UC Davis
UC Davis Previously Published Works

Title

The PGS1 basic helix-loop-helix protein regulates Fl3 to impact seed growth and grain yield in cereals

Permalink

https://escholarship.org/uc/item/4sn4072c

Journal

Plant Biotechnology Journal, 20(7)

ISSN

1467-7644

Authors

Guo, Xiaojiang
Fu, Yuxin
Lee, Yuh-Ru Julie
et al.

Publication Date

2022-07-01

DOI

10.1111/pbi.13809

Peer reviewed

The PGS1 basic helix-loop-helix protein regulates *Fl3* to impact seed growth and grain yield in cereals

Xiaojiang Guo[1,2,3,†] (iD), Yuxin Fu[1,4,†], Yuh-Ru Julie Lee[3], Mawsheng Chern[5], Maolian Li[1], Mengping Cheng[1], Huixue Dong[1], Zhongwei Yuan[1], Lixuan Gui[1], Junjie Yin[2], Hai Qing[2], Chengbi Zhang[1], Zhien Pu[1], Yujiao Liu[1], Weitao Li[2], Wei Li[1], Pengfei Qi[1] (iD), Guoyue Chen[1], Qiantao Jiang[1] (iD), Jian Ma[1], Xuewei Chen[2], Yuming Wei[1,6], Youliang Zheng[1,6], Yongrui Wu[4] (iD) and Jirui Wang[1,2,6,*] (iD)

[1]*Triticeae Research Institute, Sichuan Agricultural University, Chengdu, China*

[2]*State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Sichuan Agricultural University, Chengdu, China*

[3]*Department of Plant Biology, University of California, Davis, CA, USA*

[4]*National Key Laboratory of Plant Molecular Genetics, 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*

[5]*Department of Plant Pathology, University of California, Davis, CA, USA*

[6]*Ministry of Education Key Laboratory for Crop Genetic Resources and Improvement in Southwest China, Sichuan Agricultural University, Chengdu, China*

Summary

Plant transcription factors (TFs), such as basic helix-loop-helix (bHLH) and AT-rich zinc-binding proteins (PLATZ), play critical roles in regulating the expression of developmental genes in cereals. We identified the bHLH protein TaPGS1 (*T. aestivum* Positive Regulator of Grain Size 1) specifically expressed in the seeds at 5–20 days post-anthesis in wheat. *TaPGS1* was ectopically overexpressed (OE) in wheat and rice, leading to increased grain weight (up to 13.81% in wheat and 18.55% in rice lines) and grain size. Carbohydrate and total protein levels also increased. Scanning electron microscopy results indicated that the starch granules in the endosperm of *TaPGS1* OE wheat and rice lines were smaller and tightly embedded in a proteinaceous matrix. Furthermore, TaPGS1 was bound directly to the E-box motif at the promoter of the PLATZ TF genes *TaFl3* and *OsFl3* and positively regulated their expression in wheat and rice. In rice, the *OsFl3* CRISPR/Cas9 knockout lines showed reduced average thousand-grain weight, grain width, and grain length in rice. Our results reveal that *TaPGS1* functions as a valuable trait-associated gene for improving cereal grain yield.

Introduction

Higher grain production is essential for breeding cereal crops because they are the primary energy source for humans and livestock (USDA, 2018). Grain weight is a significant component of grain yield in cereals and is determined by grain size and seed-filled rate (Su *et al.*, 2011; Xing and Zhang, 2010). Molecular pathways that regulate seed development are often associated with grain weight and filling (Gaur *et al.*, 2011; Wang *et al.*, 2015; Xing and Zhang, 2010). Therefore, it is important to understand the molecular mechanisms of seed development to increase the cereal yield.

The basic helix-loop-helix (bHLH) family transcription factors (TFs) have been implicated in the seed development of cereals. For example, POSITIVE REGULATOR OF GRAIN LENGTH 1 (PGL1) and PGL2 form heterodimers that suppress the function of *ANTAGONIST OF PGL1* (*APG*), affecting cell length and positively regulating rice grain length (Heang and Sassa, 2012a, 2012b). PHYTOCHROME INTERACTING FACTOR-LIKE 15 (OsPIL15) regulates grain size by directly targeting the purine permease gene *OsPUP7* of *Oryza sativa* (Ji *et al.*, 2019). Maize opaque11 (o11) affects endosperm development and nutrient metabolism by increasing starch and protein accumulation (Feng *et al.*, 2018).

The maize PHYTOCHROME INTERACTING FACTOR 1 (ZmPIF1) increases grain yield in transgenic rice by increasing tiller and panicle numbers (Gao *et al.*, 2018).

We identified 225 wheat bHLH TFs and separated them into seven groups according to their expression patterns in the endosperm, seedlings, heading-stage spikes, flag leaves, shoots, and roots in our previous studies (Chen *et al.*, 2015; Guo and Wang, 2017). One of the genes, *T. aestivum POSITIVE REGULATOR of GRAIN SIZE 1 (TaPGS1)*, is homologous to *OsRc* (*brown pericarp and seed coat*) in rice (Gu *et al.*, 2011) and *AtTT8 (Transparent Testa 8)* in *Arabidopsis (Nathalie Nesi, 2000)*. *OsRc* is a pleiotropic gene that participates in flavonoid synthesis to control pericarp colour and has also been reported to increase grain weight in rice (Gu *et al.*, 2011). However, it is not clear how *OsRc* influences grain weight. *AtTT8* also regulates flavonoid biosynthesis in *Arabidopsis* (Nathalie Nesi *et al.*, 2000). This study investigated the involvement of the *TaPGS1* gene in grain weight and the molecular mechanism involved.

The results showed that *TaPGS1* plays an essential role in seed development. Overexpression (OE) of *TaPGS1* in wheat and rice resulted in altered expression levels of seed-development-related genes and increased grain weight and grain size in field tests. In addition, the OE lines had increased starch, soluble sugar, and

protein in seed. TaPGS1 directly binds to the promoter of PLATZ TFs *TaFl3* (*TraesCS3A02G497900*) and *OsFl3* (*Os01g0517800*) and regulates their expression at early seed developmental stages. Furthermore, CRISPR/Cas9 knockout lines of *OsFl3* showed higher flat grains rates than the WT lines.

Results

TaPGS1 is specifically expressed in the seed tissues

TaPGS1 encodes a protein that contains 19 highly conserved amino acids in the 60-amino acid bHLH domain (Figure 1a). Based on the phylogenetic analysis of the protein sequence, TaPGS1 and its homologs in plants were classified into two clades: dicot and monocot (Figure 1b; Table S1). It was mainly expressed at the early stage of the aleurone layer (part of the endosperm), pericarp, and 15 days post-anthesis (DPA) in wheat caryopsis (Figure S1; Table S2). The expression of *TaPGS1* in developing seeds increased from 5 to 15 DPA, peaked at 20 DPA, and then declined to lower levels at 25 DPA and 30 DPA. *TaPGS1* was not expressed in leaves or roots (Figure 1c). RNA in situ hybridisation assays indicated strong *TaPGS1* transcript signals in the aleurone layer, starchy endosperm, pericarp, and embryo (Figure 1d). These results suggested that *TaPGS1* is specifically expressed in seed tissues and may participate in seed development in wheat.

TaPGS1 increased grain weight and size in wheat and rice

To explore the function of *TaPGS1* in seed development, we transformed the wheat cultivar Fielder (WT) with *TaPGS1* cDNA driven by the *Maize* ubiquitin1 promoter (Cornejo et al., 1993). Ten independent lines were obtained, and two stable OE lines (166-6-OE and 166-39-OE) were obtained (Figure 2a). The expression patterns of *TaPGS1* were examined at different stages of seed development. Compared with the WT, the OE lines showed distinctly higher transcription levels of *TaPGS1* (Figure 2b). TaPGS1 protein levels were significantly increased in 166-6-OE and 166-39-OE compared to WT at 10-15 DPA (Figure 2c).

We measured the grain traits of the *TaPGS1* OE and WT lines in two field trials and found that the thousand-grain weights of 166-6-OE and 166-39-OE were significantly increased by 5.18%–10.04% in trial 1 (Wenjiang, China, 2018) and 11.58%–17.24% in trial 2 (Wenjiang, China, 2019; Figure 2d; Table S3). The average grain length and width (grain size) of the 166-6-OE and 166-39-OE lines also significantly increased by 2.19%–21.07% in trials in Wenjiang (2018) and Wenjiang (2019), except for line 166-6-OE in Wenjiang (2018) (Figure 2e,f; Table S3).

We also generated *TaPGS1* homozygous OE lines using the Kitaake rice cultivar. Two stable OE lines, OE-3 and OE-11, were selected for further analysis (Figure 3a). The *TaPGS1* expression level in OE lines was significantly higher than that in WT (Figure S2c). The TaPGS1 RNA and protein levels at 10 and 15 DPA in OE-3 and OE-11 were higher than those in WT (Figure 3b, c).

We conducted multi-environment trials over three consecutive years to compare rice OE lines with WT regarding grain traits, including average grain weight, length, and width. OE-3 and OE-11 lines displayed significant increases in grain size by 1.29%–18.84% in trials in Wenjiang (2017), Lingshui (2018), and Wenjiang (2019), except for line OE-3 in Lingshui (2018) (Figure 3d,e; Table S4). We

obtained the average grain weight of each line by weighing 1000 seeds from each line. Consistent with the increase in grain size, the average thousand-grain weights of OE-3 and OE-11 were higher than those of the WT by 10.09% (Wenjiang 2017), 9.555% (Lingshui 2018), and 7.22% (Wenjiang 2019), respectively (Figure 3f; Table S4). Compared with the WT, the plant height, ear length, number of grains per panicle, number of filled grains per panicle, and effective number per rice plant of the OE lines did not show no significant (Figure S2a,b; Table S5). However, the grain weights per plant of the OE-3 and OE-11 lines were significantly increased by 4.55% and 8.75%, respectively, compared with WT (Figure 3g; Table S5). These results indicated that higher levels of *TaPGS1* increased the thousand-grain weight and seed size in rice and wheat.

Carbohydrate and total protein contents increased in *TaPGS1* OE lines

To study the causes of the increase in grain weight and size, we examined the seed structures of the OE and WT lines (wheat and rice) using scanning electron microscopy. The results indicated that the endosperm of the WT contained loosely packed starch granules with relatively large spaces in between. In contrast, the endosperms of 166-6-OE and 166-39-OE contained smaller starch granules tightly embedded in a proteinaceous matrix with little space (Figure 4a). The starch granules in rice OE-3 and OE-11 lines were compactly packed and rounder, whereas the WT endosperm contained more gaps between the starch granules (Figure 4b).

