The PLATZ transcription factor GL6 affects grain length and number in rice

Ahong Wang, Qingqing Hou, Lizhen Si, Xuehui Huang, Jianghong Luo, Danfeng Lu, Jingjie Zhu, Yingying Shangguan, Jiashun Miao, Yifan Xie, Yongchun Wang, Qiang Zhao, Qi Feng, Congcong Zhou, Yan Li, Danlin Fan, Yiqi Lu, Qilin Tian, Zixuan Wang and Bin Han

1National Center for Gene Research, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China.
2University of Chinese Academy of Sciences, Beijing, China.
3College of Life Sciences, Shanghai Normal University, Shanghai, China.
†These authors contributed equally to this work.
*Corresponding author: bhan@ncgr.ac.cn

Short title: GL6 determines grain length and number in rice

One-sentence summary: The plant-specific protein GL6 determines grain length and spikelet number in rice by affecting cell proliferation through gene expression regulation via the RNAPIII transcription machinery.

Author contributions
B.H. designed studies and contributed to the original concept of the project; A.W. and Q.H. performed most of the experiments; Q.Z, Y. L contributed to genome data analysis; Q.F., C.Z., D.F., Q.T. and Y. Lu performed the genome sequencing; L.S., X.H., J.L., D.L., J.Z., Y.S., J.M., Y.X., Y.W. and Z.W provided field genetic analysis, and A.W., Q.H. and B.H. analyzed whole data and wrote the paper.
Funding information:

This work was funded by the Ministry of Science and Technology of China (2016YFD0100902), the National Natural Science Foundation of China (31630055&31788103), and the Chinese Academy of Sciences (XDB27010301).

Abstract

Grain size is one of the key determinants of grain yield. Although a number of genes that control grain size in rice (*Oryza sativa*) have been identified, the overall regulatory networks behind this process remain poorly understood. Here, we report the map-based cloning and functional characterization of the QTL *GL6*, which encodes a plant-specific PLATZ (plant AT-rich sequence- and zinc-binding) transcription factor that regulates rice grain length and spikelet number. *GL6* positively controls grain length by promoting cell proliferation in young panicles and grains. The null *gl6* mutant possesses short grains, whereas overexpression of *GL6* results in large grains and decreased grain number per panicle. We demonstrate that GL6 participates in RNA polymerase III (RNAPIII) transcription machinery by interacting with RPC53 and TFC1 to regulate the expression of genes involved in rice grain development. Our findings reveal a further player involved in the regulation of rice grain size that may be exploited in future rice breeding.

Key words

Rice, grain length, *GL6*, PLATZ, RNAPIII

Introduction

Rice (*Oryza sativa*) is one of the three major cereal crops in the world, and the most important staple food in Asia. Rice has served as a model monocot plant for molecular genetic dissection since its reference genome sequence was generated in 2005 (International Rice Genome Sequencing Project, 2005). The exploitation of rice genetics to increase grain-yield and improve plant architecture are the focus of current
Grain size, one of the most important determinate factors of grain yield, is specified by grain length, width and thickness. In recent years, a number of genes and quantitative trait loci (QTLs) that control grain size have been identified and functionally characterized in rice. These grain size genes have been found to be involved in signaling pathways mediated by G-proteins, proteasomal degradation, phytohormones, protein kinases and transcriptional factors that control cell division and/or cell expansion during seed development (Zuo and Li, 2014). GS3, a major gene controlling grain length and weight, encodes a putative Gγ protein that functions in G-protein signaling (Fan et al., 2006; Mao et al., 2010). GS3 suppresses the interaction between Gβ protein RGB1 and two other Gγ proteins, namely DEP1 and GGC2, to antagonistically regulate grain size (Sun et al., 2018). These Gβγ subunits are also colocalized in the nucleus and interact directly with the downstream effector OsMADS1 to regulate rice grain length and yield (Liu et al., 2018). GW2 encodes a RING-type E3 ubiquitin ligase that functions in protein degradation via the ubiquitin-proteasome pathway, and negatively regulate cell division of spikelet hull, thus affecting grain width (Song et al., 2007). GLW7, a target gene of OsmiR156, encodes the plant-specific transcription factor OsSPL13 that positively regulates grain length and yield. The large-grain allele of GLW7 in tropical japonica varieties was identified to be introgressed from indica varieties under artificial selection (Si et al., 2016). Copy number variation at the GL7 locus affects the expression of two linked genes to regulate grain length (Wang et al., 2015b). GW8, encoding the further SBP-domain transcription factor OsSPL16, positively regulates grain width and directly binds to the GW7/GL7 promoter to repress its expression (Wang et al., 2012; Wang et al., 2015a). Some genes involve in auxin response and brassinosteroid (BR) signaling, such as TGW6 and GW5, also influence the grain size and yield in rice (Ishimaru et al., 2013; Liu et al., 2017). Most of these genes regulate grain size by increasing or decreasing cell number (Qi et al., 2012; Zhang et al., 2012; Liu et al., 2015; Song et al., 2015; Yu et al., 2018; Zhao et al., 2018), and a few genes control grain length by regulating cell size or length (Wang et al., 2015b; Si et al., 2016). In
the case of GS2 and GS5, they affect grain shape by both controlling cell number and size (Li et al., 2011; Che et al., 2015; Hu et al., 2015).

