Theor Appl Genet (2011) 123:1173–1181
DOI 10.1007/s00122-011-1657-x

ORIGINAL PAPER

Seed size is determined by the combinations of the genes controlling different seed characteristics in rice

Song Yan · Guihua Zou · Sujuan Li · Hua Wang · Heqin Liu · Guowei Zhai · Peng Guo · Hongmiao Song · Changjie Yan · Yuezhi Tao

Received: 16 February 2011 / Accepted: 9 July 2011 / Published online: 31 July 2011

Abstract Rice seed size is an important agronomic trait in determining the yield potential, and four seed size related genes (*GS3*, *GW2*, *qSW5*/*GW5* and *GIF1*) have been cloned in rice so far. However, the relationship among these four genes is still unclear, which will impede the process of gene pyramiding breeding program to some extent. To shade light on the relationship of above four genes, gene expression analysis was performed with *GS3*-RNAi, *GW2*-RNAi lines and CSSL of *qSW5* at the transcriptional level. The results clearly showed that *qSW5* and *GW2* positively regulate the expression of *GS3*. Meanwhile, *qSW5* can be down-regulated by repression of *GW2* transcription. Additionally, *GIF1* expression was found to be positively regulated by *qSW5* but negatively by *GW2* and *GS3*. Moreover, the allelic effects of *qSW5* and *GS3* were detailedly characterized based on a natural population consisting of 180 rice cultivars. It was indicated that mutual interactions exist between the two genes, in which, *qSW5* affecting seed length is masked by *GS3* alleles, and *GS3* affecting seed width is masked by *qSW5* alleles. These findings provide more insights into the molecular mechanisms underlying seed size development in rice and are likely to be useful for improving rice grain yield.

Communicated by A. Paterson.

S. Yan and G. Zou contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1657-x) contains supplementary material, which is available to authorized users.

S. Yan · G. Zou · S. Li · H. Wang · H. Liu · G. Zhai · P. Guo · H. Song · Y. Tao (✉)
State Key Laboratory Breeding Base for Zhejiang Sustainable Pest and Disease Control, Institute of Crop and Nuclear Technology Utilization, Zhejiang Academy of Agricultural Sciences, 198 Shiqiao Road, Hangzhou 310021, People's Republic of China
e-mail: taoyz@zaas.org

S. Yan
Rice Research Institute, Jiangxi Academy of Agricultural Sciences, 602 Nanlian Road, Nanchang 330200, People's Republic of China

C. Yan
The Key Laboratory of Plant Functional Genomics, Ministry of Education of China; Jiangsu Key Laboratory of Crop Genetics and Physiology, Agricultural College of Yangzhou University, 12 East Wenhui Road, Jiangsu 225009, People's Republic of China

Introduction

Rice yield is determined by three key components: number of panicles per unit area, number of filled grains per panicle and seed weight. Seed or grain size, the major determinant of seed weight, is affected by four parameters: seed length, width, thickness and filling degree. Seed size, after long-time extensive studies, has been widely accepted as a complex trait controlled by polygenes, and numerous QTLs associated with seed or grain shape have been identified in rice (Lin et al. 1995; Lin and Wu 2003; Huang et al. 1997; Redona and Mackill 1998; Tan et al. 2000; Kubo et al. 2001; Xing et al. 2001; Yamagishi et al. 2002; Xu et al. 2002; Thomson et al. 2003; Yan et al. 2003; Li et al. 2004; Rabiei et al. 2004; Tian et al. 2005; Wan et al. 2006, 2008; Xie et al. 2006; Yoon et al. 2006; Lei et al. 2008; Bai et al. 2010).

With technological advances of functional genomics, great progress has been achieved in the clarification of the

molecular mechanisms of seed size in recent years. To date, four genes associated with seed size and weight have been identified and cloned. The first, *GS3*, cloned from a major QTL for grain length and weight, encodes a protein with four putative domains: a plant-specific organ size regulation (OSR) domain in the N terminus, a transmembrane domain, a tumor necrosis factor receptor/nerve growth factor receptor (TNFR/NGFR) family cysteine-rich domain, and a von Willebrand factor type C (VWFC) in the C terminus. These domains function differentially in grain size regulation (Fan et al. 2006; Takano-Kai et al. 2009; Mao et al. 2010). The second, *GW2*, controlling grain width and weight, encodes a RING-type protein with E3 ubiquitin ligase activity (Song et al. 2007). The third, *qSW5* or *GW5*, identified from another QTL controlling seed width and weight, encodes a nuclear polyubiquitin-binding protein (Shomura et al. 2008; Weng et al. 2008). The above three genes all function as negative regulators of grain size and organ size. The last, *GIF1*, encodes a cell-wall invertase required for carbon partitioning during early grain-filling. Over-expression of *GIF1* under its native promoter leads to large grains (Wang et al. 2008).

It is expected that these findings would facilitate the application of favorable alleles in marker assisted selection (MAS) breeding to increase yield potential in rice. However, there are two main problems which need to be addressed. Firstly, as phenotypic variation is always associated with allelic variation, it is necessary to identify the natural allelic variation to determine which one is the suitable allele in MAS programs. Secondly, due to the additive and epistatic effects of alleles across multiple genes, some combinations of favorable alleles result in enhanced traits, whereas other combinations do not (Benfey and Mitchell-Olds 2008). It is critical to discover the interactions among genes and consequently optimize gene pyramiding.

