Theoretical and Applied Genetics (2022) 135:51–64
https://doi.org/10.1007/s00122-021-03951-7

ORIGINAL ARTICLE

# Fine mapping and candidate gene analysis of *qGSN5*, a novel quantitative trait locus coordinating grain size and grain number in rice

**Hua Yuan[1] · Peng Gao[2] · Xiaoling Hu[2] · Min Yuan[1] · Zhengyan Xu[2] · Mengya Jin[2] · Wencheng Song[2] · Shijie Zhan[2] · Xiaobo Zhu[1] · Bin Tu[2] · Ting Li[1] · Yuping Wang[2] · Bingtian Ma[2] · Peng Qin[2] · Weilan Chen[2] · Shigui Li[1,2]**

Received: 5 August 2021 / Accepted: 14 September 2021 / Published online: 23 October 2021


## Abstract

***Key message*** ***qGSN5*, a novel quantitative trait locus coordinating grain size and grain number in rice, was fine-mapped to an 85.60-kb region. *GS3* may be a suppressor of *qGSN5*.**

**Abstract** Grain size and grain number are two factors that directly determine rice grain yield; however, the underlying genetic mechanisms are complicated and remain largely unclear. In this study, a chromosome segment substitution line (CSSL), CSSL28, which showed increased grain size and decreased grain number per panicle, was identified in a set of CSSLs derived from a cross between 93-11 (recipient) and Nipponbare (donor). Four substitution segments were identified in CSSL28, and the substitution segment located on chromosome 5 was responsible for the phenotypes of CSSL28. Thus, we defined this quantitative trait locus (QTL) as *grain size and grain number 5* (*qGSN5*). Cytological and quantitative PCR analysis showed that *qGSN5* regulates the development of the spikelet hull by affecting cell proliferation. Genetic analysis showed that *qGSN5* is a semi-dominant locus regulating grain size and grain number. Through map-based cloning and overlapping substitution segment analysis, *qGSN5* was finally delimited to an 85.60-kb region. Based on sequence and quantitative PCR analysis, *Os05g47510*, which encodes a P-type pentatricopeptide repeat protein, is the most likely candidate gene for *qGSN5*. Pyramiding analysis showed that the effect of *qGSN5* was significantly lower in the presence of a functional *GS3* gene, indicating that *GS3* may be a suppressor of *qGSN5*. In addition, we found that *qGSN5* could improve the grain shape of hybrid rice. Together, our results lay the foundation for cloning a novel QTL coordinating grain size and grain number in rice and provide a good genetic material for long-grain hybrid rice breeding.

## Introduction

Rice (*Oryza sativa* L.) is one of the most important staple foods in the world, feeding more than half of the world's population; thus, continuously increasing rice grain yield is still one of the major goals of breeding programs. Grain size and grain number per panicle are two factors that directly determine rice grain yield (Xing et al. 2010; Li et al., 2021a). Grain size is characterized by grain length, grain width, and grain thickness (Zuo et al. 2014). Many genes/quantitative trait loci (QTLs) associated with grain size have been cloned and shown to be involved in several signaling/regulatory pathways, including G-protein signaling, mitogen-activated protein kinase (MAPK) signaling, the ubiquitin–proteasome degradation pathway, transcriptional regulation, and phytohormone signaling (Fan et al., 2019; Li et al., 2019; Li et al., 2021a). For example, *GS3*, which encodes an atypical Gγ protein, was the first identified major negative regulator of

Communicated by Jiankang Wang.

Hua Yuan, Peng Gao and Xiaoling Hu have contributed equally to this work.

✉ Weilan Chen
chenweilan@sicau.edu.cn

✉ Shigui Li
lishigui@sicau.edu.cn

[1] State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Chengdu 611130, Sichuan, China

[2] Rice Research Institute, Sichuan Agricultural University, Chengdu 611130, Sichuan, China

grain size (Fan et al., 2006; Mao et al., 2010). In addition, other G-protein subunits also regulate grain size in rice (Sun et al., 2018), including *RGA1* (Gα subunit), *RGB1* (Gβ subunit), and *RGG1*, *RGG2*, *DEP1*, and *OsGGC2* (Gγ subunits). The OsMKKK10-OsMKK4-OsMPK6 signaling cascade positively regulates grain size in rice (Duan et al., 2014; Liu et al., 2015b; Xu et al., 2018a), and *GRAIN SIZE AND NUMBER1* (*GSN1*), which encodes a MAPK phosphatase (OsMKP1), negatively regulates grain size by directly interacting with and deactivating OsMPK6 (Guo et al., 2018; Xu et al., 2018b). To date, several components of the ubiquitin pathway have been shown to be involved in the regulation of grain size in rice, such as *GRAIN WIDTH 2* (*GW2*) and *Chang Li Geng 1* (*CLG1*), which both encode RING-type E3 ubiquitin ligases; *GW2* negatively regulates grain width and *CLG1* positively regulates grain length (Song et al., 2007; Yang et al., 2021). *WIDE AND THICK GRAIN 1*/*OsOTUB1* and *large grain1*/*OsUBP15*, which encode proteases with deubiquitination activity, are important factors that determine grain size and shape (Huang et al., 2017; Wang et al., 2017; Shi et al., 2019). Many transcription factors are involved in regulating grain size, such as members of the SQUAMOSA promoter binding protein-like (SPL) family (*GW8/OsSPL16*, *GLW7/OsSPL13*, *OsSPL18*) (Wang et al., 2012; Si et al., 2016; Yuan et al., 2019) and basic helix-loop-helix (bHLH) family (*An-1*, *OsbHLH107*, *OsbHLH079*) (Luo et al., 2013; Yang et al., 2018; Seo et al., 2020). Phytohormone signaling genes also play important role in regulating grain size; these include *GW5/GSE5* (Duan et al., 2017; Liu et al., 2017) and *GS2/GL2/OsGRF4* (Che et al., 2015; Duan et al., 2015; Hu et al., 2015), which regulate grain size through brassinosteroid signaling, *TGW6* (Ishimaru et al., 2013) and *BG1* (Liu et al., 2015a), which regulate grain size by affecting indole-3-acetic acid homeostasis, and *BG3/OsPUP4* (Xiao et al., 2019) and *AGO2* (Yin et al., 2020), which regulate grain size by affecting cytokinin (CK) distribution.

Grain number per panicle is determined mainly by the number of panicle branches (Li et al., 2021b). Many genes/QTLs associated with grain number have been identified in rice, and most of them have been found to be involved in phytohormone signaling. For example, *Grain number 1a* (*Gn1a*), which encodes a cytokinin oxidase/dehydrogenase (OsCKX2), is a major QTL for grain number (Ashikari et al., 2005), and *DROUGHT AND SALT TOLERANCE* (*DST*) directly regulates its expression in the reproductive meristem (Li et al., 2013); thus, the *DST-Gn1a* pathway negatively regulates grain number by affecting CK accumulation in the young panicle. *NUMBER OF GRAINS 1*, which encodes an enoyl-CoA hydratase/isomerase, positively regulates grain number per panicle and may be involved in the jasmonic acid pathway (Huo et al., 2017). *Grain Number per Panicle1* and *Semi Dwarf1*, which encode gibberellin 20 oxidase 1 (GA20ox1) and GA20ox2, respectively, are involved in gibberellic acid biosynthesis and both genes positively regulate grain number in rice (Wu et al., 2016; Su et al., 2021).

Usually, there is a coordination between the grain size and grain number in plants (Sadras 2007; Fan et al., 2019). For example, *GSN1/OsMKP1*, *BG3/OsPUP4*, and *SGDP7/FZP* have been shown to coordinate the trade-off between grain size and grain number in rice; the *gsn1* mutant and *bg3-D* mutant produce larger but fewer grains (Guo et al., 2018; Xiao et al., 2019). On the contrary, *FZP* RNA-interference lines produce more grains but with significantly decreased grain size (Bai et al., 2017). Although there is clear evidence for a trade-off between grain size and grain number, the molecular mechanism underlying the coordination of this trade-off remains largely unclear.

In the present study, using a chromosome segment substitution line (CSSL) population derived from a cross between 93-11 and Nipponbare (Nip), we identified a novel QTL, *qGSN5*, that coordinates grain size and grain number in rice. And *qGSN5* was finally delimited to an 85.60-kb region. Candidate gene analysis showed that *Os05g47510*, which encodes a P-type pentatricopeptide repeat (PPR) protein, was the most likely candidate gene responsible for *qGSN5*. In addition, we found that the effect of *qGSN5* was significantly inhibited by the functional *GS3* gene, indicating that *GS3* may be a suppressor of *qGSN5* in regulating grain size. Thus, our results lay the foundation for cloning a novel QTL and studying the mechanism of coordination of grain size and grain number.

## Materials and methods

### Plant materials and growth conditions

CSSL28, a CSSL with increased grain size and decreased grain number per panicle, was identified in a set of CSSLs ($BC_5F_5$) derived from a cross between 93-11 (recipient) and Nip (donor). CSSL28 was crossed with 93-11 to generate a segregating $BC_6F_2$ population for primary mapping. SSSL-*qGSN5*, a single segment substitution line (SSSL) harboring the *qGSN5* locus, was identified in the $BC_6F_2$ population using marker-assisted selection (MAS). Then, SSSL-*qGSN5* was crossed with 93-11 to generate a segregating $BC_7F_2$ population for fine mapping. Simultaneously, recombinants harboring substitution segments of different lengths were selected in the $BC_7F_2$ population to perform fine mapping using the overlapping substitution segment method. SSSL-*GS3*, an SSSL harboring the functional *GS3* gene from Nip, was also identified in our CSSL population, and SSSL-*GS3* was crossed with SSSL-*qGSN5* to generate a double segment substitution line (DSSL), DSSL-*GS3/qGSN5*, using MAS. All plants were grown in the experimental fields of

Sichuan Agricultural University in Wenjiang (Sichuan Province, China) or Lingshui (Hainan Province, China) under normal cultivation conditions.

