

ORIGINAL ARTICLE

Open Access



Dissection of the Genetic Basis of Rice Panicle Architecture Using a Genome-wide Association Study

Shaoxing Bai^{1†}, Jun Hong^{1†}, Ling Li¹, Su Su¹, Zhikang Li^{3,4}, Wensheng Wang^{3,5}, Fengli Zhang¹, Wanqi Liang¹ and Dabing Zhang^{1,2*}

Abstract

Panicle architecture is one of the major factors influencing productivity of rice crops. The regulatory mechanisms underlying this complex trait are still unclear and genetic resources for rice breeders to improve panicle architecture are limited. Here, we have performed a genome-wide association study (GWAS) to analyze and identify genetic determinants underlying three panicle architecture traits. A population of 340 rice accessions from the 3000 Rice Genomes Project was phenotyped for panicle length, primary panicle number and secondary branch number over two years; GWAS was performed across the whole panel, and also across the japonica and indica sub-panels. A total of 153 quantitative trait loci (QTLs) were detected, of which 5 were associated with multiple traits, 8 were unique to either indica or japonica sub-panels, while 37 QTLs were stable across both years. Using haplotype and expression analysis, we reveal that genetic variations in the *OsSPL18* promoter significantly affect gene expression and correlate with panicle length phenotypes. Three new candidate genes with putative roles in determining panicle length were also identified. Haplotype analysis of *OsGRRP* and *LOC_Os03g03480* revealed high association with panicle length variation. Gene expression of *DSM2*, involved in abscisic acid biosynthesis, was up-regulated in long panicle accessions. Our results provide valuable information and resources for further unravelling the genetic basis determining rice panicle architecture. Identified candidate genes and molecular markers can be used in marker-assisted selection to improve rice panicle architecture through molecular breeding.

Keywords: Rice, Panicle architecture, Panicle length, Genome-wide association study, Quantitative trait loci, Hormone, Natural variation

Background

Rice (*Oryza sativa* L.) is a staple food crop. It has been estimated that an additional 116 million tons of rice per year will be needed by 2035 to feed growing populations (Seck et al. 2012). Panicle size and architecture are key agronomic traits that greatly affect yield, so

understanding the molecular and genetic mechanisms underlying panicle development is of great importance to both plant biologists and modern breeders.

Panicle architecture has various quantitative characteristics, such as panicle length (PL), and the numbers of primary branches (PBN), secondary branches (SBN), and spikelets/grain, that affect yield (Crowell et al. 2016). Over the past two decades, a number of genes regulating panicle development have been identified and functionally characterized, e.g., *SP1* (Li et al. 2009), *DEP2* (Li et al. 2010; Zhu et al. 2010) and *DEP3* (Qiao et al. 2011) affect panicle length; and *OsSPL14* (Jiao et al. 2010; Miura et al. 2010), *ASPI* (Yoshida et al. 2012), and *APO1* (Terao et al.

*Correspondence: zhangdb@sjtu.edu.cn

[†]Shaoxing Bai and Jun Hong have contributed equally to this work

¹ Joint International Research Laboratory of Metabolic and Developmental Sciences, State Key Laboratory of Hybrid Rice, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

Full list of author information is available at the end of the article

2010) affect branch number. Many panicle trait genes are also pleiotropic, e.g., *MOC1* (Li et al. 2003), *LAX1* (Komatsu et al. 2001), and *OsSPL14* (Lu et al. 2013) affect not only branch number, but also tillering; while defects in *OsSPL18* decreased grain size and number and panicle length, but increased tillering (Yuan et al. 2019).

Panicle architecture QTLs commonly associate with other growth traits, in particular, heading date, e.g., *Hd1* (Zhang et al. 2019c), *EHD4* (Gao et al. 2013), *Ghd7* (Xue et al. 2008), *Ghd7.1* (Liu et al. 2013), and *DTH8* (Wei et al. 2010). Plant hormones can also play crucial roles in panicle development, and factors controlling hormone metabolism can affect panicle architecture, e.g., *Gn1a/OsCKX2* (Ashikari et al. 2005) and *LP/EP3* (Piao et al. 2009; Li et al. 2011) control spikelet number per panicle by regulating cytokinin; *OsGRF4* affects panicle architecture by mediating gibberellin and brassinosteroid responses (Che et al. 2015; Tang et al. 2018); while *OsPIN5b* determines panicle length by participating in auxin homeostasis (Lu et al. 2015).

With the advent of next-generation sequencing, genome-wide association studies (GWAS) that link genotype with phenotype in a natural population have rapidly become powerful gene/QTL mapping tools to detect complex agronomic traits, superseding the traditional time-consuming, imprecise QTL mapping techniques used in the past (Huang et al. 2010, 2012; Wang et al. 2018b). GWAS has been successfully used to dissect complex traits in multiple crop species, including barley (Cockram et al. 2010), maize (Wang et al. 2016), sorghum (Zhou et al. 2019), wheat (Luján Basile et al. 2019), potato (Okada et al. 2019), and cotton (Hinze et al. 2017). Dissection of QTL and identification of candidate genes associated with panicle development have recently been assisted by GWAS in rice. However, most of them have been conducted with low-density SNPs and do not have sufficient resolution to provide precise and complete information about the numbers and locations of the QTLs controlling the traits of interest (Reig-Valiente et al. 2018; Ta et al. 2018; Zang et al. 2015; Bai et al. 2016; Crowell et al. 2016; Rebolledo et al. 2016).

The recent 3000 Rice Genomes Project (3 K RGP) has provided a high-density single nucleotide polymorphisms (SNPs) database that is becoming popular for mining useful genetic information in rice using GWAS. For example, Jiang et al. (2020) detected 13 genetic loci related to development of leaf hairs; Zhang et al. (2019a,b,c) found 27 suggestively associated loci for sheath blight; while Shi et al. (2017) detected 22 significant salt tolerance-associated SNPs at the seed germination stage.

Here, we have used a panel of 340 diverse rice accessions selected from the 3 K RGP to perform GWAS on three panicle traits—PL, PBN, and SBN. We report 153

QTLs that significantly associate with variations in panicle architecture, and characterize the nucleotide diversity and expression levels of major *OsSPL18* haplotypes. In addition, we have used haplotype and gene expression analysis to identify three novel candidate genes (*OsGRRP*, *LOC_Os03g03480* and *DSM2*) that associate with panicle length. Our findings provide the basis for further elucidating mechanisms underlying panicle size and shape in rice.

Results

Panicle Phenotype Variation Across Rice Accessions

Panicle phenotype was analyzed in a panel of 340 japonica and indica accessions from the 3 K Rice Genomes Project with diverse genotypes, origins, and subpopulations. These rice varieties originate from 48 countries across Asia, Australia, North and South America, Europe, and Africa (Additional File 2: Table S1), and phylogenetic and principal component analyses confirmed that they fell into two main clades: indica (Xian, 161 accessions) and japonica (Geng, 179 accessions; Additional File 1: Figure S1a,b). The genome-wide linkage disequilibrium (LD) decay rates along chromosomes were found to be similar across the whole panel and indica and japonica sub-panels (Additional File 1: Figure S1c), and consistent with previous reports (Wang et al. 2018b).

Panicle length (PL), primary branch number (PBN), and secondary branch number (SBN) for the main panicle in all 340 varieties were evaluated for 2 growing seasons (2015 and 2017 in Hainan, China; Table 1). SBN displayed the largest phenotypic variation, with coefficients of variance (CVs) from 30–35%, while the PL and PBN traits were less variable. While all three traits exhibited normal distribution (Fig. 1a–c, Additional File 1: Figure S2a–c), there was an obvious difference between the two years. Panicles in 2015 were generally longer and more branched (Table 1), indicating that these quantitative traits are controlled by genetic and environmental factors. In general, panicle architecture traits showed a high heritability, with broad-sense heritability (H^2) from 75 to 84% in 2017 samples. These populations thus exhibited extensive variation in panicle architecture traits strongly linked to genotype, making them suitable for genome-wide association studies (GWAS).