The previous studies showed that there are three haplotypes at the *qSW5* locus, i.e. Kasalath-type, *Indica* II-type and Nipponbare-type, and that a 1,212-bp deletion occurred in Nipponbare-type *qsw5*, compared with the Kasalath-type *qSW5*, contributing to the increase of seed width, and that the average of seed width of *Indica* II-type landraces was between those of Kasalath-type and Nipponbare-type landraces carriers (Shomura et al. 2008). Additionally, Takano-Kai et al. (2009) conducted an association study between *GS3* genotype and seed morphology of 157 accessions of *O. sativa*, and revealed that the A allele of *GS3* (Minghui-type *gs3*) masked the effects of other seed length associated genes and that it differentially interacted with seed width genes in different subpopulations. Therefore, it is obvious that the multiple alleles existed in rice germplasms and functionally differentiated. However, the exactly genetic effects of the alleles at the target gene locus, as well as the interactions among different target genes need to be characterized.

In order to gain an understanding of the relationship among the four rice seed size genes, *GS3*, *GW2*, *qSW5* and *GIF1*, the gene expressions were investigated based on the *GS3*-RNAi, *GW2*-RNAi lines and one chromosome segment substitution line (CSSL) of *qSW5* at the transcriptional level. Moreover, we detailedly characterized the allele effects at *qSW5* and *GS3* loci in a natural population. And the genetic interaction between *qSW5* and *GS3* loci was also clarified in the present study.

Materials and methods

Plant materials

To down-regulate the expression of *GS3*, RNAi-*GS3* vector was constructed using a p1301UbiNOS binary vector, as described in Shi et al. (2007). One 578-bp Nipponbare *GS3* fragment was PCR-amplified using the primer pair, GSRiBamF/GSRiSpeR (Fig. S1A, Table S2). The hairpin structure with two inverted repeat fragments was then cloned into the p1301UbiNOS binary vector. RNAi-*GW2* vector was also constructed using the p1301UbiNOS vector to suppress the expression of *GW2*. The RNAi fragment of 373-bp in length was amplified with the primer pair, GWRiBamF/GWRiSpeR (Fig. S1B; Table S2). The two constructs were introduced into Zhonghua 11 respectively, using the *Agrobacterium tumefaciens*-mediated method described in Hiei et al. (1994). The T_1 transgenic lines and the wild-type plants were planted in the rice-growing season in 2010.

A total of 180 rice cultivars including 96 *indica* and 84 *japonica* accessions collected from several countries (Table S1), and a set of CSSLs derived from Sasanishiki/Habataki (Ando et al. 2008), were used for this study. All the materials were planted in the rice-growing season of 2009 in an experimental field of Zhejiang Academy of Agricultural Sciences, Hangzhou, China. Each cultivar/line was planted in two rows, with eight plants per row, spaced at 15×25 cm^2. The field management was same as described by Yan et al. (2007).

Phenotypic measurements

The seed (grain with hull) length, seed width and seed weight were measured after the seeds were harvested and air-dried. In detail, thirty full seeds were randomly selected from each cultivar/line and divided into three groups equally. All seeds from each group were lined up lengthwise to measure the seed length and then arranged breadthwise to measure the seed width. Seed length and width were determined by averaging three measurements. Seed thickness was measured individually by using a vernier cal-

iper. Hundred-seed weight was determined and converted to 1,000-seed weight. For the measurement of panicle traits, three medium-sized panicles were obtained from each transgenic and wild-type plant. We measured the panicle length, the number of primary rachis branches, the number of secondary rachis branches, and the number of grains per panicle. Duncan or Dunnett test was performed to compare the means of seed traits for different allelic groups or cultivars/lines using SPSS 19.0 (SPSS Inc., IBM Company). *t* test or *F* statistic was estimated for each of allelic groups using Excel 2007 (Microsoft Corporation, Redmond, USA).

DNA extraction, PCR and genotyping

Genomic DNA was extracted from fresh leaves of each cultivar by using the CTAB method (Rogers and Bendch 1988). Two FNP (Functional Nucleotide Polymorphism) markers, GS3-*Pst* I and N1212del, kindly provided by Yan et al. (2009) and Shomura et al. (2008), were employed for genotyping *GS3* and *qSW5* alleles respectively. PCR was performed using the procedures described by Yan et al. (2009) and Shomura et al. (2008). PCR products or restriction enzyme-digested fragments were separated on 1% agarose gels and visualized with an AlphaImager (Alpha Innotech Corporation, San Leandro, CA).

Sequence analysis

PCR products amplified from four rice varieties Zhenshan 97, Habataki, 02428 and Sasanishiki, using the N1212del marker, were cloned into pMD19-T vector (Takara Biotech, Dalian) and sequenced by HuaDa Biotechnology Co. of Hangzhou. The sequence data were then analyzed using Vector NTI 9.0 software.