### Agronomic trait investigation and statistical analysis

For systematic investigation of agronomic traits, plants were grown in a random block design with three replicates. At maturity, five plants in the middle of each block were harvested. Then, the following agronomic traits were investigated: plant height, grain length, grain width, 1000-grain weight, number of tillers per plant, panicle length, numbers of primary and secondary branches, number of grains per panicle, seed setting rate, grain yield per plant, percentage of grains with chalkiness, and chalkiness degree. Grain length, grain width, and 1000-grain weight were measured using a Mini 1600 automatic analysis system (Jie Lai Mei Technology Co., LTD., Chengdu, China). The percentage of grains with chalkiness and the chalkiness degree were measured using SC-E software (Wanshen Detection Technology Co., LTD., Hangzhou, China). A Student's *t* test was conducted for each trait to analyze the significance of differences. The mean value and standard deviation (SD) were calculated using Microsoft Excel 2019 (Microsoft, Redmond, USA). Multiple comparisons and two-way analysis of variance (ANOVA) were conducted using data processing system (Tang et al., 2013).

### Scanning electron microscopy

Cytological analysis of spikelet hulls was performed as described previously (Yuan et al., 2017). In brief, spikelet hulls from 93-11 and SSSL-*qGSN5* plants were collected before anthesis and fixed in 2.5% glutaraldehyde. Then, gradient ethanol dehydration (from 30 to 100%), critical drying, and gold spraying were performed, and the samples were observed by scanning electron microscopy (SEM) (JSM-7500F, JEOL, Tokyo, Japan). Cell length was measured using Image-Pro Plus 6.0 (Media Cybernetics, Rockville, USA). To compare the starch granules of 93-11 and SSSL-*qGSN5* grains, transverse sections of mature grains were sputter-coated with platinum and observed by SEM.

### Development of markers

For genotyping and fine mapping, 39 polymorphic markers were developed according to the sequence differences between 93-11 and Nip: six tetra-primers and 33 InDel (Insertion/Deletion) markers (Table S1). Tetra-primers based on single nucleotide polymorphisms (SNPs) were developed using an online system (http://primer1.soton.ac.uk/primer1.html), and InDel markers were designed using Primer Premier 6.0 (www.PremierBiosoft.com). The marker for genotyping of *GS3* was developed based on the functional SNP (C/A) located in exon 2 (Fan et al., 2006).

### Total RNA extraction and quantitative real-time PCR assay

Total RNA was extracted from young panicles using the Plant Total RNA Isolation Kit (FOREGENE, Chengdu, China). Total RNA (500 ng) was used to synthesize first-strand cDNA with the HiScript® III 1st Strand cDNA Synthesis Kit (+ gDNA wiper) (Vazyme, Nanjing, China). Quantitative real-time PCR (qRT-PCR) was performed using 2 × SYBR Green qPCR Master Mix (Bimake, Houston, USA) on a qTOWER$^3$G Real-Time PCR thermocycler (Analytik Jena AG, Jena, Germany). Three independent replicates were performed for each assay, and *OsActin* was used as an internal control. The primers for qRT-PCR assays are listed in Table S2.

### Dual-luciferase transient assay

To compare the promoter activity of candidate gene *Os05g47510*, the ~ 2 kb promoter fragments of *Os05g47510* were separately amplified from 93-11 and Nip using the primers listed in Table S1 and then subcloned into a pGreenII 0800-LUC vector, which contains the LUC (luciferase) reporter gene and REN (renilla luciferase) internal control. The resulting vectors were transfected into rice protoplasts for transient expression, and the relative luciferase activity LUC/REN was measured using the Dual-Luciferase Reporter Gene Assay Kit (Beyotime, Shanghai, China).

## Results

### Identification of a QTL associated with grain size and grain number

To identify QTLs for grain size and grain number in rice, two sequenced cultivars, 93-11 (*indica*) and Nip (*japonica*), which had significant differences in grain size and grain number per panicle (Fig. S1), were respectively selected as the recipient parent and donor parent to construct a set of CSSLs (Fig. S2). Among the CSSLs in the $BC_5F_5$ population (Fig. S2), one, CSSL28, showed no difference in plant architecture compared with 93-11 (Fig. S3a), but showed increased grain size and 1000-grain weight (Fig. S3b–e) and decreased grain number per panicle (Fig. S3f). To quickly and precisely identify the substitution segment in CSSL28, CSSL28 was subjected to genotyping with the Rice 9 K SNP array, and four substitution segments were identified on chromosomes 5 (25.23–27.83 Mb), 8 (0.04–0.71 Mb,

18.78–22.00 Mb), and 10 (0.09–0.76 Mb) (Fig. S3g). To determine which substitution segment was responsible for the phenotypes of CSSL28, we crossed CSSL28 with the recurrent parent 93-11 to generate $BC_6F_1$ plants, and a $BC_6F_2$ segregating population comprising 872 plants was developed by self-pollination of $BC_6F_1$ (Fig. S2, Fig. S4a). As shown in Fig. S4a, the $BC_6F_1$ plants and most of the individuals in the $BC_6F_2$ population had a grain length intermediate between that of the two parents, indicating that the large-grain phenotype was semi-dominant. Subsequently, four SSSLs were identified in the $BC_6F_2$ population using MAS and named SSSL28-5, SSSL28-8.1, SSSL28-8.2, and SSSL28-10 (Fig. S3g). Phenotypic investigation showed that only SSSL28-5 had phenotypes similar to those of CSSL28, including increased grain size and decreased grain number (Fig. S3c-f). These results indicated that the QTL for the grain size and grain number phenotypes of CSSL28 was located in the substitution segment on chromosome 5. Thus, we named this QTL *grain size and grain number 5* (*qGSN5*), and SSSL28-5 was renamed SSSL-*qGSN5*.

## Characterization of SSSL-*qGSN5*

The genetic background of SSSL-*qGSN5* was reconfirmed using the Rice 9 K SNP array; this SSSL had only one substitution segment on the long arm of chromosome 5 (Fig. 1a). There was no difference in plant architecture or height between SSSL-*qGSN5* and 93-11 (Fig. 1b, Table 1), but the grains were significantly longer (about 10% longer), and the grains were also slightly wider, leading to significantly higher 1000-grain weight (> 5 g) (Fig. 1c, Table 1). However, the grain number per panicle was significantly decreased (by more than 20%) because of the smaller number of secondary branches (Fig. 1d, Table 1). In addition, we compared change in caryopsis fresh and dry weight during grain filling between 93-11 and SSSL-*qGSN5* and found that the fresh and dry weights of SSSL-*qGSN5* were significantly higher than those of 93-11 starting at 9 days after fertilization, with the differences leveling off at 21 days after fertilization, indicating that SSSL-*qGSN5* had a higher grain filling rate (Fig. 1e–g). Importantly, the increased 1000-grain weight and higher grain filling rate of SSSL-*qGSN5* did not affect the appearance of the rice grains (Table 1, Fig. S5), suggesting that SSSL-*qGSN5* may have the potential for use in rice breeding.

Grain size is determined by cell number and cell size (Li et al., 2018), SSSL-*qGSN5* had a larger spikelet hull before flowering (Fig. 2a, f); therefore, we compared the inner and outer epidermal cells in spikelet hulls between 93-11 and SSSL-*qGSN5* using SEM (Fig. 2b–e). There was no difference in cell length between 93-11 and SSSL-*qGSN5*; however, the cell number of both the inner and outer epidermal cells was significantly higher in SSSL-*qGSN5* (Fig. 2g–j). We then investigated the expression level of cell cycle-related genes in 93-11 and SSSL-*qGSN5* and found that most of the genes facilitating cell proliferation were significantly up-regulated in SSSL-*qGSN5*; however, four ICK/KRP family genes, which are inhibitors of cyclin-dependent kinases and negatively regulate cell number in plants (Torres Acosta et al., 2011), were significantly down-regulated in SSSL-*qGSN5* (Fig. 2k). Taken together, these results indicated that *qGSN5* regulates the development of the spikelet hull by affecting cell proliferation.

## Fine mapping of *qGSN5*

As described above, we had determined that *qGSN5* was located on the substitution segment on chromosome 5 (Fig. S3, Fig. 3a). To map the *qGSN5* locus, we first performed map-based cloning using the 93-11/CSSL28 $BC_6F_2$ population (see above) (Fig. S4a), and *qGSN5* was initially mapped to a 395.45-kb region between markers ARMS7 and Y101 (Fig. 3b). To further narrow down the interval of *qGSN5*, we generated a larger mapping population derived from a cross between 93-11 and SSSL-*qGSN5* (Fig. S4b), and *qGSN5* was further mapped to a 111.25-kb region between markers Y83 and ARMS4 based on the number of recombinants (Fig. 3c). Moreover, we identified ten recombinants (L1–L10) harboring substitution segments of different lengths in the 93-11/SSSL-*qGSN5* $BC_7F_2$ population using MAS (Fig. 3d), and these recombinants were selfed to obtain the $BC_7F_3$ generation. The grain length of L1–L5, L9, and L10 was similar to that of 93-11; however, the grains of L6–L8 were significantly longer than those of 93-11. Based on the phenotypes of the recombinants with overlapping substitution segments, *qGSN5* was finally delimited to an 85.60-kb region flanked by markers Y83 and ARMS3 (Fig. 3d).

