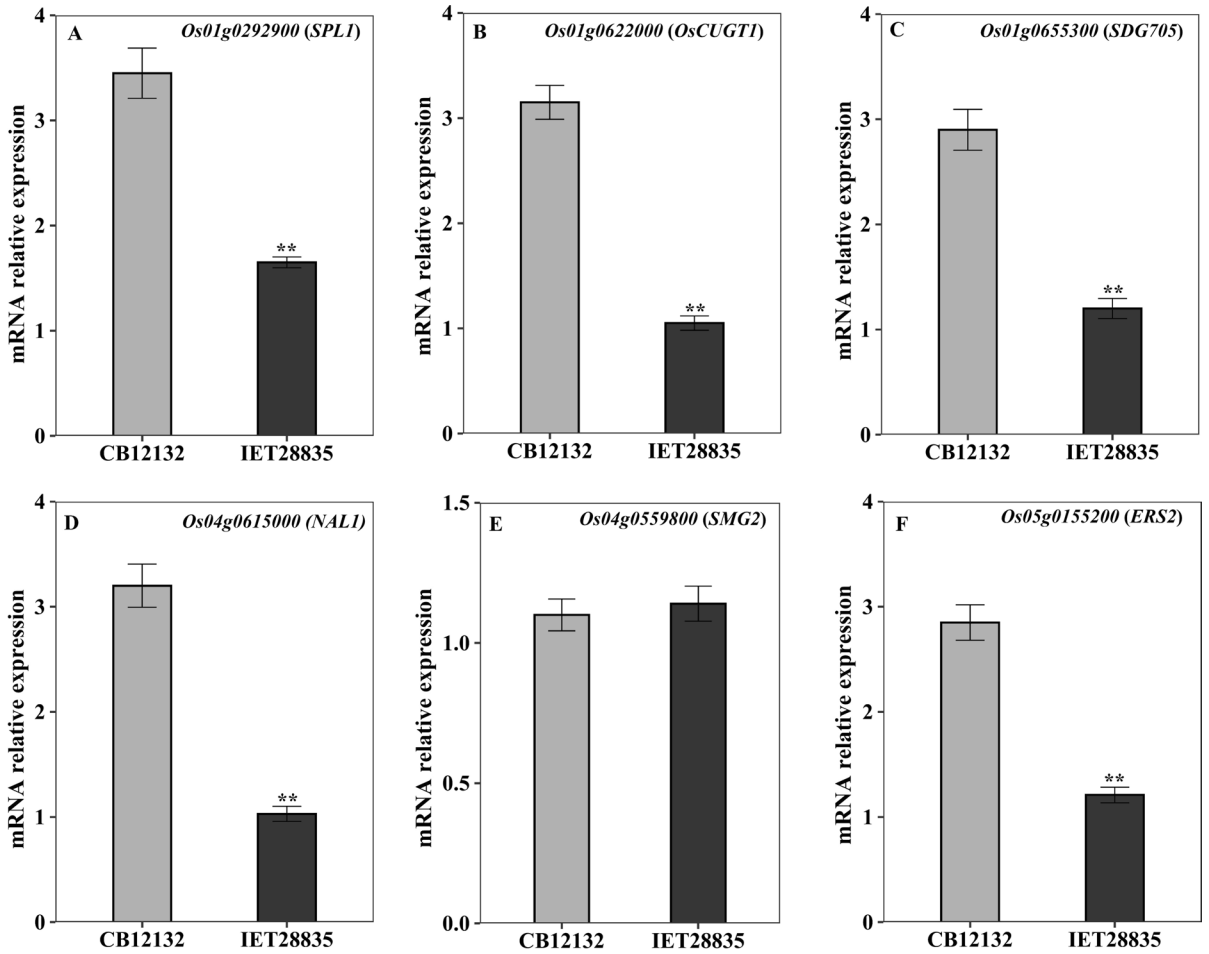

Fig. 6 The expression patterns of six candidate genes were analyzed in high and low grain number per panicle parents (CB12132 and IET28835) using qRT-PCR. Data are presented as the means of three biological replicates, with error bars indicating the standard deviation (SD)

genes: *Os05g0158500* (*GS5*), *Os05g0157300*, *Os05g0166600* (*SDRLK-36*), *Os05g0155200* (*ERS2*), *Os05g0168800*, and *Os05g0137200* (*MDR3*). Notably, *ERS2* (*Os05g0155200*) encodes an ethylene receptor, previously reported to regulate grain number per panicle. Ethylene plays a crucial role in grain filling, with higher levels at anthesis associated with poorer grain filling (Panigrahi et al. 2023). Another notable candidate gene is *GS5* (*Os05g0158500*), encoding serine carboxypeptidase, which positively regulates grain size and grain yield. These genes may play important roles in regulating grain number per panicle and grain yield (Li et al. 2011b).