These findings have enriched our knowledge of the regulatory mechanisms behind grain size in rice. However, how these genes are integrated into signaling pathways and the cross-talk between them, as well as into the regulatory networks behind grain development, remain poorly understood. Therefore, identification and molecular characterization of new QTLs/genes involved in grain size will help to comprehensively describe regulatory networks and serve for future improvement of rice yield (Duan et al., 2015).

Here, we report the identification and characterization of the rice grain length QTL GL6. GL6 encodes a PLATZ protein and is preferentially expressed in young panicles. A mutant version of GL6 that carries a premature stop codon results in short grain via impaired RNA polymerase III-mediated transcription. We demonstrate that GL6 functions in a further molecular mechanism that modulates grain length and grain weight.
Results

Map-based cloning of GL6

In our previous work, we mapped a major QTL, GL6, which explained 20.5% of the phenotypic variation ($R^2$) in grain length within a set of 271 lines derived from a cross between a cultivated rice variety *O. sativa* ssp. *indica* Guangluai4 (GLA4) and a wild rice accession *Oryza rufipogon* W1943 (W1943) (Huang et al., 2012). The GL6 locus was initially mapped to the region between the recombination bins Bin_1024 (53.3 cM) and Bin_1028 (55.5 cM) on chromosome 6.

To fine map this locus, one backcross inbred line BIL219 with shorter grain length, harboring the GL6 locus and carrying several other segments from W1943 (Supplemental Figure S1), was selected and backcrossed to GLA4. We then carried out high-resolution mapping using 2,181 BC1F5 individuals, and the locus was finally delimited to a 6.1-kb region between the two markers GL4350 and GL4411 (Figure 1B-D).

According to the RAP annotation (http://rapdb.dna.affrc.go.jp/), this region contains only one candidate gene, Os06g0666100, which encodes a zinc binding PLATZ protein. Compared to GLA4, the coding region of GL6 from BIL219 contained a “C” to “T” substitution in the third exon at nucleotide 352 that introduced a stop codon resulting in premature termination of translation (Figure 1E and Supplemental Figure S2). Interestingly, the original donor parent W1943 represented a “T/C” heterozygote at the premature stop codon allele; however, all other single nucleotide polymorphisms (SNPs) were identical to those in BIL219. Thus, we considered that BIL219 inherited the “T” allele from the generation of progeny. In addition, we isolated three genotypes of GL6 in a filial generation by self-pollinating the heterozygous W1943, and found that the grain size of W1943 was related to different GL6 genotypes (Supplemental Figure S3). Thus, we hypothesized that this premature stop codon might cause reduction or loss of function of GL6, thus affecting rice grain length.

Characterization and validation of GL6 function in regulating grain length

To further investigate the function of GL6, we developed a near isogenic line
6 (NIL-gl6), containing a 17-kb “T-type” W1943 chromosomal region at the GL6 locus in the GLA4 genetic background (Figure 1F). Compared with GLA4, NIL-gl6 showed...
shorter grains and less 1,000-grain weight, specifically decreases of 10.4% and 14%, respectively. However, the width of the grains was not affected (Figure 1G-H).

In order to verify the function of the candidate gene GL6, we generated a genetic complementation construct in which the GL6 locus from GLA4, containing the entire gene region, the 9277-bp promoter region and the 728-bp downstream sequence, was introduced into NIL-gl6. In comparison with NIL-gl6, the transgenic genetic complementation lines NIL-gl6-CP#1 and #2 showed obvious increases in grain length and weight, even greater than in GLA4, alongside a slight increase in grain width (Figure 2A-B and Supplemental Figure S4). In addition, we carried out gene mutation of GL6 by using the CRISPR/Cas9 genome editing system (Ma et al., 2015) in GLA4 (Figure 2C) and a japonica variety Nipponbare (Supplemental Figure S5A). In GLA4, the resulting GL6 loss-of-function CRISPR-Cas9-derived mutants, GLA4-gl6<sup>CRISPR</sup>#1 and #2, displayed reduced grain length and weight when compared with the wild-type plants (Figure 2D-E and Supplemental Figure S5B-E). Altogether, these results demonstrate that Os06g0666100 is the causative gene for the QTL GL6, and functions in the positive regulation of rice grain length and weight.

