



Fine mapping of the panicle length QTL *qPL5* in rice

Pengfei Wang¹ · Ling Ma¹ · Daoyang Li¹ · Bo Zhang¹ · Tianhao Zhou¹ · Xiangchun Zhou¹ · Yongzhong Xing¹

Received: 8 November 2023 / Accepted: 2 December 2023 / Published online: 17 January 2024
© The Author(s) 2024

Abstract

Panicle length is a crucial trait tightly associated with spikelets per panicle and grain yield in rice. To dissect the genetic basis of panicle length, a population of 161 recombinant inbred lines (RILs) was developed from the cross between an *aus* variety Chuan 7 (C7) and a tropical Geng variety Haoboka (HBK). C7 has a panicle length of 30 cm, 7 cm longer than that of HBK, and the panicle length was normally distributed in the RIL population. A total of six quantitative trait loci (QTLs) for panicle length were identified, and single QTLs explained the phenotypic variance from 4.9 to 18.1%. Among them, three QTLs were mapped to the regions harbored *sd1*, *DLT*, and *Ehd1*, respectively. To validate the genetic effect of a minor QTL *qPL5*, a near-isogenic F₂ (NIF₂) population segregated at *qPL5* was developed. Interestingly, panicle length displayed bimodal distribution, and heading date also exhibited significant variation in the NIF₂ population. *qPL5* accounted for 66.5% of the panicle length variance. The C7 allele at *qPL5* increased panicle length by 2.4 cm and promoted heading date by 5 days. Finally, *qPL5* was narrowed down to an 80-kb region flanked by markers M2197 and M2205 using a large NIF₂ population of 7600 plants. LOC_Os05g37540, encoding a phytochrome signal protein whose homolog in *Arabidopsis* enlarges panicle length, is regarded as the candidate gene because a single-nucleotide mutation (C1099T) caused a premature stop codon in HBK. The characterization of *qPL5* with enlarging panicle length but promoting heading date makes its great value in breeding early mature varieties without yield penalty in rice.

Keywords *aus* · Panicle length · QTL · Fine mapping · Early mature

✉ Yongzhong Xing
yzxing@mail.hzau.edu.cn

¹ National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan 430070, China

Introduction

Rice is not only a staple food in Asia but also a crucial model crop for biological research. Rice yield is mainly contributed by spikelets per panicle. Panicle length, a crucial trait controlling rice panicle architecture, exerts a substantial influence on spikelets per panicle and yield per plant (Bai et al. 2016). Panicle length is coordinated by a complex gene cluster associated with panicle development, plant morphology, and heading date. Importantly, these genes shape panicle length via different molecular mechanisms.

In the past two decades, a series of important genes controlling panicle length have been identified. *APO1* encodes an F-box protein, and its genetic mutations cause premature cessation of meristem proliferation, culminating in shortened panicle length, which concurrently impacts both the number of branches and yield (Ikeda et al. 2005). *APO2*, like the Arabidopsis gene *FLY*, functions as a homologous counterpart to hinder the meristematic transition from inflorescences to spikelets, while the loss of *APO2* function results in reduced panicle lengths and branches (Rao et al. 2008). *LARGE2* interacts with both *APO1* and *APO2* and negatively modulates the stability of these proteins, while loss of *LARGE2* function leads to the accumulation of *APO1* and *APO2*, thereby yielding enlarged panicle architecture (Huang et al. 2021). *SP1* acts as a polypeptide transporter and affects panicle length (Li et al. 2009). *DEP1* encodes the γ subunit of a trimeric G protein, while loss of *DEP1* functions influences both panicle length and grain density (Huang et al. 2009). The ideal plant architecture determinant gene, *IPA1*, positively regulates *DEP1*, resulting in substantial alterations in panicle length (Miura et al. 2010; Jiao et al. 2010). Furthermore, genes such as *sd1*, *Ghd7*, *Ghd7.1*, and *Ghd8*, which show significant effects on plant architecture or heading date, also exhibit a minor impact on panicle length (Sasaki et al. 2002; Xue et al. 2008; Yan et al. 2013; Yan et al. 2011). The previous studies are mainly focused on the populations derived from limited rice subgroups, which exhibit a broader geographical distribution and wider production application.

Asian cultivated rice has evolved into two subspecies: Xian/*indica* and Geng/*japonica* (Ting 1949). Generally, Xian varieties exhibit longer panicle length than Geng varieties. The two subspecies are further categorized into multiple subgroups, such as East Asian Xian, IRRI Xian, South Asia Xian, Southeast Asia Xian, *aus*, Temperate Geng, Tropical Geng, and Basmati, demarcated by morphological characteristics, geographical distribution, and pedigree connections (Wang et al. 2018). Over the past two decades, numerous studies have explored the genetic basis of quantitative traits in rice. But the *aus* rice, which is regarded as a valuable resource for mining favorable genes/alleles (Casarelli et al. 2018), has not been paid enough attention in QTL mapping.

Aus rice, primarily cultivated in challenging environment in Bangladesh and India, has undergone significant evolutionary adaptation due to prolonged natural selection and artificial domestication (Londo et al. 2006). This has conferred remarkable adaptability to environmental stress and unproductive conditions upon *aus* varieties. Notably, the *aus* variety N22 stands out as one of

the most thermotolerant rice cultivars (Li et al. 2015). Moreover, the *aus* subgroup has contributed valuable genetic resources, including the flood tolerance gene *Sub1* and the phosphorus starvation tolerance gene *OsPSTOL1* (Xu et al. 2006; Gamuyao et al. 2012). The *aus* variety Dular contains wide compatibility alleles like a bridge connecting Xian and Geng subspecies, which accelerates the advancement of inter-subspecies hybridization breeding programs (Wang et al. 1998). The *aus* genetic resources harbor numerous unique genetic resources distinct from the conventional Xian and Geng varieties, while the exploration and recognition of these noteworthy genetic resources hold substantial promise for molecular breeding.

In this study, we developed a RIL population through the cross between the *aus* variety C7 with long panicles and the tropical Geng variety HBK with short panicles. The population was genotyped by resequencing, and a comprehensive genome-wide bin map was constructed. The genetic basis of panicle length was dissected, and a total of 6 QTLs were identified. To validate and fine map the QTL *qPL5* with a moderate effect, a near-isogenic F₂ population was developed. *qPL5* was finally narrowed down to an interval of 80-kb region. The favorable allele from the *aus* variety exhibits valuable potential for breeding early mature varieties without yield penalty in rice.

Results

High polymorphism between the parents

The sequencing depth of two parents was exceeded 30 ×, and more than 2 G of high-quality resequencing data were generated for each of 161 RILs. Following stringent quality control, a total of 1,711,474 high-quality SNPs were identified (Fig. 1A). These variants exhibited a distinct distribution across chromosomes.