We examined the carbohydrate and protein storage contents. The starch content and soluble sugars based on dry seed flour were significantly increased by 3.85%–32.00% in lines 166-6-OE (wheat) and OE-3 (rice) (Figure 4c–f; Tables S6 and S7). Although storage carbohydrates of the other two lines did not reach significant levels compared to the WT, the overall contented still increased (Figure 4c, d, e, f; Tables S6 and S7). Compared with the WT, the total protein content of all OE lines, except 166-6-OE, increased by 1.42%–8.95%. (Figure 4g, and h; Tables S6 and S7). Based on these results, we conclude that overexpression of the *TaPGS1* gene led to an increase in seed carbohydrate and total protein, increasing the thousand-grain weight and seed size.

TaPGS1 regulated the expression of *OsFl3* and *TaFl3*

We carried out RNA-seq to identify significantly differentially expressed genes (DEGs) using samples of pooled seeds at 10, 15, and 20 DAP; root; and leaved from rice OE-11 line and WT plants. Three independent biological replicates were used. In OE-11, the expression of 7686 genes was significantly altered (*P* < 0.05; false discovery rate <0.2; Table S8).

Subsequently, we performed *in vitro* chromatin immunoprecipitation sequencing (ChIP-seq) analysis using GST antibodies against GST-tagged TaPGS1 proteins to determine which rice DEGs are directly regulated by TaPGS1. A total of 3881 peaks distributed in 2459 rice genes were identified (Table S9). Among the detected peaks, 20.5% (793 peaks) were located upstream of the related genes, core promoter regions (−2 kb of the transcription start site, TSS), 68.4% (2660 peaks) in regions from the TSS to the transcription end site (TES), and 11.1% (431 peaks) in the downstream region (+2k bp of the TES, Figure 5a, b). TaPGS1 binding sites were concentrated in the upstream and downstream regions (Figure 5b). Then we identified the cis-regulatory
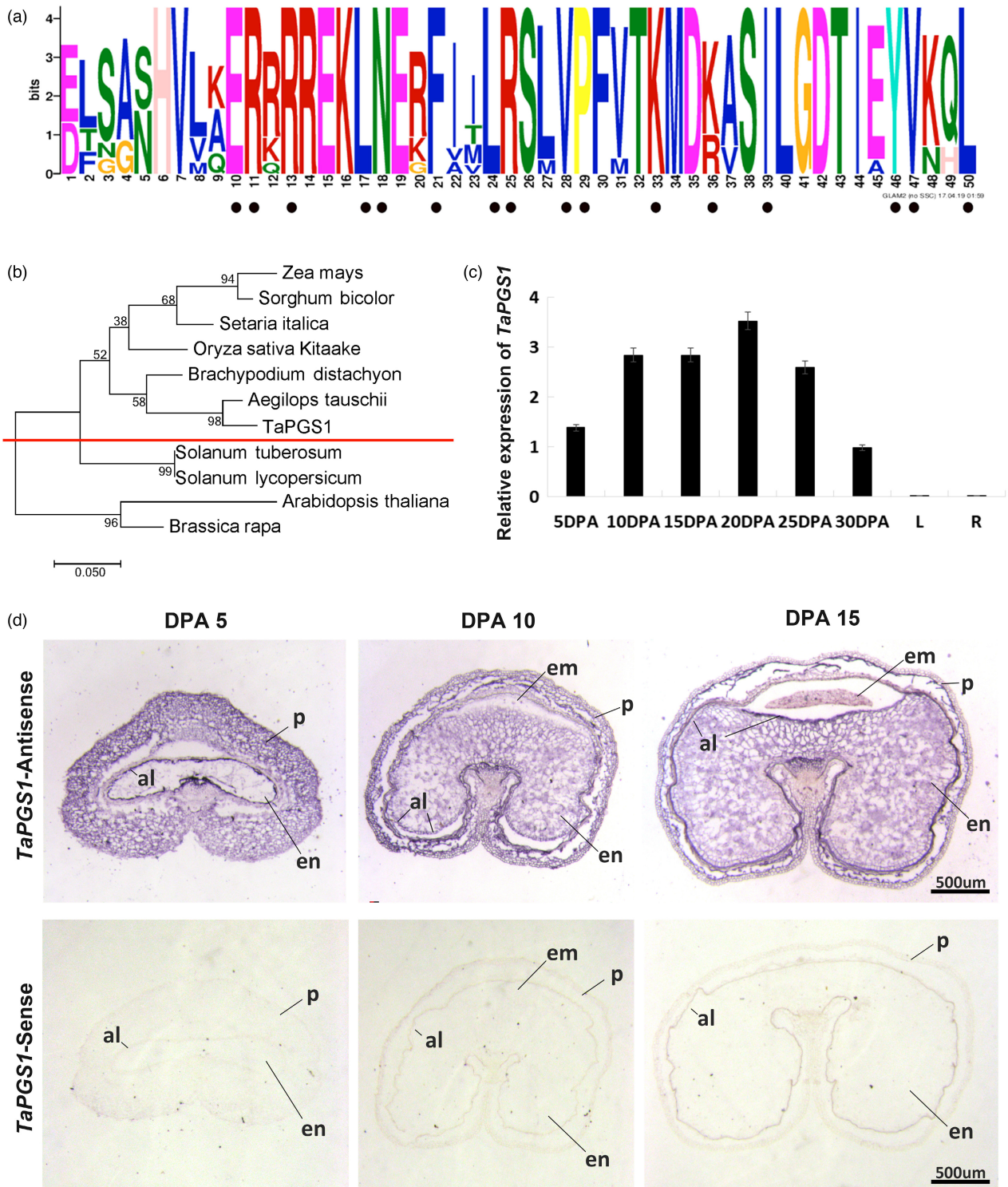
Figure 1 Evolution, expression, and characterisation of TaPGS1. (a) Multiple sequence alignment at the conserved bHLH domain of TaPGS1 and ten homologous bHLH proteins. (b) Evolutionary relationships of TaPGS1 homologs proteins (*Aegilops tauschii* L., *Zea mays* L., *Arabidopsis thaliana* L., *Brachypodium distachyon* L., *Solanum tuberosum* L., *Solanum lycopersicum* L., *Setaria italica* L., *Sorghum bicolor* L., *Brassica rapa* L., *Oryza sativa* L.). The evolutionary history was inferred using the neighbour-joining method. (c) Relative RNA levels of *TaPGS1* at 5, 10, 15, 20, 25, and 30 days post-anthesis (DPA), L (leaf), and R (root) in wheat cultivar CN 16. (d) Expression pattern of TaPGS1 detected by RNA in situ hybridisation. Seed tissue was magnified 10x by OLYMPUS BX53 microscope. Endosperm (en), Embryo (em), Aleurone cell (al), Pericarp (p), Scale bar = 500 μm.

sequences that mediate TaPGS1 binding and lead to the regulation of TaPGS1 target genes. An E-box (5′-CANNTG-3′) in the binding sites was the top-scoring motif (55.98%, E-value = 6.7E-5; Figure 5c, d), indicating that TaPGS1 might transactivate the promoters of its target genes by binding to the E-box motif.
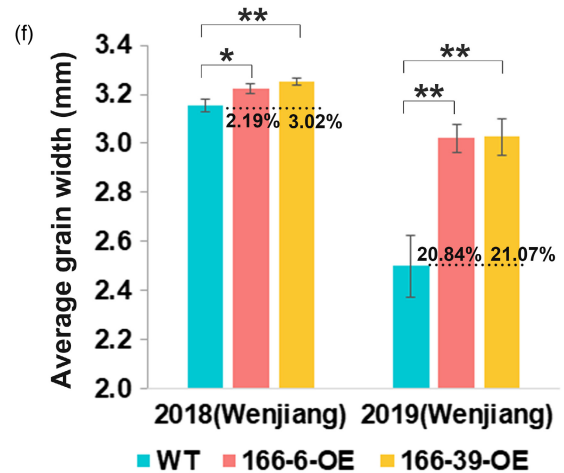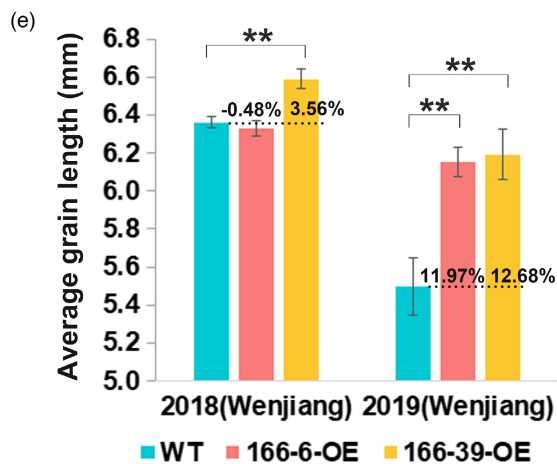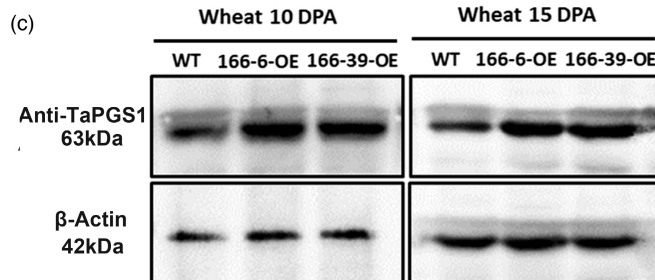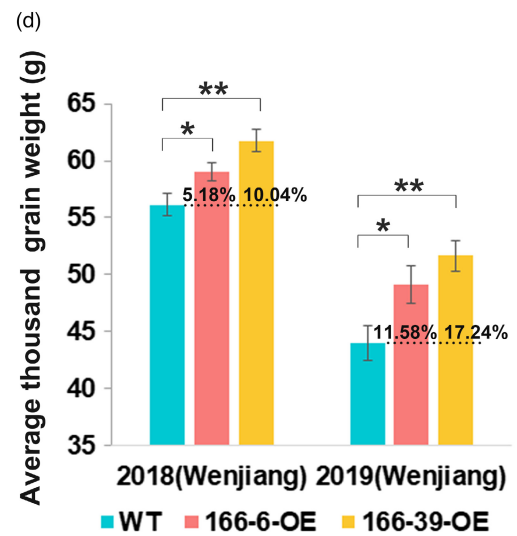
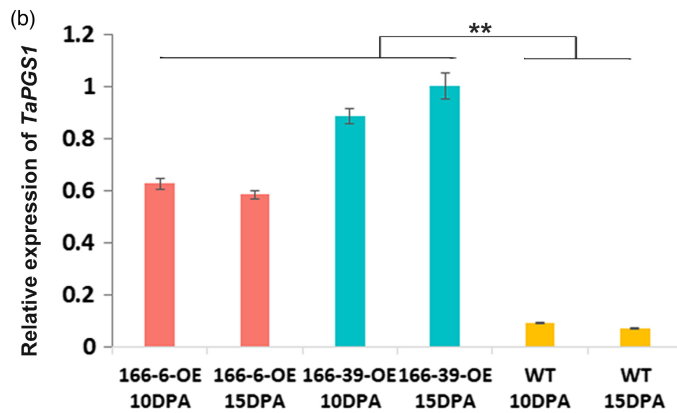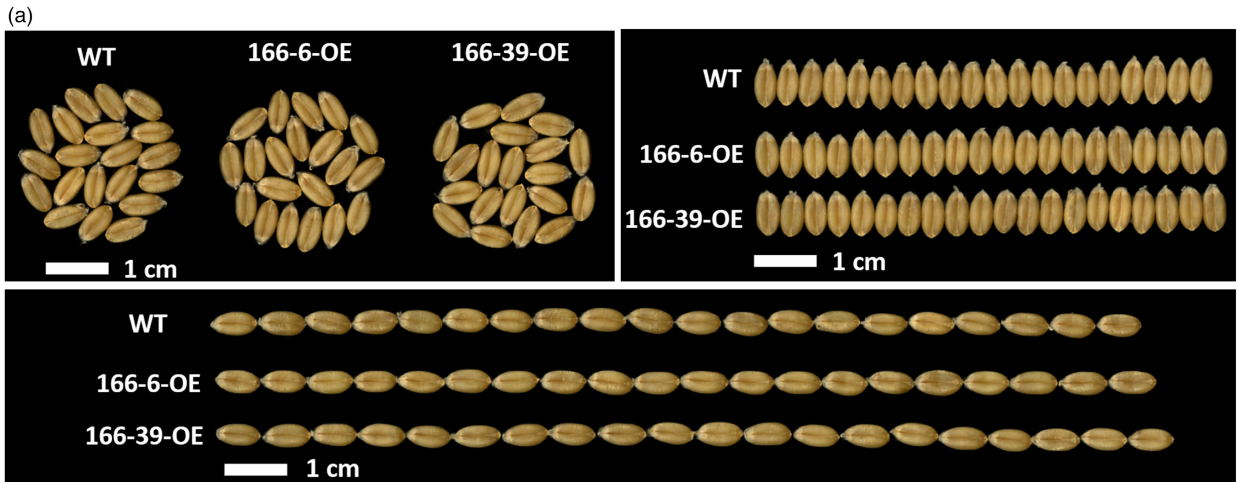(a)