**GL6 influences cell proliferation to regulate grain length**

Since cell division and cell expansion are responsible for altering grain size and grain size is restricted by the size of the spikelet hull, we compared the epidermal cells of GLA4 and NIL-gl6 by scanning electron microscope (SEM). Observations of the outer glume surface showed that total cell number along the longitudinal axis in NIL-gl6 was reduced by 17.38% compared with that in GLA4, without any significant difference in single cell length (Figure 3A-D). Similarly, total cell number of the outer epidermal cells were also decreased by 20.72% in GLA4-gl6<sup>CRISPR</sup>#1 and increased by 16.54% in NIL-gl6-CP#1 (Figure3A-D). The CRISPR-Cas9-derived gl6 mutant in Nipponbare, NIP-gl6<sup>CRISPR</sup>#1, also showed reduced cell number (Supplemental Figure S5F-I). Consistent with this, comparative measurement of GL6 expression level among these plants showed that NIL-gl6-CP#1 exhibited elevated GL6 transcripts, whereas decreased GL6 expression levels observed in GLA4-gl6<sup>CRISPR</sup>#1 and NIP-gl6<sup>CRISPR</sup>#1 (Figure 3E and Supplemental Figure S5J). Together, these results
Figure 2. Validation of GL6 controlling grain length. A. Grain morphology of GLA4, NIL-g6, and two independent complemented transgenic lines (NIL-g6-CP1 and NIL-g6-CP2). Scale bar, 5mm. B. Comparisons of grain length, width and 1,000-grain weight among lines shown in (A). C. The sequence of the CRISPR mutant alleles. The wild-type sequence is shown at the top with the target sites underlined in black and the PAM sequence highlighted in blue. Deletions are shown as red dashes. D. Grain phenotype of GLA4, and two independent CRISPR transgenic lines (GLA4-g6(F132P) and GLA4-g6(A133P)). Scale bar, 5mm. E. Comparisons of grain length, width and 1,000-grain weight among lines shown in (D). Values are given as the mean ± SD. Different letters indicate significant differences (P<0.05) determined by Duncan’s multiple range test.
spikelet hull, thus resulting in large grains. As expected, expression levels of 12 cell cycle-related genes, including CAK1A, CYCD4, CYCT1, E2F2 and H1, were significantly down-regulated in NIL-gl6 (Figure 3F), indicating that the reduced cell
number in NIL-gl6 might result from decreased expression of genes that promote cell proliferation. Hence, these results suggest that GL6 positively regulates grain length by altering cell division instead of cell expansion to control cell number of the glume during spikelet development.

**GL6 negatively regulates grain number per panicle**

In addition to the grain size and weight, we also compared other agronomic traits among GLA4, NIL-gl6, and transgenic lines in field trials. We found that larger seed had less grains per panicle (Figure 4A). Compared to that in GLA4 (107.25 ± 15.16), the number of grains per panicle was increased by 32.9% and 33.4% in NIL-gl6 (141.53 ± 19.89) and GLA4-gl6CRISPR#1 (143.10 ± 20.82), respectively. However, the complemented transgenic line NIL-gl6-CP#1 showed dramatically reduced grain number (84.16 ± 14.74) (Figure 4E). We then measured the panicle length, primary branches and secondary branches in these lines (Figure 4B-D). NIL-gl6 and GLA4-gl6CRISPR#1 exhibited more secondary branches whereas NIL-gl6-CP#1 had fewer secondary branches compared to GLA4 (Figure 4D). The increased grain number was mainly attributed to the increased number of secondary branches. Other panicle phenotypes such as panicle length and primary branches showed only a slight effect on grain number. Moreover, we observed that the complemented line showed significantly increased tiller number per plant compared to other lines (Figure 4F).

Taking all these effects into account, the grain yield per plant of NIL-gl6, GLA4-gl6CRISPR#1 and NIL-gl6-CP#1 were all lower than that in GLA4 (Figure 4G). These results reveal that GL6 influences both panicle and spikelet development and that a balance between grain number and grain size may exist that determines grain yield.

**Spatial expression pattern of GL6**

The expression pattern of GL6 was detected by reverse transcription quantitative PCR (RT-qPCR) analysis. GL6 expression was found in all organs and tissues examined, with higher expression levels observed in young panicles (<1 cm) which gradually decreased during panicle development (Figure 5A). Furthermore, we investigated the specific temporal and spatial expression pattern of GL6 by RNA in
situ hybridization (Figure 5B-M). *GL6* transcripts were initially detected when the secondary branch primordia were formed (Figure 5B). During subsequent growth,
GL6 transcripts were abundantly observed both in the lemma and palea primordia (Figure 5C and Figure 5E-F). With the development of floral organs, lemma, palea,
and stamen primordia all exhibited strong expression signal of *GL6* at stage Sp6 (Figure 5G). Subsequently, the expression of *GL6* was restricted to the stamen and carpel primordia, and gradually decreased in the primordia of lemma and palea (Figure 5H-I). Finally, after floral organ differentiation, *GL6* expression signals disappeared during late stage Sp8 (Figure 5J). No signals were detected with the sense probe (Supplemental Figure S6A). By contrast, compared to expression of the positive control *HISTONE H4*, we could barely detect *GL6* expression in NIL-*gl6* at any stage (Figure 5K-M and Supplemental Figure S6B).