Expression analysis

Total RNA was extracted from panicles collected just before heading stage using RNAiso reagent according to the manufacturer's protocol (Takara). Total RNA (~1 μg) was treated with Promega RQ DNase (Promega, Madison, WI), and the first-strand cDNA was then synthesized by *TransScript* II Reverse Transcriptase (Beijing. TransGen Biotech Co. Ltd.). Subsequently, the first-strand cDNA was used as a template for semi-quantitative PCR analysis after normalized with rice *Actin1* gene using the marker Act1. The specific markers GSRT and GWRT were used to detect *GS3* and *GW2* expression levels, respectively. The PCR procedure is following: 94°C for 4 min, followed by 30 cycles of 94°C 20 s, 55°C for 30 s and 72°C for 50 s and an elongation step at 72°C for 5 min. The PCR products were analyzed on 1% agarose gels. Real-time PCR analysis was performed using a LightCycler (Roche). Amplification reactions were performed using Thunderbird SYBR qPCR Mix (TOYOBO), according to manufacturer's instructions. As control, the *Actin1* cDNA was amplified using the marker Act2. Each reaction was made in triplicate. The primer sequences were listed in Table S2.

Results

RNAi of *GS3* and *GW2*

To suppress the expression of *GS3* and *GW2*, two constructs, RNAi-*GS3* and RNAi-*GW2*, were independently transformed into rice cultivar Zhonghua 11. For RNAi-*GS3*, twenty-one independent transformants were obtained. The agronomic traits of four homozygous transgenic lines from T_1 generation were investigated, both of the seed length and weight of the transgenic lines were significantly larger than that of the wild-type ($P < 0.001$), however, the seed width was decreased when compared to that of the wild-type. The other agronomic traits were not altered obviously in *GS3*-RNAi lines (Table 1). The results further confirmed that *GS3* has a significantly negative effect on seed length and weight and a minor positive effect on seed width. For RNAi-*GW2*, sixteen independent transformants were obtained, and three independent homozygous transgenic lines of T_1 generation were used for agronomical trait measurement. The transgenic lines showed a significant increase in seed length, width, thickness and weight ($P < 0.001$), whereas the tillers, secondary rachis branches and grain number per panicle were all slightly reduced in *GW2*-RNAi lines compared with those of the wild-type (Table 1), indicating that *GW2* may have some effects on plant morphology, but significantly negative effects on seed traits.

SL416 contains *Indica* II type *qsw5*

The previous study showed that there were three allelic types at *qSW5* locus, including Kasalath-type, *Indica* II-type and Nipponbare-type (Shomura et al. 2008), of which, Kasalath-type allele is a functional allele, Nipponbare-type allele is a loss-of-function allele, but the genetic nature of *Indica* II type allele yet remains to be determined. In the present study, the *Indica* II type allele was detected in many *indica* varieties, including Zhenshan 97, Teqing, Habataki, etc. Sequencing results showed a 950-bp deletion occurred in *Indica* II type allele compared to the *qSW5* region of Kasalath (Fig. 1). Sasanishiki was found harboring Nipponbare-type *qsw5*. Hence, we screened four CSSLs, SL415, SL416, SL417 and SL418, covering rice chromosome 5, using the N1212del marker, and found SL416 has the same

Table 1 Agronomic performances of the T_1 transgenic lines compared to wild-type plants

	PH (cm)	NT	PL (cm)	1RB	2RB	NGP	SL (mm)	SW (mm)	ST (mm)	SWT (g)
WT	95.17 ± 1.76	11.00 ± 2.00	23.43 ± 0.64	11.33 ± 0.58	38.67 ± 5.51	192.00 ± 21.38	7.59 ± 0.09	3.41 ± 0.01	2.24 ± 0.04	26.70 ± 0.28
GS3-RNAi-L1	96.50 ± 3.12	10.67 ± 1.53	24.40 ± 0.46	12.67 ± 2.52	42.33 ± 7.02	201.00 ± 21.70	8.31 ± 0.01***	3.35 ± 0.01	2.24 ± 0.04	28.70 ± 0.14***
GS3-RNAi-L2	97.83 ± 2.75	11.33 ± 1.53	23.00 ± 0.50	11.00 ± 0.00	39.00 ± 3.61	197.33 ± 13.05	8.47 ± 0.07***	3.33 ± 0.03*	2.25 ± 0.03	27.60 ± 0.14**
GS3-RNAi-L3	95.17 ± 5.35	11.00 ± 2.00	24.27 ± 0.93	13.00 ± 0.00	38.60 ± 1.53	209.00 ± 8.00	8.38 ± 0.02***	3.30 ± 0.05*	2.22 ± 0.04	28.25 ± 0.07***
GS3-RNAi-L4	96.33 ± 5.03	11.00 ± 2.00	23.30 ± 0.36	11.67 ± 0.58	36.67 ± 2.89	182.00 ± 7.55	8.58 ± 0.09***	3.34 ± 0.01	2.27 ± 0.07	29.45 ± 0.07***
GW2-RNAi-L1	96.17 ± 2.84	9.50 ± 0.71	22.77 ± 0.32	12.00 ± 2.00	35.00 ± 2.00	172.67 ± 20.03	7.88 ± 0.02***	3.75 ± 0.05***	2.38 ± 0.03***	30.50 ± 0.36***
GW2-RNAi -L2	94.50 ± 1.80	9.33 ± 0.58	22.13 ± 0.99	12.33 ± 1.53	36.00 ± 2.65	189.67 ± 16.74	8.22 ± 0.06***	3.92 ± 0.08***	2.40 ± 0.06***	31.03 ± 0.06***
GW2-RNAi -L3	92.00 ± 2.65	9.50 ± 0.71	23.83 ± 0.47	10.33 ± 1.15	36.67 ± 4.93	177.33 ± 22.03	8.22 ± 0.08***	3.75 ± 0.05***	2.39 ± 0.05***	31.30 ± 0.36***