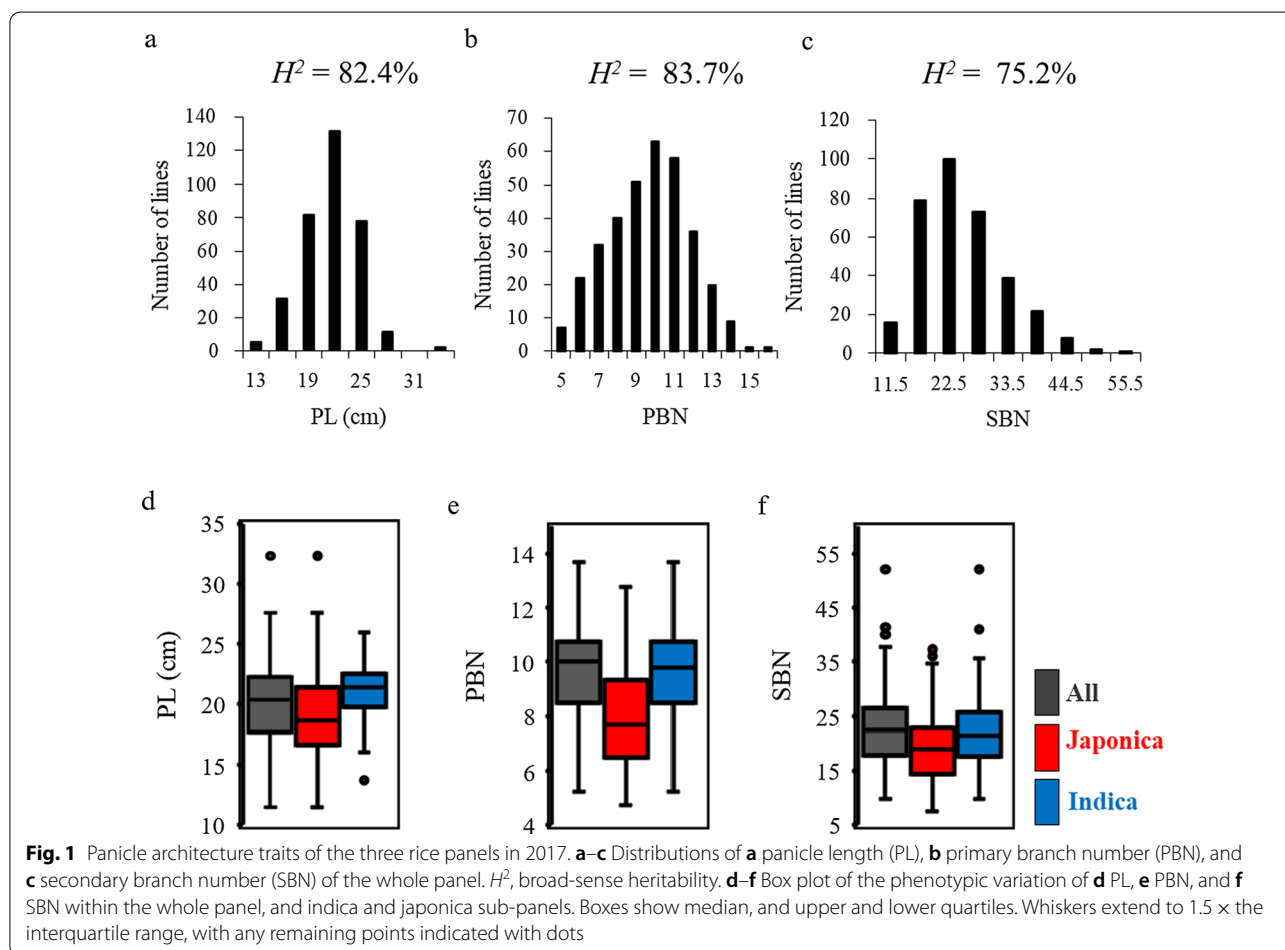
Analysis of the japonica and indica sub-panels revealed that values for the three panicle traits largely overlapped, though values for indica varieties were generally higher than for japonica varieties (Fig. 1d–f, Additional File 1: Figure S2d–f), consistent with previous reports (Ta et al. 2018). The three panicle traits only weakly correlated with one another in each year (Additional File 1: Figure S3a, b). Generally, correlations were strongest between PBN and SBN, as secondary branches originate

Table 1 Phenotypic variations of panicle traits

| Trait | 2015 | | | | 2017 | | | |
|------------------------|------------|------|------|----|------------|------|------|----|
| | Mean | Min | Max | CV | Mean | Min | Max | CV |
| Panicle length (cm) | 22.1 ± 3.3 | 13.5 | 37.1 | 15 | 20.1 ± 3.2 | 11.5 | 32.4 | 16 |
| Primary branches (#) | 10.7 ± 2.0 | 6.0 | 17.0 | 19 | 9.0 ± 2.2 | 4.8 | 16.0 | 24 |
| Secondary branches (#) | 29.0 ± 9.0 | 8.0 | 61.0 | 30 | 22.0 ± 7.8 | 7.6 | 52.0 | 35 |

Branch numbers per panicle; mean ± SD of 3 replicates

CV coefficient of variation (%)

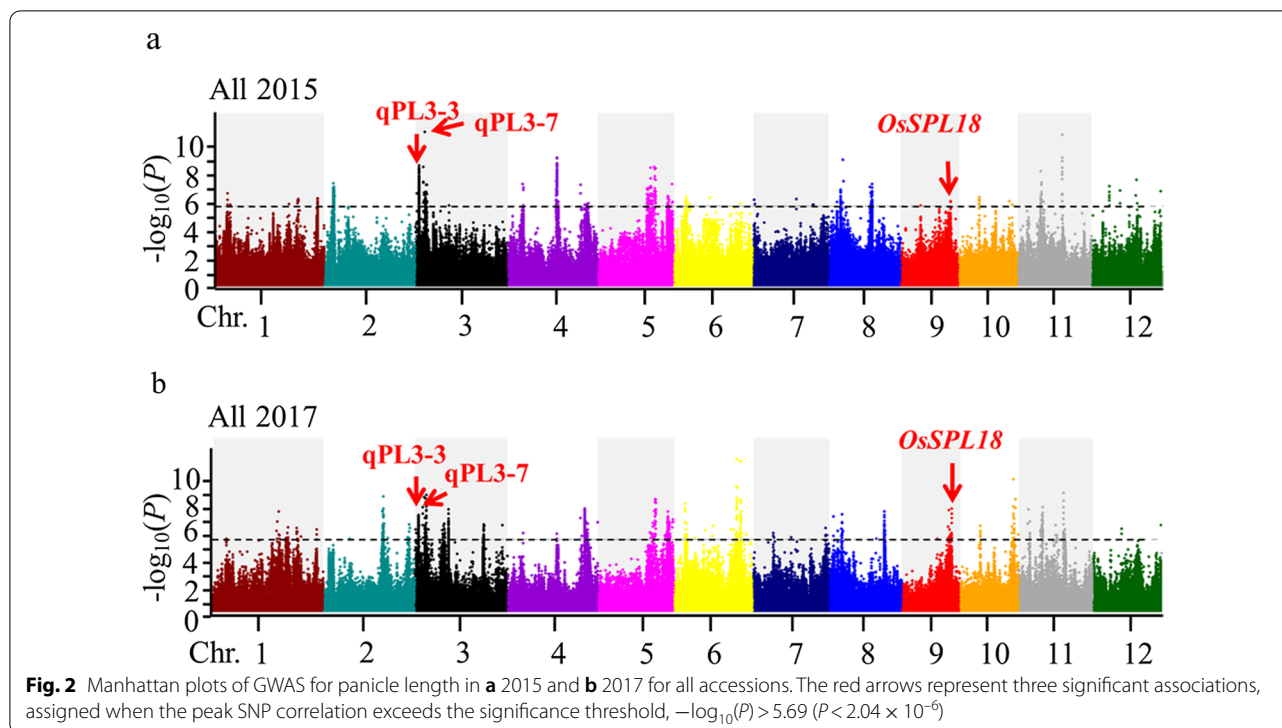


from primary branches; with the next highest correlation between PL and PBN, again, as primary branches come from the rachis of the panicle.