(b)

(c)

(d)

(e)

(f)

Figure 2 Phenotypes and expression levels of TaPGS1 overexpression lines (166-6-OE, 166-6-39-OE) in wheat. (a) Photographs from the WT and *TaPGS1* overexpression lines 166-OE-6 and 166-39-OE. (b) Relative RNA expression levels of *TaPGS1* in seeds of WT and overexpression lines (166-6-OE, 166-6-39-OE) 10–15 days post-anthesis (DPA) were measured by RT-qPCR. Bars represent means ± SEM, **P < 0.01. (c) TaPGS1 protein levels in WT and overexpression lines (166-6-OE, 166-6-39-OE). Equal loading was verified by probing with a β-actin antibody. The average thousand-grain weight (d), average grain length (e), and grain width (f) in WT and OE lines (166-6-OE and 166-39-OE). Bars represent means ± SEM (n = 10). *P < 0.05, **P < 0.01.

To obtain high-confidence primary targets of TaPGS1, we searched for overlaps among the 793 genes bound by TaPGS1 in the promoter regions and 7686 DEGs specific to OE-11 in the RNA-seq data. In total, we identified 164 high-confidence potential targets of TaPGS1 (Figure 5e; Table S10), including 13 seed (10, 15, 20 DPA) specific DEGs (Figure 5f; Table S11). Two genes (*Os01g0567200* and *Os01g0517800*) showed expression patterns similar to *TaPGS1* (Figure 1c; Figure 5g; Table S12). Only *Os01g0517800* was upregulated at all stages of seed development in the OE lines (Table S12) and is homologous to the gene product of *ZmFl3,* which encodes a PLATZ TF specifically expressed in the endosperm, regulating seed development (Li et al., 2017). These results suggested that *Os01g0517800* (*OsFl3*) might be a target of TaPGS1 and might be involved in seed development.

The RNA-seq and reverse transcription-quantitative polymerase chain reaction (RT-qPCR) analyses showed that *OsFl3* was specifically and highly expressed in the early stages of seed development (Figure 5g, h). The expression levels of *OsFl3* were 12-fold higher at 5 DPA and 1.34-fold higher at 10 DPA in the OE-11 line than in the WT (Figure 5h). RNA in situ hybridisation assays showed that *OsFl3* had the same expression pattern as *TaPGS1* in the seed (Figure 1d; Figure S3a). The peak signal was enriched at the *OsFl3* (*Os01g0517800*) promoter (Figure S4a), and *OsFl3* promoter fragments were pulled down 2–5-fold more frequently than the control without antibodies (Figure 6a). In addition, *OsFl3* contained an E-box in the putative promoter region (−2 kb to +100 of the TSS; Figure 6b; Figure S4b), suggesting that TaPGS1 potentially regulates *OsFl3* expression by binding to the E-box of the promoter.

To validate the binding of TaPGS1 to the *OsFl3* promoter, we performed a yeast one-hybrid assay, in which TaPGS1 was fused to GAL4 AD and the *OsFl3* promoter fused to the *HIS2* reporter gene. The results showed that the TaPGS1-GAL4 AD protein was bound to the *OsFl3* promoter and enhanced reporter expression (Figure 6c), indicating that TaPGS1 directly regulates *OsFl3* expression. Moreover, we carried out electrophoretic mobility shift assays (EMSAs) to confirm the binding and determine where the TaPGS1 promoter binds. A TaPGS1-His-tag fusion protein was expressed and purified in *E. coli*. For the probe, 5′ FAM-labelled (on a single strand) complementary oligonucleotides (abbreviated as probe-OsFl3) were synthesised according to the *OsFl3* promoter binding region (Figure S4a, b; Table S13) and annealed. TaPGS1-His bound to the probe-OsFl3. In contrast, histidine repeats alone did not bind to the OsFl3 probe (Figure 6e). The binding specificity was further confirmed by competitors, where the WT competitor significantly reduced binding to the probe, whereas the mutated competitor had minor effects (Figure 6e). These results showed that the TaPGS1 protein binds to OsFl3 probe, indicating that TaPGS1 directly binds to the *OsFl3* promoter to regulate its expression.

For the two wheat genes (*TaPLATZ26* and *TaPLATZ27*) homologous to *OsFl3* (Fu et al., 2020), the yeast one-hybrid assay showed that only TaFl3 (TaPLATZ27) was activated by TaPGS1

(Figure S5, Figure 6d). We performed a tobacco transient trans-activation assay, in which the *TaFl3* promoter was fused with a luciferase (LUC) reporter (TaFl3$_{Pro}$:LUC) and the 35S promoter to effector *TaPGS1* (35S$_{pro}$:TaPGS1) or a blank control (35S$_{pro}$). The effector and reporter constructs were transiently co-expressed in the tobacco leaves. The results showed that TaPGS1 effectively activated *TaFl3* (Figure 6f, g). In addition, *TaFl3* expression levels were higher in 166-6-OE plants than in WT plants (Figure S3c). *TaFl3* showed the same expression pattern as *TaPGS1* in the seeds with the RNA in situ hybridisation assay (Figure S3b). These results suggested that TaPGS1 in wheat regulates *TaFl3* expression.

The *Fl3* mutant showed defective endosperm and a decline in grain yield

To confirm *OsFl3* function, we generated *OsFl3* knockout lines in the Kitaake rice cultivar using the CRISPR/Cas9 technology. We selected and confirmed two frameshift knockout lines (KO-831-2 and KO-836-4), one with a single-base deletion and the other with a two-base deletion (Figure S6a, b).

The *OsFl3* knockout lines displayed a reduced average thousand-grain weight, grain width, and grain length than the WT (Kitaake) (Figure 7a–d; Table S14). In addition, the average grain-filling rates of KO-831-2 and KO-836-4 (65.85% and 76.32%) were significantly lower than those of the WT (94.74%) (Figure 7e). KO-831-2 and KO-836-4 also had lower total grain numbers per panicle and filled grains per panicle than the WT. In contrast, KO-831-2 and KO-836-4 had significantly higher numbers of flat and empty grains per panicle (P < 0.001; Figure 7f; Table S15). Therefore, *OsFl3* knockout reduced grain yield per panicle. Starch granules in the endosperm were relatively tightly packed in the WT but loosely packed with more gaps in the knockout lines (Figure 7g). These results suggested that *OsFl3* is involved in endosperm development, influencing rice grain size and grain filling rate.

Discussion

Domestication involves artificial selection through human action, resulting in genetic modification and phenotypic evolution of cereals. Orthologous loci among species usually control similar traits. Seed size contributes to a higher yield (Asano et al., 2011). Wheat, rice, and maize have the same ancestor, and multiple orthologous loci were identified. For example, *GW2* regulates seed development to influence grain yield in rice, and the wheat and maize orthologous genes, *TaGW2* and *ZmGW2-CHR4/5* control similar traits (Li et al., 2010a; Song et al., 2007; Su et al., 2011). The *GS3* orthologs in maize and rice are both involved in seed development (Fan et al., 2006; Li et al., 2010b). In addition, rice *Gn1a/OsCKX2* and wheat *TaCKX6-D1* are associated with grain weight by controlling cytokinin levels (Li et al., 2013; Zhang et al., 2012).