## Analysis of *qGSN5* candidate genes

According to the Rice Genome Annotation Project (http://rice.uga.edu), 15 predicted genes are located in the 85.60-kb region (Table S3). First, we investigated the expression patterns of these genes using our previously published R527 RNA-sequencing data (Qin et al., 2021) and found that the expression of most genes was extremely low or not detectable in all investigated tissues, and only six genes (*Os05g47510*, *Os05g47530*, *Os05g47540*, *Os05g47550*, *Os05g47560*, *Os05g47630*) were expressed in young panicles (Table S3). Among these genes, *Os05g47560* and *Os05g47630* had no difference in amino acid sequence between 93-11 and SSSL-*qGSN5*. Thus, the remaining four genes (*Os05g47510*, *Os05g47530*, *Os05g47540*, *Os05g47550*) were prioritized as candidate genes for analysis.

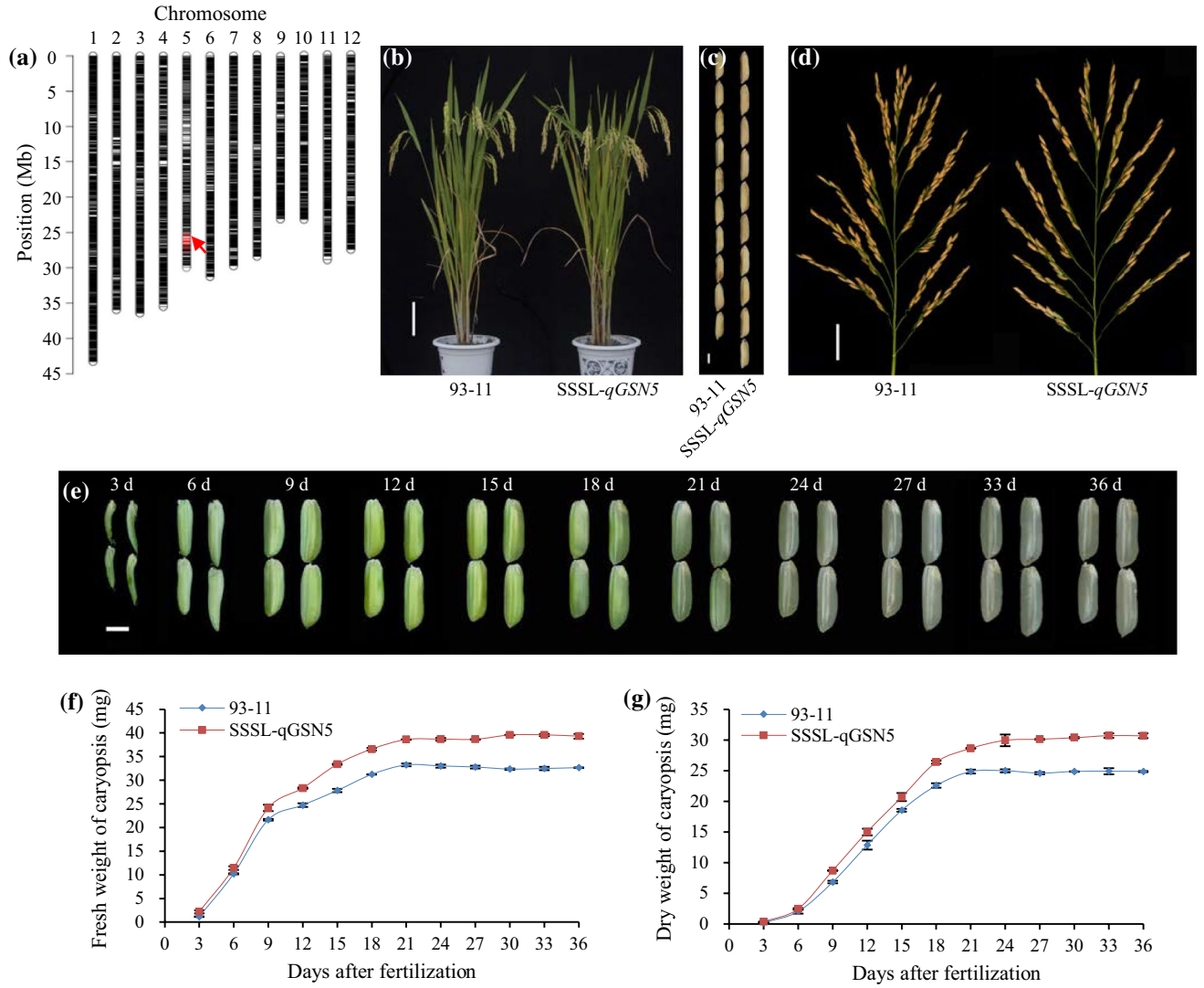


**Fig. 1** Comparison of phenotypes between 93-11 and SSSL-*qGSN5*. **a** Diagram showing the substitution segment in SSSL-*qGSN5*. Black and red lines indicate 93-11 and Nip genotypes, respectively, as determined by genotyping with the Rice 9 K SNP array. The red arrow indicates the location of the substitution segment in SSSL-*qGSN5*. **b** Plant architecture of 93-11 and SSSL-*qGSN5* at maturity. Scale bar, 10 cm. **c** Comparison of grain length between 93-11 and SSSL-*qGSN5*. Scale bar, 3 mm. **d** Panicle architecture of 93-11 and SSSL-*qGSN5*. Scale bar, 3 cm. **e** Dynamic comparison of caryopsis development between 93-11 and SSSL-*qGSN5*. Scale bar, 3 mm. **f** Dynamic comparison of caryopsis fresh weight between 93-11 and SSSL-*qGSN5*. **g** Dynamic comparison of caryopsis dry weight between 93-11 and SSSL-*qGSN5*. Data are given as means ± SD ($n = 90$ grains)

Sequence analysis showed that all four genes had differences in the coding regions between the two parents (Fig. 4a, Fig. S6a–c). *Os05g47510*, which encodes a P-type PPR protein, had four SNPs, resulting in one synonymous and three nonsynonymous mutations, and two of the nonsynonymous mutations were located in the PPR domain (Fig. 4a). *Os05g47530*, encoding an acid phosphatase/vanadium-dependent haloperoxidase, had a 6-bp deletion and three SNPs in 93-11, resulting in deletion of two amino acids and two nonsynonymous substitutions (Fig. S6a). *Os05g47540* encodes a phosphoethanolamine N-methyltransferase, OsPEAMT2, and *Os05g47550* encodes a clathrin assembly protein with an ENTH (Epsin N-terminal homology) domain. Both *Os05g47540* and *Os05g47550* had three SNPs, and one SNP in *Os05g47540* and two SNPs in *Os05g47550* resulted in nonsynonymous mutations; however, no mutation was located in a predicted domain (Fig. S6b, c). We also investigated the expression of these four genes in developing young panicles and found that only *Os05g47510* showed a consistent difference in expression between 93-11 and SSSL-*qGSN5*, with significantly higher expression levels observed in SSSL-*qGSN5* (Fig. 4b, Fig. S6d). Consistent with the difference in expression levels, there were several polymorphisms, namely 13 SNPs and

**Table 1** Comparison of agronomic traits between different plants

| Trait | Wenjiang in 2018 | | | Wenjiang in 2019 | | | Wenjiang in 2020 | |
|---|---|---|---|---|---|---|---|---|
| | 93-11 | SSSL-*qGSN5* | 93-11/SSSL-*qGSN5* | 93-11 | SSSL-*qGSN5* | 93-11/SSSL-*qGSN5* | 93-11 | SSSL-*qGSN5* |
| Plant height (cm) | 114.08 ± 2.16 | 111.77 ± 3.69 | 111.62 ± 2.81 | 111.67 ± 0.67 | 113.33 ± 1.20 | 114.89 ± 3.24 | 112.00 ± 3.87 | 114.44 ± 2.35 |
| Grain length (mm) | 9.41 ± 0.06 | 10.44 ± 0.10** | 9.95 ± 0.05** | 9.54 ± 0.02 | 10.63 ± 0.05** | 10.14 ± 0.12** | 9.49 ± 0.05 | 10.43 ± 0.07** |
| Grain width (mm) | 2.85 ± 0.02 | 2.97 ± 0.01* | 2.89 ± 0.03 | 2.92 ± 0.01 | 3.08 ± 0.03* | 2.98 ± 0.01 | 2.79 ± 0.01 | 2.92 ± 0.04* |
| 1000-grain weight (g) | 31.58 ± 0.67 | 38.28 ± 0.22** | 34.07 ± 0.18** | 32.41 ± 0.29 | 38.01 ± 1.19** | 35.58 ± 0.30** | 32.33 ± 0.21 | 37.65 ± 0.17** |
| Number of tillers per plant | 7.20 ± 0.60 | 7.67 ± 0.95 | 7.53 ± 0.64 | 7.53 ± 0.12 | 7.40 ± 1.25 | 6.53 ± 0.42 | 6.33 ± 0.90 | 6.13 ± 0.12 |
| Panicle length (cm) | 22.18 ± 0.44 | 22.93 ± 0.81 | 22.42 ± 0.59 | 21.99 ± 0.60 | 23.31 ± 0.70 | 22.41 ± 0.21 | 22.91 ± 0.55 | 23.85 ± 0.83 |
| Number of primary branches | 11.08 ± 0.79 | 11.42 ± 0.51 | 11.38 ± 0.52 | 10.63 ± 0.12 | 10.83 ± 0.42 | 10.90 ± 0.26 | 12.00 ± 0.00 | 11.92 ± 0.51 |
| Number of secondary branches | 39.75 ± 3.55 | 28.17 ± 2.55** | 35.13 ± 2.36 | 39.80 ± 3.60 | 32.50 ± 5.36* | 36.87 ± 5.38 | 41.00 ± 3.07 | 28.67 ± 4.27** |
| Number of grains per panicle | 170.88 ± 3.90 | 123.29 ± 6.40** | 146.85 ± 11.19* | 171.55 ± 10.88 | 130.16 ± 7.49** | 160.40 ± 7.50* | 165.60 ± 7.06 | 130.37 ± 5.21** |
| Seed setting rate (%) | 93.68 ± 0.83 | 84.29 ± 1.34** | 89.21 ± 1.32* | 84.44 ± 2.57 | 74.78 ± 2.54** | 86.34 ± 6.84 | 93.67 ± 1.24 | 84.56 ± 3.06** |
| Grain yield per plant (g) | 36.39 ± 3.05 | 31.04 ± 2.64* | 33.85 ± 4.20 | 37.77 ± 3.79 | 26.43 ± 5.60* | 34.08 ± 4.37 | 33.40 ± 1.91 | 26.00 ± 1.22** |
| Percentage of grain with chalkiness (%) | / | / | / | 30.33 ± 3.79 | 33.00 ± 1.41 | / | 41.01 ± 2.64 | 42.44 ± 1.36 |
| Chalkiness degree (%) | / | / | / | 13.47 ± 1.52 | 12.34 ± 0.01 | / | 35.44 ± 2.60 | 36.49 ± 1.42 |