PH plant height, *NT* number of tillers, *PL* panicle length, *1RB* number of primary rachis branches per panicle, *2RB* number of secondary rachis branches per panicle, *NGP* number of grains per panicle, *SL* seed length, *SW* seed width, *ST* seed thickness, *SWT* 1,000-seed weight

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. The level of significance based on Dunnett test

allele with Habataki (Fig. 2). Therefore, the SL416 and Sasanishiki could be served as a set of nearly isogenic lines for the investigation of the genetic effect of *qSW5* gene. Interestingly, the seed width of SL416 (3.53 ± 0.05 mm) was much wider than that of Habataki (2.80 ± 0.02 mm), which was close to that of Sasanishiki (3.49 ± 0.02 mm) (Table 2), indicating that *qSW5* in Habataki is a loss-of-function allele. In addition, we compared Sasanishiki with SL416 on the other seed traits, and found that SL416 showed a little longer, wider and thicker seeds than those of Sasanishiki, but the difference did not reach the significant level ($P > 0.05$), whereas the seed weight of SL416 were significantly increased compared to Sasanishiki ($P < 0.05$) (Table 2). These results suggested that the *Indica* II type *qsw5* might has stronger effects of enhancing seed weight than Nipponbare-type *qsw5* in *japonica* background.

mRNA expression of seed size genes

Semi-quantitative RT-PCR results showed that the *GS3* and *GW2* expressions were totally suppressed in the above analyzed *GS3*-RNAi and *GW2*-RNAi lines, respectively (Fig. 3a, e). The transcription levels of *qSW5*, *GW2*, *GS3* and *GIF1* were then examined by quantitative RT-PCR in the following three groups: wild-type and *GS3*-RNAi line 2, wild-type and *GW2*-RNAi line 1, Sasanishiki and SL416. Panicle samples from plants before heading stage were checked. The result showed that no significantly different expressions of *qSW5* and *GW2* were found between *GS3*-RNAi line 2 and its recipient (Fig. 3b, c), but a much higher expression level of *GIF1* was observed as compared with Zhonghua 11 (Fig. 3d). In *GW2*-RNAi line 1, a slight decrease in *qSW5* expression (Fig. 3f), an increase with about 2.5-folds in *GIF1* expression (Fig. 3g), and a significant decrease in *GS3* expression in relative to Zhonghua 11 (Fig. 3h) were detected. As for SL416, a significant increase in the expression of *qSW5* and *GS3* (Fig. 3i, j), and a slight increase in *GIF1* expression compared to Sasanishiki (Fig. 3k) were found, however, the expression of *GW2* did not change between the two lines (Fig. 3l). These results suggested that *GW2* may work upstream of *qSW5*, and that *GW2* and *qSW5* function to positively regulate the expression of *GS3*, and that *GIF1* may be negatively regulated by *GW2* and *GS3*, and be positively regulated by *qSW5*.

Interaction between *qSW5* and *GS3* and their effects on seed traits

In the present study, gene expression analysis clearly showed a complex relationship among the seed size-related genes at the transcriptional level. To determine how the interactions affect seed traits, we analyzed the gene effects

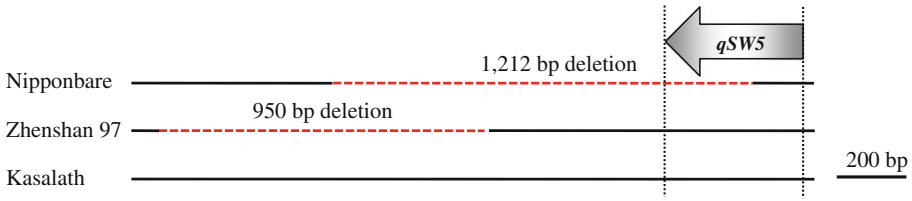

Fig. 1 Polymorphisms among Nipponbare, Zhenshan 97 and Kasalath at the *qSW5* locus. Nipponbare has a 1,212-bp deletion (*dashed line*), and Zhenshan 97 has a 950-bp deletion compared to Kasalath. Zhenshan 97 harbored *qSW5* represents the *Indica* II-type allele

