Phyton-
International Journal of
Experimental Botany

Tech Science Press

DOI: 10.32604/phyton.2021.014880

ARTICLE

Fine Mapping and Candidate Gene Prediction of the Quantitative Trait Locus *qPL8* for Panicle Length in Rice

Lin Zhang[1,2,*], Yiting Zou[1], Zhong Bian[1], Dong Xie[1], Hadi Yeilaghi[1], Xiaolei Fan[1,2], Dongsheng Zhao[1,2] and Qiaoquan Liu[1,2,*]

[1]Joint International Research Laboratory of Agriculture and Agri-Product Safety of the Ministry of Education, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Yangzhou University, Yangzhou, 225009, China

[2]Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, Jiangsu Key Laboratory of Crop Genetics and Physiology, Yangzhou University, Yangzhou, 225009, China

[*]Corresponding Authors: Qiaoquan Liu. Email: qqliu@yzu.edu.cn; Lin Zhang. Email: zhangl@yzu.edu.cn

Received: 04 November 2020 Accepted: 24 November 2020

ABSTRACT

Rice panicle is the sink organ where assimilation product accumulates, and its morphology determines the rice yield. Panicle length has been suggested as a yield-related trait, but the genetic factor for its control is still limited. In this study, we carried out fine-mapping of *qPL8*, a QTL identified for panicle length in our previous work. Near isogenic line (NIL) with *qPL8* exhibited elongated panicle without obvious effect on other panicle elements. With five key recombinants from NIL population, the locus was finally narrowed down to a 278-kb region, where 44 genes are annotated. By comparing the genomic sequence of two parents, 17 genes were identified with SNPs or InDels variations in the coding region. Expression analysis showed that eight genes were up-regulated in the NIL with *qPL8*. Considering both the coding variation and expression status, several candidate genes for the locus were identified, and *OsMADS37* was raised as the most possible candidate. Interestingly, an expression QTL (eQTL) also resides in the locus, leading to a cluster of gene expression variation in the region. This study will facilitate the application of *qPL8* locus in rice breeding for yield potential.

KEYWORDS

Rice; QTL; panicle length; fine mapping; *qPL8*

1 Introduction

Rice (*Oryza sativa* L.) is a major food crop in the world, feeding over half of the global population. As sink organ, panicles are the key determinant for yield improvement. Several components make up the panicle, including panicle length, panicle branch number (primary and secondary) and panicle density, determining the final grain number per panicle [1,2]. In natural varieties, the panicle morphology is highly diversified and belongs to the quantitative traits, which is contributed by multiple loci/genes [3]. Breeding elite variety relies on those loci. However, excellent panicle traits are mainly selected empirically by the visual observation in traditional breeding, which is laborious and difficult to combine elite loci together. Therefore, it is valuable to unveil the genetic loci responsible for those panicle elements and use the information to facilitate marker-assistant selection (MAS) breeding.

The release of both *japonica* and *indica* genome information provided rich sequence polymorphisms for marker development [4,5], and in combination of robust analytical algorithm, the genetic dissection of quantitative trait loci (QTL) has become a common approach [6]. Up to now, many QTL loci have been identified to confer panicle morphology [7], however, most of them are located in large genomic region and the genetic effect is ambiguous in primary mapping populations, limiting the utilization of those loci in rice MAS. Therefore, it is necessary to confirm the effects of relative loci and then finely map and clone those genes. Developing near-isogenic lines has been a good strategy for QTL confirmation and for fine mapping with derived populations [8]. Even for the minor QTL, the NIL population could enlarge its genetic effect and the variance it explained, and map the QTL as single mendelian factor [9]. With this strategy, many QTLs have been finely mapped and constrained to a narrow genomic region, providing essential information for exploring the candidate genes [10].

By now, several important QTLs for panicle morphology have been cloned, such as *Gn1a* for grain numbers, *Ghd7* and *Ghd8*/*DTH8* for panicle size, *dep1* for panicle density and erection and *IPA1*/*WFP* for panicle primary branch number, all of which showed the potential to increase rice yield [11–17]. *Gn1a* encodes a cytokinin oxidase that degrades the phytohormone cytokinin, and its reduced expression causes cytokinin accumulation in inflorescence meristems and increases the number of reproductive organs [11]. *Ghd7* encodes a CCT domain protein, which delays flowing time and increases panicle size by suppressing downstream *Ehd1 e*xpression under long-day conditions, so does *Ghd8*/*DTH8*, a gene encoding the OsHAP3 subunit of HAP complex [12–14]. *dep1* is a gain-of-function mutation, causing truncation of a phosphatidylethanolamine-binding protein-like domain protein. The mutant can enhance meristematic activity, resulting in a reduced length of the inflorescence internode, an increased number of grains per panicle and a consequent increase in grain yield [15]. *IPA1* encodes the OsSPL14, an SBP domain transcription factor, and higher expression of *OsSPL14* in the reproductive stage promotes panicle branching and rice grain yield [16,17].

The mechanistic explanation of those genes facilitates our understanding of panicle development and the utilization of them in rice breeding [3]. However, identification of new locus should enrich the pathway and provide more choices for panicle shape adjustment. In our previous report, we identified two QTLs for panicle length from the BC_3F_3 populations, namely *qPL6* and *qPL8*, and finely mapped the *qPL6* locus [18]. Here, we report the validation and fine-mapping of *qPL8* using the NIL population, and the candidate gene(s) was predicted according to the coding sequence variance and gene expression difference.

2 Materials and Methods

2.1 Plant Materials

As shown in our previous report, near-isogenic lines for *qPL8* were developed from the BC_3F_6 generation derived from the cross of Nipponbare and WS3, a landrace bearing large panicle and strong culm [18]. The NILs with reciprocal *qPL8* alleles were planted in the Shanghai experimental station for QTL confirmation and trait evaluation. For easy citation of the two lines, they are named NIL^{qPL8+} and NIL^{qPL8-} hereafter, and plus and minus indicate the presence and absence of the superior *qPL8* allele. To narrow down the QTL region, progenies of heterozygous NIL plants were planted for recombinant identification. Then, selected recombinants and their progenies were planted for phenotype confirmation. To evaluate the impact of *qPL8* and *qPL6* on expression of MADS-box genes, we also planted the NILs with four allele combinations of the two locus as previously reported [18].

2.2 Agronomic Trait Evaluation

To confirm the effect of *qPL8* on agronomic traits, eight agronomic traits were evaluated by NIL comparison, including panicle length (PL), panicle primary branches number (PBN), panicle secondary branches number (SBN), spikelet number per main panicle (SPP), seed setting rate (SR), plant height (PH), tiller number (TN)

and stem diameter of internode 3 (SD). For the progenies of key recombinants, PL from the main culm of each individual was obtained. The methods for these traits evaluation followed the previous report [18].

2.3 Marker Development for QTL Fine Mapping

Eight new InDel markers that evenly distributed in the *qPL8* region were designed for QTL fine mapping with the software Primer premier 5.0 according to the InDel information from two parents, WS3 and Nipponbare. With these markers, five recombinants dividing the QTL region to different parts were identified. To clarify the borders of key recombinants, one additional InDel and three dCaps markers were developed. The progenies of those recombinants were genotyped with markers located in the heterozygous region, and homozygous plants with reciprocal genotypes were obtained for phenotype confirmation. The markers used in QTL fine mapping are listed in Tab. S1. To clarify the NILs genetic background, 276 polymorphic markers were used as previously reported [18].