In summary, by adopting QTL-seq approach, we identified five candidate genomic regions associated with grain number per panicle using an F_2 population derived from CB12132×IET28835 cross. Additionally, we identified 23 possible candidate genes through analysis of gene sequences, expression, and annotation data sets. Of these, six genes were validated by qRT-PCR analysis. We recommend further investigation of these candidate genes using functional genomics approaches. The genomic regions and genes identified in this study could be valuable for enhancing grain yield in rice breeding programs.

Consent for publication

All authors have read and agreed to the published version of the manuscript. All authors read and approved the final manuscript.

Ethics approval

This article does not contain any studies with human or animal subjects.

Acknowledgements The authors express their gratitude to the Department of Rice, Centre for Plant Breeding and Genetics, and the Department of Plant Molecular Biology & Bioinformatics, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, for providing the necessary facilities to conduct the field experiment and sequencing analysis.

Author contribution Conceptualization GA, SG, RSA and MR; Field and laboratory experiment: GA, data analysis GA,AK,LA,ND; resources and supervision SG, RSA, KKS, KA, RSU; Original draft preparation: GA, Review and editing: GA,AK,SG.

Funding Authors acknowledged the partial funding support of the Indian Council of Agricultural Research - Consortium Research Platform (CRP) on hybrid rice scheme.

Data availability The raw sequence data has been deposited in the NCBI Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra) under the following Bio project and SRA accession numbers; PRJNA1094905, and SRX24115668- SRX24115671.

Declarations

Conflict of interest The authors declare no competing interests.

References

Arikit S, Wanchana S, Khanthong S, Saensuk C, Thianthavon T, Vanavichit A, Toojinda T (2019a) QTL-seq identifies cooked grain elongation QTLs near soluble starch synthase and starch branching enzymes in rice (Oryza sativa L.). Sci Rep 9(1):8328

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309(5735):741–745

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120

Bommisetty R, Chakravartty N, Hariprasad K, Rameshbabu P, Sudhakar P, Bodanapu R, Naik JB, Reddy BB, Lekkala SP, Gupta S (2023) Identification of a novel QTL for grain number per panicle employing NGS-based QTL-seq approach in rice (Oryza sativa L). Plant Biotechnol Rep. 17(2):191–201

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms. SnpEff Fly 6(2):80–92

Delteil A, Gobbato E, Cayrol B, Estevan J, Michel-Romiti C, Dievart A, Kroj T, Morel J-B (2016) Several wall-associated kinases participate positively and negatively in basal defense against rice blast fungus. BMC Plant Biol 16:1–10

Fujita D, Trijatmiko KR, Tagle AG, Sapasap MV, Koide Y, Sasaki K, Tsakirpaloglou N, Gannaban RB, Nishimura T, Yanagihara S (2013) NAL1 allele from a rice landrace greatly increases yield in modern indica cultivars. Proc Natl Acad Sci 110(51):20431–20436

Gao Q, Wang H, Yin X, Wang F, Hu S, Liu W, Chen L, Dai X, Liang M (2023) Identification of salt tolerance related candidate genes in 'Sea Rice 86'at the seedling and reproductive stages using QTL-Seq and BSA-Seq. Genes 14(2):458

Gunasekaran A, Seshadri G, Ramasamy S, Muthurajan R, Karuppasamy KS (2023) Identification of newer stable genetic sources for high grain number per panicle and understanding the gene action for important panicle traits in rice. Plants 12(2):250

Guo T, Chen K, Dong N-Q, Shi C-L, Ye W-W, Gao J-P, Shan J-X, Lin H-X (2018) Grain size and number1 negatively regulates the OsMKKK10-OsMKK4-OsMPK6 cascade to coordinate the trade-off between grain number per panicle and grain size in rice. Plant Cell 30(4):871–888

Han Y, Hu Q, Gong N, Yan H, Khan NU, Du Y, Sun H, Zhao Q, Peng W, Li Z (2024) Natural variation in MORE GRAINS 1 regulates grain number and grain weight in rice. J Integr Plant Biol 66(7):1440–1458

Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu C, Li J, Fu X (2009) Natural variation at the DEP1 locus enhances grain yield in rice. Nat Genet 41(4):494–497

Huang L, Tang W, Wu W (2022a) Optimization of BSA-seq experiment for QTL mapping. G3 12(1):jkab370

Huo X, Wu S, Zhu Z, Liu F, Fu Y, Cai H, Sun X, Gu P, Xie D, Tan L (2017) NOG1 increases grain production in rice. Nat Commun 8(1):1–11

Ikeda-Kawakatsu K, Yasuno N, Oikawa T, Iida S, Nagato Y, Maekawa M, Kyozuka J (2009) Expression level of ABERRANT PANICLE ORGANIZATION1 determines rice inflorescence form through control of cell proliferation in the meristem. Plant Physiol 150(2):736-747I

Jia X, Wang S, Zhao H, Zhu J, Li M, Wang G (2023) QTL mapping and BSA-seq map a major QTL for the node of the first fruiting branch in cotton. Front Plant Sci 14:1113059

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nat Genet 42(6):541–544

Kaur G, Yadav IS, Bhatia D, Vikal Y, Neelam K, Dhillon NK, Praba UP, Mangat GS, Singh K (2022) BSA-seq identifies a major locus on chromosome 6 for root-knot nematode (*Meloidogyne graminicola*) resistance from Oryza glaberrima. Front Genet 13:871833

Kurakawa T, Ueda N, Maekawa M, Kobayashi K, Kojima M, Nagato Y, Sakakibara H, Kyozuka J (2007) Direct control of shoot meristem activity by a cytokinin-activating enzyme. Nature 445(7128):652–655

Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25(14):1754–1760

Li S, Qian Q, Fu Z, Zeng D, Meng X, Kyozuka J, Maekawa M, Zhu X, Zhang J, Li J (2009) Short panicle1 encodes a putative PTR family transporter and determines rice panicle size. Plant J 58(4):592–605

Li M, Tang D, Wang K, Wu X, Lu L, Yu H, Gu M, Yan C, Cheng Z (2011a) Mutations in the F-box gene larger panicle improve the panicle architecture and enhance the grain yield in rice. Plant Biotechnol J 9(9):1002–1013

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J (2011b) Natural variation in GS5 plays an important role in regulating grain size and yield in rice. Nat Genet 43(12):1266–1269

Luo X, Ji S-D, Yuan P-R, Lee H-S, Kim D-M, Balkunde S, Kang J-W, Ahn S-N (2013) QTL mapping reveals a tight linkage between QTLs for grain weight and panicle spikelet number in rice. Rice 6:1–10

Luo XD, Jian L, Jun Z, Dai LF, Chen YL, Zhang L, Zhang F, Hu BL, Xie JK (2018a) Rapid mapping of candidate genes for cold tolerance in Oryza rufipogon Griff by QTL-seq of seedlings. J Integr Agric 17(2):265–275

Ma Y, Mackon E, Mackon JDE, G C, Zhao, Y., Li, Q., Dai, X., Yao, Y., Xia, X., Nong, B., Liu, P. (2022b) Combined Analysis of BSA-Seq Based Mapping, RNA-Seq, and metabolomic unraveled candidate genes associated with panicle grain number in rice (Oryza sativa L). Biomolecules 12(7):918

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M (2010) The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res 20(9):1297–1303

Michelmore RW, Paran I, Kesseli R (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88(21):9828–9832

Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. Nat Genet 42(6):545–549

Murray M, Thompson W (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4326

Nakagawa M, Shimamoto K, Kyozuka J (2002) Overexpression of RCN1 and RCN2, rice terminal flower 1/centroradialis homologs, confers delay of phase transition and altered panicle morphology in rice. Plant J 29(6):743–750

Nubankoh P, Wanchana S, Saensuk C, Ruanjaichon V, Cheabu S, Vanavichit A, Toojinda T, Malumpong C, Arikit S (2020) QTL-seq reveals genomic regions associated with spikelet fertility in response to a high temperature in rice (*Oryza sativa* L.). Plant Cell Rep 39:149–162

Oladosu Y, Rafii MY, Abdullah N, Magaji U, Miah G, Hussin G, Ramli A (2017) Genotype× Environment interaction and stability analyses of yield and yield components of established and mutant rice genotypes tested in multiple locations in Malaysia. Acta Agr Scand Acta 67(7):590–606