In this study, both grain size and grain weight were increased in the *TaPGS1* OE wheat and rice lines (Figure 2d–f; Table S3;
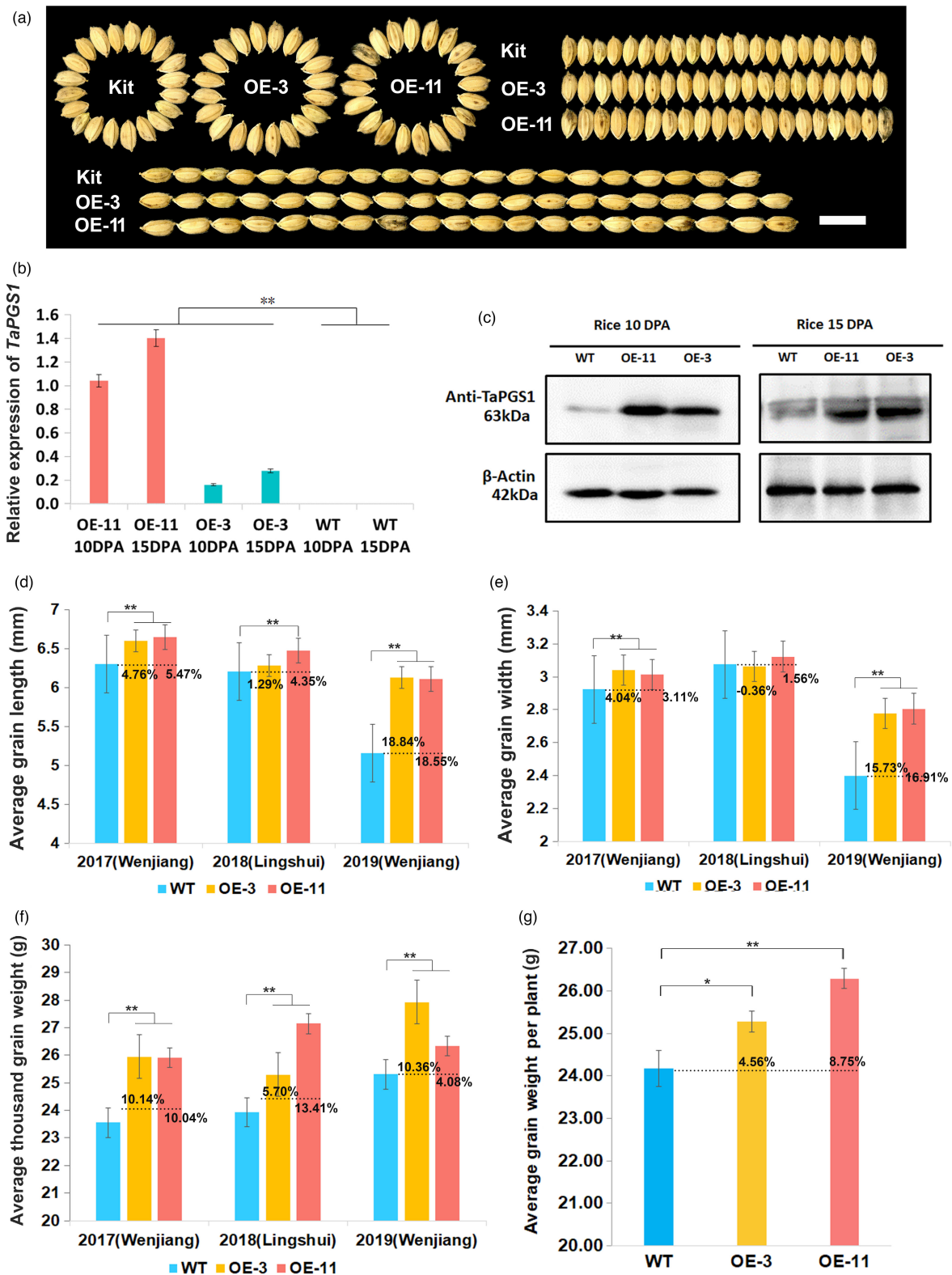
Figure 3 Phenotypes and expression levels of TaPGS1 overexpression lines (OE-3, OE-11) in rice. (a) Photographs of grains from the WT and *TaPGS1* overexpression lines (OE-3 and OE-11). (b) RT-qPCR was used to analyse of *TaPGS1* transcript levels in seeds of WT, OE-3, and OE11 10-15 DPA. Bars represent means ± SEM, **P < 0.01. (c) Protein levels of *TaPGS1* in WT, OE-11, and OE-3. Equal loading was verified by using a β-actin antibody. Average grain length (d), grain width (e), thousand-grain weight (f), and grain weight per plant (g) in WT rice and *TaPGS1* OE-3 and OE-1. All data are presented as means ± SEM (n = 5). *P < 0.05, **P < 0.01.
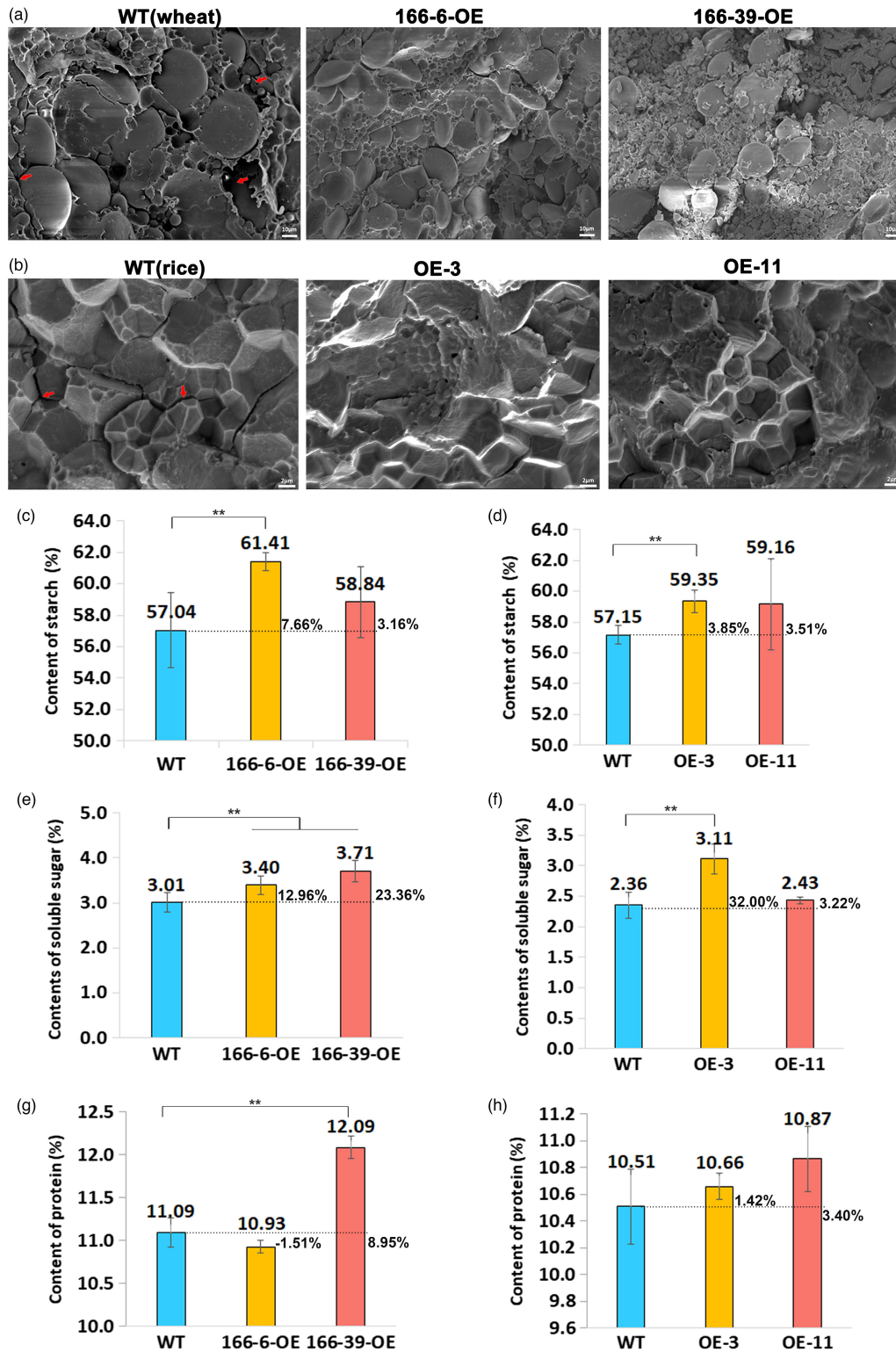
Figure 4 Scanning electron microscopy of starch granules and dry matter contents. (a) Scanning electron microscopy of starch granules of mature endosperm in wheat WT and overexpression lines (166-6-OE, 166-6-39-OE). Bar = 10 μm. (b) Scanning electron microscopy of starch granules of mature endosperm in WT rice and overexpression lines (OE-2 and OE-11). Bar = 10 μm. The red arrows indicate the gaps between structures. (c,e,g) Starch, soluble sugars, and protein content in WT and OE lines of wheat. (d,f,h) Starch, soluble sugars, and protein content in WT and OE lines of rice. All data are presented as means ± SEM (n = 5). *P < 0.05, **P < 0.01.

Figure 3d–f; Table S4). The *TaPGS1* homologous gene (*OsRc* or *OsPGS1*) increased grain weight in rice (Gu *et al.*, 2011). Furthermore, carbohydrate and total protein contents were increased in

the *TaPGS1* OE lines of wheat and rice (Figure 4c–h; Tables S6 and S7), and electron microscopy revealed dense endosperms in the OE lines (Figure 4a,b). The compact endosperm structure in OE lines

Figure 5 DEGs revealed by RNA-Seq analysis and genome-wide binding profiles from ChIP-Seq analysis. (a) Distribution of TaPGS1 binding regions in the rice genome. Up2K (Orange), −2 kb to the transcription start site (TSS); Down2K (Brown), −2 kb to the transcription end site (TES); Exon (Light blue), Intron (Green), Intergenic (Blue), regions from TSS to TES. (b) Distribution of TaPGS1 binding sites per 100 bp bin corresponding to the −2000 bp of TSS (Up2K) to +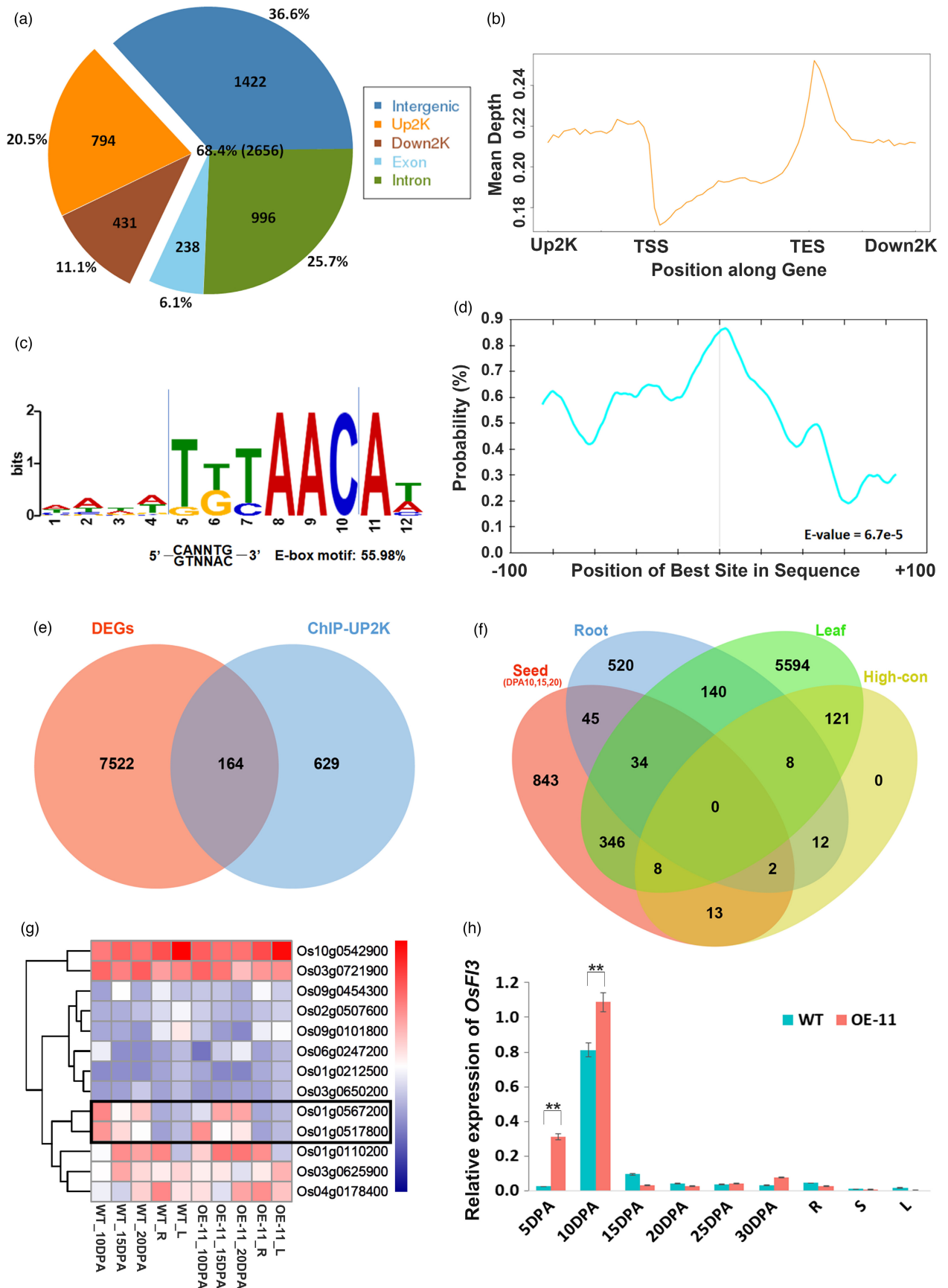2000-bp region of TES(Down2K). (c) TaPGS1 binding motif identified by MEME-ChIP in the 200 bp flanking sequences around the genic peak summits and the density plot of this motif around the summits of the peaks. The 5′-CANNTG-3′ (E-box) motif was identified as the most prominent TaPGS1 binding motif (55.98%). (d) Probability of the best binding site. The *x*-axis indicates the binding motif position in the sequence, and the *y*-axis indicates the possibility of the identified motif. The *E*-value is the enrichment *P*-value times the number of candidate motifs tested. The enrichment *P*-value was calculated using Fisher's exact test. e indicates ten raised to a power in scientific notation. (e) Venn diagram of genes whose promoters are bound by PGS1 and genes whose expression is regulated by PGS1. 164 genes out of 793 genes bound by PGS1 in the promoter region overlap with the 7686 DEGs from RNA-seq data comparing OE-11 with wild type. (f) Venn diagram of numbers of high-confidence target genes from the intersection of (e) and DEGs from RNA-seq data comparing OE-11 with wild type using different tissues. (g) Heatmap of expression at 10, 15, 20 days post-anthesis (DPA), root(R), and leaf(L) for 13 candidate genes from RNA-seq data (f) in WT rice and TaPGS1 overexpression line (OE-11). (h) Relative RNA levels of *Os01g0517800* (*OsFl3*) in seeds at 5, 10, 15, 20, 25, and 30 DPA, root(R), stem(S), and leaf(L) were measured using RT-qPCR in rice WT and TaPGS1 OE-11 line. Bars represent means ± SEM, **P* < 0.01.