Data are shown as means ± SD ($n = 15$ plants). * indicates $P < 0.05$ and ** indicates $P < 0.01$ (Student's $t$ test). "/" indicates not investigated

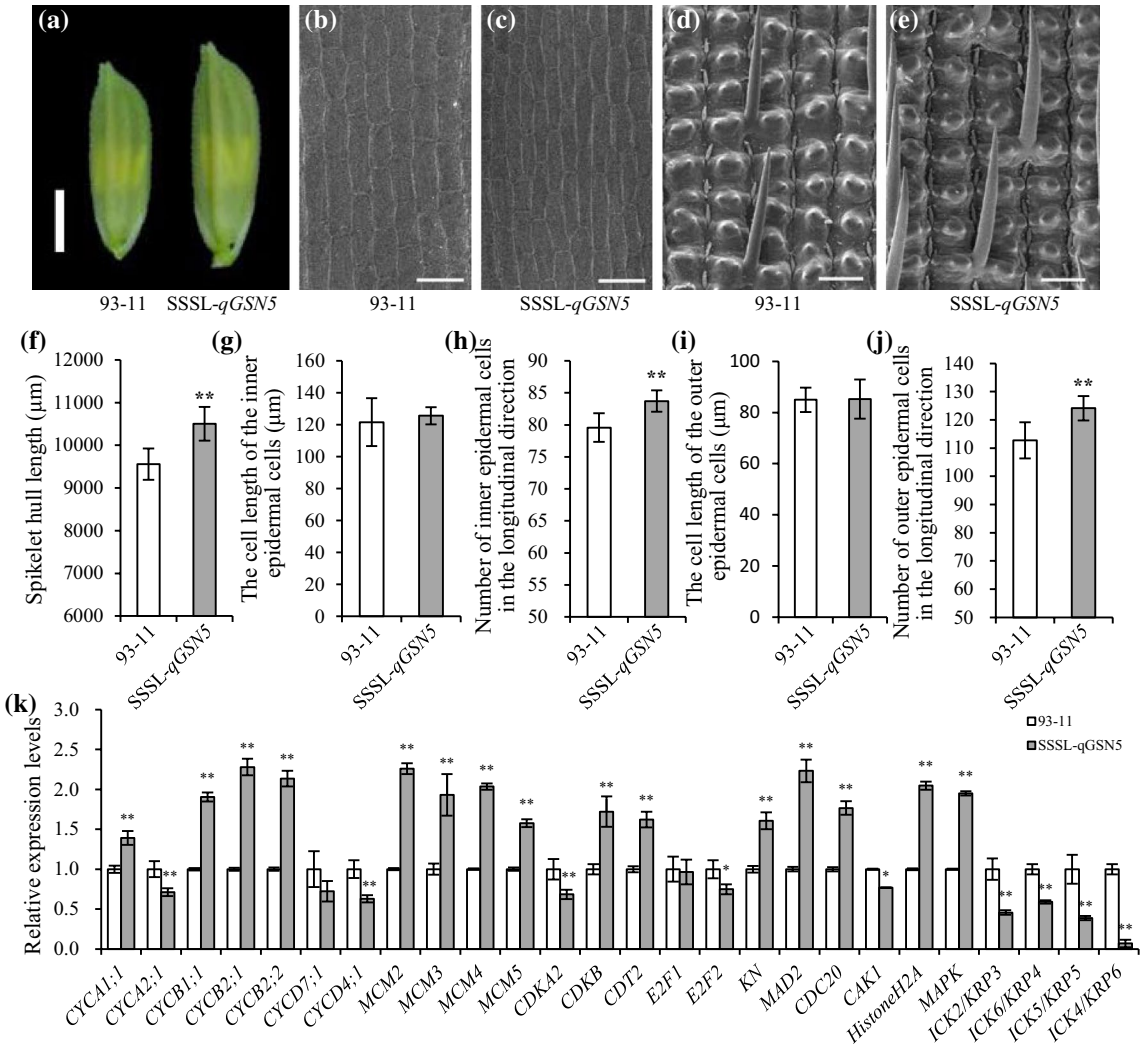


**Fig. 2** *qGSN5* regulates spikelet hull development by affecting cell proliferation. **a** Spikelet hulls of 93-11 and SSSL-*qGSN5* before flowering. Scale bar, 3 mm. **b–e** SEM analysis of the inner (b and c) and outer (d and e) epidermal cells of the lemma. Scale bar, 100 μm. **f** Spikelet hull length of 93-11 and SSSL-*qGSN5* before flowering ($n=30$). **g–h** Statistical analysis of the length (g) and number (h) of the inner epidermal cells ($n=30$). **i–j** Statistical analysis of the length (i) and number (j) of the outer epidermal cells ($n=30$). **k** Comparison of the expression levels of cell cycle-related genes between 93-11 and SSSL-*qGSN5* in young panicles. The values of genes in 93-11 were set to one. * indicates $P<0.05$ and ** indicates $P<0.01$ (Student's *t* test)

two InDels, in the 2-kb promoter region of *Os05g47510* between the two parents (Fig. 4c). We speculated that these polymorphisms, especially the 72-bp deletion in the promoter of 93-11, would lead to decreased promoter activity in 93-11. As expected, dual-luciferase transient expression assays in rice protoplasts showed that the Nip-type promoter of *Os05g47510* had significantly higher relative activity than the 93-11-type (Fig. 4d, e). Taken together, we consider *Os05g47510*, in which there are large differences in the promoter region and coding region sequences, to be the best candidate gene for *qGSN5*, although we cannot rule out the possibility that other genes in this mapping region may be the causal genes.

## The effect of *qGSN5* was inhibited by *GS3*

*GS3* is a major QTL negatively regulating grain length in rice (Fan et al., 2006; Mao et al., 2010; Sun et al., 2018). Compared with Nip, 93-11 had a C to A substitution in the second exon of *GS3*, leading to a premature termination codon (TGA) in 93-11 (Fig. S7a). Thus, 93-11 has a non-functional *GS3* gene and produces long grains (Fig. S1b). SSSL-*GS3*, an SSSL harboring the functional *GS3* gene from Nip (Fig. 5a), was identified in our CSSL population. As expected, SSSL-*GS3* showed significantly reduced grain length compared with 93-11, and we confirmed that *GS3* was responsible for the phenotype of SSSL-*GS3* by performing co-segregation analysis (Fig. S7b). To study