**GL6 interacts with OsRPC53 and OsTFC1**

RNAPIII is a multi-subunit complex eukaryotic RNA polymerase that transcribes tRNA genes, 5S rRNA, RNase P and other noncoding RNAs to regulate RNA and protein synthesis for multiple cellular developmental processes (Abascal-Palacios et al., 2018). The maize PLATZ protein FL3 was revealed to interact with RPC53 and TFC1 of the RNAPIII complex to modulate the RNAPIII transcription machinery (Li et al., 2017). Amino acid sequence analysis showed that GL6 shared high sequence similarity with other PLATZ proteins and phylogenetic analysis revealed that GL6 belonged to the PLATZ family (Supplemental Figure S7 and Supplemental Figure S8). Considering the conserved PLATZ domain and potential functional conservation among PLATZ family, we wondered whether *GL6* could affect the function of RNAPIII in rice. We tested the interaction between GL6 and three rice homologs of RNAPIII transcription machinery, namely OsRPC53, OsBRF1, and OsTFC1, respectively. Yeast two-hybrid assays indicated that GL6 interacts with OsRPC53 and OsTFC1, but not OsBRF1 (Figure 6A). Further in vivo interactions were validated using a bimolecular luciferase complementation (BiLC) assay, and we found that the interaction between GL6 and OsRPC53 was much stronger than that between GL6 and OsTFC1 (Figure 6B). Additionally, RNAPIII-dependent transcripts of tRNAs and 5S rRNA were decreased in NIL-*gl6* compared with GLA4. These results suggest that *GL6* might participate in mediating RNAPIII to coordinate ribosome biogenesis (Figure 6C).

**RNA-seq analysis of the *GL6* downstream regulatory network**
To further explore the regulatory mechanism of GL6, we carried out RNA sequencing analysis of young panicles from GLA4 and NIL-gl6 to investigate the downstream gene regulatory networks. Principle component analysis (PCA) of the gene expression data showed close clustering of biological replicates and clear differentiation of separate samples (Supplemental Figure S9A). A total of 2,711 differentially expressed genes (DEGs) were detected ($P < 0.05$), of which 52.6%
(1,426 genes) up-regulated and 47.4% (1,285 genes) down-regulated genes were found in NIL-gl6 relative to GAL4 (Supplemental Table S1). Gene Ontology (GO) analysis of these DEGs showed significant enrichment (FDR < 0.05) of biological process associated with transcription, transport, translation and hormone stimulus (Supplemental Figure S9B). Similarly, molecular functional categories associated with DNA binding, transferase activity, transcription regulator activity, protein binding and signal transducer activity were also highly enriched (Supplemental Figure S9C-D). In addition, we found that OsMADS1 (a known gene controlling grain size) (Liu et al., 2018), eIF1 (protein translation factor) and TFIIS (transcription elongation factor) were down-regulated in NIL-gl6 and further RT-qPCR analyses verified the differential expressions of these genes (Supplemental Figure S9E). These results support the hypothesis that GL6 functions as a transcription factor to regulate downstream gene expression.

Intriguingly, we found that a series of genes including TAW1, OsMADS22, OsMADS47, and OsMADS55, which function in regulatory pathways that control inflorescence architecture by delaying the developmental meristem phase from inflorescence meristem to spikelet meristem transition (Yoshida et al., 2013), were significantly up-regulated in NIL-gl6, and these expression changes were confirmed by RT-qPCR (Supplemental Figure S9F). High expression of these five genes resulted in more secondary branches and spikelets, which was consistent with the increased grain number in NIL-gl6 compared with GLA4. These results imply that the null gl6 mutation might affect grain number by extending the activity of the inflorescence meristem to produce more seeds in NIL-gl6.

Natural variation in the GL6 gene

To investigate the variation of GL6 in diverse rice germplasms, 67 accessions from the rice pan-genome dataset (Zhao et al., 2018) alongside an additional GL6 sequences of 52 wild rice accessions (O. rufipogon) were selected (Huang et al., 2012). The resulting 119 rice accessions consisted of 54 cultivated rice and 65 wild rice accessions (Supplemental Table S2). Although the null gl6 allele was from wild rice W1943, comparisons of nucleotide diversity and neutrality test between cultivars
and wild rice revealed that GL6 was not a locus targeted by human selection during domestication (Supplemental Table S3).

We then compared the GL6 ORF to mine other natural variations in the GL6 gene, and a total of 20 polymorphic sites were observed, including 16 SNPs and 4 insertions or deletions (InDels). Among these variations, 13 SNPs were synonymous mutations, whereas no other variety except W1943 carried the identical SNP (premature stop codon), “T” allele or “T/C” allele. The remaining 4 InDels and 2 missense mutation SNPs were located in the last exon of GL6. A total of 19 haplotypes based on these variations were generated, named H1 to H19 (Figure 7A and Supplemental Table S4).