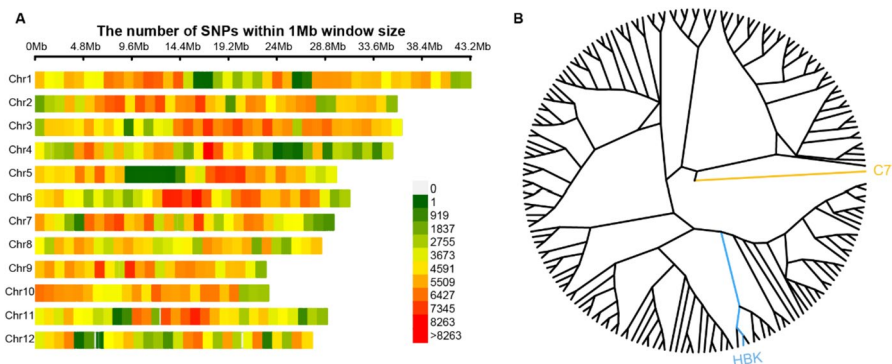


Fig. 1 Distribution of SNPs on chromosomes and phylogenetic tree in the RIL population. **A** Distribution of SNPs within 1-Mb interval on 12 chromosomes. **B** Neighbor-joining tree of the two parents and 161 RILs, the two parents C7 and HBK are marked

Chromosome 10 exhibited the highest SNP density of 5.3 SNPs/kb, while chromosome 4 displayed the lowest density of 3.6 SNPs/kb. Approximately 2.7% of the identified SNPs were heterozygous across the population, slightly lower than the expected frequency (3.2%) for the F_6 generation. A neighbor-joining phylogenetic tree constructed by 1354 evenly distributed SNPs with low linkage disequilibrium (LD) ($r^2 \leq 0.2$) revealed that the RIL population has no distinct population structure (Fig. 1B).

Construction of high-density whole-genome bin map in the RIL population

A total of 2477 bins were identified in the RIL population (Fig. 2A), which is 1.5-fold to that detected in a larger RIL population with 197 individuals directly

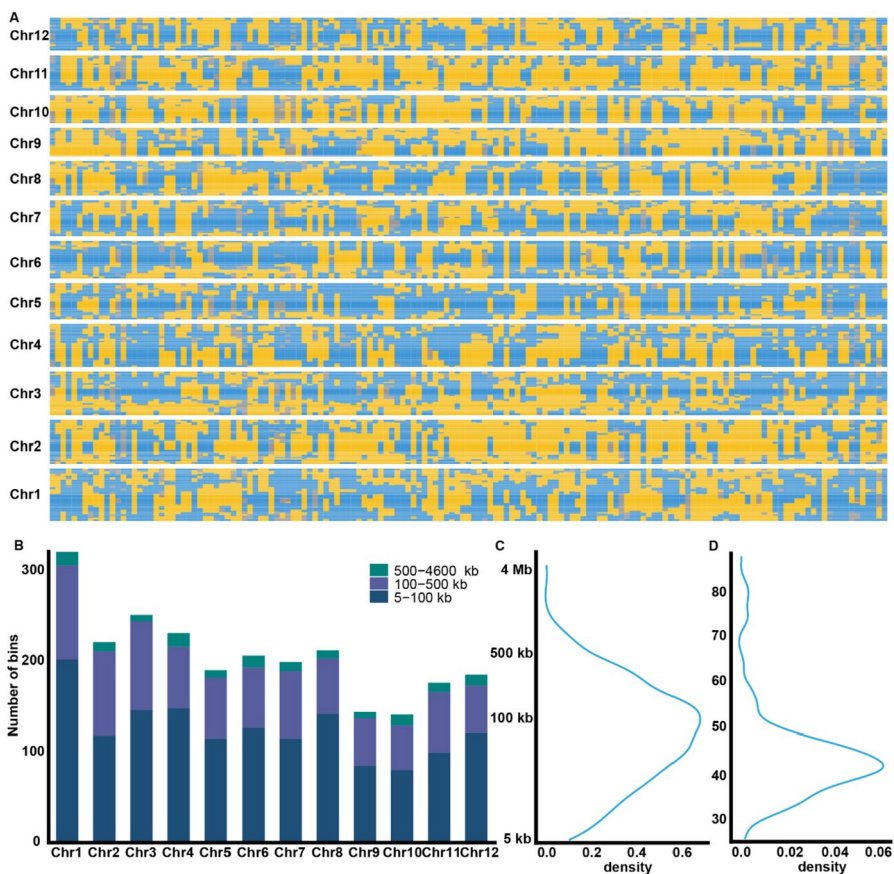


Fig. 2 Whole-genome bin construction in RIL population. **A** The bin map of the RIL population, each column represents a RIL, blue is HBK, yellow is C7, and gray is heterozygous. **B** Distribution of bin length across 12 chromosomes at population level. **C** Density plot of bin length at population level, the y axis is the length of bin. **D** Density plot of bin number at individual level, the y axis is the number of bins per individual

derived from F_1 plants (Tan et al. 2013). This indicated the strategy constructing RIL by intercrossing F_2 plants has increased the number of meiosis, resulting in a high-resolution map. At the population level, the bin lengths ranged from 5 to 4597 kb, with an average length of 150 kb (Fig. 2C). Most bins are smaller than 500 kb in length, while 60.4% of the bins are smaller than 100 kb (Fig. 2B). The longest bin was located on the centromeric region of chromosome 2 with a length of 4.6 Mb. Remarkably, the number of bins at the individual plants ranged from 25 to 88, with an average of 44 bins per line (Fig. 2D). The genetic composition of the RIL population included 49.2% homozygous *aus* genome, 46.7% homozygous Geng genome, and 4.1% heterozygous regions. On average, there were 7 recombinant events per Mb, ranging from 1 to 23. Recombination repression was observed in the centromeric regions of all 12 chromosomes, and one recombination hotspot was detected on chromosome 8 (sup Table 1). The genetic linkage map was developed covering 1397-cM regions, and genetic distance between adjacent bins was approximately 0.56 cM. The markers' position in the physical map aligned very well with those in the genetic map (Fig. 3A).