2.4 DNA Extraction and PCR Condition

Genomic DNA of WS3 was extracted by the CTAB method with slight modification. DNA for NIL genotyping and recombinants selection was extracted by the previous method [19]. PCR for InDels and dCaps markers followed the same procedure described as before [18].

2.5 Identification of Sequence Variation

The WS3 genomic DNA was subjected to library construction for paired-end sequencing, and clean data was obtained after removing adapter and low quality sequence. Then the clean reads were aligned to the Nipponbare reference genome (version_7.0), and SNPs and InDels information were extracted by SOAPsnp [20] and SAMtools [21] separately. According to the genome annotation information, SNPs and InDels in coding sequence of mapping region were identified.

2.6 RNA Extraction, RT-PCR and qRT-PCR

To analyze gene expression in the mapping region, young panicles (1–2 cm long) from the NILs were sampled and total RNA was extracted by the TRIpure Regent (Aidelai, Beijing) following the manufacturer's instruction, which was then reverse-transcribed by iscriptTM cDNA sythesis kit (Bio-Rad). Primers spanning the introns of candidate genes were designed for RT-PCR analysis at proper cycles, and *Ubiquitin* was amplified as internal control. The genomic DNA was also amplified to confirm the primer quality. The candidate genes *LOC_Os08g41950* (*OsMADS7*) and *LOC_Os08g41960* (*OsMADS37*) were subjected to qRT-PCR analysis, and *OsActin* was amplified as internal control. All the primers used for expression analysis are listed in Tab. S2.

3 Results

3.1 Effect Validation of qPL8

In our previous report, the *qPL8* was identified as a QTL for panicle length, and it has no pleiotropic effect on other traits. To further validate the effect of *qPL8*, we planted the NILs with reciprocal alleles and performed a detailed trait analysis. As shown in Fig. 1, no visual morphological difference can be observed for the whole plant between the NIL lines (Fig. 1a), but difference in panicle length can be observed by close-up inspection of the panicle (Fig. 1b). The observation was confirmed by the statistical analysis, and in addition to panicle length, difference in plant height was also identified at significant level, but no difference was detected for other traits (Tab. 1). With 276 polymorphic markers evenly distributed on 12 chromosomes, genotypes of two NILs were clarified. The result showed that all the genetic backgrounds of the NILs were homozygous and only the region at *qPL8* locus was different. The region covers about 2000-kb physical interval flanked by markers RM23466 and RM3120, laying the foundation for fine mapping (Fig. 2).

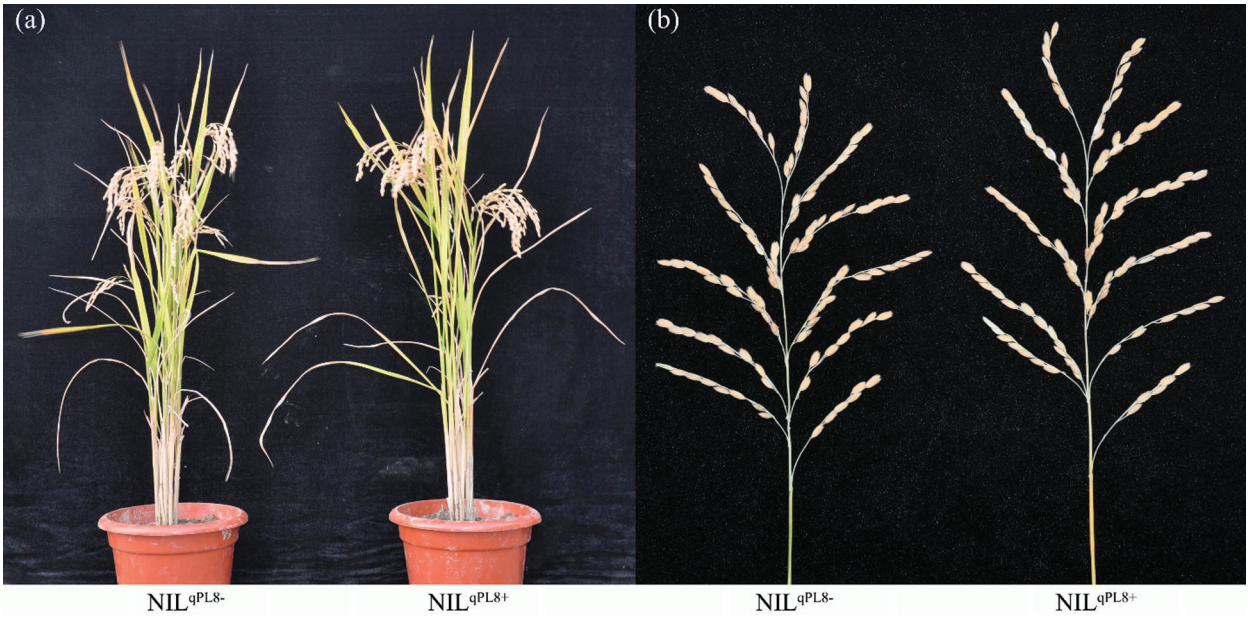

Figure 1: Phenotype of NIL^{qPL8-} and NIL^{qPL8+}. (a) The gross view of the whole plants at mature stage. (b) Close-up view of the panicles from main culms

Table 1: Agronomic traits comparison between NIL^{qPL8-} and NIL^{qPL8+}

Traits[a]	PL	PH	TN	SD	PBN	SBN	SPP	SR
NIL^{qPL8-}	18.56 ± 1.27	68.8 ± 2.9	9.3 ± 2.1	3.97 ± 0.39	11.4 ± 0.7	17.2 ± 4.9	115.6 ± 17.5	93.15% ± 2.96%
NIL^{qPL8+}	19.84 ± 1.22	71.9 ± 2.6	10.2 ± 2.9	4.07 ± 0.29	11.6 ± 0.7	17.1 ± 4.0	114.9 ± 14	92.43% ± 1.81%
p-value	0.0042**	0.002**	0.3098	0.3979	0.3382	0.9475	0.919	0.4768

Note:[a]The full names of different traits are panicle length (PL), plant height (PH), tiller number (TN), stem diameter of internode 3 (SD), panicle primary branches number (PBN), panicle secondary branches number (SBN), spikelet number per main panicle (SPP) and seed setting rate (SR).