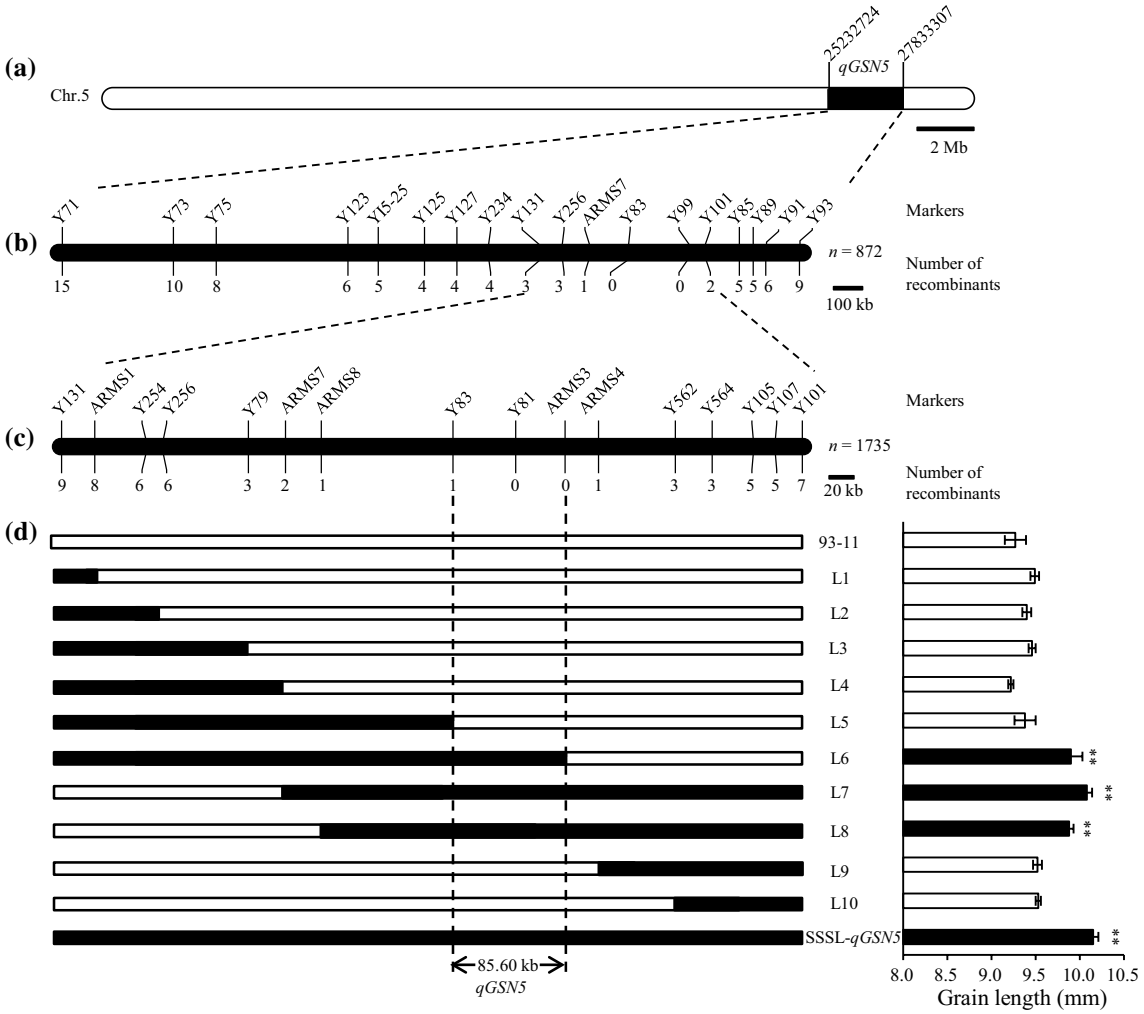


**Fig. 3** Fine mapping of *qGSN5* **a** Diagram of the genomic region of *qGSN5* genotyped using the Rice 9 K SNP array. White bar indicates the 93-11 background, and black rectangle indicates the substitution segment from Nip. **b** Primary mapping of *qGSN5* using the 93-11/CSSL28 $BC_6F_2$ population ($n = 872$). **c** Fine mapping of *qGSN5* using the 93-11/SSSL-*qGSN5* $BC_7F_2$ population ($n = 1735$). Numbers of recombinants are shown under the bar. **d** Fine mapping of *qGSN5* based on the phenotype of homozygous recombinants. White and black rectangles indicate the homozygous 93-11 genotype and homozygous *qGSN5* genotype, respectively. Data are shown as means ± SD ($n = 15$ plants). ** indicates $P < 0.01$ (Student's *t* test).

the genetic relationship between *qGSN5* and *GS3* in grain size regulation, we crossed SSSL-*GS3* with SSSL-*qGSN5*, and obtained a DSSL, DSSL-*GS3/qGSN5*, using MAS (Fig. 5a, Fig. S7c).

The grain lengths of the recurrent parent 93-11 (genotype *gs3*/*qgsn5*), SSSL-*qGSN5* (genotype *gs3*/*qGSN5*), SSSL-*GS3* (genotype *GS3*/*qgsn5*), and DSSL-*GS3/qGSN5* (genotype *GS3*/*qGSN5*), which have different allele combinations of *qGSN5* and *GS3*, were investigated (Fig. 5b, c). By comparing SSSL-*qGSN5* (*gs3*/*qGSN5*) with 93-11 (*gs3*/*qgsn5*), we determined that *qGSN5* increased the grain length by 10.98% in the non-functional *GS3* background; however, when comparing SSSL-*GS3* (*GS3*/*qgsn5*) with DSSL-*GS3/qGSN5* (*GS3*/*qGSN5*), which has a functional *GS3* background, *qGSN5* only increased the grain length by 2.28%. When comparing SSSL-*GS3* (*GS3*/*qgsn5*), which has a non-functional *qGSN5* allele, with 93-11 (*gs3*/*qgsn5*), the loss of *GS3* increased the grain length by 11.59%; however, when comparing DSSL-*GS3/qGSN5* (*GS3*/*qGSN5*), which has a functional *qGSN5* allele, with SSSL-*qGSN5* (*gs3*/*qGSN5*), the loss of *GS3* increased the grain length by 21.08%. These results indicated that the effect of *qGSN5* is significantly inhibited by functional *GS3*. Consistent with this, two-way ANOVA for grain length and genotype showed that the *qGSN5* × *GS3* interaction effect was significant ($P = 9.56 \times 10^{-6}$) (Table S4). Similar results were also observed for 1000-grain weight (Fig. 5d, Table S4). Taken together, our results suggest that *GS3* may be a suppressor of *qGSN5* (Fig. 5e).

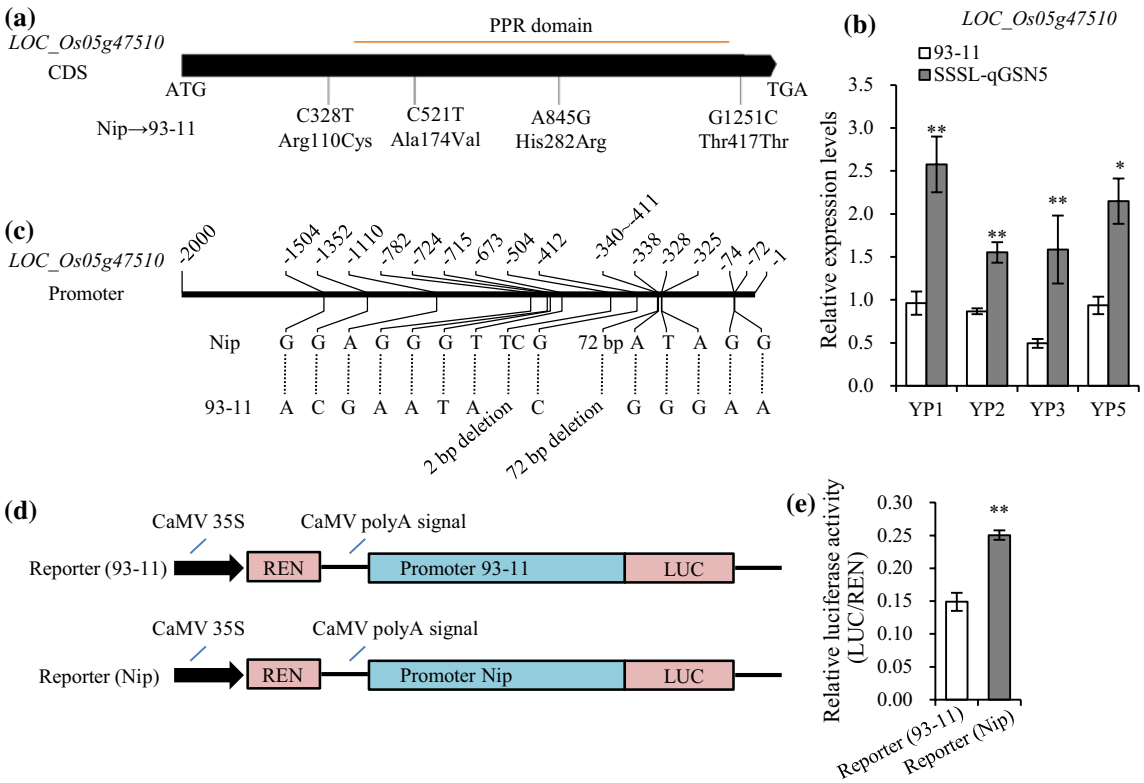


**Fig. 4** *LOC_Os05g47510* is the most likely candidate gene for *qGSN5*. **a** Diagram of the polymorphisms in the coding region of *LOC_Os05g47510* between the two parents. The orange line indicates the PPR domain. **b** Comparison of the expression level of *LOC_Os05g47510* in young panicles of 93-11 and SSSL-*qGSN5*. YP1, YP2, YP3, and YP5 indicate young panicles that are 1, 2, 3, and 5 cm in length, respectively. ** indicates $P < 0.01$ (Student's *t* test). **c** Diagram of the polymorphisms in the promoter region (2 kb) of *LOC_Os05g47510* between the two parents. **d** Schematic diagram of two reporter constructs for dual-luciferase transient assays of *LOC_Os05g47510*. **e** Relative luciferase activity (LUC/REN) was measured in rice protoplasts. ** indicates $P < 0.01$ (Student's *t* test)

### *qGSN5* could improve the grain shape of hybrid rice

Considering that 93-11, the recurrent parent of SSSL-*qGSN5*, is a famous restorer line in China, and that the large-grain phenotype of *qGSN5* is semi-dominant (Table 1), we wondered what the effect of *qGSN5* would be in hybrid rice. Therefore, we crossed 93-11 and SSSL-*qGSN5* with two male sterile lines, Y58S and Lu56S, to generate hybrid progeny. Compared with Y58S/93-11 and Lu56S/93-11, the grain width of Y58S/SSSL-*qGSN5* and Lu56S/SSSL-*qGSN5* showed no obvious difference (Fig. 6a); however, the grain length significantly increased by 6.56% and 6.20% (Fig. 6b), respectively, and the resulting 1000-grain weight increased by 14.23% and 15.65%, respectively (Fig. 6c). Consistent with the performances of SSSL-*qGSN5* and 93-11/SSSL-*qGSN5* ($BC_7F_1$) (Table 1), we also observed significant decreases in grain number per panicle in the Y58S/SSSL-*qGSN5* (-10.61%) and Lu56S/SSSL-*qGSN5* (-9.08%) plants (Fig. 6d). There were no significant differences in other agronomic traits, namely days to heading, plant height, panicle length, number of tillers per plant, and seed setting rate (Fig. 6e–i). Also, there were no significant differences in the final grain yield per plant because the increase in 1000-grain weight was balanced by the decrease in grain number (Fig. 6j). Moreover, we found that the percentage of grains with chalkiness and the degree of chalkiness in Y58S/SSSL-*qGSN5* and Lu56S/SSSL-*qGSN5* were not different compared with those in Y58S/93-11 and Lu56S/93-11 (Fig. 6k, l), indicating that the large increase in grain size and 1000-grain weight in the hybrid rice did not affect grain quality.