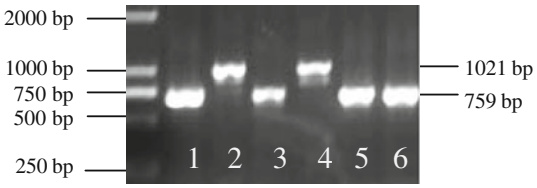

Fig. 2 PCR assay of *qSW5* alleles in *1* Sasanishiki, *2* Habataki, *3* SL415, *4* SL416, *5* SL417 and *6* SL418 using the N1212del marker

on seed traits in a natural population consisted of 180 rice cultivars, including 96 *indica* and 84 *japonica* accessions from several countries. The seed width ranged from 2.26 to 3.5 mm in *indica* subspecies, and from 2.56 to 4.14 mm in 84 *japonica* subspecies. The seed length ranged from 7.32 to 11.05 mm in *indica* subspecies, and from 5.83 to 10.55 mm in *japonica* subspecies (Table S1). The wide variations indicated the dramatic diversity and complexity of the heredity of seed shapes.

We investigated the allelic distribution of three key seed size genes, *qSW5*, *GS3* and *GW2*, in 180 cultivars by using the FNP markers, N1212del, GW2-*Hap* I and GS3-*Pst* I, respectively. Three allelic types, Kasalath-type *qSW5*, *Indica*-II type *qsw5* and Nipponbare-type *qsw5*, were identified at the *qSW5* locus (Fig. S2A); two allelic types, the Minghui-type *gs3* and Zhenshan-type *GS3*, were identified at the *GS3* locus (Fig. S2B), while no allelic variations were detected at the *GW2* locus in our survey as reported by other groups previously (Takano-Kai et al. 2009; Yan et al. 2009). Therefore, we analyzed the genetic effects of *qSW5* and *GS3*.

For the *qSW5* locus, 180 cultivars could be divided into three allelic groups, of which, 24 (20 *indica* and 4 *japonica*) belonged to Kasalath-type *qSW5* group, 76 (75 *indica* and 1 *japonica*) belonged to *Indica* II-type *qsw5* group, and 80 (1 *indica* and 79 *japonica*) belonged to Nipponbare-type *qsw5* group (Table S1). This suggests that the Kasalath-type *qSW5* and the *Indica* II-type *qsw5* are originated from *indica* germplasm and the Nipponbare-type *qsw5* from *japonica* germplasm, which is consistent with the findings of Shomura et al. (2008). Among the three allelic groups, significant differences were observed in averages of both seed width and seed length ($P < 0.01$). The Nipponbare-type *qsw5* group had the shortest and widest seeds, followed by *Indica* II-type *qsw5* group (Fig. 4a, b). As for the seed weight, the average of the Kasalath-type allelic group was significantly lighter than the ones of the other two groups ($P < 0.05$), but no significant difference was found between the two allelic groups of the Nipponbare-type and the *Indica* II-type (Fig. 4c).

As for the *GS3* locus, 180 cultivars could be divided into two allelic groups, of which, 34 (26 *indica* and 8 *japonica*) belonged to Minghui-type *gs3* group and 146 (70 *indica* and 76 *japonica*), belonged to Zhenshan-type *GS3* group (Table S2; Fig. S2B). Minghui-type *gs3* group had significantly longer, thinner and heavier seeds than Zhenshan-type *GS3* group ($P < 0.001$) (Fig. 4d, e, f). The significant differences were also found in either *indica* or *japonica* cultivars for average seed length, width and weight between two *GS3*/*gs3* groups (Table S3).

Furthermore, we analyzed the *qSW5* effects in different *GS3* alleles background, the significant differences in seed length were found at 5% level, except that the Nipponbare-type *qsw5* group had significantly shorter seeds than the other two *qSW5*/*qsw5* groups in Zhenshan-type *GS3* allelic cultivars at 1% level (Fig. 4a). Significant differences in

Table 2 Comparison of means and variances for seed morphology among Habataki, Sasanishiki and SL416 (means ± SD)

	SL (mm)	SW (mm)	ST (mm)	SWT (g)
Habataki	7.75 ± 0.05 A	2.80 ± 0.02 A	1.79 ± 0.05 A	19.67 ± 0.21 A
Sasanishiki	7.44 ± 0.05 Ba	3.49 ± 0.02 Ba	2.20 ± 0.06 Ba	25.63 ± 0.55 Ba
SL416	7.46 ± 0.04 Ba	3.53 ± 0.05 Ba	2.24 ± 0.07 Ba	26.33 ± 0.12 Bb