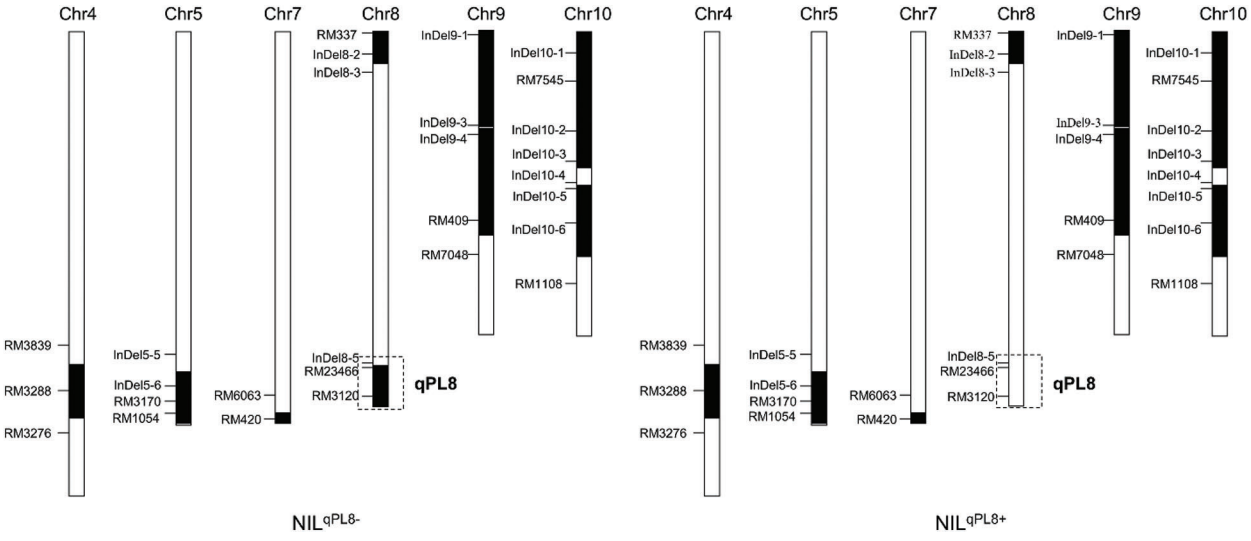

Figure 2: Graphical genotype of NIL^{qPL8-} and NIL^{qPL8+}. All the chromosomes that include introgression from parent WS3 are shown. The white and black bars indicate the region of Nipponbare and WS3, respectively. Markers that confine the borders of introgression are labeled. Dotted boxes indicate the position of *qPL8*

3.2 Fine Mapping of qPL8

By genotyping the NIL derived population, five recombinants were identified to divide the *qPL8* region to several parts (Fig. 3a). The progeny lines of these recombinants were subjected to trait measurement. If the locus resides in the heterozygous region, the progeny with different genotypes would exhibit difference in panicle length, otherwise, the locus should reside in the homozygous region. Accordingly, R1 denoted *qPL8* to the left of the marker qPL8ID-7, and R2 and R3 confirmed the result and further denoted the locus to left of the marker qPL8ID-6. Finally, R4 and R5 confined the locus between markers qPL8ID-1 and qPL8ID-4, harboring a ~318-kb genomic region. Then, the breaking point of R4 and R5 were clarified by four new markers, and the *qPL8* was finally mapped to a 287-kb interval flanked by markers qPL8dCaps1 and qPL8dCaps3 (Fig. 3b).

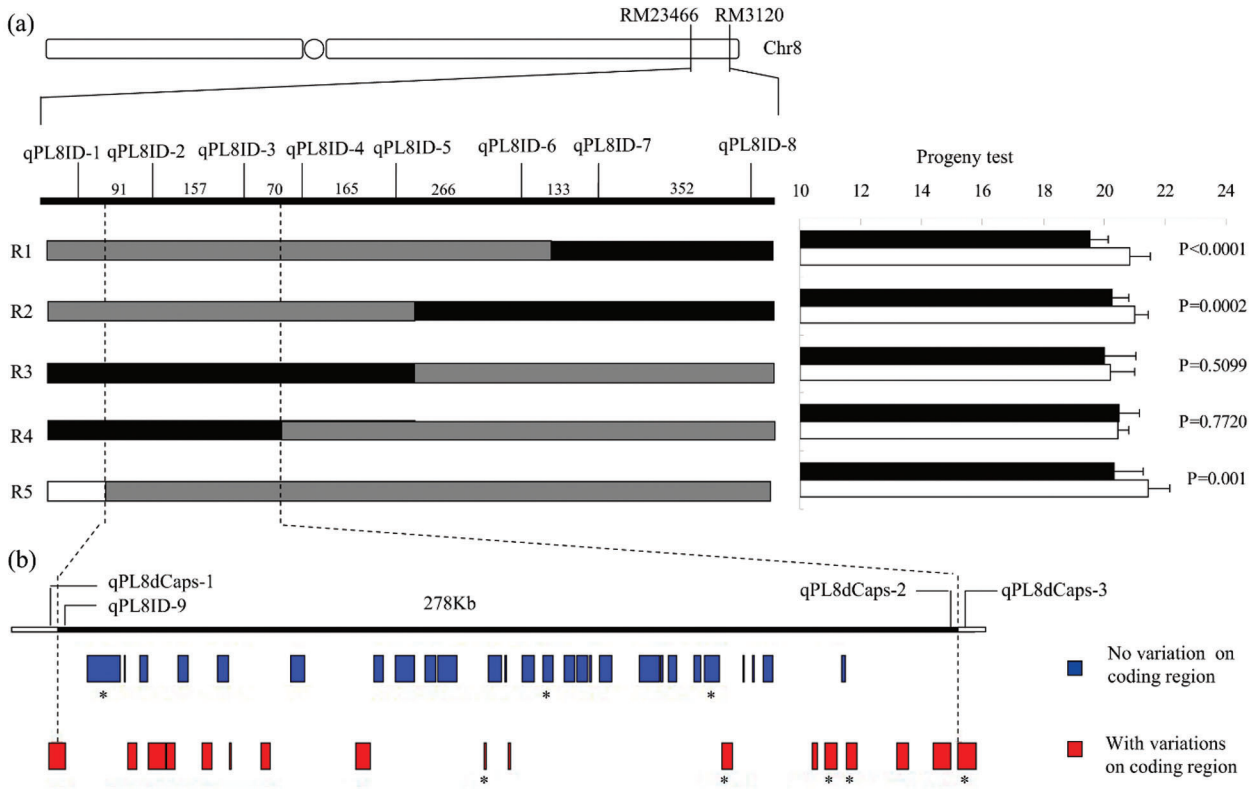

Figure 3: Fine mapping and candidate gene prediction of *qPL8*. (a) R1-R5 indicate the recombinants that divide the *qPL8* region. Eight InDels markers were used to denote the breaking point, and the numbers between markers indicate physical interval in kb. Black and white bars indicate the homozygous region from parents WS3 and Nipponbare, and gray bar indicates the heterozyougs region. The progeny sibling lines of each recombinants with reciprocal genotypes from the heterozyous region were measured for panicle length (cm), and *t*-test were performed for difference comparison of two genotypes (n = 16). The resulting *p*-values were showed on the right with $p < 0.01$ as significance threshold. (b) Candidate genes in the 287-kb mapping region. All the 44 annotated genes are plotted along the mapping region according to the gene position and gene size, and genes with and without sequence variation are highlighted by the red and blue boxes respectively. Genes with expression difference were labeled with asterisks under corresponding boxes

3.3 Coding Variance of Candidate Genes

According to MSU rice genome annotation project database (http://rice.plantbiology.msu.edu), 44 genes are predicted in the mapping region (Fig. 3b). Mutations on coding regions usually have effect on the phenotype. We then compared the mapped genomic region between two parents to identify sequence variations that lead to amino acid change. The sequencing depth of parent WS3 is about 39×, which provides high fidelity for the sequence variation. In all, 23 SNPs and 5 Indels were identified on the protein coding region (Tabs. S3 and S4), which distributed on 17 genes of the mapping region (Fig. 3b). None of the SNPs produces mutation as premature termination. Among the five InDels, two of them are not in triplex form, which should change the predicted protein largely. A 2-bp deletion was identified in one of the two transcripts from *LOC_Os08g41950*, a gene encoding the MADS-box family protein (OsMADS7), and an 8-bp insertion was identified in the gene *LOC_Os08g42060*, which was annotated as expressed protein.