## Discussion

### *qGSN5* is a novel QTL coordinating grain size and grain number in rice

In plants, there is usually a negative correlation between grain size and grain number (Sadras 2007; Fan et al., 2019), and *GSN1* coordinates the trade-off between grain number and grain size in rice (Guo et al., 2018). In this study, SSSL-*qGSN5*, which showed significantly increased grain size and decreased grain number (Table 1), was identified in our 93-11/Nip CSSL population. Consistent with the phenotypes

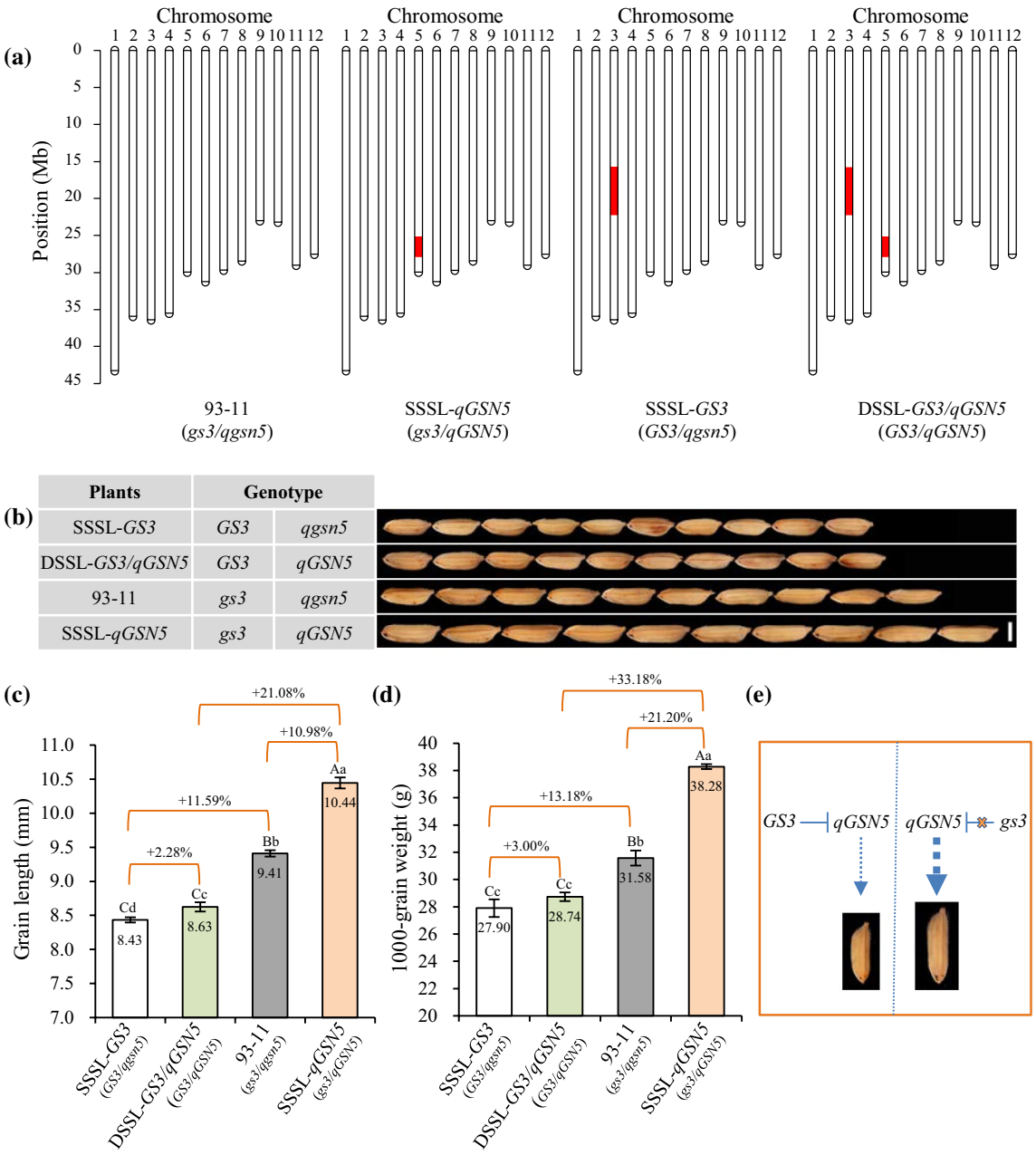


**Fig. 5** *GS3* may be a suppressor of *qGSN5*.**a** Schematic diagram of substitution segments in different plants identified by the Rice 9 K SNP array. The white bars indicate the 93-11 background, and red rectangles indicate substitution segments from Nip. **b** Comparison of grain length in 93-11, SSSL-*qGSN5*, SSSL-*GS3*, and DSSL-*GS3/qGSN5*. Scale bar, 3 mm. **c–d** Statistical analysis of the grain length **c** and 1000-grain weight **d** in 93-11, SSSL-*qGSN5*, SSSL-*GS3*, and DSSL-*GS3/qGSN5*. Multiple comparisons were performed. Lowercase and capital letters indicate differences at the 0.05 and 0.01 levels, respectively. **e** A proposed model of the genetic relationship between *GS3* and *qGSN5*

of SSSL-*qGSN5*, the hybrid rice lines Y58S/SSSL-*qGSN5* and Lu56S/SSSL-*qGSN5* also produced larger and fewer grains (Fig. 6a–d), indicating that *qGSN5* is a QTL coordinating grain size and grain number in rice. Rice 9 K SNP array analysis showed that SSSL-*qGSN5* harbored a 2.6-Mb substitution segment from Nip on the long arm of chromosome 5 (25.23–27.83 Mb) (Fig. 0.1a), which was responsible for the phenotypes of SSSL-*qGSN5*. Furthermore, through map-based cloning and overlapping substitution segment analysis, *qGSN5* was finally delimited to an 85.60-kb region flanked by markers Y83 and ARMS3 (Fig. 3). To our knowledge, there is only one recently reported grain size-related gene, *CLG1* (*Os05g47780*), located in the substitution segment of SSSL-*qGSN5*, but it is not located in the final fine

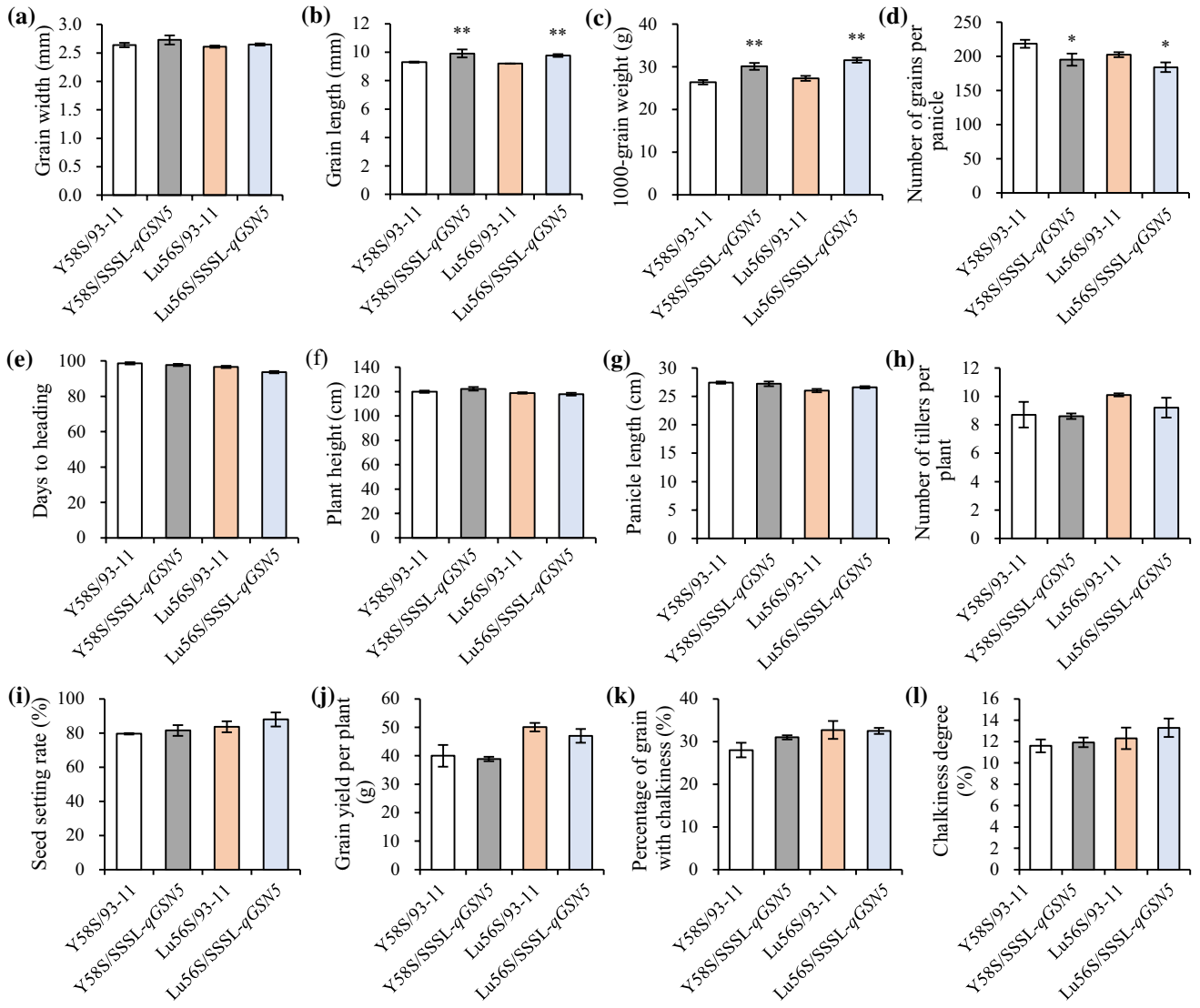


**Fig. 6** *qGSN5* has potential value in long-grain hybrid rice breeding. Comparison of agronomic traits between Y58S/93-11 and Y58S/SSSL-*qGSN5* and between Lu56S/93-11 and Lu56S/SSSL-*qGSN5*. **a** Grain width. **b** Grain length. **c** 1000-grain weight. **d** Number of grains per panicle. **e** Days to heading. **f** Plant height. **g** Panicle length. **h** Number of tillers per plant. **i** Seed setting rate. **j** Grain yield per plant. **k** Percentage of grains with chalkiness. **l** Chalkiness degree. Data are shown as means ± SD ($n = 15$ plants). * indicates $P < 0.05$ and ** indicates $P < 0.01$ (Student's *t* test)

mapping interval. Moreover, the homozygous recombinant L9, carrying the *CLG1* allele from Nip, showed no difference in grain size compared with 93-11 (Fig. 3d). Consistent with our result, Yang et al. also identified a CSSL carrying the *CLG1* allele from Nip in the 93-11 background and found that it showed no obvious difference compared with 93-11 (Yang et al., 2021). Thus, *CLG1* is not the causal gene of *qGSN5*, and *qGSN5* is a novel QTL for grain size and grain number in rice.