Association Mapping of Three Panicle Architecture Traits

Across the three traits in two years, 1087 significant SNPs yielded 153 QTLs in 148 genomic regions across the three rice panels (Fig. 2, Additional File 1: Figures S4–S6 and Additional File 2: Table S2). GWAS mapping resulted in 129 individual QTLs (115 for PL, 12 for PBN, and 2 for

SBN) for the whole panel; 11 QTLs (7 for PL, 1 for PBN, and 3 for SBN) for the japonica sub-panel; and 21 QTLs (6 for PL, 6 for PBN, and 9 for SBN) for the indica panel. Only 5 QTLs (qPL3-14/qPBN3-1, qPL5-3/qPBN5-1, qPL5-16/qPBN5-2, qPL6-1/qPBN6-1, qPL11-3/qSBN11-1) were common across more than one trait. Intriguingly, there were no common QTLs between PBN and SBN, suggesting different regulatory mechanisms underlie these two levels of panicle branching. Only 37 QTLs (31 for PL and 6 for SBN) were stable across 2015 and 2017,



suggesting that these loci contain genes that function in panicle architecture without the influence of environmental factors (Additional File 2: Table S3).

Among the 37 stable QTLs, comparison with previous literature revealed overlaps with 8 known genes, and 10 QTLs from other GWAS analyses and QTARO (Additional File 2: Tables S3 and S4, respectively). The remaining 19 QTLs (15 for PL and 4 for SBN) are the novel, and thus we did further analysis (Additional File 2: Table S5).

OsSPL18, the Candidate Gene of qPL9-6

One stable QTL, qPL9-6, was found to contain *OsSPL18* gene, which had previously been reported to affect panicle length and grain number via new regulatory pathways (Yuan et al. 2019); but the influence of genetic sequence diversity within this gene was not examined. We analyzed SNPs in the 2 kb upstream promoter and genomic *OsSPL18* coding region to reveal six major haplotypes. Indica varieties generally contained a higher proportion of HapB, HapC, and HapD types, while japonica varieties contained a higher proportion of HapA, HapE, and HapF sequences (Fig. 3a). The panicle length of accessions containing HapE and HapF was significantly lower than for other haplotypes (Fig. 3b). Therefore, this result is useful for clarifying the panicle length variation within japonica varieties because haplotype E and F exhibited smaller panicle length than other haplotypes and were only found in japonica group except three indica accessions (Fig. 3a).

A large proportion of SNPs were identified in the promoter and introns compared with the coding sequence, suggesting that these genetic variations may impact *OsSPL18* gene expression. Gene expression analysis from rachis revealed that HapE varieties had significantly higher levels of *OsSPL18* expression than HapA/B/C/D varieties; HapF expression levels were intermediate, and not significantly different to either of the other groups (Fig. 3c, Additional file 2: Table S6). In addition, haplotype analysis showed that two non-synonymous SNPs variation (Chr9_19647656, base G-T, amino acid Gly-Val, Chr9_19647782, base C-T, amino acid Thr-Met) were also significantly associated with panicle length (data not shown). Therefore, both polymorphisms in the coding and promoter regions of *OsSPL18* may cause the variation of panicle length.

OsGRRP, the Candidate Gene of qPL3-7

The lead SNP (Chr3_3720893(T/C), $P = 1.1 \times 10^{-11}$ in 2015 and 1.19×10^{-09} in 2017) in qPL3-7 caused a missense mutation from cysteine to arginine in the first exon of *OsGRRP* (*LOC_Os03g07310*), predicted to encode a growth regulator-related protein (Fig. 4a–b). Panicle length correlated significantly with gene haplotype at this location in the whole panel and japonica sub-panel in both years. In the indica sub-panel, there was no difference in the panicle length of two haplotype (Fig. 4c, Additional file 1: Figure S7). However, an obvious difference in

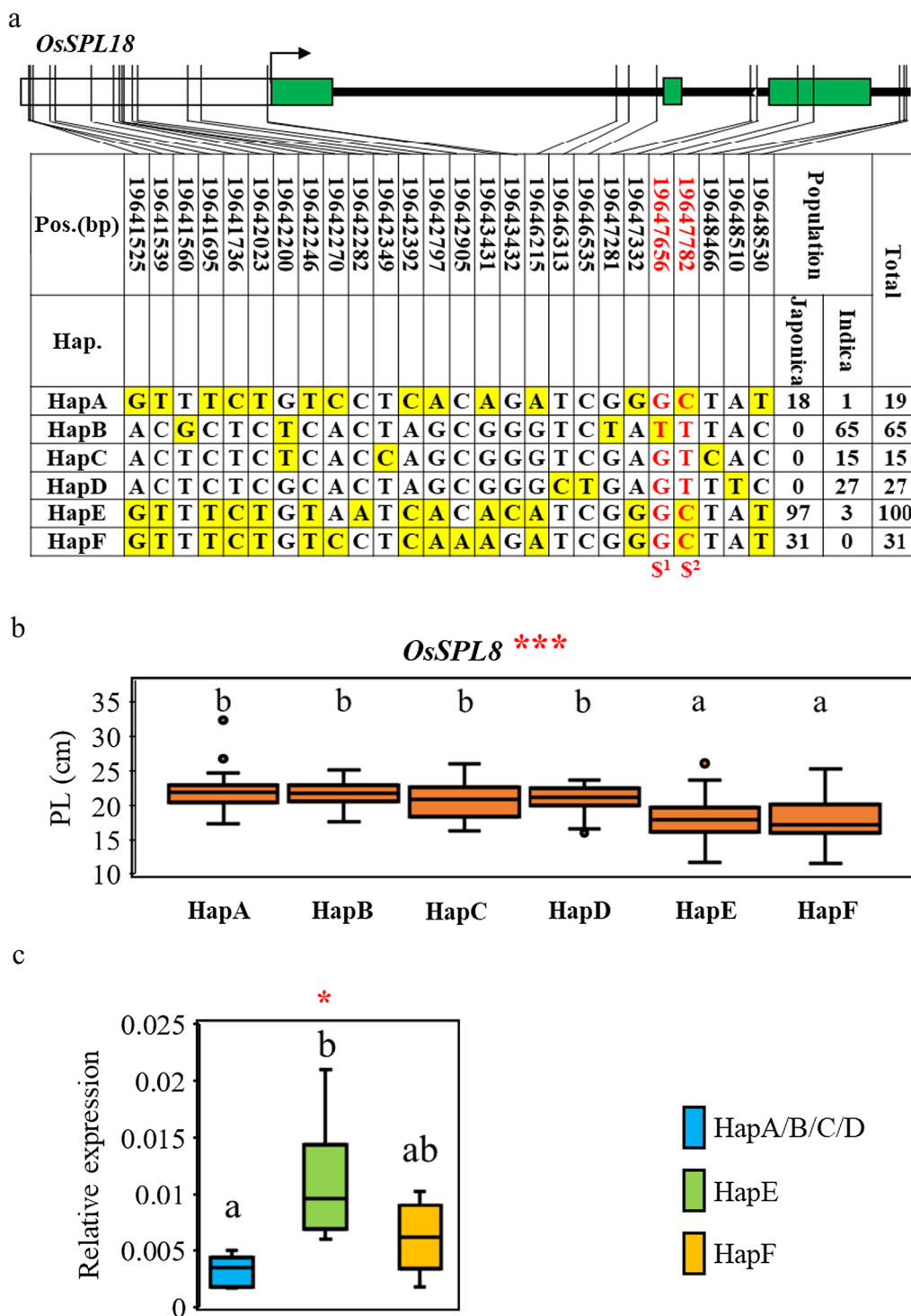
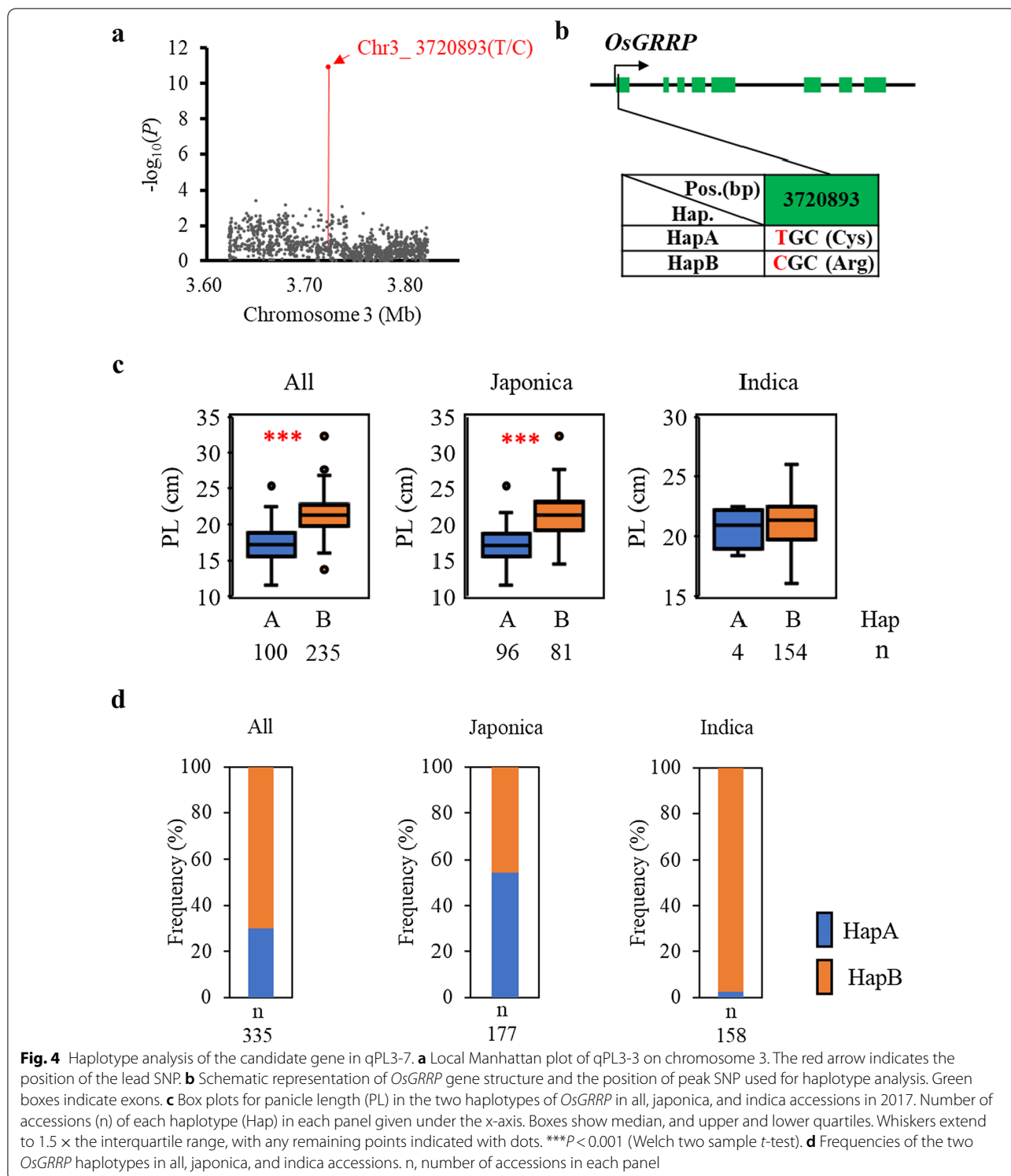