further suggested that overexpression of *TaPGS1* promotes accumulation of dry matter and smaller gaps in seeds, affecting the weight and size of seeds. We hypothesise that PGS1 may be involved in a conserved regulatory pathway that affects grain weight and size in cereals.

We identified the *Fl3* gene as a PGS1 downstream target using ChIP-seq coupled with transcriptomic differential expression analysis (Figure 5a–g; Figure 6a, b; Tables S8-S12). We showed that TaPGS1 activated *Fl3* genes from wheat and rice in yeast one-hybrid and tobacco transient expression experiments (Figure 6c, d). Meanwhile, the EMSAs verified that *TaPGS1* was specifically bound to the E-box of the *OsFl3* promoter (Figure 5c; Figure 6b, f; Figure S4a, b).
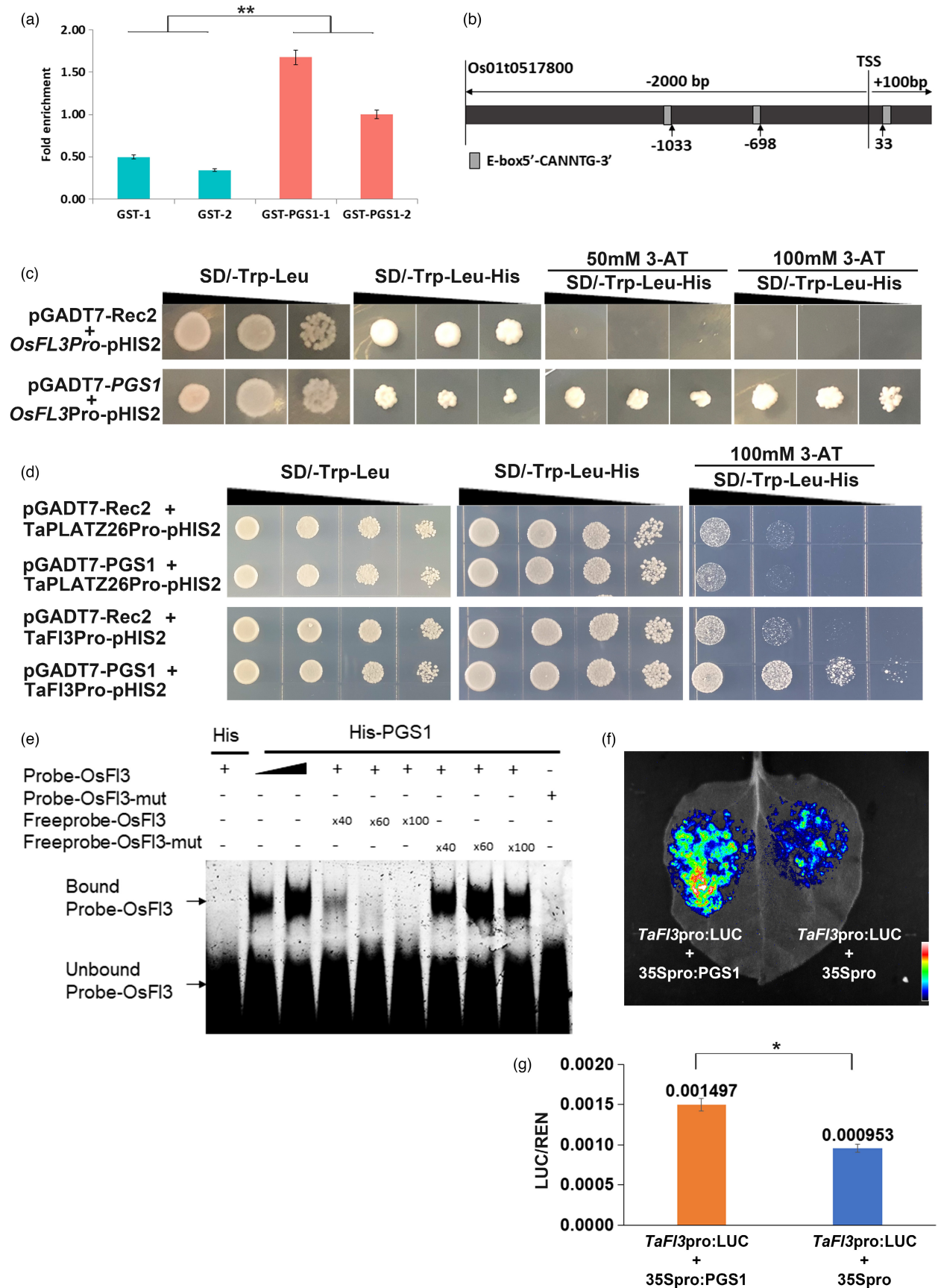
Many studies showed that different genes in the same regulatory pathway have similar or opposite expression patterns with either positive or negative regulatory functions (Dorweiler, 2008; Shirley et al., 2019; Sreenivasulu et al., 2006; Zhu et al., 2017). *TaPGS1* exhibits an expression pattern similar to that of *OsRc* (*OsPGS1*), where it is specifically expressed in seed tissues (Gu et al., 2011; Figure 1c, d). *TaPLATZ27 (TaFl3)* and *OsFl3* also show similar expression patterns in early developing wheat and rice seeds (Figure 5h; Figure S3c). Furthermore, *TaPGS1* enhanced the expression levels of *TaFl3* and *OsFl3* at early seed developmental stages in wheat and rice OE lines, respectively (Figure 5h; Figure S3c). Although maize *ZmFl3* is specifically expressed in the endosperm (Li et al., 2017), its expression pattern is similar to that of *Fl3* in wheat and rice seeds (Figure S3a, b). Moreover, *OsFl3* knockout rice lines showed an opposite phenotype to OE lines (Figure 7b–d, g; Table S14). In maize, the mutants that lost the homologous gene (*ZmFl3*) showed seed defects, resulting in lower grain weight (Li et al., 2017). These

data suggest that PGS1 and Fl3 have a conserved regulatory relationship in their expression patterns and confer similar seed phenotypes in different cereals. In other words, parallel evolution of *PGS1* and *Fl3* orthologous genes, gene networks, and similar phenotypes occurred in different crops under artificial selection. Thus, the bHLH-PLATZ TF (PGS1-Fl3) regulatory system representing a seed developmental pathway may be selected simultaneously under parallel evolution in different species.

An increase in seed size and weight is often accompanied by a decrease in seed number or other negative phenotypes, which is an undesirable result. For example, loss-of-function mutant SMOS1/2 decreases the size of various organs due to reduced cell size and abnormal microtubule orientation (Aya et al., 2014; Hirano et al., 2017). Phenotypes of GSK2 overexpression lines were dark green and dwarfed with a compact structure and fewer tillers (Tong et al., 2012). In the *Gramineae* rice model plant, the seed size, thousand-grain weight, and total grain yield were significantly increased in *TaPGS1* OE lines (Figure 3d–g; Tables S4, S5), with no significant differences in other traits (Figure S2b; Table S5). In our OE lines, the *TaPGS1* gene driven by the ubiquitin promoter showed the highest expression in stems and leaves compared to the WT (Figure S2c). However, its target gene (*OsFl3*) was specifically expressed in the tissues of early developing seeds (Figure 5h; Figure S3a). TaPGS1 only activated *OsFl3* expression in developing seeds but not in other organs, such as stems and leaves. How this selective activation of *OsFl3* is achieved remains unknown.

Plant PLATZ family proteins were first identified in pea (Nagano et al., 2001). To date, only three of these functions have been reported. *Fl3, GL6,* and *SG6* are all involved in the seed development pathway (Li et al., 2017; Wang et al., 2019; Zhou

Figure 6 Screening and identification of candidate target genes for TaPGS1 in rice. (a) *In vitro* pull-down of target DNA by TaPGS1. The samples of GST-TaPGS1(Red graphics) or GST(Cyan graphics) alone were subjected to quantitative PCR to detect OsFl3. The fold enrichment was normalised against the 18s, G-3PD, and actin promoters. Error bars indicate SEM, **P* < 0.01. (b) Distribution of E-box (marked by grey box) in the core promoter region (black box represents −2000 bp to +100 bp of TSS) of TaPGS1 binding gene *OsFl3* (*Os01g0517800*). (c) Yeast one-hybrid assay to test the interaction between TaPGS1 and the *OsFl3* promoter. (d) Yeast one-hybrid assay to test the interaction between TaPGS1 and the wheat *TaPLATZ26 and TaPLATZ27(TaFl3)* promoters. (e, g) Firefly luciferase (LUC) and Renilla luciferase (REN) activities were measured using the Dual-Luciferase Assay Reagents (Promega). (f) Electrophoretic mobility shift assay (EMSA) for specific binding of TaPGS1 to the OsFl3. Wildtype (Probe-OsFl3, 5′-ATGCCTTATGCCATCAATACCATGTGTCGCGTGAGTTGCATGCAACATC-3′) or mutant (Probe-OsFl3-mut, 5′-ATGCCTTATGCCCTCCCTACCCTTTTTCGCGTTAGTTTCCTTCAACATC-3′) oligonucleotides was used for competitive binding in 40-, 60-, and 100-fold excess of the wild-type probe (Table S13).