Through expression pattern, sequence, and quantitative PCR analysis, four genes (*Os05g47510*, *Os05g47530*, *Os05g47540*, and *Os05g47550*) were prioritized as candidate genes, none of which have been cloned in rice. Among them, *Os05g47510*, which encodes a PPR protein with eight canonical P-type PPR repeats (Fig. 4a), aroused our interest. Firstly, the sequence of the promoter region and coding region of *Os05g47510* and its expression level in young panicles were significantly different between 93-11 and SSSL-*qGSN5* (Fig. 4). In addition, many PPR proteins have been cloned in rice and shown to be involved in diverse biological processes (Kim et al., 2009; Huang et al., 2015; Xiao et al., 2018; Qiu et al., 2021; Zheng et al., 2021), including regulation of grain size. For example, *OGR1* encodes a PPR-DYW protein, and the *ogr1* mutant showed decreased grain width and thickness (Kim et al., 2009); *FLO10* and OsPPR5 encode canonical P-type PPR proteins, and both the *flo10*

and *Osppr5* mutants exhibited significantly decreased grain length, grain width, and grain thickness (Wu et al., 2019; Zhang et al., 2021). Thus, we suggest that *Os05g47510* is the most likely candidate gene for *qGSN5*, although we cannot rule out other candidate genes, and further transgenic complementation experiments need to be performed.

### *GS3* may be a suppressor of *qGSN5*

Many grain size-related genes have been cloned in rice and shown to be involved in several signaling pathways (Fan et al., 2019; Li et al., 2019); however, it is difficult to study the genetic relationship between these cloned genes because of differences in genetic background. SSSLs are ideal materials for studying genetic relationships between genes because they effectively minimize the interference of genetic background. In this study, the SSSLs SSSL-*qGSN5* and SSSL-*GS3* with the same 93-11 genetic background were identified, and a pyramiding line, DSSL-*GS3/qGSN5*, was generated to analyze the genetic relationship between *qGSN5* and *GS3* in the regulation of grain size (Fig. 5a). As shown in Fig. 5c, the functional *GS3* allele from Nip had a negative effect on grain length (-10.41%, SSSL-*GS3* vs 93-11), and the *qGSN5* allele from Nip had a positive effect on grain length (+ 10.98%, SSSL-*qGSN5* vs 93-11); however, the positive effect of *qGSN5* was significantly inhibited by *GS3*; DSSL-*GS3/qGSN5* showed a grain length similar to that of SSSL-*GS3* (-8.30%, DSSL-*GS3/qGSN5* vs 93-11). Thus, according to our results, we speculate that *GS3* may be a suppressor of *qGSN5*, and loss of function of *GS3* leads to a synergistic increase in grain length (Fig. 5e). Moreover, our results highlight the fact that (1) the small-grain parent (Nip) has hidden large-grain genes, which may be masked by the major QTL *GS3*, and (2) utilization of *qGSN5* depends on the presence of a non-functional allele of *GS3*.

### *qGSN5* has potential value in long-grain hybrid rice breeding

93-11 is an elite *indica* restorer line in China, and the genetic background of SSSL-*qGSN5*, which harbors only one substitution segment derived from Nip, is highly similar to that of 93-11 (Fig. 1a). Thus, SSSL-*qGSN5* has the potential to be a new restorer line; in particular, the large-grain phenotype of SSSL-*qGSN5* is semi-dominant (Table 1), and thus, this line can be utilized in hybrid rice breeding. To test this hypothesis, we crossed 93-11 and SSSL-*qGSN5* with two male sterile lines, Y58S and Lu56S. Compared with Y58S/93-11 and Lu56S/93-11, Y58S/SSSL-*qGSN5* and Lu56S/SSSL-*qGSN5* showed significantly increased grain size and 1000-grain weight (Fig. 6b, c). However, there were no differences in the final grain yield because there was a decrease in grain number (Fig. 6d, j). In spite of this, *qGSN5* may play a potentially important role in breeding long-grain hybrid rice, as grain shape is an important quality trait affecting the market value of rice grain; for example, consumers in Southern China generally prefer long grains (Huang et al., 2013). In addition, the grain number of hybrid rice carrying *qGSN5* may be improved by crossing with sterile lines producing a large number of grains.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00122-021-03951-7.

**Acknowledgements** This work was supported by grants from the Major Research program of the National Natural Science Foundation of China (92535301), the National Natural Science Foundation of China (32001489), the Open project of Sichuan Provincial Key Laboratory of Crop Resources and Genetic Improvement, and the Outstanding Young Scientific and Technological Talents Project in Sichuan Province (2020JDJQ0040).

**Author contribution statement** HY and SL conceived and designed all the experiments. HY, PG, and XH performed most of the experiments and contributed equally to this work. MY, ZX, MJ, WS, and SZ participated in construction of plant materials, agronomic trait investigation, and primary mapping. XZ, BT, TL, and YW performed cytological analysis and fine mapping. BM, PQ, and WC were involved in candidate gene analysis and exploration of breeding utilization. HY wrote the manuscript, and all authors approved the final manuscript.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309:741–745

Bai X, Huang Y, Hu Y, Liu H, Zhang B, Smaczniak C, Hu G, Han Z, Xing Y (2017) Duplication of an upstream silencer of FZP increases grain yield in rice. Nat Plants 3:885–893

Che R, Tong H, Shi B, Liu Y, Fang S, Liu D, Xiao Y, Hu B, Liu L, Wang H (2015) Control of grain size and rice yield by GL2-mediated brassinosteroid responses. Nat Plants 2:15195

Duan P, Rao Y, Zeng D, Yang Y, Xu R, Zhang B, Dong G, Qian Q, Li Y (2014) SMALL GRAIN 1, which encodes a mitogen-activated protein kinase kinase 4, influences grain size in rice. Plant J 77:547–557

Duan P, Ni S, Wang J, Zhang B, Xu R, Wang Y, Chen H, Zhu X, Li Y (2015) Regulation of OsGRF4 by OsmiR396 controls grain size and yield in rice. Nat Plants 2:15203

Duan P, Xu J, Zeng D, Zhang B, Geng M, Zhang G, Huang K, Huang L, Xu R, Ge S, Qian Q, Li Y (2017) Natural variation in the promoter of gse5 contributes to grain size diversity in rice. Mol Plant 10:685–694

Fan Y, Li Y (2019) Molecular, cellular and Yin-Yang regulation of grain size and number in rice. Mol Breed 39:163

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (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. Theor Appl Genet 112:1164–1171

Guo T, Chen K, Dong N-Q, Shi C-L, Ye W-W, Gao J-P, Shan J-X, Lin H-X (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

Hu J, Wang Y, Fang Y, Zeng L, Xu J, Yu H, Shi Z, Pan J, Zhang D, Kang S, Zhu L, Dong G, Guo L, Zeng D, Zhang G, Xie L, Xiong G, Li J, Qian Q (2015) A rare allele of GS2 enhances grain size and grain yield in rice. Mol Plant 8:1455–1465

Huang R, Jiang L, Zheng J, Wang T, Wang H, Huang Y, Hong Z (2013) Genetic bases of rice grain shape: so many genes, so little known. Trends Plant Sci 18:218–226

Huang W, Yu C, Hu J, Wang L, Dan Z, Zhou W, He C, Zeng Y, Yao G, Qi J, Zhang Z, Zhu R, Chen X, Zhu Y (2015) Pentatricopeptide-repeat family protein RF6 functions with hexokinase 6 to rescue rice cytoplasmic male sterility. Proc Natl Acad Sci U S A 112:14984–14989

Huang K, Wang D, Duan P, Zhang B, Xu R, Li N, Li Y (2017) WIDE AND THICK GRAIN 1, which encodes an otubain-like protease with deubiquitination activity, influences grain size and shape in rice. Plant J 91:849–860

Huo X, Wu S, Zhu Z, Liu F, Fu Y, Cai H, Sun X, Gu P, Xie D, Tan L, Sun C (2017) NOG1 increases grain production in rice. Nat Commun 8:1497

Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N, Onodera H, Kashiwagi T, Ujiie K, Shimizu B, Onishi A, Miyagawa H, Katoh E (2013) Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. Nat Genet 45:707–711

Kim SR, Yang JI, Moon S, Ryu CH, An K, Kim KM, Yim J, An G (2009) Rice OGR1 encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. Plant J 59:738–749