Most of the japonica varieties showed no differences from the Nipponbare genome, belonging to the H1 haplotype, whereas the indica varieties were distributed among several haplotypes. Interestingly, 60% of the wild rice and most aus accessions, such as Kasalath, shared one identical insertion (+595) in the last GL6 exon, which resulted in an insertion of 3–6 histidines. The geological distribution of these accessions originated mainly from India, indicating that this insertion was fixed in this region (Figure 7B). However, whether the Kasalath allele and other InDels have effects on the determination of grain size needs more investigation in the future.
PLATZ family proteins are a class of plant-specific transcription factors with widespread distribution in dicots, monocots, mosses and algae. All PLATZ family members share a conserved PLATZ domain (for plant AT-rich sequence- and

---

**Figure 7.** Natural variation in GL6 ORF. A, Haplotype network of GL6 ORF among 119 rice accessions. Each circle represented one of the nineteen different haplotypes and circle size was proportional to haplotype frequency. Different colors refer to different rice subpopulation. The red dashed square enclosed rice lines with the same insertion (+595), and the purple one refered to W1943. B, Comparisons of distribution among accessions with the insertion (+595).

**Discussion**

PLATZ family proteins are a class of plant-specific transcription factors with widespread distribution in dicots, monocots, mosses and algae. All PLATZ family members share a conserved PLATZ domain (for plant AT-rich sequence- and
zinc-binding) that is approximately 82 amino acids in length and consists of two noncanonical zinc finger domains (Nagano et al., 2001; Wang et al., 2018). The first PLATZ gene was isolated from pea (*Pisum sativum*), which binds to A/T-rich sequence to negatively regulate the enhancer element of the petE gene (Nagano et al., 2001). At present, only a few PLATZ genes have been isolated. *Arabidopsis thaliana* *AtPLATZ1* and *AtPLATZ2* were confirmed to enhance seed desiccation tolerance (Gonzalez-Morales et al., 2016). The maize (*Zea mays*) PLATZ protein FL3 interacts with RNA polymerase III for biogenesis of tRNA and 5S rRNA to regulate endosperm storage filling (Li et al., 2017). Moreover, *AtORE15*, an Arabidopsis ortholog of GL6, was identified recently to be involved in the regulation of leaf growth and suppression of senescence via cooperation with the GRF/GIF regulatory pathway (Kim et al., 2018).

There are 15 PLATZ genes in rice, however, there has been little functional characterization of rice PLATZ genes. In this study, we firstly identified the rice PLATZ gene GL6 that affects grain length and yield in rice. GL6 positively regulates cell division to increase cell numbers of the spikelet hull, resulting in larger grains.

RNA polymerase III synthesizes various small non-coding RNAs that are essential for general biological activities. Dysregulation of the RNAPIII machinery results in multiple reported instances of reduced cellular and organismal growth (Dauwerse et al., 2010; Borck et al., 2015; Johnson et al., 2016; Soprano et al., 2013; Soprano et al., 2017). However, the regulation of RNAPIII transcriptional activity remains poorly understood in plants. Our data shows that GL6 interacts with two critical RNAPIII subunits, RPC53 and TFC1, to regulate tRNAs and 5S rRNA, indicating that rice GL6 might have a similar function as that of the maize PLATZ protein FL3.

Seed size is crucial for evolutionary fitness in plants. Considering rice yield potential, there is a balance between panicle architecture and grain size. Larger grains are often associated with reduced grain number and smaller grain size indicates that more seeds are produced (Guo et al., 2018). For example, the rare allele *gw2* increases grain size and weight but also reduces grain number per panicle (Song et al., 2007). Similarly, in our study, we observed that NIL-*gl6* and GLA4-*gl6*<sub>Crispr</sub> lines that had
shorter grains showed increased grain number per panicle compared to that in GLA4, consistent with the observed dramatic reduction in grain number per panicle in NIL-CP lines that had large grains. These phenotypes demonstrate the negative correlation between grain size and grain number due to a trade-off between inflorescence and spikelet development.

In conclusion, we identified a new QTL, *GL6*, which functions in the RNAPIII transcription machinery to affect rice grain length and number. Although the detailed mechanism by which GL6 regulates expression of tRNAs and 5S rRNA is still unclear, our data provided here indeed contribute towards understanding this process. The extended regulatory pathway surrounding GL6 remains to be investigated in further detail. Moreover, the exploitation of *GL6* may be a potential approach to manipulate the molecular balance between grain size and number for the development of elite rice varieties with improved grain productivity.

**Materials and Methods**

**Plant materials and trait measurement**

BIL219 (small grain) and GLA4 (large grain) were used as two parents for QTL mapping, and Nipponbare was used for transgenic confirmation. All rice plants were cultivated under field conditions with transplant spacing of 20 x 20 cm in Shanghai and Hainan, China. The measurement of grain yield-related field agronomic traits was conducted with edge lines excluded. Plant height, panicle morphology and grain number were obtained from main culm. Grain yield per plant and tiller number were measured from the whole plant. Fully filled dry grains were used for measuring grain length, width and weight by image analysis method provided with SC-E software (Hangzhou Wanshen Detection Technology Co., Ltd., Hangzhou, China). All trait measurements were repeated at least 3 times.