3.4 Expression Variation of the Candidate Genes

In addition to coding variation, expression difference would also contribute to the phenotype change, and we further compared the expression of candidate genes between the NILs. As panicle length was the major difference between the NILs, the gene expression analysis was performed with young panicles in the stage 5 of eight rice inflorescence stages defined by Ding Ying, the father of rice in China. Among the 44 genes, eight were annotated as transposon elements (TEs) and two were hypothetical proteins without expression evidence. The ten predicted genes were removed from the candidate list. For the rest genes, we succeeded in designing primers except for *LOC_Os08g42240*, which is rich in repeat element and could also be a putative TE. The RT-PCR results showed that eight genes were up-regulated in the NIL with superior *qPL8* allele (Fig. 4), and five of them also bear coding sequence variation (Fig. 3b). Therefore, the candidate genes for *qPL8* are suggested to be among the 19 genes with either coding variation or expression difference (Tab. 2). We also found that four of the 19 candidate genes including the gene with 8-bp insertion were not expressed in the young panicle (Fig. 4 and Tab. 2), reducing their possibility as the causative genes. Among the eight differentially expressed genes, expression difference of *LOC_Os08g42310* and *LOC_Os08g42320* was obvious between the NILs, and the expression could be hardly detected in NIL^{qPL8-} (Fig. 4). Interestingly, five of the six additional genes with expression difference are located tightly around the two genes (Fig. 3b), making it possible that this region affects gene expression and functions as a cis-expression quantitative trait locus (cis-eQTL).

3.5 MADS-Box Genes Might Function in Panicle Length Control

Although many candidate genes were identified, we suggested that the MADS-box genes might be the most likely candidate genes for phenotype variation, as such genes have been repeatedly reported to participate in inflorescence and floret development [22]. In the mapping region, two MADS-box genes, *LOC_Os08g41950* (*OsMADS7*) and *LOC_Os08g41960* (*OsMADS37*), were identified. Although there was a large deletion in the predicted transcript 1 of *OsMADS7*, we could not detect this transcript in both NILs. Amplification of transcript 2 of *OsMADS7* was successful, but no difference was detected between the NILs (Fig. 4). These results suggested the deletion in *OsMADS7* might have little effect on gene function. For *OsMADS37*, no coding variation was identified, but its expression was elevated significantly in NIL^{qPL8+}. Expression profile of two MADS-box genes were confirmed by qRT-PCR, and NILs with *qPL6* background were also subjected to analysis (Fig. 5). Interestingly, expression of *OsMADS7* was elevated by *qPL6* background, suggesting a role of *OsMADS7* in mediating the panicle length control by *qPL6*. In an analogous manner, the up-regulation of *OsMADS37* might contribute to the *qPL8*-mediated panicle length control.

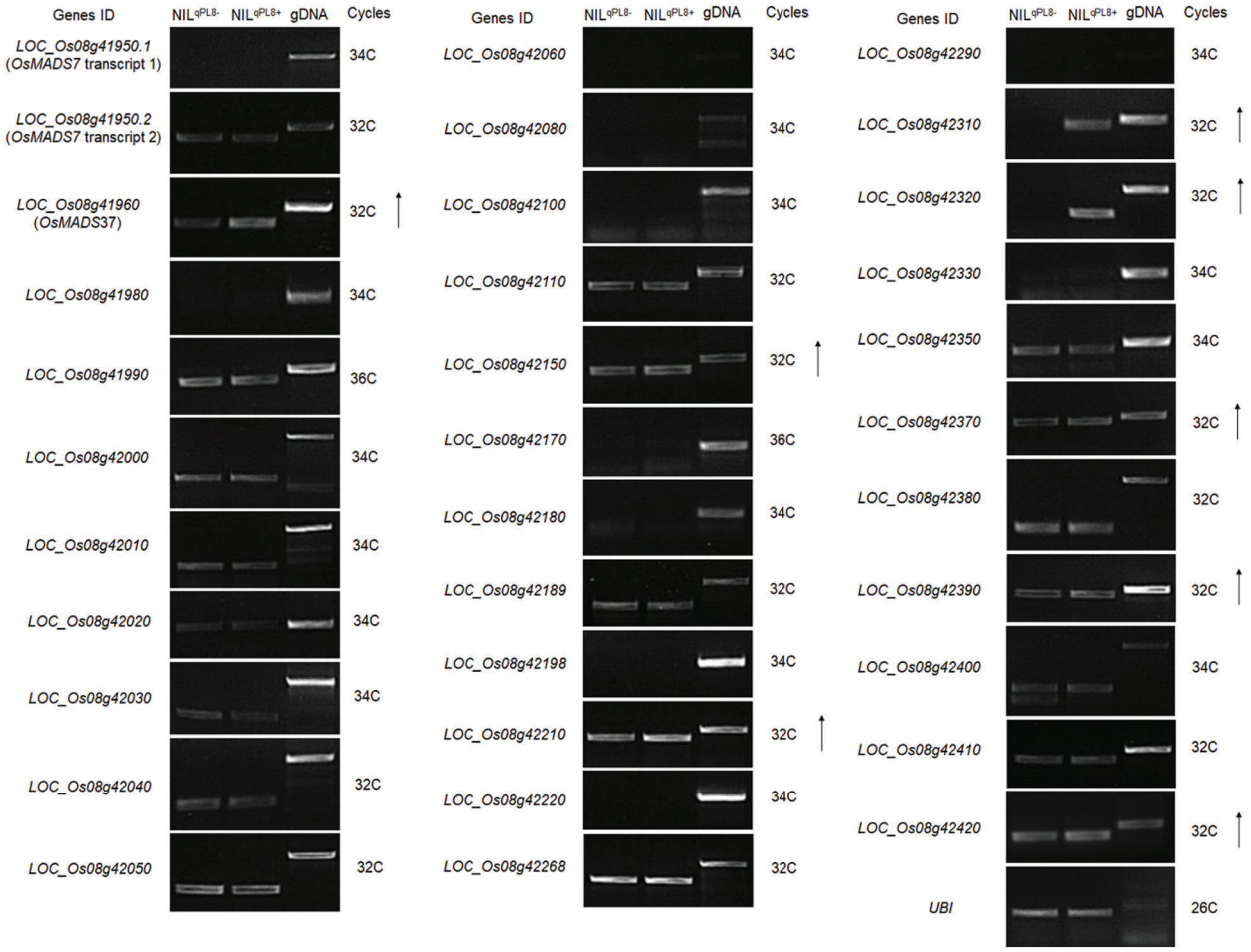

Figure 4: RT-PCR analysis for 33 genes in the mapping region. The result is confirmed by four independent experiments and representative one is showed with gene IDs. cDNAs from young panicles of NIL^{qPL8-} and NIL^{qPL8+} were amplified for different genes at suitable cycles as showed. *UBI* denotes the amplification of *ubiquitin*. Arrows designate the genes with expression difference

Table 2: The information of candidate genes with sequence variation or expression difference

Gene IDs	Putative Functions	Coding variation	Expression status
LOC_Os08g41950	OsMADS7–MADS-box family gene with MIKCc type-box	yes	no difference
LOC_Os08g41960	OsMADS37–MADS-box family gene with MIKC* type-box	no	up-regulation
LOC_Os08g41990	aminotransferase	yes	no difference
LOC_Os08g42010	nodulin	yes	no difference
LOC_Os08g42020	zinc ion binding protein	yes	no difference
LOC_Os08g42040	LTPL80 - Protease inhibitor/seed storage/LTP family protein precursor	yes	no difference

(Continued)

Table 2 (continued).