Li S, Zhao B, Yuan D, Duan M, Qian Q, Tang L, Wang B, Liu X, Zhang J, Wang J, Sun J, Liu Z, Feng YQ, Yuan L, Li C (2013) Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression. Proc Natl Acad Sci U S A 110:3167–3172

Li N, Xu R, Duan PG, Li YH (2018) Control of grain size in rice. Plant Reprod 31:237–251

Li N, Xu R, Li Y (2019) Molecular networks of seed size control in plants. Annu Rev Plant Biol 70:435–463

Li G, Tang J, Zheng J, Chu C (2021a) Exploration of rice yield potential: Decoding agronomic and physiological traits. Crop J 9:577–589

Li G, Zhang H, Li J, Zhang Z, Li Z (2021b) Genetic control of panicle architecture in rice. Crop J 9:590–597

Liu L, Tong H, Xiao Y, Che R, Xu F, Hu B, Liang C, Chu J, Li J, Chu C (2015a) Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. Proc Natl Acad Sci U S A 112:11102–11107

Liu S, Hua L, Dong S, Chen H, Zhu X, Jiang J, Zhang F, Li Y, Fang X, Chen F (2015b) OsMAPK6, a mitogen-activated protein kinase, influences rice grain size and biomass production. Plant J 84:672–681

Liu J, Chen J, Zheng X, Wu F, Lin Q, Heng Y, Tian P, Cheng Z, Yu X, Zhou K, Zhang X, Guo X, Wang J, Wang H, Wan J (2017) GW5 acts in the brassinosteroid signalling pathway to regulate grain width and weight in rice. Nat Plants 3:17043

Luo J, Liu H, Zhou T, Gu B, Huang X, Shangguan Y, Zhu J, Li Y, Zhao Y, Wang Y, Zhao Q, Wang A, Wang Z, Sang T, Wang Z, Han B (2013) An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. Plant Cell 25:3360–3376

Mao H, Sun S, Yao J, Wang C, Yu S, Xu C, Li X, Zhang Q (2010) Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc Natl Acad Sci U S A 107:19579–19584

Qin P, Lu H, Du H, Wang H, Chen W, Chen Z, He Q, Ou S, Zhang H, Li X, Li X, Li Y, Liao Y, Gao Q, Tu B, Yuan H, Ma B, Wang Y, Qian Y, Fan S, Li W, Wang J, He M, Yin J, Li T, Jiang N, Chen X, Liang C, Li S (2021). Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. Cell 184:3542–3558 e3516

Qiu T, Zhao X, Feng H, Qi L, Yang J, Peng YL, Zhao W (2021) OsNBL3, a mitochondrion-localized pentatricopeptide repeat protein, is involved in splicing nad5 intron 4 and its disruption causes lesion mimic phenotype with enhanced resistance to biotic and abiotic stresses. Plant Biotechnol J. https://doi.org/10.1111/pbi.13659

Sadras VO (2007) Evolutionary aspects of the trade-off between seed size and number in crops. Field Crops Res 100:125–138

Seo H, Kim SH, Lee BD, Lim JH, Lee SJ, An G, Paek NC (2020) The rice basic helix-loop-helix 79 (OsbHLH079) determines leaf angle and grain shape. Int J Mol Sci 21:2090

Shi C, Ren Y, Liu L, Wang F, Zhang H, Tian P, Pan T, Wang Y, Jing R, Liu T, Wu F, Lin Q, Lei C, Zhang X, Zhu S, Guo X, Wang J, Zhao Z, Wang J, Zhai H, Cheng Z, Wan J (2019) Ubiquitin specific protease 15 has an important role in regulating grain width and size in rice. Plant Physiol 180:381–391

Si L, Chen J, Huang X, Gong H, Luo J, Hou Q, Zhou T, Lu T, Zhu J, Shangguan Y, Chen E, Gong C, Zhao Q, Jing Y, Zhao Y, Li Y, Cui L, Fan D, Lu Y, Weng Q, Wang Y, Zhan Q, Liu K, Wei X, An K, An G, Han B (2016) OsSPL13 controls grain size in cultivated rice. Nat Genet 48:447–456

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

Su S, Hong J, Chen X, Zhang C, Chen M, Luo Z, Chang S, Bai S, Liang W, Liu Q, Zhang D (2021) Gibberellins orchestrate panicle architecture mediated by DELLA–KNOX signalling in rice. Plant Biotechnol J. https://doi.org/10.1111/pbi.13661

Sun S, Wang L, Mao H, Shao L, Li X, Xiao J, Ouyang Y, Zhang Q (2018) A G-protein pathway determines grain size in rice. Nat Commun 9:851

Tang QY, Zhang CX (2013) Data Processing System (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research. Insect Sci 20:254–260

Torres Acosta JA, Fowke LC, Wang H (2011) Analyses of phylogeny, evolution, conserved sequences and genome-wide expression of the ICK/KRP family of plant CDK inhibitors. Ann Bot 107:1141–1157

Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, Zhang G, Fu X (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet 44:950–954

Wang S, Wu K, Qian Q, Liu Q, Li Q, Pan Y, Ye Y, Liu X, Wang J, Zhang J, Li S, Wu Y, Fu X (2017) Non-canonical regulation of SPL transcription factors by a human OTUB1-like deubiquitinase defines a new plant type rice associated with higher grain yield. Cell Res 27:1142–1156

Wu Y, Wang Y, Mi XF, Shan JX, Li XM, Xu JL, Lin HX (2016) The QTL GNP1 encodes GA20ox1, which increases grain number and yield by increasing cytokinin activity in rice panicle meristems. PLoS Genet 12:e1006386

Wu M, Ren Y, Cai M, Wang Y, Zhu S, Zhu J, Hao Y, Teng X, Zhu X, Jing R, Zhang H, Zhong M, Wang Y, Lei C, Zhang X, Guo X, Cheng Z, Lin Q, Wang J, Jiang L, Bao Y, Wang Y, Wan J (2019) Rice FLOURY ENDOSPERM10 encodes a pentatricopeptide repeat protein that is essential for the trans-splicing of mitochondrial nad1 intron 1 and endosperm development. New Phytol 223:736–750

Xiao HJ, Xu YH, Ni CZ, Zhang QN, Zhong FY, Huang JS, Liu W, Peng LL, Zhu YG, Hu J (2018) A rice dual-localized pentatricopeptide repeat protein is involved in organellar RNA editing together with OsMORFs. J Exp Bot 69:2923–2936

Xiao Y, Liu D, Zhang G, Gao S, Liu L, Xu F, Che R, Wang Y, Tong H, Chu C (2019) Big Grain3, encoding a purine permease, regulates grain size via modulating cytokinin transport in rice. J Integr Plant Biol 61:581–597

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

Xu R, Duan P, Yu H, Zhou Z, Zhang B, Wang R, Li J, Zhang G, Zhuang S, Lyu J, Li N, Chai T, Tian Z, Yao S, Li Y (2018a) Control of grain size and weight by the OsMKKK10-OsMKK4-OsMAPK6 signaling pathway in rice. Mol Plant 11:860–873

Xu R, Yu H, Wang J, Duan P, Zhang B, Li J, Li Y, Xu J, Lyu J, Li N, Chai T, Li Y (2018b) A mitogen-activated protein kinase phosphatase influences grain size and weight in rice. Plant J 95:937–946

Yang XM, Ren YL, Cai Y, Niu M, Feng ZM, Jing RN, Mou CL, Liu X, Xiao LJ, Zhang X, Wu FQ, Guo XP, Jiang L, Wan JM (2018). Overexpression of OsbHLH107, a member of the basic helix-loop-helix transcription factor family, enhances grain size in rice (Oryza sativa L.). Rice 11:41

Yang W, Wu K, Wang B, Liu H, Guo S, Guo X, Luo W, Sun S, Ouyang Y, Fu X, Chong K, Zhang Q, Xu Y (2021) The RING E3 ligase CLG1 targets GS3 for degradation via the endosome pathway to determine grain size in rice. Mol Plant. https://doi.org/10.1016/j.molp.2021.1006.1027

Yin W, Xiao Y, Niu M, Meng W, Li L, Zhang X, Liu D, Zhang G, Qian Y, Sun Z, Huang R, Wang S, Liu CM, Chu C, Tong H (2020) ARGONAUTE2 enhances grain length and salt tolerance by activating big grain3 to modulate cytokinin distribution in rice. Plant Cell 32:2292–2306

Yuan H, Fan S, Huang J, Zhan S, Wang S, Gao P, Chen W, Tu B, Ma B, Wang Y, Qin P, Li S (2017) 08SG2/OsBAK1 regulates grain size and number, and functions differently in Indica and Japonica backgrounds in rice. Rice 10:25

Yuan H, Qin P, Hu L, Zhan S, Wang S, Gao P, Li J, Jin M, Xu Z, Gao Q, Du A, Tu B, Chen W, Ma B, Wang Y, Li S (2019) OsSPL18 controls grain weight and grain number in rice. J Genet Genomics 46:41–51

Zhang L, Qi Y, Wu M, Zhao L, Zhao Z, Lei C, Hao Y, Yu X, Sun Y, Zhang X, Guo X, Ren Y, Wan J (2021) Mitochondrion-targeted PENTATRICOPEPTIDE REPEAT5 is required for cis-splicing of nad4 intron 3 and endosperm development in rice. Crop J 9:282–296

Zheng P, Liu Y, Liu X, Huang Y, Sun F, Wang W, Chen H, Jan M, Zhang C, Yuan Y, Tan BC, Du H, Tu J (2021) OsPPR939, a nad5 splicing factor, is essential for plant growth and pollen development in rice. Theor Appl Genet 134:923–940

Zuo J, Li J (2014) Molecular genetic dissection of quantitative trait loci regulating rice grain size. Annu Rev Genet 48:99–118