QTLs for panicle length

The panicle lengths of the two parents C7 and HBK were 30 cm and 23 cm, respectively. Panicle length exhibited a normal distribution ranged from 17.2 to 39.5 cm with a mean of 26.4 cm in the RIL population (sup Fig 1). Linkage analysis based on the ICIM model identified 6 QTLs on chromosomes 1, 5, 6, 9, 10, and 12 (Fig. 3B; Table 1). The major QTL, *qPL1* with an additive effect of 1.6 cm, accounted for 18.1% of the phenotypic variance and the C7 allele increasing panicle length. Its confidence interval harbored the well-documented major gene, *sd1*, known for its impact on plant architecture and panicle length (Sasaki et al. 2002). The minor QTL, *qPL6*, explained 8.0% of the phenotypic variance, and the HBK allele increased panicle length by 1.0 cm. The gene *DLT*, which affected the internode length and panicle length by modulating BR level, was located in the confidence interval (Tong et al. 2009). The confidence interval for another minor QTL, *qPL10*, contained *Ehd1* with an additive effect of 0.8 cm, which was reported to affect heading date and panicle length (Hu et al. 2019). Previously reported functional polymorphisms in the three genes were identified between two parents (Table 2). The missense variant in *sd1* coding region resulted in lower gibberellin biosynthetic activity and a shorter culm (Asano et al. 2011). The 5' UTR variant in *DLT* affected its expression level, and *DLT* modulated panicle length by regulating gene expression of BR synthesis (Sun et al. 2013; Tong et al. 2009). The 3' UTR variant in *Ehd1* was reported as its causal mutation (Wei et al. 2021). Therefore, these three genes are very likely underlying the three QTLs. The minor QTL, *qPL5* with an additive effect of 1.0 cm, accounted for 6.6% of the phenotypic variance, and the C7 allele increased panicle length. A previously characterized gene, *EUII*, in the confidence interval of *qPL5* was documented for its role in influencing panicle length (Zhu et al. 2006).

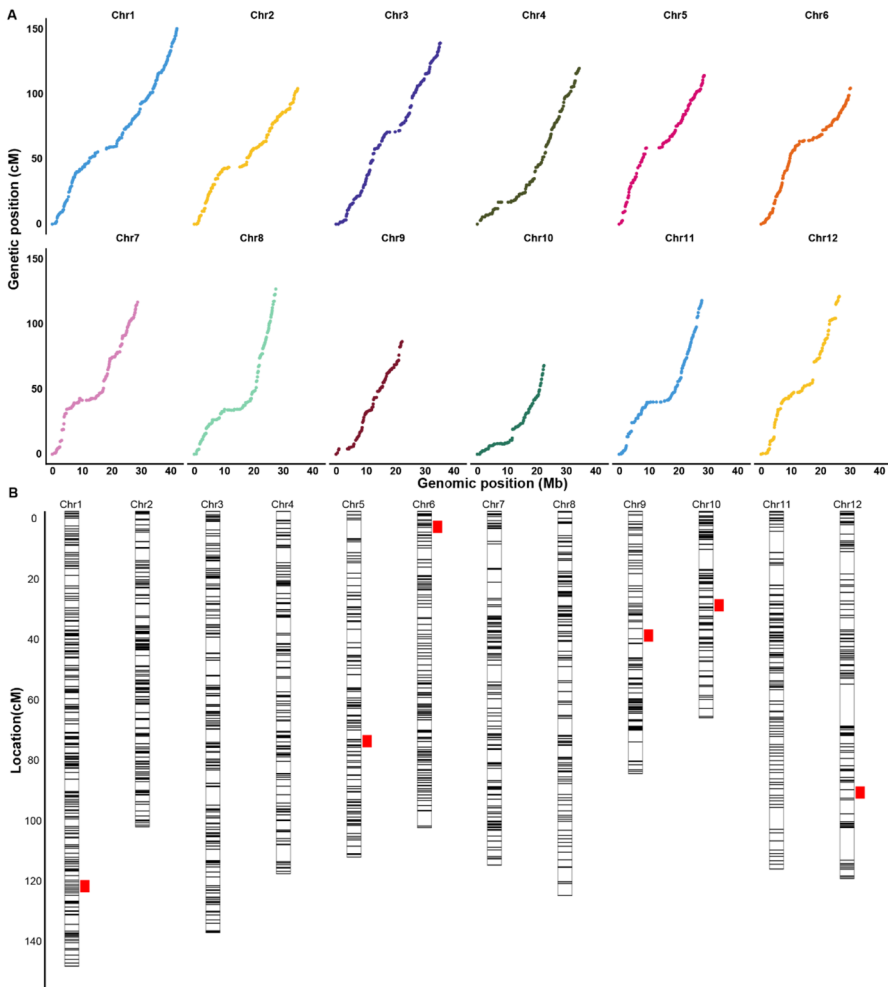


Fig. 3 Collinearity analysis and genetic linkage map with the RIL population. **A** Collinearity analysis of genetic linkage map and physical map. **B** The genetic map showing QTLs for panicle length, black line represents each bin. Red box is the confidence interval of QTLs

Table 1 Putative QTLs for panicle length detected in the RIL population

QTL	Chr	Start	End	LOD	PVE %	Add
<i>qPL1</i>	1	38019364	38610747	11.2	18.1	1.6
<i>qPL5</i>	5	19403806	20167654	4.4	6.6	1.0
<i>qPL6</i>	6	1347846	2163965	5.3	8.0	-1.0
<i>qPL9</i>	9	12220801	14075636	5.5	8.6	1.1
<i>qPL10</i>	10	16069736	17494721	3.4	4.9	0.8
<i>qPL12</i>	12	21736638	22964216	4.7	7.0	1.0

LOD, logarithm of odds. *Add*, the additive effect. Positive values indicate that alleles from C7 increase the trait scores. *PVE*, phenotypic variance explained

Table 2 Functional polymorphisms between parents in known genes located in three QTL regions

QTL	Gene	Chr	Pos	C7	HBK	Anno	Reference
<i>qPL1</i>	<i>sd1</i>	Chr1	38385057	G	A	missense	Asano et al. (2011)
<i>qPL6</i>	<i>DLT</i>	Chr6	1468526	TCTCTC ACTCTC A	T	5_prime_UTR	Tong et al. (2009)
<i>qPL10</i>	<i>Ehd1</i>	Chr10	17076177	C	T	3_prime_UTR	Wei et al. (2021)

Anno, the functional annotation of the variant

Validation and fine mapping of *qPL5*

To validate the genetic effect of *qPL5*, we constructed a single-segment introgression line, FZB7, against the HBK background (Fig. 4A, B). Analysis of rice 40 K SNP array indicated a background recovery rate of 97.1%. Subsequent phenotypic evaluations revealed that the growth and developmental stages of FZB7 were nearly same as that of HBK (Fig. 4C). FZB7 exhibited a longer panicle length and an increased number of second branches (Fig. 4F, G), indicating the presence of the *qPL5* effect in the introgression fragment. Interestingly, significant increases were observed in plant height, seed setting rate, and spikelets per panicle in FZB7, while HBK showed a 5-day delayed heading date, 1 more tiller per plant, and a heavier thousand-grain weight (Fig. 4D, E, G). Consequently, no difference in yield per plant was detected between two genotypes. To exclude the possibility of *EU11* as the candidate, Sanger sequencing was performed for *EU11* in both parents. No variations in the *EU11* coding region were detected between the parents (sup Data 1). Simultaneously, the expression patterns of *EU11* were conducted in FZB7 and its recurrent parent HBK. No significant differences in expression pattern were ascertained (sup Fig 2). Therefore, *EU11* was unlikely the causal gene for *qPL5*.