Gene IDs	Putative Functions	Coding variation	Expression status
LOC_Os08g42060	expressed protein	yes	not expressed
LOC_Os08g42080	ACR5	yes	not expressed
LOC_Os08g42100	ACT domain containing protein	yes	not expressed
LOC_Os08g42150	zinc transporter 2 precursor	yes	up-regulation
LOC_Os08g42180	expressed protein	yes	not expressed
LOC_Os08g42210	expressed protein	no	up-regulation
LOC_Os08g42310	expressed protein	no	up-regulation
LOC_Os08g42320	expressed protein	yes	up-regulation
LOC_Os08g42370	zinc finger DHHC domain-containing protein	yes	up-regulation
LOC_Os08g42390	glycerophosphoryl diester phosphodiesterase family protein	yes	up-regulation
LOC_Os08g42400	no apical meristem protein	yes	no difference
LOC_Os08g42410	transketolase	yes	no difference
LOC_Os08g42420	expressed protein	yes	up-regulation

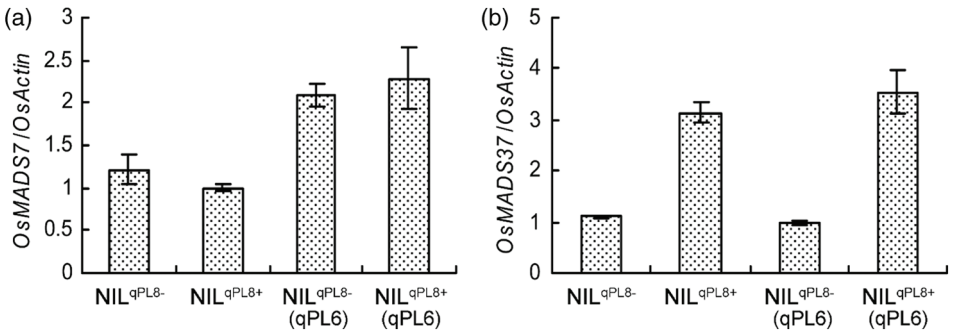

Figure 5: Quantitative RT-PCR for *OsMADS7* (a) and *OsMADS37*. (b) Characters in parentheses indicate the NILs with superior *qPL6* allele. Error bars represent S.D. of three replicates

4 Discussion

As a panicle component, panicle length is usually measured as the yield-related trait, and many QTL for this trait was identified throughout the 12 chromosomes in rice [23,24], but few of them have been finely mapped. In our previous study, we identified a PL QTL on chromosome 8, and named it *qPL8* [18]. However, the QTL interval for *qPL8* was too large, making it hard to identify the candidate gene and apply it to rice breeding. In this study, we clarified the background of two NILs, and made sure that only the QTL region is different, therefore the difference between the NILs should only be caused by the QTL region. In addition to the panicle length, we also found the region had slight effect on the plant height. The field test confirmed that the PL trait conferred by *qPL8* is stable among different seasons, making fine mapping of the QTL feasible. During NIL development, the heterozygous plants were used to develop population for QTL fine mapping directly. Similar to the process of NIL development, each recombinant was selfed to produce sibling lines with reciprocal genotypes, which made phenotypes

determination more accurate. Accordingly, we succeeded in narrowing down the QTL to 278-kb interval, demonstrating the availability of the strategy. However, we could not further narrow down the mapping interval due to low recombination rate in the region, similar to the case of *Ghd7* [12]. Nevertheless, the present result still facilitates the prediction of the underlying gene for *qPL8*.

Recent advance in sequencing technology has enabled easy re-sequencing of multiple rice genomes, which facilitates identification of functional variations in elite varieties [25–27], and it has been used to predict candidate genes in genome wide association studies (GWAS) [28]. In this study, we provided high-quality coding variation in the mapping region to predict the candidate genes by re-sequencing the parent WS3, and found rich SNPs and deletions in this region. Although we could not exclude the possibility that small variant might lead to phenotype change as showed by some study [29], expression variation is likely the alternative mechanism under the *qPL8* phenotype, given that obvious expression difference are usually associated with the tissues showing phenotypes [30,31]. Therefore, we further qualify the candidate gene list of *qPL8* by expression analysis.

By surveying the published function, we suggested the two MADS-box genes as the most likely gene(s) for PL control. Although the expression of *OsMADS7* was not affected by *qPL8*, it is under control of *qPL6*, reflecting the roles of MADS-box genes in panicle elongation. Similarly, the *OsMADS37* was greatly up-regulated by the *qPL8*, reminiscent of the *OsMADS7* up-regulation by *qPL6*. Several transgenic studies have been performed to elucidate the function of *OsMADS7* [32,33]. In the *OsMADS7*-RNAi lines, no obvious phenotype was observed, but silencing of both *OsMADS7* and its homolog *OsMADS8* caused severe morphological alterations of floral organs. Interestingly, PL was greatly decreased by knocking down *OsMADS1/5/7/8* simultaneously [32], and ectopic over-expression of *OsMADS7* led to early flowering and dwarf phenotype [33]. *OsMADS37* is one of five MADS-box genes that belong to the MIKC* subgroup, however, its biological function is still unknown [34]. It would be worthy to clarify the effect of two MADS-box genes on PL by over-expressing them with native promoter.

It is interesting that a cluster of genes were up-regulated in the fine mapping region of *qPL8*. We therefore postulate that the *qPL8* region harbors an eQTL that affects gene expression in cis, which was conceptualized recently [35]. Genome-wide eQTL analysis has been performed in rice, linking the transcript variation to the genomic loci [36]. It has been reported that cis-control of gene expression has important roles in crop domestication [30,31,37,38], and understanding its molecular mechanism will facilitate the manipulation of relative traits. The NILs developed in our study could be ideal plant materials for further exploration of the cis-eQTL.

Fine-mapping of the *qPL8* locus will facilitate its application in rice breeding for yield potential. Although the grain number per panicle was not significantly improved, the locus can be used to modify the panicle shape, which may improve final yield by changing canopy structure [39]. It is also notable that the rice yield could be improved by pyramiding minor yield-related genes even though each of them contributes little to the total yield [40]. In this scenario, *qPL8* can be selected to pyramid with other loci to increase yield potential, and we did find the better effect of *qPL8* in another background. Moreover, the markers developed in this study will simplify the application of *qPL8* locus.

5 Conclusion

In this study, we confirmed the effect of *qPL8* in shaping panicle length by detailed trait analysis of the NILs. Five recombinants were identified from the NIL populations using 12 newly-developed InDel/Caps markers, and the *qPL8* locus was finally confined to a 278 kb interval covering 44 genes. Facilitated by whole genome sequencing and RT-PCR, 17 genes with sequence variation and eight genes with expression variation were identified among the candidate gene list. Especially, expression of *OsMADS37* and *OsMADS7* was responsive to the allele variation of *qPL8* and *qPL6*, respectively, suggesting that

MADS-box genes might mediate the control of panicle length. This work lays the foundation for gene cloning and breeding utilization of *qPL8* in the future.