Panigrahi S, Kariali E, Dash SK, Sahu BB, Mohapatra PK (2023) Ethylene sensitivity underscores the yield advantage of high-grain numbers in cylinder-shaped rice panicles. Environ Exp Bot 214:105466

Priyanka A, Gnanamalar R, Banumathy S, Senthil N, Hemalatha G (2019) Genetic variability and frequency distribution studies in F_2 segregating generation of rice. Electron J Plant Breed 10(3):988–994

Pujol M, Alexiou KG, Fontaine A-S, Mayor P, Miras M, Jahrmann T, Garcia-Mas J, Aranda MA (2019b) Mapping cucumber vein yellowing virus resistance in cucumber (Cucumis sativus L.) by using BSA-seq analysis. Front Plant Sci 10:1583

Resilience B (2017) The state of food security and nutrition in the world: Building resilience for peace and food security.

Sekhar S, Kumar J, Mohanty S, Mohanty N, Panda RS, Das S, Shaw BP, Behera L (2021) Identification of novel QTLs for grain fertility and associated traits to decipher poor grain filling of basal spikelets in dense panicle rice. Sci Rep 11(1):13617

Standard Evaluation System (SES) for Rice. p 57, 5th edition. Los Banos, Philippines

Singh VK, Ellur RK, Singh AK, Nagarajan M, Singh BD, Singh NK (2018) Effect of qGN4 1 QTL for grain number per panicle in genetic backgrounds of twelve different mega varieties of rice. Rice 11(1):1–13

Singh V, Sinha P, Obala J, Khan AW, Chitikineni A, Saxena RK, Varshney RK (2022) QTL-seq for the identification of candidate genes for days to flowering and leaf shape in pigeonpea. Hered 128(6):411–419

Sugihara Y, Young L, Yaegashi H, Natsume S, Shea DJ, Takagi H, Booker H, Innan H, Terauchi R, Abe A (2022) High-performance pipeline for MutMap and QTL-seq. PeerJ 10:e13170

Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J 74(1):174–183

Takai T, Adachi S, Taguchi-Shiobara F, Sanoh-Arai Y, Iwasawa N, Yoshinaga S, Hirose S, Taniguchi Y, Yamanouchi U, Wu J (2013) A natural variant of NAL1, selected in high-yield rice breeding programs, pleiotropically increases photosynthesis rate. Sci Rep 3(1):2149

Vogel G, LaPlant KE, Mazourek M, Gore MA, Smart CD (2021) A combined BSA-Seq and linkage mapping approach identifies genomic regions associated with Phytophthora root and crown rot resistance in squash. Theor Appl Genet 134:1015–1031

Wang C, Tang S, Zhan Q, Hou Q, Zhao Y, Zhao Q, Feng Q, Zhou C, Lyu D, Cui L (2019) Dissecting a heterotic

gene through GradedPool-Seq mapping informs a rice-improvement strategy. Nat Commun 10(1):2982

Wang Y, Zhai L, Chen K, Shen C, Liang Y, Wang C, Zhao X, Wang S, Xu J (2020) Natural sequence variations and combinations of GNP1 and NAL1 determine the grain number per panicle in rice. Rice 13:1–15

Wang X, Han Y, Zhang Y-x, Deng B, Wu B-q, Guo X-y, Qin Y-f, Fang Y-y, Liu F, Qin B-x (2022c) QTL mapping integrated with BSA-Seq analysis identifies a novel gene conferring resistance to brown planthopper from common wild rice (Oryza rufipogon Griff.). Euphytica 218(3):34

Wang Y, Wang X, Zhai L, Zafar S, Shen C, Zhu S, Chen K, Xu J (2024) A novel effective panicle number per plant 4 haplotype enhances grain yield by coordinating panicle number and grain number in rice. Crop J 12(1):202–212

Weng Z, Yang Y, Wang X, Wu L, Hua S, Zhang H, Meng Z (2021) Parentage analysis in giant grouper (*Epinephelus lanceolatus*) using microsatellite and SNP markers from genotyping-by-sequencing data. Genes 12(7):1042

Wu Y, Wang Y, Mi X-F, Shan J-X, Li X-M, Xu J-L, Lin H-X (2016) The QTL GNP1 encodes GA20ox1, which increases grain number and yield by increasing cytokinin activity in rice panicle meristems. PLoS Genet 12(10):e1006386