Abbreviations are the same as in Table 1

A, B: ranked by Duncan test at $P < 0.01$

a, b: ranked by Duncan test at $P < 0.05$

Same letter within a column represents no significant difference

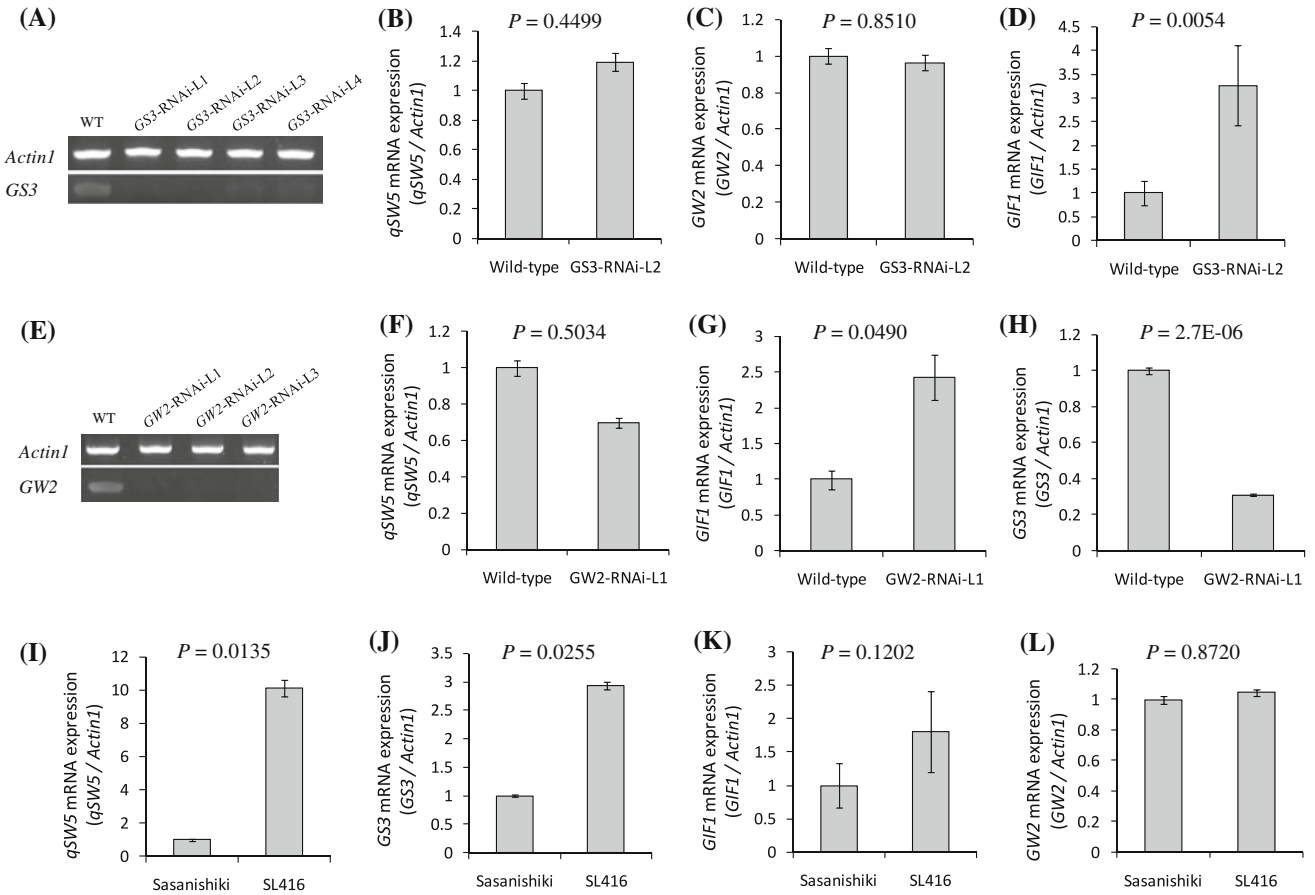

Fig. 3 Expression analysis of seed size related genes in panicles before heading stage. **a** Semi-quantitative RT-PCR detection of *GS3* expression levels in four RNAi lines compared to the wild-type. Real-time RT-PCR detection of *qSW5* (**b**), *GW2* (**c**) and *GIF1* (**d**) transcription levels in the wild-type and *GS3*-RNAi-L2. **e** Semi-quantitative RT-PCR detection of *GW2* expression levels in three RNAi lines compared to the wild-type. Real-time RT-PCR detection of *qSW5* (**f**), *GIF1* (**g**) and *GS3* (**h**) transcription levels in the wild-type and *GW2*-RNAi-L1. Real-time RT-PCR detection of *qSW5* (**i**), *GS3* (**j**), *GIF1* (**k**) and *GW2* (**l**) transcription levels in Sasanishiki and SL416. A Student's *t* test was used to generate the *P* values. The experiment was repeated three times with similar results

seed width among three *qSW5* allelic groups were found similar to the results from all cultivars ($P < 0.01$) (Fig. 4b). Similarly, the *GS3* effects on seed traits were also analyzed in three *qSW5* allelic groups, respectively. The most obvious change was that the differences in seed width almost disappeared (Fig. 4e). These findings clearly showed mutual interactions between *qSW5* and *GS3*, *qSW5* affecting seed length was masked by *GS3* alleles, and *GS3* affecting seed width was masked by *qSW5* alleles.