Funding Statement: This work was supported by grants from the National Natural Science Foundation of China (31871217 and 32072037), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (20KJA210002) and the Key Research and Development Program of Jiangsu Province (BE2018357).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Liu, T. M., Li, L. Z., Zhang, Y. S., Xu, C. G., Li, X. H. et al. (2011). Comparison of quantitative trait loci for rice yield, panicle length and spikelet density across three connected populations. *Journal of Genetics, 90(2),* 377–382. DOI 10.1007/s12041-011-0083-9.
2. Chongloi, G. L., Prakash, S., Vijayraghavan, U. (2019). Regulation of meristem maintenance and organ identity during rice reproductive development. *Journal of Experimental Botany, 70(6),* 1719–1736. DOI 10.1093/jxb/erz046.
3. Xing, Y. Z., Zhang, Q. F. (2010). Genetic and molecular bases of rice yield. *Annual Review of Plant Biology, 61(1),* 421–442. DOI 10.1146/annurev-arplant-042809-112209.
4. Wang, W. S., Mauleon, R., Hu, Z. Q., Chebotarov, D., Tai, S. S. et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature, 557(7703),* 43–49. DOI 10.1038/s41586-018-0063-9.
5. Zhao, H., Yao, W., Ouyang, Y. D., Yang, W. N., Wang, G. W. et al. (2015). RiceVarMap: A comprehensive database of rice genomic variations. *Nucleic Acids Research, 43(D1),* D1018–D1022. DOI 10.1093/nar/gku894.
6. Meng, L., Li, H. H., Zhang, L. Y., Wang, J. K. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop Journal, 3(3),* 269–283. DOI 10.1016/j.cj.2015.01.001.
7. Yamamoto, T., Yonemaru, J., Yano, M. (2009). Towards the understanding of complex traits in rice: Substantially or superficially? *DNA Research, 16(3),* 141–154. DOI 10.1093/dnares/dsp006.
8. Ding, X. P., Li, X. K., Xiong, L. H. (2011). Evaluation of near-isogenic lines for drought resistance QTL and fine mapping of a locus affecting flag leaf width, spikelet number, and root volume in rice. *Theoretical and Applied Genetics, 123(5),* 815–826. DOI 10.1007/s00122-011-1629-1.
9. Zhang, Y. S., Luo, L. J., Liu, T. M., Xu, C. G., Xing, Y. Z. (2009). Four rice QTL controlling number of spikelets per panicle expressed the characteristics of single Mendelian gene in near isogenic backgrounds. *Theoretical and Applied Genetics, 118(6),* 1035–1044. DOI 10.1007/s00122-008-0960-7.
10. Bai, X. F., Wu, B., Xing, Y. Z. (2012). Yield-related QTLs and their applications in rice genetic improvement. *Journal of Integrative Plant Biology, 54(5),* 300–311. DOI 10.1111/j.1744-7909.2012.01117.x.
11. Ashikari, M., Sakakibara, H., Lin, S. Y., Yamamoto, T., Takashi, T. et al. (2005). Cytokinin oxidase regulates rice grain production. *Science, 309(5735),* 741–745. DOI 10.1126/science.1113373.
12. Xue, W. Y., Xing, Y. Z., Weng, X. Y., Zhao, Y., Tang, W. J. et al. (2008). Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nature Genetics, 40(6),* 761–767. DOI 10.1038/ng.143.
13. Yan, W. H., Wang, P., Chen, H. X., Zhou, H. J., Li, Q. P. et al. (2011). A Major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Molecular Plant, 4(2),* 319–330. DOI 10.1093/mp/ssq070.
14. Wei, X. J., Xu, J. F., Guo, H. N., Jiang, L., Chen, S. H. et al. (2010). DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiology, 153(4),* 1747–1758. DOI 10.1104/pp.110.156943.
15. Huang, X. Z., Qian, Q., Liu, Z. B., Sun, H. Y., He, S. Y. et al. (2009). Natural variation at the DEP1 locus enhances grain yield in rice. *Nature Genetics, 41(4),* 494–497. DOI 10.1038/ng.352.
16. Jiao, Y. Q., Wang, Y. H., Xue, D. W., Wang, J., Yan, M. X. et al. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature Genetics, 42(6),* 541–544. DOI 10.1038/ng.591.

17. Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M. et al. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics, 42(6),* 545–549. DOI 10.1038/ng.592.
18. Zhang, L., Wang, J. J., Wang, J. M., Wang, L. Y., Ma, B. et al. (2015). Quantitative trait locus analysis and fine mapping of the qPL6 locus for panicle length in rice. *Theoretical and Applied Genetics, 128(6),* 1151–1161. DOI 10.1007/s00122-015-2496-y.
19. Xu, Y. H., Zhu, Y. Y., Zhou, H. C., Li, Q., Sun, Z. X. et al. (2004). Identification of a 98-kb DNA segment containing the rice Eui gene controlling uppermost internode elongation, and construction of a TAC transgene sublibrary. *Molecular Genetics and Genomics, 272(2),* 149–155. DOI 10.1007/s00438-004-1019-3.
20. Li, R. Q., Li, Y. R., Fang, X. D., Yang, H. M., Wang, J. et al. (2009). SNP detection for massively parallel whole-genome resequencing. *Genome Research, 19(6),* 1124–1132. DOI 10.1101/gr.088013.108.
21. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J. et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25(16),* 2078–2079. DOI 10.1093/bioinformatics/btp352.
22. Gao, X. C., Liang, W. Q., Yin, C. S., Ji, S. M., Wang, H. M. et al. (2010). The SEPALLATA-Like Gene OsMADS34 is required for rice inflorescence and spikelet development. *Plant Physiology, 153(2),* 728–740. DOI 10.1104/pp.110.156711.
23. Zhu, Z. Z., Li, X. Q., Wei, Y., Guo, S. B., Sha, A. H. (2018). Identification of a novel QTL for panicle length from wild rice (Oryza minuta) by specific locus amplified fragment sequencing and high density genetic mapping. *Frontiers in Plant Science, 9,* 1492. DOI 10.3389/fpls.2018.01492.
24. Zhang, B., Shang, L. G., Ruan, B. P., Zhang, A. P., Yang, S. L. et al. (2019). Development of three sets of high-throughput genotyped rice chromosome segment substitution lines and QTL mapping for eleven traits. *Rice, 12(1),* 27. DOI 10.1186/s12284-019-0293-y.
25. Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F. Y. et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology, 30(1),* 105–111. DOI 10.1038/nbt.2050.
26. Xie, W. B., Wang, G. W., Yuan, M., Yao, W., Lyu, K. et al. (2015). Breeding signatures of rice improvement revealed by a genomic variation map from a large germplasm collection. *Proceedings of the National Academy of Sciences of the United States of America, 112(39),* E5411–E5419. DOI 10.1073/pnas.1515919112.
27. Li, X. X., Chen, Z., Zhang, G. M., Lu, H. W., Qin, P. et al. (2020). Analysis of genetic architecture and favorable allele usage of agronomic traits in a large collection of Chinese rice accessions. *Science China Life Sciences, 63(11),* 1688–1702. DOI 10.1007/s11427-019-1682-6.
28. Huang, X. H., Zhao, Y., Wei, X. H., Li, C. Y., Wang, A. et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics, 44(1),* 32–39. DOI 10.1038/ng.1018.
29. Jin, J., Huang, W., Gao, J. P., Yang, J., Shi, M. et al. (2008). Genetic control of rice plant architecture under domestication. *Nature Genetics, 40(11),* 1365–1369. DOI 10.1038/ng.247.
30. Huo, X., Wu, S., Zhu, Z. F., Liu, F. X., Fu, Y. C. et al. (2017). NOG1 increases grain production in rice. *Nature Communications, 8(1),* 25. DOI 10.1038/s41467-017-01501-8.
31. Huang, Y. Y., Zhao, S. S., Fu, Y. C., Sun, H. D., Ma, X. et al. (2018). Variation in the regulatory region of FZP causes increases in secondary inflorescence branching and grain yield in rice domestication. *Plant Journal, 96(4),* 716–733. DOI 10.1111/tpj.14062.
32. Cui, R. F., Han, J. K., Zhao, S. Z., Su, K. M., Wu, F. et al. (2010). Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). *Plant Journal, 61(5),* 767–781. DOI 10.1111/j.1365-313X.2009.04101.x.
33. Jeon, J. S., Lee, S., Jung, K. H., Yang, W. S., Yi, G. H. et al. (2000). Production of transgenic rice plants showing reduced heading date and plant height by ectopic expression of rice MADS-box genes. *Molecular Breeding, 6(6),* 581–592. DOI 10.1023/A:1011388620872.
34. Arora, R., Agarwal, P., Ray, S., Singh, A. K., Singh, V. P. et al. (2007). MADS-box gene family in rice: Genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics, 8(1),* 242. DOI 10.1186/1471-2164-8-242.