To figure out the candidate gene, a BC₄F₂ population of 472 individuals was then generated through the cross between HBK and FZB7. The panicle length distributed with bimodal model ranging from 19.5 to 25.6 cm, and a large variation ranging from 90 to 105 days was identified in heading date in the BC₄F₂ population (sup Fig 3). To facilitate the fine mapping work, five polymorphic InDel markers and 11 polymorphic KASP markers were developed. Finally, *qPL5* was mapped to a 2.3-Mb interval delimited by markers RID518 and ML70 (Fig. 5A). The logarithm of odds (LOD) score attributed to *qPL5* showed a remarkable elevation to 112.2. Its additive effect on panicle length was 1.2 cm; it accounted for 66.5% of the panicle length variance (Table 3). Moreover, *qPL5* had effects on heading date and spikelets per panicle but no effects on yield per plant (Table 3).

To further fine map *qPL5*, a BC₄F₃ population of 7600 plants was developed. From the large population, we screened 11 and 38 recombinants between *qPL5* and two markers RID518 and ML70, respectively (Fig. 5B). Additionally, 6 polymorphic InDel markers were developed in the target region between the markers RID518 and ML70. After progeny test of 15 recombinants screened by markers

Fig. 4 Genome constitution of FZB7 and phenotypic comparison between *qPL5* near-isogenic lines. **A** Genome constitution of HBK (**A**) and FZB7 (**B**) identified by rice 40K SNP array, blue is HBK genotype, white is the genome not covered by SNP markers, and yellow is C7 genotype. The comparison of whole plant (**C**), panicle architecture (**D**), internode length (**E**), and panicle length (cm) (**F**) between HBK and FZB7. **G** The comparisons of agronomic traits between HBK and FZB7. HD, heading date; PH, plant height (cm), TP, tillers per plant; SR, seed setting rate; GL, grain length (mm), GW, grain width (mm); TGW, thousand-grain weight (g); NPB, the number of primary branches; NSB, the number of second branches; SPP, spikelets per panicle; YD, yield per plant (g)

RID519 and M2205, *qPL5* was mapped to an 80-kb interval delimited by markers M2197 and M2205 (Fig. 5C), in which *EUII* was not located.

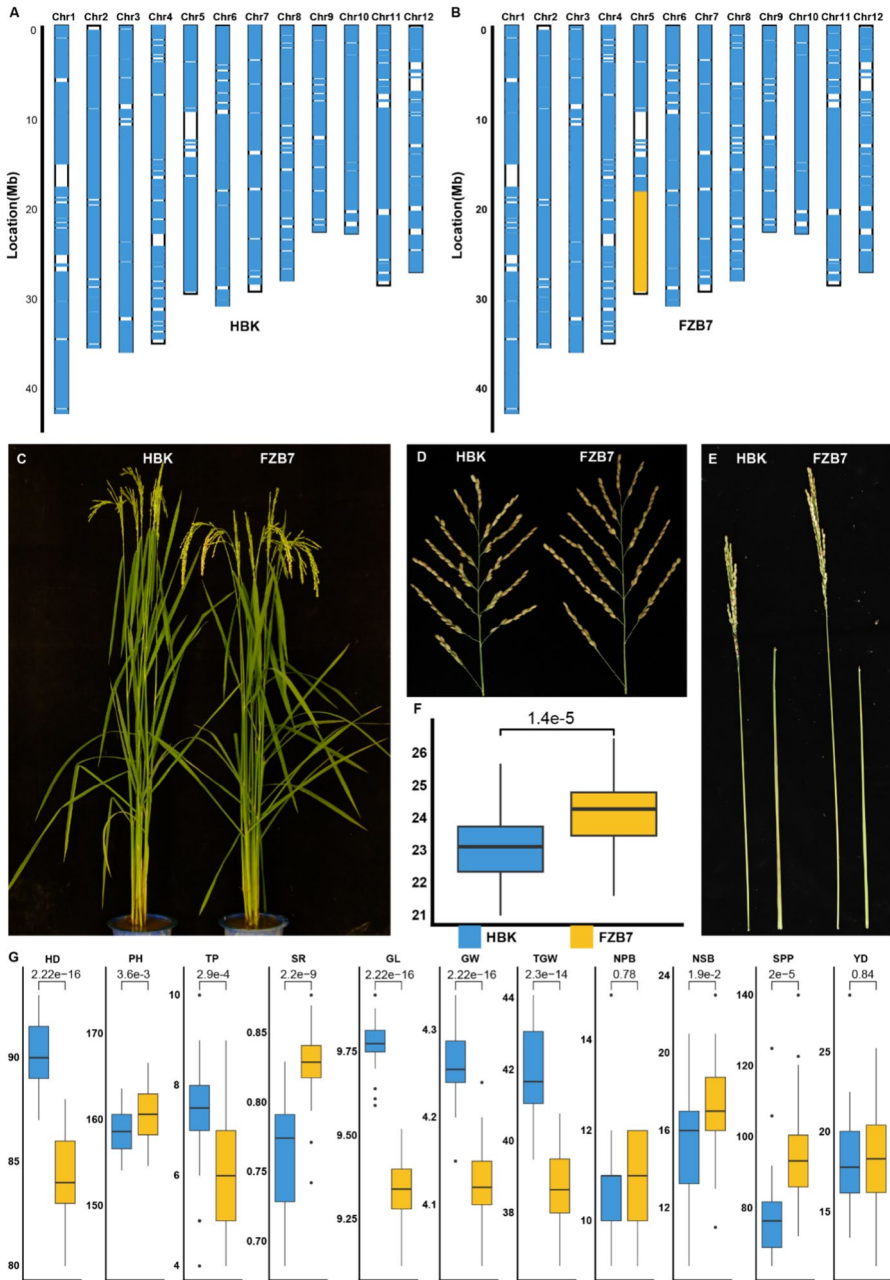
Possible candidate gene for *qPL5*

There were a total of 16 genes in the targeted 80-kb region in RGI database (Table 4) (Yu et al. 2023). Among them, the gene LOC_Os05g37540 was annotated to encode the protein of Far-red impaired Response 1 (FAR1), an essential transcriptional factor for phytochrome A signaling. In *Arabidopsis*, FAR1 directly interacts with three Squamosa-promoter binding protein-like (SPL) transcription factors, SPL3, SPL4, and SPL5, and inhibits their binding to the promoter of key flowering gene *FUL*, *LFY*, *API*, and *MIR172C*, thereby delaying flowering and inducing the development of enlarged panicles (Xie et al. 2020). In rice, mutations in phytochrome genes have been associated with a moderate acceleration of flowering under long day conditions (Su et al. 2016). A mutation C1099T was identified in the LOC_Os05g37540 exon of HBK, leading to a premature termination. The mutation was infrequently detected in tropical Geng rice based on the ricevarmap database (Zhao et al. 2021). Thus, LOC_Os05g37540 is a possible candidate gene for *qPL5*.