**Fine mapping of GL6**

The backcross inbred line BIL219 was selected to carry out the backcross with GLA4 to generate the BC₁F₂ population, which containing 288 lines. Then we performed marker assisted selection on each substitution segment to purify the genetic
background. Further fine mapping using 2181 BC1F5 narrowed the *GL6* locus down to a 6.1-kb region between markers GL4350 and GL4411.

**Primers**

All primers used in this study are listed in Supplementary Table 5.

**Transgene constructs and plant transformation**

The entire 14365-bp *GL6* genomic region was digested with *SacI* from GLA4 BAC clone osigba0159g02 and then inserted into the binary vector pCAMBIA1301 to generate the complementation construct. The gene editing constructs of *GL6* via CRISPR/Cas9 were designed as previously described (Ma et al., 2015). All these constructs were introduced into *Agrobacterium tumefaciens* strain EHA105 and subsequently transferred into Nipponbare, GLA4 and NIL-GL6 by *Agrobacterium*-mediated transformation. More than 10 independent transgenic lines were generated, respectively. All analyzed phenotypes were measured in the T2 generation of transgenic plants.

**Scanning electron microscopy**

Mature seeds were first cleaned ultrasonically several times to remove epidermal hairs and dust. The samples were then dried in a critical point drier and coated with gold sputter. For glume cell observation, the outer surfaces of the spikelet glumes were observed by scanning electron microscope (Hitachi). Cell size and cell number were calculated along the longitudinal axis.

**Phylogenetic Analysis**

Protein sequences of PLATZ family members in rice and other organisms were obtained by BLAST from NCBI database. Multiple sequence alignments of protein were performed by ClustalW program. The phylogenetic tree of aligned sequence was constructed by MEGA7 using neighbor-joining tree with 1000 replicates bootstrap.

**Neutrality Test**

Multiple sequences of *GL6* genomic DNA were aligned with ClustalW. Nucleotide diversity and Tajima’s *D* test were calculated and performed using DnaSP v6.12.03 (Rozas et al., 2017).

**Haplotype Network**
Multiple sequences of GL6 ORF were aligned with ClustalW. Haplotype-frequency data were processed with DnaSP v6.12.03 and visualized Median-joining networks were generated by PopART with some modifications (each continuous indel was considered as one site) (Leigh et al., 2015).

**Yeast two-hybrid assays**

For the two-hybrid assay, the full-length coding region of GL6 was amplified and fused in frame with the GAL4 DNA-binding domain via cloning into pGBK7 as the bait plasmid (Clontech). The entire coding regions of GL6, OsRPC53, OsBRF1 and OsTFC1 were introduced into the pGADT7 (Clontech) prey vector. The resulting constructs were then transformed into yeast strain AH109. The co-transformants were diluted (1, 1/10, 1/100, 1/1000) and spotted on control medium (SD/-Trp/-Leu) and selective medium (SD/-Trp/-Leu/-His/-Ade) and incubated at 30°C for 3 days.

**BiLC assay**

The coding region sequences of GL6 was cloned to the N-terminal Luc fusion vector JW771-NLUC, and OsRPC53, OsBRF1 and OsTFC1 were cloned to the C-terminal Luc fusion vector JW772-CLUC, respectively. The BiLC assay procedure was performed as previously described (Gou et al., 2011). Agrobacterium tumefaciens (strain GV3101 (pSoup-p19)) transformants containing the testing split LUC fusion constructs were co-transfected into Nicotiana benthamiana leaves via infiltration, and LUC activity was captured using a cooled CCD-image system (Tanon 5200) with injecting 0.94 mM luciferin (PerkinElmer 122799) after growing for 48 h under 16 h light/d.

**RNA extraction and RT-qPCR analysis**

Total RNA was extracted using the Trizol reagent (Invitrogen) according to the manufacturer’s instructions. Then 500 ng of total RNA was used to synthesize first-strand cDNA using ReverTra Ace® qPCR RT Master Mix with gDNA Remover (toyobo). Real-time PCR was performed on the Applied Biosystems QuantStudio 5 PCR system with diluted cDNA was used as a template using THUNDERBIRD® SYBR® qPCR Master Mix (toyobo). Rice gene UBQ5 (Os01g0328400) was used as the internal control to normalize all data. Each set of experiments was repeated three times...
times.

**RNA in situ hybridization**
Freshly young panicles of GLA4 and NIL-\textit{gl6} were collected and fixed in FAA solution at 4°C overnight, dehydration by series ethanol procedures, and embedded in paraplast. The tissues were sliced into 8-μm sections with a microtome (Leica). And then gene specific region of \textit{GL6} were amplified from FL-cDNA and used to generate digoxigenin-labeled RNA probes (Roche). In situ hybridization was performed as described (Luo et al., 1996).