35. Jansen, R. C., Nap, J. P. (2001). Genetical genomics: The added value from segregation. *Trends in Genetics, 17(7),* 388–391. DOI 10.1016/S0168-9525(01)02310-1.
36. Wang, J., Yu, H. H., Xie, W. B., Xing, Y. Z., Yu, S. B. et al. (2010). A global analysis of QTLs for expression variations in rice shoots at the early seedling stage. *Plant Journal, 63(6),* 1063–1074. DOI 10.1111/j.1365-313X.2010.04303.x.
37. Clark, R. M., Wagler, T. N., Quijada, P., Doebley, J. (2006). A distant upstream enhancer at the maize domestication gene tb1 has pleiotropic effects on plant and inflorescent architecture. *Nature Genetics, 38(5),* 594–597. DOI 10.1038/ng1784.
38. Zhu, Z. F., Tan, L. B., Fu, Y. C., Liu, F. X., Cai, H. W. et al. (2013). Genetic control of inflorescence architecture during rice domestication. *Nature Communications, 4(1),* 630. DOI 10.1038/ncomms3200.
39. Li, F., Liu, W. B., Tang, J. Y., Chen, J. F., Tong, H. N. et al. (2010). Rice DENSE AND ERECT PANICLE 2 is essential for determining panicle outgrowth and elongation. *Cell Research, 20(7),* 838–849. DOI 10.1038/cr.2010.69.
40. Cheung, F. (2014). Yield: The search for the rice of the future. *Nature, 514(7524),* S60–S61. DOI 10.1038/514S60a.

Appendix

Table S1: Markers developed for qPL8 fine mapping

Marker name	Forward primer (5'-3')	Reverse primer (5'-3')	Enzyme
qPL8ID-1	GAAAAAAAATCACAAAGCAC	CAGGCACCACACTACAACTA	
qPL8ID-2	CGACAGACGAGACGACAGG	ACGGAAGACGGGAGCTAAG	
qPL8ID-3	GAAAGAGGGTAGACTAACGCA	ACCAAAGACAAAGCAACAGGA	
qPL8ID-4	AACTTGCTCTAAGTAATGATTGG	ATAATATGTGGGTAATTTGTTGC	
qPL8ID-5	CAAGTAGGCGTGTATGGTCA	GAGATGCTTCAATCCTGTCC	
qPL8ID-6	AGCCTCTCTTAATTCGATCTC	AACCAGGAACCTATGTTTGTA	
qPL8ID-7	GTTTGTGCCACCCAACTTTA	ACCACCCATCCACTCTTACC	
qPL8ID-8	CCATAAATCAACACAGTCTATTT	AAGGTCTACTATCGCCAGGT	
qPL8ID-9	CCCTTCAGATTGGGTATGCT	GCACTGTTCATCGCCTCAT	
qPL8dCaps-1	AATGTGCCATGCCCCTATTGTAGATCTGAA	TTCGGGTCCTGCGTAACTGC	HaeIII
qPL8dCaps-2	TCAATAACATGTATGTAATGGAGCTCCATG	ACATTTCTGGATAACAGGACC	HinfI
qPL8dCaps-3	CTCGAGCAGGCTGTACGAGTGTAGCTCGAC	TAACCTGAGGCAAACCCATT	HinfI

Table S2: Primers used for semi-quantitative and quantitative RT-PCR analysis

Genes ID	Forward primers (5'-3')	Reverse primers (5'-3')
LOC_Os08g41950.1	CGATGTGACTGAAAAGTTGTG	CAAACAGGCTACATACGAAAC
LOC_Os08g41950.2	TAATTGGCTATGAACGTCAGC	GGAATGAAAATGATAATAATACGC
LOC_Os08g41960	AACTGATGATGAAGGTGGCT	GTTGTTGAAATGGCTGGTAC
LOC_Os08g41980	AGCTGCAGGGGTGGGAGACG	CGGCACGAGCACGAACAACA
LOC_Os08g41990	GAAGGGTTCTCGTATGTGGG	TCATTGTATGGTGCCGTTAGA
LOC_Os08g42000	TGCCACCACGACATCAACAC	CGCCCATAGACTGAAACCAA
LOC_Os08g42010	ATTCAGTGAGGTGGAGGATG	TGCCCAAATGCCATAAATAC

(Continued)

Table S2 (continued).

Genes ID	Forward primers (5'-3')	Reverse primers (5'-3')
LOC_Os08g42020	CTTGCTCCTGAACGCACCC	TCGCATCAATTCGGCACTTA
LOC_Os08g42030	GAGTCCATCGTGCGGTATGAG	TGGTCGGAGGTGAAGAGGC
LOC_Os08g42040	CTGTCTTCGCCAACGCCTCC	ATCTTCGTGTAGCACATCACAAACC
LOC_Os08g42050	CCCAAACTGACCGACAAG	GATGGAAGCTGCATAGTGAA
LOC_Os08g42060	GTTAGCTTCCCTAATTATTTTCACCCTT	GTTCTGCGTGCCGTTCCTCC
LOC_Os08g42080	ATCCAACGCCGCAACACC	TTCTCCGACCGTGACCTGAT
LOC_Os08g42100	CTATGTCCGGCACGTCGAC	CCCCATTCTCCATTTATTTCCT
LOC_Os08g42110	TGAAGTGCCAGTTGATGAGA	TCTTGGAAGCCAACGAGTA
LOC_Os08g42150	AGGACGATGACGATGAGGGTTTCTT	AACGGGTACTGGTTGGTAGTGAGGC
LOC_Os08g42170	TGACCGTCGTCTACTCCCTCG	GGCGCAGGAATTGAAGAAATG
LOC_Os08g42180	GGGCTCTTATCCCCTCCTTCTCCTC	CGCCTCCTCATCCCCACCGT
LOC_Os08g42189	TCTGTGAAGCTGCTGAGGC	AAAGGTGAATCGGATGGAAA
LOC_Os08g42198	GTGAGGAACAAGCAGGGAGTGA	CGGAGCAACCAGCATGAGGT
LOC_Os08g42210	CGTTGTTGGACCTAGCAGTCAT	GTTTCCGTTCATGCTCATTGTC
LOC_Os08g42220	TCCGTGGTTCCCAGGGTT	GCCATCTGAAGTCCGAGTGAT
LOC_Os08g42268	GTCTGTGAAGCTGCTGAGGC	AAAGGTGAATCGGATGGAAA
LOC_Os08g42290	TATCAAAAGCTAAGACAAGTGAG	AATCACAATCCATACATCAAAA
LOC_Os08g42310	AGGTCACAAGCGTACTTTTAGG	TTCTCGATACTTTTGGTGGTTTA
LOC_Os08g42320	ACTTCATGTCATCTGCGGTAC	CACATAAACCCACAAACTCAAT
LOC_Os08g42330	GGCTGGACCGACCATAGAGGGC	GTCCCGCTCTTCCCCATCCG
LOC_Os08g42350	GACCAAGAGCGTCGAGTTCA	TCTTCATCAGTAGCGAGAATCTGT
LOC_Os08g42370	TTGTATTCTCGACCACGTTAC	ATGCGTTAGTCTACCCTCTTT
LOC_Os08g42380	ATGGAGAACTCCGCCGAGAT	GTTGGTCAGGCAAGTCGTTT
LOC_Os08g42390	GGTCTGCCTAGAGTACGAGTT	TAGCCAATGATGAGTGGAGTT
LOC_Os08g42400	AGGGCATGGTGTTCGTCC	TCAGTGCTTCCTTCCCGTTA
LOC_Os08g42410	GAGGATTGCTGGAGCTGATG	GGTTTGGTTTGCCGACTTAT
LOC_Os08g42420	ATTCACTGATGGTCCTTTACTT	ATCCTCGTCAATGCTTGTTA
Ubiquitin	GACGGACGCACCCTGGCTGACTAC	TGCTGCCAATTACCATATACCACGAC
LOC_Os08g41950 qRT	TCAGAAATGCAGTTACGCAGGAC	TTGCGGCTAGCTTTCAATTGCTC
LOC_Os08g41960 qRT	ATGTCAACGAGCTGAACATCGC	TCTTCCTCTTCTTCTCGCTTTGGG
Actin	CGGGAAATTGTGAGGGACAT	AGGAAGGCTGGAAGAGGACC