Discussion

LOC_Os05g37540 encoding a FAR1-related protein is a strong candidate gene of *qPL5*

In this study, *qPL5* was validated in a NIF₂ population and fine mapped to an 80-kb region flanked by markers M2197 and M2205. There are 16 genes in the target region. The functional annotation highlighted three genes associated with plant development as potential candidates underlying *qPL5*. LOC_Os05g37600 encodes a member of the glycerol-3-phosphate acyltransferase family, which has been previously reported to be related to cellular metabolism associated with plant growth and development (Safder et al. 2021). LOC_Os05g37640 encodes an AP2/ERF transcription factor, while the AP2 transcription factors FZP, OsSNB, OsIDS1, and MFS1 were reported to affect panicle architecture (Ren et al. 2013; Lee and An 2012; Bai et al. 2017). But both genes have no functional polymorphism between parents (Table 4); they are excluded for the candidate. LOC_Os05g37540 encodes a FAR1-related protein. *FAR1* in *Arabidopsis*



was associated with delayed flowering, which led to enlarged panicles (Xie et al. 2020). Moreover, *qPL5* has effects on heading date (Fig. 4G). More important is that there is a premature termination variation in *LOC_Os05g37540* gene between the parents (sup Data 2). Taken together, *LOC_Os05g37540* is strongly

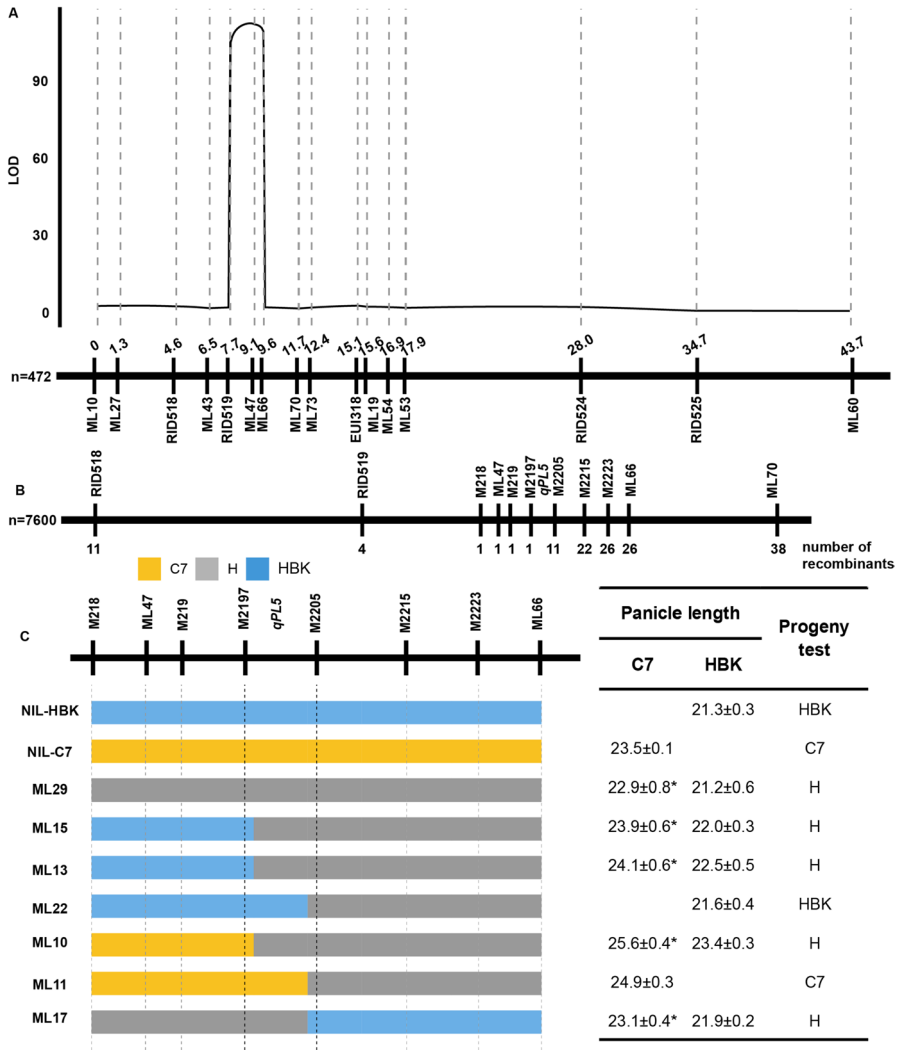


Fig. 5 Fine mapping of *qPL5*. **A** QTL scanning for panicle length across the target region on chromosome 5 in the BC₄F₂ population. X axis is the genetic position, and Y axis is the LOD score. **B** The number of recombinants between *qPL5* and molecular markers in the BC₄F₃ population. **C** Progeny test of the recombinants. *, *P* < 0.05. More than 10 plants each two homozygous were used for comparison in the progeny test

Table 3 Pleiotropic effects of *qPL5* on panicle length, heading date, and spikelets per panicle in the NIF₂ population

Trait	QTL	Confident interval	LOD	PVE %	Add	Dom
PL	<i>qPL5</i>	RID519-ML47	112.2	66.5	1.2	0.5
HD			132.5	72.3	-2.7	1.1
SPP			95.4	52.1	12.6	4.2

Add, additive effect; *Dom*, dominant effect

Table 4 Candidate genes in the 80-kb target region of *qPL5*

Gene	Annotation	C7	HBK
LOC_Os05g37530	Expressed protein	-	-
LOC_Os05g37540	Protein FAR1-RELATED SEQUENCE 5	-	StGV
LOC_Os05g37550	Expressed protein	-	-
LOC_Os05g37560	Expressed protein	StGV, MsV	-
LOC_Os05g37570	Putative; Ty3-gypsy subclass	-	-
LOC_Os05g37580	Putative; unclassified; expressed	-	-
LOC_Os05g37590	Transposon Ty3-I Gag-Pol polyprotein	-	-
LOC_Os05g37600	Glycerol-3-phosphate 2-O-acyltransferase 6	-	-
LOC_Os05g37610	Expressed protein	-	-
LOC_Os05g37620	Retrovirus-related Pol polyprotein from transposon RE1	StGV, MsV	-
LOC_Os05g37630	Expressed protein	MsV	-
LOC_Os05g37640	AP2/ERF transcription factor	-	-
LOC_Os05g37650	Expressed protein	MsV	-
LOC_Os05g37660	Uncharacterized acetyltransferase	MsV	-
LOC_Os05g37670	Putative; unclassified; expressed	StGV, MsV	-
LOC_Os05g37680	Putative; unclassified; expressed	FsV, MsV	-

StGV, stop gained variant; *MsV*, missense variant; *FsV*, frameshift variant

suggested as the candidate gene of *qPL5*, which will be eventually confirmed by gene transformation.