(a)

(b)

Os01t0517800

-2000 bp

TSS

+100bp

-1033

-698

33

E-box5'-CANNTG-3'

(c)

SD/-Trp-Leu   SD/-Trp-Leu-His   50mM 3-AT SD/-Trp-Leu-His   100mM 3-AT SD/-Trp-Leu-His

pGADT7-Rec2 + OsFL3Pro-pHIS2

pGADT7-PGS1 + OsFL3Pro-pHIS2

(d)

SD/-Trp-Leu   SD/-Trp-Leu-His   100mM 3-AT SD/-Trp-Leu-His

pGADT7-Rec2 + TaPLATZ26Pro-pHIS2

pGADT7-PGS1 + TaPLATZ26Pro-pHIS2

pGADT7-Rec2 + TaFl3Pro-pHIS2

pGADT7-PGS1 + TaFl3Pro-pHIS2

(e)

His   His-PGS1

Probe-OsFl3

Probe-OsFl3-mut

Freeprobe-OsFl3

Freeprobe-OsFl3-mut

Bound Probe-OsFl3

Unbound Probe-OsFl3

(f)

TaFl3pro:LUC + 35Spro:PGS1

TaFl3pro:LUC + 35Spro

(g)

© 2022 The Authors. *Plant Biotechnology Journal* published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 20, 1311–1326

Figure 7 Phenotypes of the *OsFL3* knockout rice lines. (a) Photographs of grains from the WT rice and OsFl3 knockout lines KO-831-2 and KO-836-4. Average thousand-grain weight (b), grain width (c), and grain length (d), seed-filled rate (e), grains per panicle, filled grain per panicle, flat grain, and empty grain (f) of WT rice and OsFl3 knockout lines KO-831-2 and KO-836-4 were quantified. All data represent means ± SEM ($n = 5$). * $P < 0.05$, ** $P < 0.01$. (g) Scanning electron microscopy of starch granules of mature endosperm in WT and OsFl3 KO-831-2 and KO-836-4 lines. Bar = 10 μm.

and Xue, 2020). *OsFl3* is homologous to *ZmFl3,* which is expressed in the endosperm and influences seed development and storage reserve filling (Li *et al.*, 2017). RNAPIII subunit RPC53, NRPC2, the TFIIIB subunit TbBRF1 and the TFIIIC subunit TFC1 physically interacts with ZmFl3 to affect RNAPIII activity and endoreduplication to control kernel development (Li *et al.*, 2017; Velez-Ramirez et al., 2015; Zhao et al., 2020). Figure S7 shows a proposed model to summarise how TaPGS1 may control seed development based on available knowledge, although many details on TaFl3 and OsFl3 remain unclear. Our study expands the understanding of the molecular mechanisms controlling cereal grain development and provides a valuable strategy for enhancing cereal yield.

Materials and methods

Plant materials

The material of common wheat (Fielder and CN 16) was used for expression pattern analysis of wheat bHLH genes by RT-qPCR at different developmental stages. CN 16 was provided by the Triticeae Institute of Sichuan Agriculture University and the Fielder by the Shandong Academy of Agricultural Sciences. Root, leaf, and seed samples were collected from materials growing in the greenhouse under a 16/8-h day-night photoperiod at 18 °C and 55%–60% relative humidity from July to September 2018. Samples of wheat WT and OE materials all consisted of seeds collected at 5, 10, 15, 20, 25, 30 DPA. Mature seeds were collected from T_3 plants, dried, and incubated for 2 days at 37 °C in the dark to germinate. Germinated seeds were then planted in the field under the natural environment. Root, leaf, and seed samples were collected from materials growing in the field under the natural environment from July to September 2016 and 2017.

Samples of rice WT and OE materials included roots, leaves at the seedling stage, and seeds collected at 5, 10, 15, 20, 25, and 30 DPA. The phenotypes of homozygous T_3, T_4, and T_5 generations were investigated. All samples were snap-frozen in liquid nitrogen and stored at −80 °C. Genomic DNA samples from all young plant leaves were extracted by the CTAB method. Total RNA was isolated with the Trizol reagent (Biofit Biotechnology, Chengdu, China) according to the manufacturer's instructions. Three independent samples from different panicles at the same time were mixed to extract RNA for transcriptome sequencing. The RNA samples were stored at −80 °C. Overexpression Kitaake rice lines (T_2 plants) were planted in the greenhouse under normal conditions (16-light at 30 °C/8-h dark at 22 °C).

Extraction of total protein from seed

Developing seeds of rice and wheat 10 and 15 days post-pollination were ground to a fine powder in liquid nitrogen. Approximately 50 mg of ground tissue per sample was used for total protein extraction. Protein extracts were obtained after the addition of lysis buffer (YEASEN, 20118ES60), rest in ice for 10 min, and centrifuged for 5 min at 13000 *g*. Total protein from the supernatant was mixed in equal volume with 2X SDS loading dye, denatured for 5 min at 95 °C, subjected to gel electrophoresis on a 10% SDS-PAGE gel, and analysed by immunoblotting.

Vector constructions and plant transformation

Various DNA fragments were amplified and introduced into different binary vectors for plant transformation. The Genome Analysis Centres (TGAC) showed full-length DNA sequence of

TaPGS1 (*Triticum aestivum* TGACv1 scaffold 080648 1DS, TRIAE_CS42_1DSTGACv1_080648_AA0251560, https://www.ncbi.nlm.nih.gov/assembly/GCA_900067645.1/#/st). In the wheat genome database IWGSC RefSeq v2.1, *TaPGS1* was annotated as two genes, numbered TraesCS1D02G094000 and TraesCS1D02G094200. For the *TaPGS1* overexpression vector construction, the full-length CDS (coding sequence, from TraesCS1D02G094000 and TraesCS1D02G094200) of *TaPGS1* was synthesised by Invitrogen Company and then introduced into the expression vector to generate pCAMBIA1300-Ubi-*TaPGS1*. The *Agrobacterium tumefaciens* strain AGL6 was used to transform rice and wheat plants. In the CRISPR/Cas9-mediated genome editing experiment, the paired primers gOsFl3F and gOsFl3R were selected based on an online analysis tool (http://crispr.dbcls.jp/) and synthesised. We used a vector system reported previously (Yan *et al.*, 2015). To produce a vector expressing both the Cas9 protein and guide RNAs, primers gOsFl3F and gOsFl3R (Table S16) were first annealed and mixed with *Bsa*I-digested *pAtU6-26-sgRNA-SK* plasmid to generate the entry plasmid. The sgRNA cassette was released from the resulting recombinant plasmid by *Nhe*I and *Spe*I digestion. The cassette was cloned into the *pCAMBIA1300-pYAO:Cas9* binary vector pre-digested by *Spe*I followed by dephosphorylation. The resulting plasmid was used for transformation into WT. Candidate genome-edited plants were screened first by PCR with the OsFl3-F and OsFl3-R primers (Table S16). Homozygous mutants were produced after self-crossing of candidate mutants and verified by sequencing.

Trait measurements

Grain length, width and thickness, and 1000-grain weight were measured when plants were completely mature. The entire spikelet hull, including the lemma and palea, was included in these measurements. For plot yield analysis, seven trial plots containing a total of 140 plants for each overexpression line were grown in a paddy field. Three plants and five replicates from each plot were randomly selected for yield measurement.

Measurement of starch, soluble sugars, and proteins

At least 60 mature grains of each WT and overexpression lines were ground into fine powder by an oscillator (60 HZ, 60S). The powder was filtered through an 80-mesh sieve to prepare samples for starch measurement. According to the manufacturer's protocol, total starch was measured using the Megazyme Total Starch Assay Kit (catalog number: 788K-TSTA-50A). Five biologically independent samples for each genotype were weighed for 50 mg of fine powder and applied to the measurement of protein and starch. The sulphuric acid-anthrone colorimetric method determined the soluble sugar content with wavelength at 620 nm according to the adapted optimised standard method (Zhiyong Zhang *et al.*, 2016). The content of each genotype was measured by weighing 25 mg powder. For assaying the total protein content, 50 mg of powder was wrapped in a special foil and used for measurement using the instrument of ELEMENTAR Rapid N exceed.

RT-qPCR

RT-qPCR was carried out with SYBR Premix Ex-Taq (Takara, Dalian, China) on a CFX96 Touch™ Deep Well Real-Time PCR Detection System (Bio-Rad, Hercules, CA). The cycling conditions included incubation for 60 s at 95 °C followed by 40 cycles of amplification (95 °C for 5 s and 60 °C for 30 s). According to the

manufacturer's instructions, the first-strand cDNA synthesis was carried out using a Prime-Script TM RT reagent kit (Takara Biotechnology, Dalian, China). Reference genes were used as an internal control to normalise the expression level of the target gene in wheat (Xiang-Yu Long et al., 2010) and rice (Bo-Ra Kim et al., 2003). The primers are listed in Table S16.

RNA in situ hybridisation

The spatial expression pattern of PGS1, TaPLATZ27, Osfl3 (Os01g0517800) in the wheat cultivar Fielder and rice cultivar kittaka was assayed by RNA in situ hybridisation. Four time-points (5 DAP, 10 DAP, 15 DAP, and 20 DAP) of development grain were fixed in 4% paraformaldehyde solution (sigma) with 0.1% TritonX-100 (Sigma) and 0.1% Tween-20 in PBS (Takara, cat #900) for 16 h. After dehydration using graded ethanol and vitrification by xylene, the samples were embedded in paraffin. 10-μm vertical sections of seeds were cut using Leica manual microtome (Leica FM2235). In situ hybridisation was carried out according to the protocol in previous work (Huang, 2019, PC). The fragment of each cDNA sequence was cloned and inserted into the pEasy-blunt-zero cloning vector. Primer sequences are listed in Table S16. The vectors used to synthesise antisense and sense RNA probes were transcribed in vitro by T7 and SP6 RNA polymerases according to the instructions for the DIG RNA labelling kit (Roche, catalog number 11175025910).

Scanning electron microscopy

The samples were transversely sectioned, air dried, sputter-coated with gold particles, observed, and photographed using a scanning electron microscope.