**RNA-seq analysis**
Total RNA was extracted from GLA4 and NIL-\textit{gl6} young inflorescences (<1 cm) with two biological replicates using the Trizol reagent (Invitrogen) according to the manufacturer’s instructions. After treatment with DNase, the mRNA was purified by NEBNext® Poly(A) mRNA Magnetic Isolation Module (E7490). Libraries were synthesized using the NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (E7770) and sequenced on an Illumina HiSeq2500. A total of 269 million 150-bp paired end reads were generated, yielding 40.34 Gb raw reads. After trimming of Illumina adaptors and low-quality reads, the filtered clean reads were aligned to MSU v7 genome assembly (http://rice.plantbiology.msu.edu) using HISAT2 (Kim et al., 2015). The aligned read files were sorted and indexed by SAMtools (Li et al., 2009) and reads of each sample were then used to calculate raw counts for each gene and transcript using the function SummarizeOverlaps within the GenomicAlignments package. The sample-to-sample distances were presented by principal-components analysis (PCA), which based on the transcript count of each sample. DESeq2 software packages (bioconductor.org/) were used to detect differentially expressed genes with the threshold of genes with Benjamini-Hochberg (BH) adjusted \textit{p}-values < 0.05 and absolute \textit{log}$_2$ Fold change > 0.6. GO enrichment analyses were performed using agriGO V2.0 (Tian et al., 2017).

**Statistical analysis**
Statistical analyses were carried out using Excel 2010 with two-tailed Student’s \textit{t}-test for comparison of two groups and R package “agricolae” with Duncan’s multiple
range tests for multiple mean comparisons.

**Accession numbers**

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers: GL6 cDNA from W1943 (MK959365), OsRPC53 (XM_015781518.2), OsTFC1 (XM_015766861.2), OsBRF1 (XM_015785297.2).

**Supplemental Data**

**Supplemental Figure S1.** Graphical genotype and phenotype of BIL219.

**Supplemental Figure S2.** Sequence alignment of GL6 in GLA4 and NIL-gl6.

**Supplemental Figure S3.** Relationship between different GL6-type and grain length in W1943.

**Supplemental Figure S4.** GL6 positively regulate grain length.

**Supplemental Figure S5.** Suppression of GL6 in Nipponbare results in smaller grains.

**Supplemental Figure S6.** RNA in situ hybridization analysis.

**Supplemental Figure S7.** Amino acid sequence alignments of GL6 and its homologs in various species.

**Supplemental Figure S8.** Phylogenetic analysis of GL6 and other related PLATZ proteins.

**Supplemental Figure S9.** GO Enrichment analysis of RNA-seq differentially expressed genes.

**Supplemental Table S1.** The DEGs and Gene Ontology enrichment in GLA4 and NIL-gl6 (excel file).

**Supplemental Table S2.** The list of 119 rice accessions in the collection.

**Supplemental Table S3.** Nucleotide diversity and Tajima's $D$ test.

**Supplemental Table S4.** Sequence variation and distribution of GL6 ORF haplotypes among 119 rice germplasms.

**Supplemental Table S5.** List of primers used in this study.

**Acknowledgements**

We thank the China National Rice Research Institute for providing the cultivated rice
germplasm and the Chinese wild rice accessions. We thank Jiawei Wang (Institute of
Plant Physiology and Ecology, SIBS, CAS) for providing the JW771 and JW772
plasmids and Yaoguang Liu (South China Agricultural University) for the
CRISPR/Cas9 vector. We thank Jiqin Li, Xiaoyan Gao and Xiaoshu Gao (Institute of
Plant Physiology and Ecology, SIBS, CAS) for technical support. This work was
funded by the Ministry of Science and Technology of China (2016YFD0100902), the
National Natural Science Foundation of China (31630055&31788103), and the
Chinese Academy of Sciences (XDB27010301).

Competing interests
The authors declare no competing financial interests.

Figure Legends

Figure 1. Map-based cloning of GL6. A, Location of GL6 on rice chromosome 6 with
271 chromosome segment substitution lines. B, High-resolution mapping of the GL6
region was performed with 2,181 BC1F5 lines. The number of recombinants between
molecular markers is indicated below the linkage map. C, GL6 was narrowed down to
a 6.1-kb genomic region. D, Grain length is shown for each recombinant line. Values
are given as the mean ± SD. E, Gene structure and allelic variations of GL6 between
GLA4 and BIL219. Filled boxes and black lines represent exons and introns,
respectively. The start and stop codons are indicated above the gene with the coding
region filled with blue color. F, Chromosome maps of NIL-gl6. NIL-gl6 contained the
“T-type” W1943 allele at GL6 in the 17-kb region on chromosome 6, showing as red
bar. G, Mature paddy grain (up) and brown rice (down) morphology of GLA4 and
NIL-gl6. Scale bar, 5 mm. H, Comparison of grain length, width and 1,000-grain
weight between GLA4 and NIL-gl6. Values are given as the mean ± SD. ***
significant difference (P < 0.001, Student’s t-test).