Fig. 3 Haplotype analysis of *OsSPL18*. **a** Six haplotypes of *SPL18* based on 26 SNPs observed in all assessed rice accessions. The schematic representation of *SPL18* gene structure (upper) shows the promoter as a white box, exons as green boxes, and introns and intergenic regions as black lines. Thin black lines indicate the genomic position of each SNP. Haplotypes with fewer than 10 accessions are not shown. Yellow highlight indicates SNP alternatives. The SNP in red and bold is a non-synonymous SNP. '\$¹' indicates a missense mutation from Gly to Val. '\$²' indicates a missense mutation from Thr to Met **b-c** Box plots for **b** panicle length based on the six haplotypes for *SPL18* in 2017 and **c** *OsSPL18* expression levels in different haplotypes relative to *OsActin1*. Boxes show median, and upper and lower quartiles. Whiskers extend to 1.5 × the interquartile range, with any remaining points indicated with dots. **P* < 0.05, ****P* < 0.001 (ANOVA). Letters indicate significant differences, *P* < 0.05 (Duncan's multiple comparison test)



allele frequency was observed between the japonica and indica varieties: nearly all (97%) of the indica lines contained the long panicle HapB allele, compared with 46% of the japonica varieties (Fig. 4d). This genetic difference

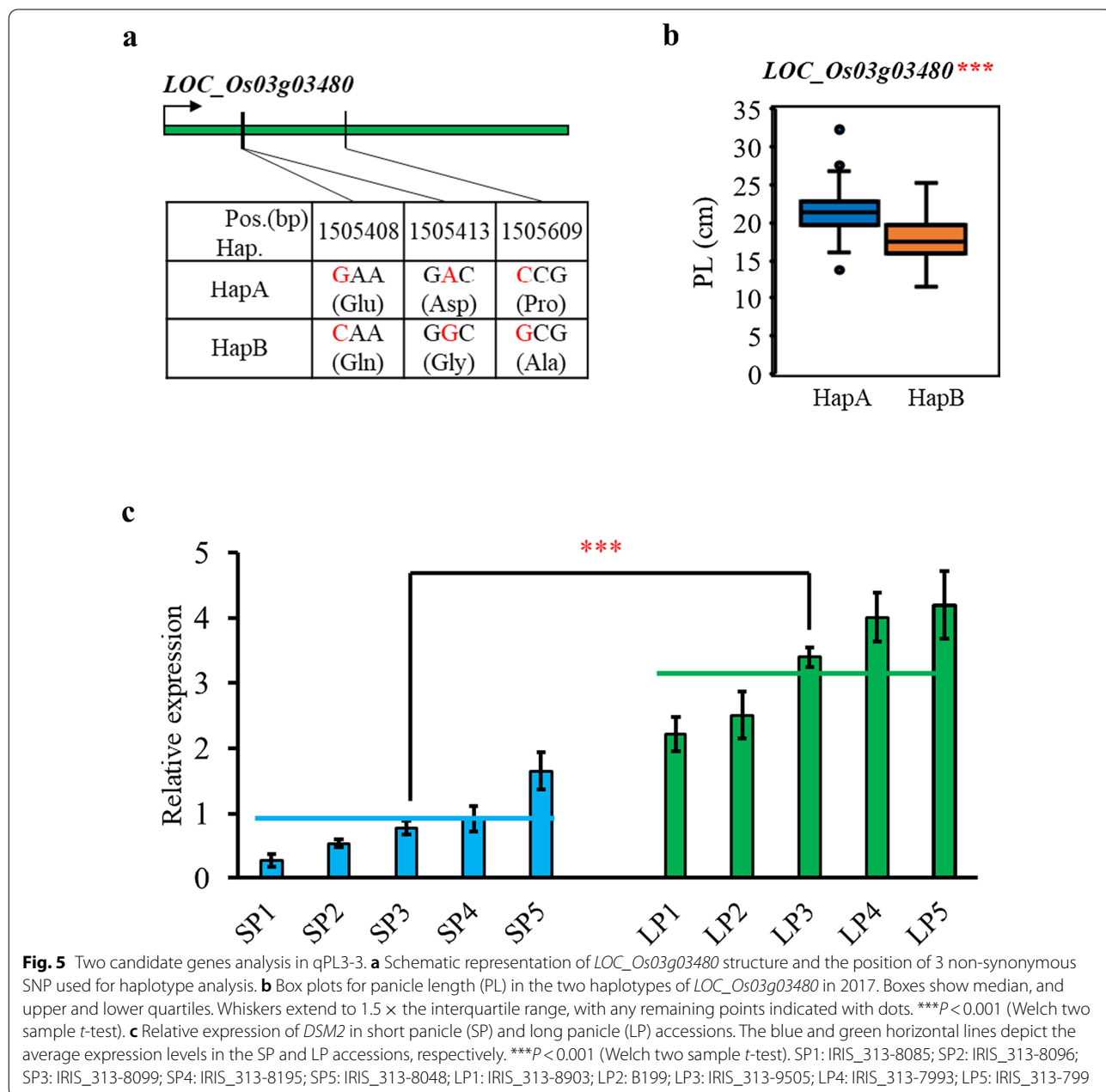
may partially explain the longer average panicle length in indica varieties (Ta et al. 2018).