Immunoblot analysis

Separation of protein by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed using the Trans-blot Turbo Transfer System (Bio-Rad) and transferred onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad, #1620177) following instructions by the manufacturer. PVDF membranes containing electroblotted proteins were then incubated with the *TaPGS1*-specific anti-peptide antibody (cat No. WG-03074D, Abclonal) followed by a goat anti-rabbit IgG-HRP secondary antibody (cat No. M21002L, Abclonal). For equal loading control, blots were probed by the beta-actin mouse primary antibody followed by an anti-mouse secondary antibody (Lot#294272, Abmart). Proteins on immuno-blots were detected using the Novex ECL HRP Chemiluminescent substrate Reagent Kit (cat No. WP200005, Lot no.2043781, Invitrogen) according to the manufacturer's instructions, and images of different exposures were acquired using Tanon 6600 Luminescent Imaging Workstation (Tanon-5200Multi, Tanon).

Protein expression in *Escherichia coli* and purification

To generate His-tag fusion protein of TaPGS1, the full-length CDS of *TaPGS1* was amplified and cloned into the *pCold* TF DNA vector (TaKaRa) using *Kpn*I and *Xba*I enzyme sites. The plasmid of the positive clone was then introduced to *E. coli* Rosetta cells (YEASEN). The TaPGS1-His fusion protein was induced by the addition of 1 mM IPTG after OD reached the range of 0.6–0.9, then incubating at 16 °C for about 12 h in the incubator. According to the manufacturer's instruction, the TaPGS1-His fusion protein was purified using a Ni-NTA prepacked chromato-graphic column (Sangon, Lot. F618DA0004). 1 mM DTT, PMSF, and other protease inhibitors were included in the disruption and elution buffer to prevent degradation. The purity and

concentration of purified TaPGS1 protein were monitored with SDS-PAGE gel.

RNA-seq analysis

RNA-seq data of wheat containing different tissues and stages were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/). Clean data were remapped to the wheat genome (IWGSC RefSeq v1.0, https://www.wheatgenome.org/) by TopHat2.0 and anal-ysed with the Cufflinks software.

In rice, three independent biological replicates from three different tassel were mixed. cDNA libraries were constructed following Illumina standard protocols and sequenced with an Illumina HiSeq 2500 by Shanghai Biotechnology. The sequence reads were trimmed using fastx (version: 0.0.13) and mapped to Oryza sativa Japonica Group (IRGSP-1.0) using TopHat (version: 2.0.9). Kallisto (version: 0.46.0) was employed to reconstruct the transcripts and to estimate gene expression levels (Bray et al., 2016). Gene Ontology analysis was performed at the PANTHER 15.0 database (http://pantherdb.org/) and the *Arabidopsis* Infor-mation Resource (https://www.arabidopsis.org/).

ChIP in vitro and ChIP-Seq

Genomic DNA of Nipponbare (Kittaka) and purified GST-PGS1 protein were used for ChIP assays. Three-week-old seedlings were used for genomic DNA extraction. The genomic DNA was sheared into 200–600 bp fragments using an ultrasonic crusher. The GST fusion protein was affinity purified on glutathione-agarose beads (BD biosciences). GST-PGS1 and DNA fragments were co-incubated for 2 h. The incubation buffer includes 50 mM Tris, 1 mM EDTA, 100 mM KCl, adjust pH to 7.0 by HCl, 5% Glycerol, 0.1% Triton X-100; add freshly-made 100 mM DTT to a final concentration of 1 mM. After co-incubation, glutathione-agarose beads were washed three times using the incubation buffer. Then 4 mL 5 M NaCl was added into the sample for each 100 mL volume and was incubated for 4 h to break up cross-linked GST-PGS1 and DNA fragments. Subsequently, DNA fragments were extracted using the phenol-chloroform method for constructing the Illumina sequencing library. An Illumina sequencing library was constructed with the above-prepared DNA samples: DNA fragments between 100 and 500 bp were recovered from the gel and amplified by PCR for 20 cycles. The amplified DNA products were collected, ligated to the PEASY-Blunt vector for a quality test, and then sequenced with Hiseq2500 (Berry Genomics, Beijing). The primers are listed in Table S16.

ChIP-qPCR

The prepared DNA in ChIP was applied for qPCR using respective primer pairs in an EVER-Green PCR Master Mix (Bio-Rad) with a Bio-Rad CFX96 real-time PCR detection system. PCR reactions were performed in triplicate for each sample, and the expression levels (Enrichment of target DNA fragments) were normalised to the input sample for enrichment detection. The fold enrichment was calculated against rice housekeeping genes 18s, G-3PD, and actin. No addition of antibodies (NoAbs) served as a negative control. The primers are listed in Table S16.

ChIP-seq data analysis

Sequencing reads for ChIP and input DNA were mapped to Os-Nipponbare-Reference-IRGSP-1.0 (Kawahara et al., 2013) using Bowtie2 with default parameters (Langmead et al., 2009). Cross-correlation metrics were calculated using phantom peak qual tools (https://encodeproject.org/ENCODE/encodeTools.

html#metrics; Landt *et al.*, 2012). Only uniquely mapped reads were used for peak identification, and GST-PGS1 binding peaks were obtained by model-based analysis of ChIP-seq with default parameters (Zhang *et al.*, 2008). The peak summits were used to define the location types in the genome by the following criteria. Suppose a peak summit was located in (1) an intergenic region between genes, (2) a promoter region of a gene (upstream 2000 bp from the TSS, (3) exons of a gene, or (4) downstream of a gene (2000 bp after the TES, it was labelled according to the first criterion it matched. Because of the detailed peak distribution on the genome and promoter region, a strict cutoff was used to identify PGS1 target genes by labelling the genes with peak summits located in the upstream 2000 bp from the TSS.

Yeast one-hybrid assays

The full-length CDS sequence of *PGS1* was amplified and fused in frame with the GAL4 activation domain in the *pGADT7-Rec2* vector. The *OsFl3*, *TaPLATZ26*, and *TaPLATZ27* (*TaFl3*) promoter fragments were amplified and inserted into the *pHIS2* vector. Sequences of primers are listed in Table S16. The various pairs of recombined *pHIS2* and *pGADT7-PGS1* plasmids were co-transformed into yeast strain *Y187* (WEIDI, China, CAT#: YC1020) following the manufacturer's manual. The empty vector pGADT7--Rec 2 and pHIS2 plasmids were co-transformed as the negative control. Yeast colonies were grown on synthetic defined (SD)/-Leu/-Trp/-His medium containing 50 and 100 mM 3-AT.

Electrophoretic mobility shift assay

Oligonucleotide probes were synthesised according to the binding site and labelled with FAM at 5'end (at a single strand). EMSAs were carried as below. The binding reaction mixture included 29 µg of the purified fusion protein (Pcold-TaPGS1), 1 µL of 1 µm FAM-labelled DNA, 1 µL of 1 mg/mL salmon sperm DNA, 4 µL of 5X binding buffer, which includes 4 µL (5X EMSA Buffer, 3.54 µL; 1 M $MgCl_2$, 0.2 µL; 1 M DTT, 0.02 µL; H20, 0.22 µL), and ultrapure water to a final volume of 20 µL. The reaction mixture was incubated for 30 min at 30 °C and electrophoresed in a 6% polyacrylamide gel for about 80 min at constant 100 volts. A FLA9000 generated the image under the FAM laser. The sequences of probes used in EMSA are listed in Table S13.

Transient expression assay

A 2 kb promoter fragment of *TaFl3* was amplified from genomic DNA and inserted into the pGreenII-0800-LUC vector to generate the reporter construct TaFl3$_{pro}$:LUC. The coding sequence of PGS1 was cloned into the PRI101-AN vector to generate the effector construct 35S$_{pro}$:PGS1. The empty vector PRI101 AN was co-transformed with the TaFl3$_{pro}$:LUC reporter as the negative control. Primers used for generating these constructs are listed in TableS 16. Bacterial suspensions were infiltrated into young, fully expanded leaves of 3-week-old *N. benthamiana* leaves with Agrobacterium strain GV3101 following the manufacturer's manual (WEIDI, China, CAT#: AC1001). Transfected plants were grown under continuous white light. Firefly luciferase (LUC) and Renilla luciferase (REN) activities were measured 2 days after transformation using the Dual-Luciferase Assay Reagents (Promega) and a GLOMAX 20/20 luminometer. Three biological replicates were measured for each sample.

Accession numbers

Acknowledgments

Conflicts of interest

The authors declare that they have no competing interests.

Author contributions

J.W. designed the research; X.G. and L.G. performed the RNA-Seq and ChIPseq analysis; X.G., Y.F., Y. Lee., M.C., and M.L. performed the plasmid construction, genetic transformation, and CRISPR/Cas9 assay. X.G., C.Z., J.Y., W.T.L., and Q.H. performed the phenotypic analysis and planted the experimental material; X.G., Y.F., and M.L. performed experiments of smFISH, gene expression, SEM observation, and EMSA assay; X.G., H.D., J.W., M.C., B.L. wrote and revised the paper; J.W., B.L., Y.R.W., Y.Z., Y.M.W., Z.Y., X.C., J.M., Q.J., G.C., P.Q., W.L., W.T.L., Z.P., and Y.L. supervised the project.

References

Asano, K., Yamasaki, M., Takuno, S., Miura, K., Katagiri, S., Ito, T., Doi, K. *et al.* (2011) Artificial selection for a green revolution gene during japonica rice domestication. *Proc. Natl Acad. Sci.* **108**, 11034–11039.

Aya, K., Hobo, T., Sato-Izawa, K., Ueguchi-Tanaka, M., Kitano, H. and Matsuoka, M. (2014) A novel AP2-type transcription factor, SMALL ORGAN SIZE1, controls organ size downstream of an auxin signaling pathway. *Plant Cell Physiol.* **55**, 897–912.

Bo-Ra Kim, H.-Y.-N., Kim, S.-U., Kim, S.-I. and Chang, Y.-J. (2003) Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotech. Lett.* **25**, 3.

Bray, N.L., Pimentel, H., Melsted, P. and Pachter, L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527.

Chen, Z.Y., Guo, X.J., Chen, Z.X., Chen, W.Y., Liu, D.C., Zheng, Y.L., Liu, Y.X. *et al.* (2015) Genome-wide characterization of developmental stage- and tissue-specific transcription factors in wheat. *BMC Genom.* **16**, 125.

Cornejo, M.-J., Luth, D., Blankenship, K.M., Anderson, O.D. and Blechl, A.E. (1993) Activity of a maize ubiquitin promoter in transgenic rice. *Plant Mol. Biol.* **23**, 567–581.

Dorweiler, T.A.M.E.H.M.J.E. (2008) A maize CONSTANS-like gene, conz1, exhibits distinct diurnal expression patterns in varied photoperiods. *Planta*, **27**, 11.

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X. *et al.* (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoret. Appl. Genet.* **112**, 1164–1171.

Feng, F., Qi, W., Lv, Y., Yan, S., Xu, L., Yang, W., Yuan, Y. *et al.* (2018) OPAQUE11 is a central hub of the regulatory network for maize endosperm development and nutrient metabolism. *Plant Cell*, **30**, 375–396.