Minor QTL is the target for gene cloning in the future

A total of six QTLs for panicle length were identified in the RIL population. Among these QTLs, only one QTL *qPL1* explained 18.1% of panicle length variance, which would be classified as a major QTL. All the other 5 QTLs are minor ones, each explaining less than 9% of the variance. Three genes, *sd1*, *DLT*, and *Ehd1*, which regulate panicle length by influencing GA biosynthesis, BR levels, and heading date, respectively (Sasaki et al. 2002; Tong et al. 2009; Hu et al. 2019), were located in the QTL regions of *qPL1*, *qPL6*, and *qPL10*, respectively. Moreover, the previously reported functional polymorphisms such as missense mutation and UTR variants were detected between the two parents (Table 2), strongly suggesting they are the genes underlying these three QTLs. In addition, the minor QTL *qPL5* was validated in a near-isogenic F₂ population. The other two QTLs, *qPL9* and *qPL12*, which exhibited comparable effects to *qPL5*, were detected in several studies (Cui et al. 2002; Kobayashi et al. 2003; Xiao et al. 1998). Thus, the minor QTLs identified in this study are reliable. A rapid QTL cloning strategy by generating a pseudo-near-isogenic F₂ population can be used to speed up the fine mapping of *qPL9* and *qPL12* without continuous advanced backcross (Sherif et al. 2023).

***qPL5* is a valuable gene for breeding early mature varieties without yield penalty in rice**

The *aus* rice is a valuable resource for mining favorable alleles because many superior alleles have been specially identified in *aus* subpopulation (Casartelli et al. 2018). The superior alleles at five of six QTLs for panicle length detected in this study were contributed by the *aus* cultivar C7, which is in accordance with the fact that C7 has the panicle length of 7 cm longer than HBK. The semidwarf gene *sd1* and *DLT* increase panicle length but significantly increase plant height which leads to lodging and substantial yield loss. *Ehd1* enlarges the panicle length by delaying heading, which influences the adaptation of varieties to cropping seasons and ecological zones (Hu et al. 2019). In contrast, *qPL5* showed negative effects on heading date and tiller number per plant and positive effects on plant height and seed setting ratio in the near-isogenic lines (Table 3). Ultimately, it has no significant impact on the yield per plant. Therefore, the *aus* C7 *qPL5* is a potential favorable allele for breeding early mature varieties without yield penalty. The closely linked markers M2197 and M2205 would be used for assistant selection of *qPL5*.

Methods

Plant materials

We developed a RIL population between an *Oryza sativa aus* inbred line (Chuan 7, C7) and an *Oryza sativa Geng* inbred line (Haoboka, HBK). The hybrid F₁ plant was obtained by crossing C7 with HBK, and then several hundred F₂ plants were obtained by self-crossing. From these, 16 F₂ plants were randomly selected to make eight pairwise crosses. Eight hybrids from each pairwise cross were used for selfing, and then 3 individuals each self-bred family were randomly selected for continuous self-pollination for 5 generations by single seed descent (sup Fig 4). Finally, a total of 161 RILs were constructed. SSR markers were employed to confirm the true hybrid identity throughout the process.

A set of CSSLs (BC₄F₁) were previously developed with HBK as the recurrent parents and C7 as the donor parent (Zhang et al. 2021). A single-segment introgression line FZB7 possessed the *qPL5* was used to generate a BC₄F₂ population of 472 plants. A BC₄F₃ population of 7600 plants was used for fine map of *qPL5*. The progenies of the recombinants between markers and *qPL5* were tested in the field.

Field experiments and panicle length measurement

Field experiments were performed at the experimental farm of Huazhong Agricultural University in Wuhan. The RIL population were sowed on May 18, 2018. Fourteen 25-day old seedlings each RIL were transplanted into a two-row plot, with a distance of 16.5 cm between plants within a row and 26.5 cm between rows. All plants except marginal and abnormal individuals were used to measure the panicle

length, and 3 panicles were recorded for each plant. The average panicle length was used for QTL mapping. Four hundred seventy-two BC₄F₂ plants was planted on May 18, 2020. Seven thousand six hundred BC₄F₃ plants and progeny test for recombinants were sowed on May 21, 2021, and May 19, 2022, respectively.

DNA extraction and re-sequencing

DNA was extracted from young leaves by the CTAB method. A DNA library was constructed using the NGS pipeline, and then paired-end sequencing was performed on the HiSeq 2000 platform in GENOSEQ (Wuhan, China) with a read length of 150 bp. The sequencing depth for the two parents and each RIL was 30× and 5×.

Identification of DNA variations

The re-sequencing raw reads were trimmed and clipped the high-quality part of read by Fastp software (Chen et al. 2018). Then, clean reads were aligned to the Nipponbare reference genome from the Rice Genome Annotation Project (<http://rice.uga.edu/>), and Samtools was used to sort and convert the SAM file to BAM file (Danecek et al. 2021). PCR duplicates were removed by the MarkDuplicates program in the Picard software (<https://broadinstitute.github.io/picard/>). SNPs and InDels were called by the HaplotypeCaller and GenotypeGVCFs program with GVCF mode in the Genome Analysis Toolkit (McKenna et al. 2010). SNPs with missing rates above 50% and minor allele frequency above 10% were used for downstream analysis.

Phylogenetic analysis

Independent SNPs were extracted using the independent pairwise method in R/ SNPRelate package, employing a linkage disequilibrium threshold of 0.2 (Zheng et al. 2012). The distance matrix was calculated, and a phylogenetic tree was constructed by SNPRelate package. The ape and gtree packages were used to display the tree and to label the two parents (Paradis and Schliep 2019; Yu et al. 2017).

Bin map construction

Using the python program SNPbinner, bin map was conducted for the RIL population based on the recalibrated SNPs genotyped by the GATK pipeline to further decrease the impact of sequencing errors and computational consumption (Gonda et al. 2019).

QTL mapping

Genetic map was constructed using R package Asmap with Kosambi mapping function (Kosambi 1943; Taylor and Butler 2017). QTL mapping was performed by the

QTL IciMapping software (Meng et al. 2015). Additive QTL was identified by inclusive composite interval mapping (ICIM) method. The threshold of the logarithm of odds (LOD) value was 2.5.