Two candidate Genes of qPL3-3

qPL3-3 was especially attractive, as it contained the highest number of SNPs (444 and 291 in 2015 and 2017, respectively; Additional File 2: Table S5). After removal of genes encoding hypothetical proteins, retrotransposons, and transposon proteins, 28 candidate genes were identified (Additional File 2: Table S7).

We observed 23 stable non-synonymous SNPs in the coding region of 6 genes detected in two years (Additional File 2: Table S7). *LOC_Os03g03480*, encoding a DUF623 domain containing protein caught more

attention. Previous studies have demonstrated that these kinds of proteins are involved in panicle length (Schmitz et al. 2015; Yang et al. 2018). Two types of haplotypes based on 3 non-synonymous SNPs and panicle length correlated significantly with gene haplotypes (Fig. 5a, b), suggesting that *LOC_Os03g03480* might be a candidate gene involved in panicle length. Besides these 6 genes, *OsSWN5* and *LOC_Os03g03260* in qPL3-3 were of interest as other genes annotated as “no apical meristem protein” and “homeobox domain containing protein” (*LOC_Os01g47710* and *LOC_Os05g48990*) had



previously been shown to have a role in controlling panicle length in rice (Wang et al. 2014; Sakuraba et al. 2015; Fang et al. 2020), so *OsSWN5* and *LOC_Os03g03260* in qPL3-3 were of interest. Similarly, panicle length is sensitive to genes involved in plant hormone biosynthesis or signaling, including ABA (Wang et al. 2015b; Hong et al. 2016; Zhang et al. 2016), so *DSM2* (*LOC_Os03g03370*), which encodes a β -carotene hydroxylase involved in the biosynthesis of the ABA precursor zeaxanthin (Du et al. 2010), was also a gene of interest in qPL3-3. These 3 genes were thus selected as candidates for expression analysis in a small group of 5 short panicle (SP) and 5 long panicle (LP) accessions (Fig. 5b, Additional file 1: Figure S8). Only *DSM2* was differentially expressed between SP and LP varieties, being more highly expressed in the LP rather than in SP lines. Thus, expression of *DSM2* was positively associated with panicle length, and also likely accounts for the association of qPL3-3 with panicle length.

Discussion

Rice panicle size and architecture are important factors for rice yield. Panicle architecture is a complex quantitative trait controlled by both genes and environment (Cubry et al. 2020). The diversity of genetic variation in natural populations is an important potential source of beneficial alleles. Here, we have performed GWAS on 340 rice accessions selected from 3 K RGP to identify new genes underlying panicle architecture traits.

Phenotypic Variations for Targeted Traits

A high heritability was observed in the three analyzed panicle architecture traits, indicating that their diversity was more under genetic than environmental control in these accessions (Fig. 1). Moreover, similar trends of positive correlation between these three traits were observed in the whole panel and indica and japonica sub-panels (Additional File 1: Figure S3), consistent with the GWAS results in which some QTLs could be detected in two years or associated with multiple traits (Additional File 2: Table S2). However, despite the high correlation between PBN and SBN, there was no single SNP that associated with both PBN and SBN, suggesting that the mechanisms related to their establishment and function are regulated by discrete genetic determinants (Ta et al. 2018).

A small number of GWAS sites in this study were detected only in indica or japonica sub-panels (Additional File 2: Table S2). Differences between panicle architecture between japonica and indica subspecies have previously been observed, driven in part by artificial human selection and geographic regions of cultivation, such as *Ghd7*, evolved from two distinct ancestral gene pools (Xue et al. 2008). Another gene from japonica cultivars, *LSCHL4* (allelic to *NAL1*), was shown to enhance

grain productivity when introduced into indica cultivars through pleiotropic effects on plant architecture (Zhang et al. 2014). These results suggest that there are different genetic networks regulating panicle architecture variation in indica and japonica sub-panels.

GWAS-Based Mapping of PL, PBN, and SBN QTLs

GWAS is an efficient technique to analyze genetic variation for multiple traits in rice. Several studies employed GWAS in the rice population for various panicle architecture traits and reported some novel QTLs (Zhang et al. 2015; Sahu et al. 2020), but such study in diverse rice landrace is limited. In the present study, our results yielded 153 QTLs for panicle architecture traits. Of these, 5 were associated with more than one trait, which suggests that these loci contain pleiotropic genes (Additional File 2: Table S2). 37 QTLs could be detected over 2 years (Additional File 2: Table S3), indicating these QTLs in rice were stable against environmental factors and supporting the notion of panicle architecture depends not only on genetic diversity but also the environmental factors. Among the 37 stable QTLs, we observed that 8 genes were known to be involved in panicle architecture (Additional File 2: Table S3), suggesting that our results are of high reliability. For example, *SPOTTED LEAF3*(*SPL3*), which is involved in ABA and ethylene signalling pathways (Wang et al. 2015a), was co-localized with qPL3-6. *LAX2*, which is involved in maintenance of apical meristems in rice, including in the panicle (Tabuchi et al. 2011), was found in qPL4-3. 10 QTLs from other studies and the remaining 19 QTLs are reported here for the first time, and represent promising targets for further analysis of panicle architecture (Additional File 2: Table S5). Overall, the QTLs detected in this study indicated that a number of as yet unknown factors may be involved in the determination of panicle architecture and helpful in molecular marker-assisted selection of rice panicle shape breeding.

Candidate Gene Analysis for Three QTLs

In rice, the SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (*SPL*) family proteins play important roles during panicle development. To date, five known *SPL* genes (*OsSPL6*, *OsLGI/OsSPL8*, *OsSPL13*, *IPA1/WFP/OsSPL14*, and *OsSPL18*) have been reported to regulate panicle architecture in rice (Jiao et al. 2010; Miura et al. 2010; Ishii et al. 2013; Zhu et al. 2013; Si et al. 2016; Wang et al. 2018a; Yuan et al. 2019). *OsSPL18* was located in qPL9-6, a QTL associated with panicle length, and we have used haplotype and expression analysis to show that nucleotide polymorphisms in the *OsSPL18* promoter significantly affect gene expression and panicle length (Fig. 3). Haplotypes with the desired phenotype may be

suitable for traditional rice breeding and are also good targets for molecular marker-assisted selection breeding.

Analysis of the leading SNP in *OsGRRP* in qPL3-7 revealed two haplotypes (Fig. 4a, b). Significant differences in mean panicle length were associated with haplotype in the whole panel and japonica accessions, with HapB directing the longer phenotype (Fig. 4c), indicating that *OsGRRP* may be responsible for panicle length variation. Differences in haplotype frequency in the two subspecies, with 97% of indica accessions containing HapB (Fig. 4d), suggests that *OsGRRP* has been subjected to selection during rice breeding and may have contributed to differentiation of the two subspecies.

Among the QTLs for panicle length that were the same across both years, we selected qPL3-3 for further analysis of causal genes due to high SNP concentrations (Additional File 2: Table S5). We, firstly, based on functional annotation and haplotype analysis, proposed that *LOC_Os03g03480*, a DUF623 domain containing protein, is one of candidate gene in qPL3-3. Furthermore, three candidate genes were selected for expression analysis based on their functional annotation (Additional File 2: Table S7). Expression of *DSM2*, a fatty acid hydroxylase involved in ABA synthesis, was found to correlate with panicle length (Fig. 5), indicating that during the rice panicle development, the altered expression of enzyme genes related to the ABA synthesis pathway contributes to the rice panicle length determination, consistent with the previous results (Wang et al. 2015b; Zhang et al. 2016).

Natural Variations of Candidate Genes

Natural variation in genes involved in the regulation of important agronomic traits has been utilized by breeders to improve rice yield. Although dozens of genes related to panicle architecture have been identified using multiple strategies, only a few reported genes, like *NALI*, have been applied in rice breeding (Fujita et al. 2013; Zhang et al. 2014).