Fu, Y., Cheng, M., Li, M., Guo, X., Wu, Y. and Wang, J. (2020) Identification and characterization of PLATZ transcription factors in wheat. *Int. J. Mol. Sci.* **21**, 8934.

Gao, Y., Wu, M., Zhang, M., Jiang, W., Ren, X., Liang, E., Zhang, D. *et al.* (2018) A maize phytochrome-interacting factors protein ZmPIF1 enhances drought tolerance by inducing stomatal closure and improves grain yield in *Oryza sativa. Plant Biotechnol. J.* **16**, 1375–1387.

Gaur, V.S., Singh, U.S. and Kumar, A. (2011) Transcriptional profiling and in silico analysis of Dof transcription factor gene family for understanding their regulation during seed development of rice *Oryza sativa* L. *Mol. Biol. Rep.* **38**, 2827–2848.

Gu, X.Y., Foley, M.E., Horvath, D.P., Anderson, J.V., Feng, J., Zhang, L., Mowry, C.R. *et al.* (2011) Association between seed dormancy and pericarp color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid synthesis in weedy red rice. *Genetics*, **189**, 1515–1524.

Guo, X.J. and Wang, J.R. (2017) Global identification, structural analysis and expression characterization of bHLH transcription factors in wheat. *BMC Plant Biol.* **17**, 90.

Heang, D. and Sassa, H. (2012a) Antagonistic actions of HLH/bHLH proteins are involved in grain length and weight in rice. *PLoS One*, **7**(2), e31325.

Heang, D. and Sassa, H. (2012b) An atypical bHLH protein encoded by POSITIVE REGULATOR OF GRAIN LENGTH 2 is involved in controlling grain length and weight of rice through interaction with a typical bHLH protein APG. *Breed. Sci.* **62**, 133–141.

Hirano, K., Yoshida, H., Aya, K., Kawamura, M., Hayashi, M., Hobo, T., Sato-Izawa, K. *et al.* (2017) SMALL ORGAN SIZE 1 and SMALL ORGAN SIZE 2/ DWARF AND LOW-TILLERING form a complex to integrate auxin and brassinosteroid signaling in rice. *Molecular Plant*, **10**, 590–604.

Ji, X., Du, Y., Li, F., Sun, H., Zhang, J., Li, J., Peng, T. *et al.* (2019) The basic helix-loop-helix transcription factor, Os PIL 15, regulates grain size via directly targeting a purine permease gene Os PUP 7 in rice. *Plant Biotechnol. J.* **17**, 1527–1537.

Kawahara, Y., de la Bastide, M., Hamilton, J.P., Kanamori, H., McCombie, W.R., Ouyang, S., Schwartz, D.C. *et al.* (2013) Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice*, **6**, 4.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E. *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831.

Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.

Li, Q., Li, L., Yang, X., Warburton, M.L., Bai, G., Dai, J., Li, J. *et al.* (2010a) Relationship, evolutionary fate and function of two maize co-orthologs of rice GW2 associated with kernel size and weight. *BMC Plant Biol.* **10**, 143.

Li, Q., Wang, J., Ye, J., Zheng, X., Xiang, X., Li, C., Fu, M. *et al.* (2017) The maize imprinted gene Floury3 encodes a PLATZ protein required for tRNA and 5S rRNA Transcription through Interaction with RNA polymerase III. *Plant Cell*, **29**, 2661–2675.

Li, Q., Yang, X., Bai, G., Warburton, M.L., Mahuku, G., Gore, M., Dai, J. *et al.* (2010b) Cloning and characterization of a putative GS3 ortholog involved in maize kernel development. *Theoret. Appl. Genet.* **120**, 753–763.

Li, S., Zhao, B., Yuan, D., Duan, M., Qian, Q., Tang, L., Wang, B. *et al.* (2013) Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression. *Proc. Natl Acad. Sci.* **110**, 3167–3172.

Nagano, Y., Furuhashi, H., Inaba, T. and Sasaki, Y. (2001) A novel class of plant-specific zinc-dependent DNA-binding protein that binds to A/T-rich DNA sequences. *Nucleic Acids Res.* **29**, 4097–4105.

Nathalie Nesi, I.D., Jond, C., Pelletier, G., Caboche, M. and Lepiniec, L. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in arabidopsis siliques. *Plant Cell*, **12**, 1863–1878.

Shirley, N.J., Aubert, M.K., Wilkinson, L.G., Bird, D.C., Lora, J., Yang, X. and Tucker, M.R. (2019) Translating auxin responses into ovules, seeds and yield: Insight from Arabidopsis and the cereals. *J. Integr. Plant Biol.* **61**, 310–336.

Song, X.J., Huang, W., Shi, M., Zhu, M.Z. and Lin, H.X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630.

Sreenivasulu, N., Radchuk, V., Strickert, M., Miersch, O., Weschke, W. and Wobus, U. (2006) Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA- regulated maturation in developing barley seeds. *Plant J.* **47**, 310–327.

Su, Z., Hao, C., Wang, L., Dong, Y. and Zhang, X. (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoret. Appl. Genet.* **122**(1), 211–223.

Tong, H., Liu, L., Jin, Y., Du, L., Yin, Y., Qian, Q., Zhu, L. *et al.* (2012) DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell*, **24**, 2562–2577.

USDA. (2018) *World Agricultural Supply and Demand Estimates*. Washington, DC: USDA.

Velez-Ramirez, D., Florencio-Martinez, L., Romero-Meza, G., Rojas-Sanchez, S., Moreno-Campos, R., Arroyo, R., Ortega-López, J., Manning-Cela, R. & Martínez-Calvillo, S. (2015) BRF1, a subunit of RNA polymerase III transcription factor TFIIIB, is essential for cell growth of Trypanosoma brucei. *Parasitology*, **142**, 1563–1573.

Wang, A., Hou, Q., Si, L., Huang, X., Luo, J., Lu, D., Zhu, J. *et al.* (2019) The PLATZ transcription factor GL6 affects grain length and number in rice. *Plant Physiol.* **180**, 2077–2090.

Wang, X., Zhou, W., Lu, Z., Ouyang, Y. and Yao, J. (2015) A lipid transfer protein, OsLTPL36, is essential for seed development and seed quality in rice. *Plant Sci.* **239**, 200–208.

Xiang-Yu Long, J.-R.-W., Ouellet, T., Rocheleau, H., Wei, Y.-M., Zhi-En, P.U., Jiang, Q.-T., Lan, X.-J. *et al.* (2010) Genome-wide identification and evaluation of novel internal control genes for Q-PCR based transcript normalization in wheat. *Plant Mol. Biol.* **74**, 4.

Xing, Y. and Zhang, Q. (2010) Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* **61**, 421–442.

Yan, L., Wei, S., Wu, Y., Hu, R., Li, H., Yang, W. and Xie, Q. (2015) High-efficiency genome editing in arabidopsis using YAO promoter-driven CRISPR/ Cas9 system. *Molecular Plant*, **8**, 1820–1823.

Zhang, L., Zhao, Y.L., Gao, L.F., Zhao, G.Y., Zhou, R.H., Zhang, B.S. and Jia, J.Z. (2012) TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *New Phytol.* **195**, 574–584.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137.

Zhao, H., Qin, Y., Xiao, Z., Li, Q., Yang, N., Pan, Z., Gong, D. *et al.* (2020) Loss of function of an RNAPIII subunit leads to impaired maize kernel development. *Plant Physiol.* **184**, 359-373.

Zhiyong Zhang, X.Z., Yang, J., Messing, J. and Yongrui, W.U. (2016) Maize endosperm-specific transcription factors O2 and PBF network the regulation of protein and starch synthesis. *PNAS*, **113**, 5.

Zhou, S.R. and Xue, H.W. (2020) The rice PLATZ protein SHORT GRAIN6 determines grain size by regulating spikelet hull cell division. *J. Integr. Plant Biol.* **62**, 847–864.

Zhu, Q., Yu, S., Zeng, D., Liu, H., Wang, H., Yang, Z., Xie, X. *et al.* (2017) Development of "Purple Endosperm Rice" by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system. *Molecular Plant*, **10**, 918–929.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Bioinformatics analysis on expression characteristics of TaPGS1.
Figure S2 The phenotype of rice OE-11 line.
Figure S3 Expression analysis of *Fl3*.
Figure S4 *In vitro* pull-down of target DNA by TaPGS1.
Figure S5 The E-box locations of *TaPLATZ26* and *TaFl3* promoter.

Figure S6 Two frameshift mutations (KO-831-2, KO-836-4) of the *OsFl3* gene were generated by the CRISPR/Cas9 system.

Figure S7 The model of bHLH-PLATZ (PGS1-Fl3) signaling pathways of seed development in cereal grain.

Table S1 Ten homologous bHLH proteins of PGS1 (*Aegilops tauschii* L., *Zea mays* L., *Arabidopsis thaliana* L., *Brachypodium distachyon* L., *Solanum tuberosum* L., *Solanum lycopersicum* L., *Setaria italica* L., *Sorghum bicolor* L., *Brassica rapa* L., *Oryza sativa* L.).

Table S2 RNAseq data of the different wheat tissues group. Transcripts Per Million (TPM) value was used to normalise the RNA-seq data.

Table S3 Thousand-grain weight, grain length, and grain width of wheat WT and TaPGS1 overexpressed lines 166-6-OE and 166-39-OE.

Table S4 Thousand grain weight, grain length, and grain width of WT rice and TaPGS1 overexpression lines OE-3 and OE-11.

Table S5 Plant height, ear length, number of grains per panicle, number of filled grains per panicle, effective spike number per plant, and grain weight per plant of WT rice and TaPGS1 overexpression lines OE-3 and OE-11.

Table S6 Starch, soluble sugar, and protein content in WT wheat and TaPGS1 overexpression lines 166-6-OE and 166-39-OE.

Table S7 Content of starch, soluble sugar, and protein content in rice WT and TaPGS1 overexpression lines OE-3 and OE-11.

Table S8 PGS1 DEGs at different tissues and stages in rice.

Table S9 PGS1 binding sites identified by ChIP-Seq and ontology annotations of the sites neighbouring gene.

Table S10 The 164 high-confidence potential targets of PGS1.

Table S11 The 13 specifically DEGs at the seed (DPA10, 15, 20) from 164 high-confidence potential targets of PGS1.

Table S12 TPM value of the 13 specifically DEGs at the seed (DPA10, 15, 20) from 164 high-confidence potential targets of TaPGS1.

Table S13 Probes list of Osfl3 promotors for electrophoretic mobility shift assay (EMSA) experiments.

Table S14 Thousand grain weight, grain length, and grain width of WT rice and *OsFl3* knockout lines KO-831-2 and KO-836-4.

Table S15 Number of grains per panicle, filled grains per panicle, flat grain, and empty grains of WT rice and *OsFl3* knockout lines KO-831-2 and KO-836-4.

Table S16 Primers in this study.