Figure 2. Validation of GL6 controlling grain length. A, Grain morphology of GLA4,
NIL-gl6, and two independent complemented transgenic lines (NIL-gl6-CP#1 and
NIL-gl6-CP#2), Scale bar, 5 mm. B, Comparisons of grain length, width and
1,000-grain weight among lines shown in (A). C, The sequence of the CRISPR mutant alleles. The wild-type sequence is shown at the top with the target sites underlined in black and the PAM sequence highlighted in blue. Deletions are shown as red dashes. D, Grain phenotype of GLA4, and two independent CRISPR transgenic lines (GLA4-gl6CRISPR#1 and GLA4-gl6CRISPR#2). Scale bar, 5 mm. E, Comparisons of grain length, width and 1,000-grain weight among lines shown in (D). Values are given as the mean ± SD. Different letters indicate significant differences ($P < 0.05$) determined by Duncan’s multiple range test.

**Figure 3.** The effect of *GL6* on cell number contributes to grain length. A and B, Scanning electron micrographs of the whole grain (A) and outer glume surfaces (B) in GLA4, NIL-gl6, GLA4-gl6CRISPR#1 and NIL-gl6-CP#1. Scale bars, 1 mm (A) and 100 μm (B). C and D, Comparisons of single cell length (C) and total cell number (D) among lines shown in (A). E, Comparisons of relative expression levels of *GL6* among lines shown in (A). Values are given as the mean ± SD. Different letters indicate significant differences ($P < 0.05$) determined by Duncan’s multiple range test. F, Relative expression levels of 26 cell cycle related genes in young panicles (0~1 cm) of GLA4 and NIL-gl6. *OsUBQ5* was used as the control and the values of expression levels in GLA4 were set to one. Values are given as the mean ± SD. Student’s *t*-test significant difference, * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$; ns, not significant.

**Figure 4.** *GL6* affects panicle morphology. A, Panicles of GLA4, NIL-gl6, GLA4-gl6CRISPR#1 and NIL-gl6-CP#1. Scale bar, 5 cm. B to G, Comparisons of panicle length (B), primary branches (C), secondary branches (D), spikelets per panicle (E), tiller number per plant (F) and grain yield per plant (G), among lines shown in (A). Values are given as the mean ± SD. Different letters indicate significant differences ($P < 0.05$) determined by Duncan’s multiple range test.

**Figure 5.** The spatial expression pattern analysis of *GL6* in rice. A, Relative expression levels of *GL6* mRNA in vegetative tissues and developmental panicles.
The abundance of *GL6* transcripts was normalized to that of *OsUBQ5*. Values are given as the mean ± SD. B to J, In situ analysis of *GL6* in GLA4 during different inflorescence stages and flower development processes. *GL6* transcripts at stages In5 (B), In6 (C), In7 (D), Sp4 (E), Sp5 (F), Sp6 (G), Sp7 (H), Sp8e (I), Sp8e (J). K to M, in situ analysis of *GL6* in NIL-*gl6* at stages In7 (K), Sp6 (L) and Sp7 (M). ca, carpel; fm, floral meristem; le, lemma; lo, lodicule; ov, ovule; pa, palea; pb, primary branch; sb, secondary branch; sp, spikelet; st, stamen; Sp8e, early stage Sp8; Sp8l, late stage Sp8; bars = 50 μm.

**Figure 6.** *GL6* interacts with RPC53 and TFC1 to regulate tRNA transcripts. A, Yeast two-hybrid assay, showing the interactions between *GL6* and OsRPC53, OsTFC1 and itself. AD, activation domain; BD, binding domain; tenfold serial dilutions shows the gradients. B, BiLC assay showing that *GL6* interacts with OsRPC53 and OsTFC1 in vivo. C, Relative expression levels of tRNAs and 5S rRNA in young panicles (0~1 cm) of GLA4 and NIL-*gl6*. *OsUBQ5* was used as the control and the values of expression levels in GLA4 were set to one. Values are given as the mean ± SD. Student’s *t*-test significant difference,* P < 0.05, ** P < 0.01 and *** P < 0.001; ns, not significant.

**Figure 7.** Natural variation in *GL6* ORF. A, Haplotype network of *GL6* ORF among 119 rice accessions. Each circle represented one of the nineteen different haplotypes and circle size was proportional to haplotype frequency. Different colors refer to different rice subpopulation. The red dashed square enclosed rice lines with the same insertion (+595), and the purple one refered to W1943. B, Comparisions of distribution among accessions with the insertion (+595).


Guo T, Chen K, Dong NQ, Shi CL, Ye WW, Gao JP, Lin HX (2018) GRAIN SIZE AND NUMBER1 Negatively Regulates the OsMKKK10-OsMKK4-OsMPK6 Cascade to Coordinate the Trade-off between Grain Number per Panicle and Grain Size in Rice. Plant Cell 30: 871-888


Copyright © 2019 American Society of Plant Biologists. All rights reserved.