GWAS enables the exploration of the large of alleles present in genetic resources and that are useful in rice improvement. The differential expression profile strategy combination of GWAS is a powerful method for target-gene mining. In this study, Expression analysis of candidate genes revealed two genes, *DSM2* and *OsSPL18*, that were up- and down-regulated, respectively, in long panicle accessions. These results suggest that transcriptional regulation of the candidate genes may through SNP variations in promoter region and then contribute to panicle length variation in rice. Modification of the core promoters of targeted genes by genome editing is a widely applicable and reliable approach for the fine-tuning of the expression of target genes (Huang et al. 2020; Zeng et al.

2020), so the further studies to understand the mechanisms involved in transcriptional regulation of these two genes will be conducted. Moreover, haplotype analysis was carried out for two genes (*OsGRRP* and *LOC_Os03g03480*) using significant non-synonymous SNPs located inside of the gene CDS region, especially a non-synonymous SNP (Chr3_3720893(T/C)) variation was observed in *OsGRRP*, which resulted in an amino-acid substitution from cysteine to arginine (Fig. 4), was most likely to be the functional site. Overall, further functional analysis of these candidate gene will help us better understand the genetic basis for natural variation in rice panicle architecture control and suggest that these genes may be used to improve the rice yield.

Conclusions

In our study, we have dissected the genetic basis underlying differences in rice panicle architecture traits in indica and japonica subspecies. We identified natural variation in the promoter of *OsSPL18* that affects gene expression level and panicle length. Using gene functional annotation, haplotype analysis, and expression analysis after GWAS, we have identified three novel candidate genes associated with panicle length. Genetic complementation, overexpression, and knockout studies of these three candidates will clarify their role in directing rice panicle length, and will be a focus for further studies. More broadly, other SNPs and genes reported here could be used for future research, gene validation, and marker-assisted selection for molecular breeding of rice with enhanced panicle architecture.

Materials and Methods

Plant Materials and Growing Conditions

A panel of 340 rice varieties were selected from the 3 K Rice Genomes Project, comprising 161 indica (Xian) and 179 japonica (Geng) accessions (Wang et al. 2018b). Detailed information regarding these accessions, including their geographical origin, is shown in Additional File 2: Table S1. All accessions were grown in the experimental fields around Sanya (Hainan province, China; 18°15'N, 109°30'E) from January to April in 2015 and 2017. Six plants of each variety were planted in a row with 17 cm between plants and 20 cm between rows. Panicle traits were measured only from the main tiller in each plant.

Population Genetic Analyses

3,214,392 SNPs (all accessions), 923,587 SNPs (japonica accessions), and 1,501,869 SNPs (indica accessions) with a minor allele frequency (MAF) of $\geq 5\%$ and a missing rate of $\leq 20\%$ were selected for population and association analyses.

K-estimation ($K=2$) was based on the subgroups of all accessions (Wang et al. 2018b). PCA were performed with GCTA software (<https://cns.genomics.com/software/gcta/#Overview>) and neighbor-joining trees were constructed with SNPhylo based on SNPs in all accessions (Lee et al. 2014).

Linkage disequilibrium (LD) measures between SNP loci at the individual chromosome level were plotted by evaluating r^2 (squared allele frequency correlation) estimators between significant marker(s) using the PopLD-decay (Zhang et al. 2019a).

Data Analysis

The three panicle related traits were measured as previously described (Crowell et al. 2016). Statistical analyses were performed in Excel 2010 (Microsoft) and SPSS software (version 18.0). Welch two sample t -test were used to analyze differences between the three rice panels. Panicle length and expression data of *OsSPL18* haplotypes were analyzed using ANOVA and Duncan's multiple test in the SPSS software.

Pearson's correlations were used to examine the correlations between traits in 2015 and 2017, and the broad-sense heritability (H_B^2) was calculated in 2017 to describe how each trait was affected by the environment as follows: $H^2 = \text{var}_{(G)} / (\text{var}_{(G)} + \text{var}_{(E)})$, where $\text{var}_{(G)}$ and $\text{var}_{(E)}$ are the genotypic and experimental variance, respectively.

GWAS for Panicle Architecture

The 3 K RGP 4.8mio SNP dataset was downloaded from the Rice SNP-Seek Database (<http://snp-seek.irri.org/>; Alexandrov et al. 2015). GWAS was performed on SNPs as described above using the factored spectrally transformed linear mixed models (FaST-LMM; Lippert et al. 2011). The significance threshold for the identification of QTLs was set to $P < 2.04 \times 10^{-6}$, and the SNP with the minimum P value was considered the lead SNP. Manhattan plots for the GWAS results were drawn using the R package 'qqman' (<https://www.r-project.org/>).

Haplotype Analysis and Identification of Candidate Genes

Candidate genes were scanned within the 200 kb region centered on the lead SNP of each QTL (using the reference Nipponbare genome (<http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/>)). The 2019 QTARO and MSU databases (<http://qtaro.abr.affrc.go.jp> and <http://rice.plantbiology.msu.edu>) were used to identify previously reported QTLs/genes present in the LD region. Candidate genes were identified based on predicted function from the rice genome annotation project.

All SNPs in the promoter region (2 kb) of *OsSPL18* and non-synonymous SNPs in *OsSPL18* exons, introns, and

3' untranslated region were selected from the Rice SNP-Seek Database (<https://snp-seek.irri.org/>).

Gene Expression Analysis

Rachis from 5 short panicle (SP) and 5 long panicle (LP) accessions, grown in the paddy field of Shanghai JiaoTong University, were collected in triplicate at the stage of 50% panicle emergence from the leaf sheath based on the PL in 2017. Average PL for SP accessions was 12.4 cm; for LP accessions, 25.9 cm in 2017.

Total RNA was extracted using the Trizol reagent (Invitrogen), following the manufacturer's instructions. cDNA was synthesized from total RNA using the Fast-Quant RT Kit (with gDNase; Tiangen). Quantitative real time PCR (qRT-PCR) was performed in a two-step reaction using SuperReal PreMix Color (SYBR Green; Tiangen) on a Roche Light Cycler 2.10 system using the $2^{-\Delta\Delta C_t}$ method (Livak and Schmittgen 2001) with three technical replicates. Expression levels were normalized to *OsActin1* (LOC_Os03g50885).

The sequences of the candidate genes were downloaded from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/analyses_search_locus.shtml). Primer sequences for candidate genes were downloaded from the qPrimerDB-qPCR Primer Database (<https://biodb.swu.edu.cn/qprimerdb/best-primers-ss>) except for *OsSWN5*, for which primers were designed by NCBI. Primer sequences are given in Additional File 2: Table S8.

Abbreviations

ABA: Abscisic acid; CV: Coefficient of variance; GWAS: Genome-wide association study; H^2 : Broad-sense heritability; Hap: Haplotype; LP: Long panicle; n: Number; PBN: Primary branch number; PL: Panicle length; QTL: Quantitative trait locus; SBN: Secondary branch number; SNP: Single nucleotide polymorphism; SP: Short panicle.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12284-021-00520-w>.

Additional file 1. Figure S1: Population structure of 340 rice accessions, comprising 161 indica and 179 japonica varieties. **Figure S2:** Panicle architecture traits of the three rice panels in 2015. **Figure S3:** Correlations between PL, PBN, and SBN from each GWAS panel in 2015 and 2017. **Figure S4:** Manhattan plots of GWAS for panicle length. **Figure S5:** Manhattan plots of GWAS for primary branch number. **Figure S6:** Manhattan plots of GWAS for secondary branch number. **Figure S7:** Panicle length in the two haplotypes of *OsGRRP* in 2015. **Figure S8:** Relative expression of *OsSWN5* and *LOC_Os03g03260* in SP and LP accessions.

Additional file 2. Table S1: Association panel accessions and phenotypes. **Table S2:** Significantly associated loci based on GWAS for PL, PBN and SBN in the whole panel, and the japonica and indica sub-panels, grown in 2015 and 2017. **Table S3:** 37 QTLs stable across 2015 and 2017. **Table S4:** Colocalization of detected associations with known GWAS-derived associations and the QTARO database. **Table S5:** 19 newly identified QTLs in our study. **Table S6:** Accessions of different *OsSPL18* haplotypes for whole panel and qRT-PCR. **Table S7:** Genes in qPL3-3. **Table S8:** Primers used in this study.

Acknowledgements

We sincerely thank Miss Mingjiao Chen and Mr Zhijing Luo for rice plant growth and field work; Jin Shi for advice; and Dr Natalie Betts for editing this manuscript.