Marker development

All markers for fine mapping were designed based on Nipponbare genome and parental differential variation. KASP mark was designed by snpway website (<http://www.snpway.com/>). All markers used in this experiment were listed in the Supplementary Table 2.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11032-024-01443-2>.

Acknowledgements The computations in this paper were run on the bioinformatics computing platform of the National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University.

Author contributions P.F.W. performed the experiments and drafted the manuscript. M.L. and D.Y.L. performed the phenotype experiments. B.Z. and T.H.Z. developed the genetic population. X.C.Z. contributed valuable suggestions in the manuscript. Y.Z.X. designed this study and edited the manuscript.

Funding This work was funded by the National Natural Science Foundation of China (32101740, 31901519) and the China Postdoctoral Science Foundation (Grant No. 2021M701360).

Data Availability All raw reads generated for the individuals in the study have been deposited in the National Genomics Data Center under BioProject PRJCA022700.

Declarations

Ethics approval We declare that these experiments comply with the ethical standards in China.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Asano K, Yamasaki M, Takuno S, Miura K, Katagiri S, Ito T, Doi K, Wu J, Ebana K, Matsumoto T (2011) Artificial selection for a green revolution gene during japonica rice domestication. *Proc Natl Acad Sci* 108(27):11034–11039
- Bai XF, Huang Y, Hu Y, Liu HY, Zhang B, Smaczniak C, Hu G, Han ZM, Xing YZ (2017) Duplication of an upstream silencer of FZP increases grain yield in rice. *Nat Plants* 3(11):885–893. <https://doi.org/10.1038/s41477-017-0042-4>

- Bai XF, Zhao H, Huang Y, Xie WB, Han ZM, Zhang B, Guo ZL, Yang L, Dong HJ, Xue WY, Li GW, Hu G, Hu Y, Xing YZ (2016) Genome-wide association analysis reveals different genetic control in panicle architecture between indica and japonica rice. *Plant Genome* 9(2). <https://doi.org/10.3835/plantgenome2015.11.0115>
- Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S (2018) Exploring traditional austere rice for metabolites conferring drought tolerance. *Rice* 11:1–16
- Chen SF, Zhou YQ, Chen YR, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34(17):884–890. <https://doi.org/10.1093/bioinformatics/bty560>
- Cui K-H, Peng S-B, Xing Y-Z, Yu S-B, Xu C-G (2002) Genetic analysis of the panicle traits related to yield sink size of rice. *Acta Genet Sin* 29(2):144–152
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. *Gigascience* 10(2). <https://doi.org/10.1093/gigascience/giab008>
- Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S (2012) The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature* 488(7412):535. <https://doi.org/10.1038/nature11346>
- Gonda I, Ashrafi H, Lyon DA, Strickler SR, Hulse-Kemp AM, Ma Q, Sun H, Stoffel K, Powell AF, Futrell S, Thannhauser TW, Fei Z, Van Deynze AE, Mueller LA, Giovannoni JJ, Foolad MR (2019) Sequencing-based bin map construction of a tomato mapping population, facilitating high-resolution quantitative trait loci detection. *Plant Genome* 12(1):180010. <https://doi.org/10.3835/plantgenome2018.02.0010>
- Hu Y, Li SL, Xing YZ (2019) Lessons from natural variations: artificially induced heading date variations for improvement of regional adaptation in rice. *Theor Appl Genet* 132(2):383–394. <https://doi.org/10.1007/s00122-018-3225-0>
- Huang LJ, Hua K, Xu R, Zeng DL, Wang RC, Dong GJ, Zhang GZ, Lu XL, Fang N, Wang DK, Duan PG, Zhang BL, Liu ZP, Li N, Luo YH, Qian Q, Yao SG, Li YH (2021) The LARGE2-APO1/APO2 regulatory module controls panicle size and grain number in rice. *Plant Cell* 33(4):1212–1228. <https://doi.org/10.1093/plcell/koab041>
- Huang XZ, Qian Q, Liu ZB, Sun HY, He SY, Luo D, Xia GM, Chu CC, Li JY, Fu XD (2009) Natural variation at the DEP1 locus enhances grain yield in rice. *Nat Genet* 41(4):494–497. <https://doi.org/10.1038/ng.352>
- Ikeda K, Nagasawa N, Nagato Y (2005) Aberrant panicle organization 1 temporally regulates meristem identity in rice. *Dev Biol* 282(2):349–360. <https://doi.org/10.1016/j.ydbio.2005.03.016>
- Jiao YQ, Wang YH, Xue DW, Wang J, Yan MX, Liu GF, Dong GJ, Zeng DL, Lu ZF, Zhu XD, Qian Q, Li JY (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat Genet* 42(6):541–U536. <https://doi.org/10.1038/ng.591>
- Kobayashi S, Fukuta Y, Sato T, Osaki M, Khush G (2003) Molecular marker dissection of rice (*Oryza sativa* L.) plant architecture under temperate and tropical climates. *Theor Appl Genet* 107:1350–1356
- Kosambi DD (1943) The estimation of map distances from recombination values. *Ann Eugenics* 12:172–175. <https://doi.org/10.1111/j.1469-1809.1943.tb02321.x>
- Lee DY, An G (2012) Two AP2 family genes, supernumerary bract (SNB) and Osindeterminate spikelet 1 (OsIDS1), synergistically control inflorescence architecture and floral meristem establishment in rice. *Plant J* 69(3):445–461
- Li SB, Qian Q, Fu ZM, Zeng DL, Meng XB, Kyozuka J, Maekawa M, Zhu XD, Zhang J, Li JY, Wang YH (2009) Short panicle1 encodes a putative PTR family transporter and determines rice panicle size. *Plant J* 58(4):592–605. <https://doi.org/10.1111/j.1365-313X.2009.03799.x>
- Li X, Lawas LMF, Malo R, Glaubitz U, Erban A, Mauleon R, Heuer S, Zuther E, Kopka J, Hincha DK, Jagadish KSV (2015) Metabolic and transcriptomic signatures of rice floral organs reveal sugar starvation as a factor in reproductive failure under heat and drought stress. *Plant Cell Environ* 38(10):2171–2192. <https://doi.org/10.1111/pce.12545>
- Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc Natl Acad Sci U S A* 103(25):9578–9583. <https://doi.org/10.1073/pnas.0603152103>
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce

- framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Meng L, Li HH, Zhang LY, Wang JK (2015) QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J* 3(3):269–283. <https://doi.org/10.1016/j.cj.2015.01.001>
- 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–U102. <https://doi.org/10.1038/ng.592>
- Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35(3):526–528. <https://doi.org/10.1093/bioinformatics/bty633>
- Rao NN, Prasad K, Kumar PR, Vijayraghavan U (2008) Distinct regulatory role for RFL, the rice LFY homolog, in determining flowering time and plant architecture. *Proc Natl Acad Sci U S A* 105(9):3646–3651. <https://doi.org/10.1073/pnas.0709059105>
- Ren D, Li Y, Zhao F, Sang X, Shi J, Wang N, Guo S, Ling Y, Zhang C, Yang Z (2013) Multi-floret SPIKELET1, which encodes an AP2/ERF protein, determines spikelet meristem fate and sterile lemma identity in rice. *Plant Physiol* 162(2):872–884
- Safder I, Shao G, Sheng Z, Hu P, Tang S (2021) Identification and analysis of the structure, expression and nucleotide polymorphism of the GPAT gene family in rice. *Plant Genome* 26:100290
- Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H, Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice - new insight into the rice variant that helped to avert famine over thirty years ago. *Nature* 416(6882):701–702. <https://doi.org/10.1038/416701a>
- Sherif A, Zhang B, Wu B, Hu Y, Li S, Zhou X, Elbadri AM, Elshenawy M, El-Badawy MEM, Hassan IO (2023) A pseudo-near isogenic F2 population strategy for rapid QTL cloning. *Mol Breed* 43(8):61
- Su L, Shan JX, Gao JP, Lin HX (2016) OsHAL3, a blue light-responsive protein, interacts with the floral regulator Hd1 to activate flowering in rice. *Mol Plant* 9(2):233–244. <https://doi.org/10.1016/j.molp.2015.10.009>
- Sun L, Li X, Fu Y, Zhu Z, Tan L, Liu F, Sun X, Sun X, Sun C (2013) GS 6, a member of the GRAS gene family, negatively regulates grain size in rice. *J Integr Plant Biol* 55(10):938–949
- Tan C, Han ZM, Yu HH, Zhan W, Xie WB, Chen X, Zhao H, Zhou FS, Xing YZ (2013) QTL scanning for rice yield using a whole genome SNP array. *J Genet Genomics* 40(12):629–638. <https://doi.org/10.1016/j.jgg.2013.06.009>
- Taylor J, Butler D (2017) R Package ASMap: efficient genetic linkage map construction and diagnosis. *J Stat Softw* 79(6):1–29. <https://doi.org/10.18637/jss.v079.i06>
- Ting Y (1949) Origination of the rice cultivation in China. *J College of Agric Sun Yat-Sen University* 7:11–24
- Tong HN, Jin Y, Liu WB, Li F, Fang J, Yin YH, Qian Q, Zhu LH, Chu CC (2009) Dwarf and low-tillering, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant J* 58(5):803–816. <https://doi.org/10.1111/j.1365-313X.2009.03825.x>
- Wang J, Liu KD, Xu CG, Li XH, Zhang QF (1998) The high level of wide compatibility of variety ‘Dular’ has a complex genetic basis. *Theor Appl Genet* 97(3):407–412. <https://doi.org/10.1007/s001220050910>
- Wang WS, Mauleon R, Hu ZQ, Chebotarov D, Tai SS, Wu ZC, Li M, Zheng TQ, Fuentes RR, Zhang F, Mansueto L, Copetti D, Sanciango M, Palis KC, Xu JL, Sun C, Fu BY, Zhang HL, Gao YM et al (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557(7703):43–+. <https://doi.org/10.1038/s41586-018-0063-9>
- Wei X, Qiu J, Yong KC, Fan JJ, Zhang Q, Hua H, Liu J, Wang Q, Olsen KM, Han B, Huang XH (2021) A quantitative genomics map of rice provides genetic insights and guides breeding. *Nat Genet* 53(2):243. <https://doi.org/10.1038/s41588-020-00769-9>
- Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative. *Oryza rufipogon Genetics* 150(2):899–909
- Xie YR, Zhou Q, Zhao YP, Li QQ, Liu Y, Ma MD, Wang BB, Shen RX, Zheng ZG, Wang HY (2020) FHY3 and FAR1 integrate light signals with the miR156-SPL module-mediated aging pathway to regulate Arabidopsis flowering. *Mol Plant* 13(3):483–498. <https://doi.org/10.1016/j.molp.2020.01.013>

- Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* 442(7103):705–708. <https://doi.org/10.1038/nature04920>
- Xue WY, Xing YZ, Weng XY, Zhao Y, Tang WJ, Wang L, Zhou HJ, Yu SB, Xu CG, Li XH, Zhang QF (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat Genet* 40(6):761–767. <https://doi.org/10.1038/ng.143>
- Yan WH, Liu HY, Zhou XC, Li QP, Zhang J, Lu L, Liu TM, Liu HJ, Zhang CJ, Zhang ZY, Shen GJ, Yao W, Chen HX, Yu SB, Xie WB, Xing YZ (2013) Natural variation in Ghd7.1 plays an important role in grain yield and adaptation in rice. *Cell Res* 23(7):969–971. <https://doi.org/10.1038/cr.2013.43>
- Yan WH, Wang P, Chen HX, Zhou HJ, Li QP, Wang CR, Ding ZH, Zhang YS, Yu SB, Xing YZ, Zhang QF (2011) A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol Plant* 4(2):319–330. <https://doi.org/10.1093/mp/ssq070>
- Yu GC, Smith DK, Zhu HC, Guan Y, Lam TTY (2017) GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 8(1):28–36. <https://doi.org/10.1111/2041-210x.12628>
- Yu ZC, Chen YM, Zhou Y, Zhang YL, Li MY, Ouyang YD, Chebotarov D, Mauleon R, Zhao H, Xie WB, McNally KL, Wing RA, Guo WL, Zhang JW (2023) Rice gene index: a comprehensive pan-genome database for comparative and functional genomics of Asian rice. *Mol Plant* 16(5):798–801. <https://doi.org/10.1016/j.molp.2023.03.012>
- Zhang B, Qi F, Hu G, Yang Y, Zhang L, Meng J, Han Z, Zhou X, Liu H, Ayaad M (2021) 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 F2 population. *Crop J* 9(6):1428–1437
- Zhao H, Li JC, Yang L, Qin G, Xia CJ, Xu XB, Su YM, Liu YM, Ming LC, Chen LL, Xiong LZ, Xie WB (2021) An inferred functional impact map of genetic variants in rice. *Mol Plant* 14(9):1584–1599. <https://doi.org/10.1016/j.molp.2021.06.025>
- Zheng XW, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28(24):3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>
- Zhu YY, Nomura T, Xu YH, Zhang YY, Peng Y, Mao BZ, Hanada A, Zhou HC, Wang RX, Li PJ, Zhu XD, Mander LN, Kamiya Y, Yamaguchi S, He ZH (2006) Elongated uppermost internode encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. *Plant Cell* 18(2):442–456. <https://doi.org/10.1105/tpc.105.038455>

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