Xu JL, Wang Y, Zhang F, Wu Y, Zheng T-Q, Wang Y-H, Zhao X-Q, Cui Y-R, Chen K, Zhang Q (2015a) SS1 (NAL1)- and SS2-mediated genetic networks underlying source-sink and yield traits in rice (Oryza sativa L.). PLoS ONE 10(7):e0132060

Xu R, Duan P, Yu H, Zhou Z, Zhang B, Wang R, Li J, Zhang G, Zhuang S, Lyu J (2018) Control of grain size and weight by the OsMKKK10-OsMKK4-OsMAPK6 signaling pathway in rice. Mol Plant 11(6):860–873

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nat Genet 40(6):761–767

Yang T, Amanullah S, Pan J, Chen G, Liu S, Ma S, Wang J, Gao P, Wang X (2021) Identification of putative genetic regions for watermelon rind hardness and related traits by BSA-seq and QTL mapping. Euphytica 217:1–18

Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P-c, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nat Genet 48(8):927–934

Yoshida A, Sasao M, Yasuno N, Takagi K, Daimon Y, Chen R, Yamazaki R, Tokunaga H, Kitaguchi Y, Sato Y (2013) TAWAWA1, a regulator of rice inflorescence architecture, functions through the suppression of meristem phase transition. Proc Natl Acad Sci 110(2):767–772

Yuan Y, Wei X, Zhang Q, Zhao Y, Jiang W, Yao Q 2015, BSA-Seq technologies identify a major QTL for clubroot resistance in Chinese cabbage (*Brassica rapa* ssp. pekinensis). In: Proceedings of the KSM Spring Meeting & KSM-ICWG-GSP Joint Clubroot Symposium, Daejeon, Korea, pp 13–15

Zhang ZY, Li JJ, Yao GX, Zhang HL, Dou HJ, Shi HL, Sun XM, Li ZC (2011) Fine mapping and cloning of the grain number per-panicle gene (Gnp4) on chromosome 4 in rice (Oryza sativa L). Agric Sci China 10(12):1825–1833

Zhang GH, Li SY, Wang L, Ye WJ, Zeng DL, Rao YC, Peng YL, Hu J, Yang YL, Xu J (2014) LSCHL4 from japonica cultivar, which is allelic to NAL1, increases yield of indica super rice 93–11. Mol Plant 7(8):1350–1364

Zhang B, Qi F, Hu G, Yang Y, Zhang L, Meng J, Han Z, Zhou X, Liu H, Ayaad M (2021a) BSA-seq-based identification of a major additive plant height QTL with an effect equivalent to that of Semi-dwarf 1 in a large rice F_2 population. The Crop Journal 9(6):1428–1437

Zhang C, Badri Anarjan M, Win KT, Begum S, Lee S (2021b) QTL-seq analysis of powdery mildew resistance in a Korean cucumber inbred line. Theor Appl Genet 134:435–451

Zhang Q, Xie J, Wang X, Liu M, Zhu X, Yang T, Khan NU, Sun C, Li J, Zhang Z (2022) Natural variation of RGN1a regulates grain number per panicle in japonica rice. Front Plant Sci 13:1097622

Zhang M, Lai L, Liu X, Liu J, Liu R, Wang Y, Liu J, Chen J (2022d) Overexpression of nitrate transporter 1/peptide gene OsNPF7 6 increases rice yield and nitrogen use efficiency. Life 12(12):1981

Zhao L, Tan L, Zhu Z, Xiao L, Xie D, Sun C (2015) PAY 1 improves plant architecture and enhances grain yield in rice. Plant J 83(3):528–536

Zhao L, Liu H, Peng K, Huang X (2023) Cold-upregulated glycosyltransferase gene 1 (OsCUGT1) plays important roles in rice height and spikelet fertility. J Plant Res 136(3):383–396

Zheng S, Li J, Ma L, Wang H, Zhou H, Ni E, Jiang D, Liu Z, Zhuang C (2019) OsAGO2 controls ROS production and the initiation of tapetal PCD by epigenetically regulating OsHXK1 expression in rice anthers. Proc Natl Acad Sci 116(15):7549–7558

Zhu M, Liu D, Liu W, Li D, Liao Y, Li J, Fu C, Fu F, Huang H, Zeng X (2017) QTL mapping using an ultra-high-density SNP map reveals a major locus for grain yield in an elite rice restorer R998. Sci Rep 7(1):10914