Authors' contributions

SB and JH performed the experiments. LL participated in data analysis. SS participated in field practice. SB wrote the manuscript. ZL and WW provide all the rice varieties. DZ designed and supervised this study. The final manuscript was read and approved by all authors. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (31861163002 and 31970803), the Innovative Research Team, Ministry of Education, and 111 Project (B14016).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Joint International Research Laboratory of Metabolic and Developmental Sciences, State Key Laboratory of Hybrid Rice, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China. ²School of Agriculture, Food, and Wine, University of Adelaide, Adelaide, SA 5064, Australia. ³Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China. ⁴Shenzhen Institute for Innovative Breeding, Chinese Academy of Agricultural Sciences, Shenzhen, China. ⁵College of Agronomy, Anhui Agricultural University, Hefei, China.

Received: 22 June 2021 Accepted: 26 August 2021

Published online: 06 September 2021

References

- Ashikari M, Hitoshi S, Shaoyang L et al (2005) Cytokinin oxidase regulates rice grain production. *Soil Sci Plant Nutr* 309:433–444. <https://doi.org/10.1080/00380768.2003.10410029>
- Bai X, Zhao H, Huang Y et al (2016) Genome-wide association analysis reveals different genetic control in panicle architecture between indica and Japonica Rice. *Plant Genome* 9:1–10. <https://doi.org/10.3835/plantgenome2015.11.0115>
- Che R, Tong H, Shi B et al (2015) Control of grain size and rice yield by GL2-mediated brassinosteroid responses. *Nat Plants* 1:1–7. <https://doi.org/10.1038/nplants.2015.195>
- Cockram J, White J, Zuluaga DL et al (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc Natl Acad Sci USA* 107:21611–21616. <https://doi.org/10.1073/pnas.1010179107>
- Crowell S, Korniliev P, Falcão A et al (2016) Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat Commun* 7:1–14. <https://doi.org/10.1038/ncomms10527>
- Cubry P, Pidon H, Ta KN et al (2020) Genome wide association study pinpoints key agronomic QTLs in African Rice *Oryza glaberrima*. *Rice* 13:1–12. <https://doi.org/10.1186/s12284-020-00424-1>
- Du H, Wang N, Cui F et al (2010) Characterization of the β -carotene hydroxylase gene DSM2 conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. *Plant Physiol* 154:1304–1318. <https://doi.org/10.1104/pp.110.163741>
- Fang F, Ye S, Tang J et al (2020) DWT1/DWL2 act together with OsPIP5K1 to regulate plant uniform growth in rice. *New Phytol* 225:1234–1246. <https://doi.org/10.1111/nph.16216>
- Fujita D, Trijatmiko KR, Tagle AG et al (2013) NAL1 allele from a rice landrace greatly increases yield in modern indica cultivars. *Proc Natl Acad Sci USA* 110:20431–20436. <https://doi.org/10.1073/pnas.1310790110>
- Gao H, Zheng XM, Fei G et al (2013) Ehd4 Encodes a novel and oryza-genus-specific regulator of photoperiodic flowering in rice. *PLoS Genet*. <https://doi.org/10.1371/journal.pgen.1003281>
- Hinze LL, Hulse-Kemp AM, Wilson IW et al (2017) Diversity analysis of cotton (*Gossypium hirsutum* L.) germplasm using the CottonSNP63K Array. *BMC Plant Biol* 17:1–20. <https://doi.org/10.1186/s12870-017-0981-y>
- Hong Y, Zhang H, Huang L et al (2016) Overexpression of a stress-responsive NAC transcription factor gene ONAC022 improves drought and salt tolerance in rice. *Front Plant Sci* 7:1–19. <https://doi.org/10.3389/fpls.2016.00004>
- Huang L, Li Q, Zhang C et al (2020) Creating novel Wx alleles with fine-tuned amylose levels and improved grain quality in rice by promoter editing using CRISPR/Cas9 system. *Plant Biotechnol J* 18:2164–2166. <https://doi.org/10.1111/pbi.13391>
- Huang X, Wei X, Sang T et al (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961–967. <https://doi.org/10.1038/ng.695>
- Huang X, Zhao Y, Wei X et al (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 44:32–39. <https://doi.org/10.1038/ng.1018>
- Ishii T, Numaguchi K, Miura K et al (2013) OsLG1 regulates a closed panicle trait in domesticated rice. *Nat Genet* 45:462–465. <https://doi.org/10.1038/ng.2567>
- Jiang HZ, Lin HY, Zhang B, et al (2020) Identification of genetic loci for leaf hair development in rice through genome-wide association study. *Plant Growth Regul* 90:101–108. <https://doi.org/10.1007/s10725-019-00550-x>
- Jiao Y, Wang Y, Xue D et al (2010) Regulation of OsSPL14 by OsMIR156 defines ideal plant architecture in rice. *Nat Genet* 42:541–544. <https://doi.org/10.1038/ng.591>
- Komatsu M, Maekawa M, Shimamoto K, Kyojuka J (2001) The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Dev Biol* 231:364–373. <https://doi.org/10.1006/dbio.2000.9988>
- Lee TH, Guo H, Wang X et al (2014) SNPPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* 15:1–6. <https://doi.org/10.1186/1471-2164-15-162>
- Li F, Liu W, Tang J et al (2010) Rice dense and erect panicle 2 is essential for determining panicle outgrowth and elongation. *Cell Res* 20:838–849. <https://doi.org/10.1038/cr.2010.69>
- Li M, Tang D, Wang K et al (2011) Mutations in the F-box gene LARGER PANICLE improve the panicle architecture and enhance the grain yield in rice. *Plant Biotechnol J* 9:1002–1013. <https://doi.org/10.1111/j.1467-7652.2011.00610.x>
- Li S, Qian Q, Fu Z et al (2009) Short panicle1 encodes a putative PTR family transporter and determines rice panicle size. *Plant J* 58:592–605. <https://doi.org/10.1111/j.1365-3113.2009.03799.x>
- Li X, Qian Q, Fu Z et al (2003) Control of tillering in rice. *Nature* 422:618–621. <https://doi.org/10.1038/nature01518>
- Lippert C, Listgarten J, Liu Y et al (2011) FaST linear mixed models for genome-wide association studies. *Nat Methods* 8:833–835. <https://doi.org/10.1038/nmeth.1681>
- Liu T, Liu H, Zhang H, Xing Y (2013) Validation and characterization of Ghd7.1, a major quantitative trait locus with pleiotropic effects on spikelets per panicle, plant height, and heading date in rice (*oryza sativa* L.). *J Integr Plant Biol* 55:917–927. <https://doi.org/10.1111/jipb.12070>
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ $\Delta\Delta$ CT method. *Methods* 25:402–408. <https://doi.org/10.1006/meth.2001.1262>
- Lu G, Coneva V, Casaretto JA et al (2015) OsPIN5b modulates rice (*Oryza sativa*) plant architecture and yield by changing auxin homeostasis, transport and distribution. *Plant J* 83:913–925. <https://doi.org/10.1111/tpl.12939>