Association between seed traits and genotypes of *qSW5* and *GS3*

Based on the cross analysis of *qSW5* and *GS3* effects on seed traits, association between seed traits and different alleles of *qSW5* and *GS3* revealed that *qSW5* was significantly associated with seed length, width and weight, explaining 15 and 37% of seed length variation, 72 and 61% of seed width variation, 14 and 11% of seed weight variation in Minghui-type *gs3* and Zhenshan-type *GS3* allelic cultivars, respectively. *GS3* was significantly associated with seed length and seed weight, and explained 57, 63 and 54% of seed length variation, 39, 33 and 6% of seed weight variation in three *qSW5* allelic cultivars of Kasalath-type, *Indica* II-type and Nipponbare-type, respectively (Table S4). These results strongly indicated that *qSW5* and *GS3* play an essential role in determining the seed shape. In the rice high-yield breeding program, it is very important to select favorable alleles at *GS3* and *qSW5* loci to achieve the high-yield goal.

Discussion

Relationships among seed size genes

As seed size is one of the most important agronomic traits in rice, it has attracted wide attention in molecular genetics

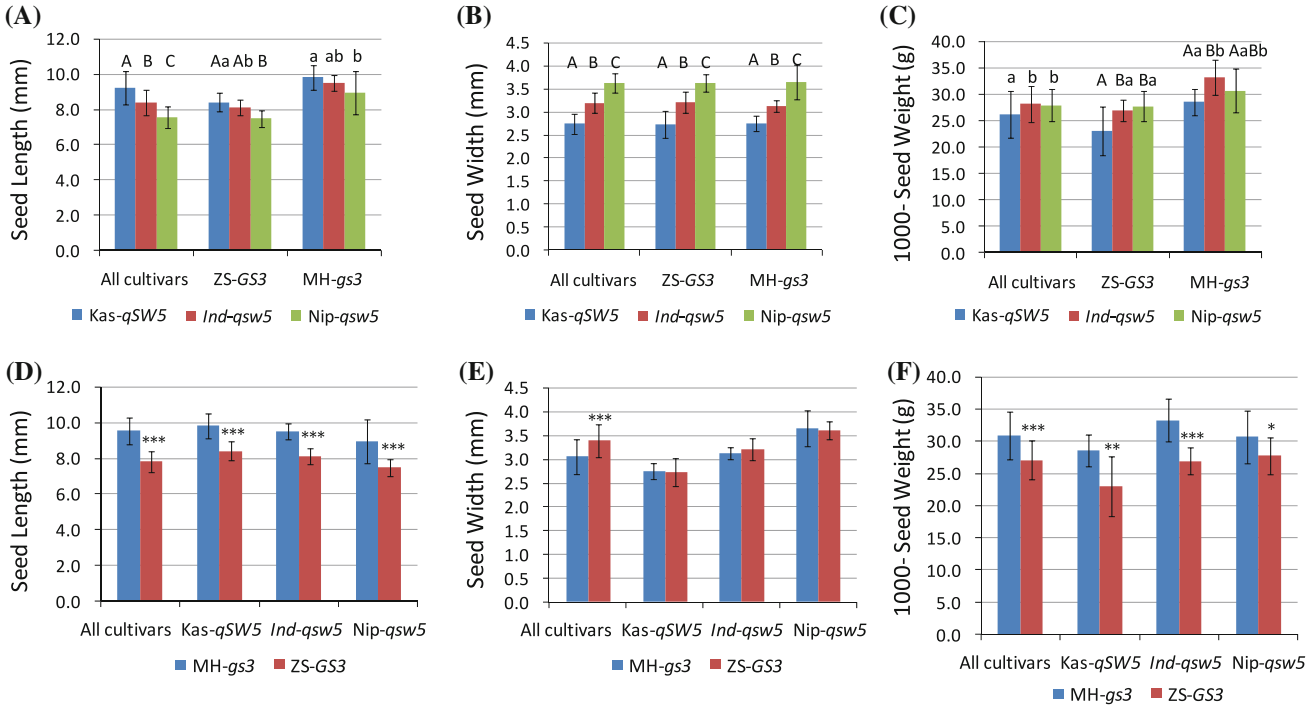

Fig. 4 Genetic effect analysis of *qSW5* and *GS3*. Mean comparison of seed length (**a**), seed width (**b**) and 1,000-seed weight (**c**) among three *qSW5* alleles in all 180 cultivars, Zhenshan type *GS3* (ZS-*GS3*) allelic cultivars and Minghui type *gs3* (MH-*gs3*) allelic cultivars, respectively. Mean comparison of seed length (**d**), seed width (**e**) and 1,000-seed weight (**f**) between two *GS3* alleles in all cultivars, Kasalath type *qSW5* (Kas-*qSW5*) allelic cultivars, *Indica* II type *qsw5* (*Ind-qsw5*) allelic cultivars and Nipponbare type *qsw5* (Nip-*qsw5*) allelic cultivars, respectively. All data are given as mean ± SD. A, B, C: ranked by Duncan test at $P < 0.01$; a, b: ranked by Duncan test at $P < 0.05$. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

study and breeding program. Up to the present, four key seed size genes have been isolated in rice, two of which, *qSW5/GW5* and *GW2* conferred both the seed or grain width and weight in rice (Song et al. 2007; Shomura et al. 2008; Weng et al. 2008). When the *GW2* expression was suppressed through RNA interference strategy, the seed length, width, thickness and weight were all highly significantly increased in the three independent homozygous RNAi transgenic lines compared with the wild type Zhonghua 11 which carried the Nipponbare-type *qsw5* (Table 1), suggesting that *gw2* may be interacted with or additive to *qsw5*. mRNA expression results suggest that *GW2* may positively regulate *qSW5*. In addition, there are some other factors between GW2 and qSW5/GW5, because GW2 was proved to have E3 ubiquitin ligase activity and GW5 may be involved in the ubiquitin–proteasome pathway (Song et al. 2007; Weng et al. 2008), but the functional relationship between the two proteins remains to be investigated.