Table S3: SNPs information between two parents on protein coding region

Chromosomes	Transcripts	cds label	SNP position	SNP	codonm change	amino acid change	cds_n start	cds_n end
Chr8	LOC_Os08g41990.1	cds_1	26533972	T<->G	TCC<->GCC;	S<->A;	26533799	26533993
Chr8	LOC_Os08g42010.1	cds_3	26542005	T<->G	TGC<->GGC;	C<->G;	26541981	26542494
Chr8	LOC_Os08g42020.1	cds_1	26544400	T<->G	TTT<->TTG;	F<->L;	26543324	26545045
Chr8	LOC_Os08g42020.2	cds_1	26544400	T<->G	TTT<->TTG;	F<->L;	26543324	26545045

(Continued)

Table S3 (continued).

Chromosomes	Transcripts	cds label	SNP position	SNP	codonm change	amino acid change	cds_n start	cds_n end
Chr8	LOC_Os08g42020.1	cds_1	26544887	C<->T	CGG<->TGG;	R<->W;	26543324	26545045
Chr8	LOC_Os08g42020.2	cds_1	26544887	C<->T	CGG<->TGG;	R<->W;	26543324	26545045
Chr8	LOC_Os08g42060.1	cds_1	26562909	G<->C	GGG<->CGG;	G<->R;	26562505	26562996
Chr8	LOC_Os08g42080.1	cds_7	26574376	A<->G	TAC<->TGC;	Y<->C;	26574357	26574707
Chr8	LOC_Os08g42100.1	cds_5	26604084	T<->C	CTG<->CCG;	L<->P;	26604056	26604682
Chr8	LOC_Os08g42100.1	cds_7	26605218	A<->G	AGC<->GGC;	S<->G;	26604957	26605304
Chr8	LOC_Os08g42100.1	cds_7	26605239	G<->T	GCC<->TCC;	A<->S;	26604957	26605304
Chr8	LOC_Os08g42150.1	cds_1	26640629	T<->C	TGG<->CGG;	W<->R;	26640473	26640655
Chr8	LOC_Os08g42180.1	cds_1	26648526	A<->G	ATG<->GTG;	M<->V;	26647843	26648526
Chr8	LOC_Os08g42320.1	cds_1	26714693	T<->C	TGC<->CGC;	C<->R;	26714649	26715161
Chr8	LOC_Os08g42320.1	cds_1	26714836	A<->G	AAT<->AGT;	N<->S;	26714649	26715161
Chr8	LOC_Os08g42360.1	cds_2	26740966	G<->A	GGC<->GAC;	G<->D;	26740932	26741165
Chr8	LOC_Os08g42360.1	cds_2	26740972	G<->A	GGG<->GAG;	G<->E;	26740932	26741165
Chr8	LOC_Os08g42360.1	cds_2	26740978	G<->A	CGT<->CAT;	R<->H;	26740932	26741165
Chr8	LOC_Os08g42370.1	cds_1	26748130	G<->T	CGC<->CTC;	R<->L;	26748063	26748152
Chr8	LOC_Os08g42390.1	cds_7	26754212	T<->C	CTG<->CCG;	L<->P;	26754016	26754247
Chr8	LOC_Os08g42400.3	cds_2	26767661	T<->G	ATG<->AGG;	M<->R;	26767591	26767877
Chr8	LOC_Os08g42400.2	cds_2	26767661	T<->G	ATG<->AGG;	M<->R;	26767591	26767895
Chr8	LOC_Os08g42400.1	cds_2	26767661	T<->G	ATG<->AGG;	M<->R;	26767591	26767877
Chr8	LOC_Os08g42410.1	cds_1	26778155	A<->G	ATC<->GTC;	I<->V;	26778146	26778184
Chr8	LOC_Os08g42420.5	cds_1	26787238	C<->T	TCG<->TTG;	S<->L;	26787177	26788301
Chr8	LOC_Os08g42420.4	cds_1	26787238	C<->T	TCG<->TTG;	S<->L;	26787177	26788301
Chr8	LOC_Os08g42420.1	cds_1	26787238	C<->T	TCG<->TTG;	S<->L;	26787177	26788301
Chr8	LOC_Os08g42420.2	cds_1	26787238	C<->T	TCG<->TTG;	S<->L;	26787177	26788301
Chr8	LOC_Os08g42420.3	cds_1	26787238	C<->T	TCG<->TTG;	S<->L;	26787177	26788301
Chr8	LOC_Os08g42420.5	cds_3	26789501	T<->C	TTT<->TCT;	F<->S;	26789392	26789517
Chr8	LOC_Os08g42420.1	cds_5	26790262	A<->C	ATT<->CTT;	I<->L;	26790100	26790602
Chr8	LOC_Os08g42420.3	cds_5	26790262	A<->C	ATT<->CTT;	I<->L;	26790100	26790506
Chr8	LOC_Os08g42420.2	cds_5	26790262	A<->C	ATT<->CTT;	I<->L;	26790100	26790506

Table S4: InDels information between two parents on protein coding region

Chromosomes	Transcripts	cds label	InDel position	InDel	InDel sequence	cds_n start	cds_n end
Chr8	LOC_Os08g41950.1	cds_1	26507346	D6	AAGAAG	26507335	26507396
Chr8	LOC_Os08g41950.1	cds_5	26509982	D2	TA	26509894	26509994
Chr8	LOC_Os08g42040.1	cds_1	26556862	I3	CGC	26556589	26556931
Chr8	LOC_Os08g42040.2	cds_1	26556862	I3	CGC	26556589	26556931
Chr8	LOC_Os08g42060.1	cds_1	26562939	I8	GAGAGAGA	26562505	26562996
Chr8	LOC_Os08g42400.1	cds_2	26767797	I6	CGGCGG	26767591	26767877
Chr8	LOC_Os08g42400.2	cds_2	26767797	I6	CGGCGG	26767591	26767895
Chr8	LOC_Os08g42400.3	cds_2	26767797	I6	CGGCGG	26767591	26767877