- Lu Z, Yu H, Xiong G et al (2013) Genome-wide binding analysis of the transcription activator IDEAL PLANT ARCHITECTURE1 reveals a complex network regulating rice plant architecture. *Plant Cell* 25:3743–3759. <https://doi.org/10.1105/tpc.113.113639>
- Luján Basile SM, Ramírez IA, Crescente JM et al (2019) Haplotype block analysis of an Argentinean hexaploid wheat collection and GWAS for yield components and adaptation. *BMC Plant Biol* 19:1–16. <https://doi.org/10.1186/s12870-019-2015-4>
- Miura K, Ikeda M, Matsubara A et al (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat Genet* 42:545–549. <https://doi.org/10.1038/ng.592>
- Okada Y, Monden Y, Nokihara K et al (2019) Genome-Wide Association Studies (GWAS) for Yield and Weevil Resistance in Sweet potato (*Ipomoea batatas* (L.) Lam). *Plant Cell Rep* 38:1383–1392. <https://doi.org/10.1007/s00299-019-02445-7>
- Piao R, Jiang W, Ham TH et al (2009) Map-based cloning of the ERECT PANICLE 3 gene in rice. *Theor Appl Genet* 119:1497–1506. <https://doi.org/10.1007/s00122-009-1151-x>
- Qiao Y, Piao R, Shi J et al (2011) Fine mapping and candidate gene analysis of dense and erect panicle 3, dep3, which confers high grain yield in rice (*Oryza sativa* L.). *Theor Appl Genet* 122:1439–1449. <https://doi.org/10.1007/s00122-011-1543-6>
- Rebolledo MC, Peña AL, Duitama J et al (2016) 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* 7:1–12. <https://doi.org/10.3389/fpls.2016.01384>
- Reig-Valiente JL, Marqués L, Talón M, Domingo C (2018) Genome-wide association study of agronomic traits in rice cultivated in temperate regions 06 Biological Sciences 0604 Genetics. *BMC Genomics* 19:1–11. <https://doi.org/10.1186/s12864-018-5086-y>
- Sahu PK, Mondal S, Sao R, et al (2020) Genome-wide association mapping revealed numerous novel genomic loci for grain nutritional and yield-related traits in rice (*Oryza sativa* L.) landraces. *3 Biotech* 10:1–22. <https://doi.org/10.1007/s13205-020-02467-z>
- Sakuraba Y, Piao W, Lim JH et al (2015) Rice ONAC106 inhibits leaf senescence and increases salt tolerance and Tiller angle. *Plant Cell Physiol* 56:2325–2339. <https://doi.org/10.1093/pcp/pcv144>
- Schmitz AJ, Begcy K, Sarath G, Walla H (2015) Rice Ovate Family Protein 2 (OFP2) alters hormonal homeostasis and vasculature development. *Plant Sci* 241:177–188. <https://doi.org/10.1016/j.plantsci.2015.10.011>
- Seck PA, Diagne A, Mohanty S, Wopereis MCS (2012) Crops that feed the world 7: Rice. *Food Secur* 4:7–24. <https://doi.org/10.1007/s12571-012-0168-1>
- Shi Y, Gao L, Wu Z et al (2017) Genome-wide association study of salt tolerance at the seed germination stage in rice. *BMC Plant Biol* 17:1–11. <https://doi.org/10.1186/s12870-017-1044-0>
- Si L, Chen J, Huang X et al (2016) OsSPL13 controls grain size in cultivated rice. *Nat Genet* 48:447–456. <https://doi.org/10.1038/ng.3518>
- Ta KN, Khong NG, Ha TL et al (2018) A genome-wide association study using a Vietnamese landrace panel of rice (*Oryza sativa*) reveals new QTLs controlling panicle morphological traits. *BMC Plant Biol* 18:1–15. <https://doi.org/10.1186/s12870-018-1504-1>
- Tabuchi H, Zhang Y, Hattori S et al (2011) LAX PANICLE2 of rice encodes a novel nuclear protein and regulates the formation of axillary meristems. *Plant Cell* 23:3276–3287. <https://doi.org/10.1105/tpc.111.088765>
- Tang Y, Liu H, Guo S et al (2018) OsmiR396d affects gibberellin and brassinosteroid signaling to regulate plant architecture in rice. *Plant Physiol* 176:946–959. <https://doi.org/10.1104/pp.17.00964>
- Terao T, Nagata K, Morino K, Hirose T (2010) A gene controlling the number of primary rachis branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. *Theor Appl Genet* 120:875–893. <https://doi.org/10.1007/s00122-009-1218-8>
- Wang QL, Sun AZ, Chen ST et al (2018a) SPL6 represses signalling outputs of ER stress in control of panicle cell death in rice. *Nat Plants* 4:280–288. <https://doi.org/10.1038/s41477-018-0131-z>
- Wang SH, Lim JH, Kim SS et al (2015a) Mutation of SPOTTED LEAF3 (SPL3) impairs abscisic acid-responsive signalling and delays leaf senescence in rice. *J Exp Bot* 66:7045–7059. <https://doi.org/10.1093/jxb/erv401>
- Wang W, Li G, Zhao J et al (2014) DWARFILLER1, a WUSCHEL-related homeobox transcription factor, is required for Tiller growth in rice. *PLoS Genet*. <https://doi.org/10.1371/journal.pgen.1004154>
- Wang W, Mauleon R, Hu Z et al (2018b) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557:43–49. <https://doi.org/10.1038/s41586-018-0063-9>
- Wang X, Wang H, Liu S et al (2016) Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. *Nat Genet* 48:1233–1241. <https://doi.org/10.1038/ng.3636>
- Wang Z, Wang Y, Hong X et al (2015b) Functional inactivation of UDP-N-acetylglucosamine pyrophosphorylase 1 (UAP1) induces early leaf senescence and defence responses in rice. *J Exp Bot* 66:973–987. <https://doi.org/10.1093/jxb/eru456>
- Wei X, Xu J, Guo H et al (2010) DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol* 153:1747–1758. <https://doi.org/10.1104/pp.110.156943>
- Xue W, Xing Y, Weng X et al (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat Genet* 40:761–767. <https://doi.org/10.1038/ng.143>
- Yang C, Ma Y, He Y et al (2018) OsOFP19 modulates plant architecture by integrating the cell division pattern and brassinosteroid signaling. *Plant J* 93:489–501. <https://doi.org/10.1111/tpj.13793>
- Yoshida A, Ohmori Y, Kitano H et al (2012) ABERRANT SPIKELET and PANICLE1, encoding a TOPESS-related transcriptional co-repressor, is involved in the regulation of meristem fate in rice. *Plant J* 70:327–339. <https://doi.org/10.1111/j.1365-3113.2011.04872.x>
- Yuan H, Qin P, Hu L et al (2019) OsSPL18 controls grain weight and grain number in rice. *J Genet Genomics* 46:41–51. <https://doi.org/10.1016/j.jgg.2019.01.003>
- Zeng D, Liu T, Ma X et al (2020) Quantitative regulation of Waxy expression by CRISPR/Cas9-based promoter and 5'UTR-intron editing improves grain quality in rice. *Plant Biotechnol J* 18:2385–2387. <https://doi.org/10.1111/pbi.13427>
- Zhang C, Dong SS, Xu JY et al (2019a) PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35:1786–1788. <https://doi.org/10.1093/bioinformatics/bty875>
- Zhang F, Zeng D, Zhang CS et al (2019b) Genome-wide association analysis of the genetic basis for sheath blight resistance in rice. *Rice* 12:1–13. <https://doi.org/10.1186/s12284-019-0351-5>
- Zhang GH, Li SY, Wang L et al (2014) LSCHL4 from japonica cultivar, which is allelic to NAL1, increases yield of Indica super rice 93–11. *Mol Plant* 7:1350–1364. <https://doi.org/10.1093/mp/ssu055>
- Zhang P, Zhong K, Zhong Z, Tong H (2019c) Genome-wide association study of important agronomic traits within a core collection of rice (*Oryza sativa* L.). *BMC Plant Biol* 19:1–12. <https://doi.org/10.1186/s12870-019-1842-7>
- Zhang Y, Lan H, Shao Q et al (2016) An A20/AN1-type zinc finger protein modulates gibberellins and abscisic acid contents and increases sensitivity to abiotic stress in rice (*Oryza sativa*). *J Exp Bot* 67:315–326. <https://doi.org/10.1093/jxb/erv464>
- Zhang Y, Ma Y, Chen Z et al (2015) Genome-wide association studies reveal new genetic targets for five panicle traits of international rice varieties. *Rice Sci* 22:217–226. <https://doi.org/10.1016/j.rsci.2015.07.001>
- Zhou Y, Srinivasan S, Mirnezami SV et al (2019) Semiautomated feature extraction from RGB images for sorghum panicle architecture GWAS. *Plant Physiol* 179:24–37. <https://doi.org/10.1104/pp.18.00974>
- Zhu K, Tang D, Yan C et al (2010) ERECT PANICLE2 encodes a novel protein that regulates panicle erectness in Indica rice. *Genetics* 184:343–350. <https://doi.org/10.1534/genetics.109.112045>
- Zhu Z, Tan L, Fu Y et al (2013) Genetic control of inflorescence architecture during rice domestication. *Nat Commun* 4:1–8. <https://doi.org/10.1038/ncomm33200>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.