In this study, we found that the expression of *GS3* was positively regulated by both *GW2* and *qSW5*, but the genetic effects of *GS3* and *qSW5* on seed size were mutually masked from each other. Previous studies revealed that *GS3* has a major effect on seed length and a minor effect on seed width, and that *qSW5* has a significant effect on seed width and a less significant effect on seed length (Fan et al. 2006; Shomura et al. 2008; Weng et al. 2008; Takano-Kai et al. 2009; Mao et al. 2010). This study showed similar results through *GS3*-RNAi lines and a natural population, however, when the effects of *GS3* and *qSW5* were cross analyzed, *qSW5* effects on seed length and *GS3* effects on seed width were all reduced sharply (Fig. 4a, e). Since the VWFC domain has been experimentally proven to be the most important one among the four domains of GS3 (Mao et al. 2010) and involved in protein–protein interactions (Colombatti et al. 1993; O'Leary et al. 2004; Zhang et al. 2007), it could be expected that further studies of the pattern of the VWFC domain of GS3 functioning with qSW5 will yield more insights into the molecular mechanisms of seed size development.

In addition, these three genes, *GS3*, *GW2*, *qSW5*, all have effects on the *GIF1* expression, co-determining the forming of seed size. *GS3*-RNAi, *GW2*-RNAi lines and SL416 all showed differences in seed thickness compared to their controls, respectively, suggesting that the three genes also are involved in seed filling.

Favorable alleles at seed size gene loci

In recent years, the molecular mechanisms of seed size are gradually being elucidated, and a set of FNP markers have

been developed for MAS in rice (Yan et al. 2009; Fan et al. 2009; Wang et al. 2011). In this study, we checked the distribution of *qSW5* and *GS3* alleles by using FNP markers in a natural population consisting of 180 rice cultivars. At the *GS3* locus, two type alleles, Minghui-type *gs3* and Zhenshan-type *GS3*, were detected. It is demonstrated by many groups that Minghui-type *gs3* contributed to significantly increasing seed length and weight (Fan et al. 2006; Takano-Kai et al. 2009; Mao et al. 2010). In this study, 26 *indica* and 8 *japonica* cultivars were found carrying Minghui-type *gs3* (Table S3), suggesting that the favorable *gs3* has been utilized in both *indica* and *japonica* cultivars.

At the *qSW5* locus, three type alleles, Kasalath-type, *Indica* II-type and Nipponbare-type, were detected. The Kasalath-type *qSW5* (20/24) and the *Indica* II-type *qsw5* (75/76) were mainly observed in *indica* rice, the Nipponbare-type *qsw5* (79/80) were mainly in *japonica* rice, showing the evidence of selection at the *qSW5* locus during rice domestication. The Nipponbare-type allelic cultivars had significantly wider seeds than the Kasalath-type allelic cultivars, and moreover, the only one *indica* cultivar, IR43, containing the Nipponbare-type *qsw5*, showed much wider seed breadth (3.06 mm) than that of Kasalath-type cultivars (2.74 mm) (Table S2), suggesting the availability of Nipponbare-type *qsw5* in *indica* breeding practice.

With regard to the *Indica* II-type *qsw5*, it is found to be another loss-of-function allele because the Habataki (seed width, 2.80 mm) fragment carrying the *Indica* II-type allele could not rescue the phenotype of Sasanishiki (seed width 3.49 mm) carrying the Nipponbare-type *qsw5* (Table 2). The *Indica* II-type allelic group of *indica* cultivars had significantly slender seeds than Nipponbare-type allelic group of *japonica* cultivars (Fig. 4b), but 2 *japonica* varieties, SL416 (Table 2) and 02428 (Table S1), harboring the *Indica* II-type *qsw5*, showed wider seeds than Nipponbare-type allelic controls, implying that the *Indica* II-type *qsw5* might be interacted with other factors in *indica* background and should be more useful than Nipponbare-type *qsw5* in *japonica* background.

Potential application of seed size gene pyramiding in yield improvement

The seed size is a complex quantitative trait in crops, genetic improvement of which by gene pyramiding would be difficult. The existing intricate and complicated interactions of seed size genes could be one main obstacle for gene pyramiding to improve seed appearance. In this study, we clarified the interactions among seed size genes, and found the combination of *Indica* II-type *qsw5* and Minghui-type *gs3* might be optimal utilization of the two genes in breeding. On the other hand, the genes pleiotropy may be another problem for gene pyramiding. Pleiotropy is one of the most common attributes of genes (Williams 1957; He and Zhang 2006), which refers to the observation that a single gene affects more than one trait simultaneously. It has been reported that all of the four cloned seed size genes have some degree of pleiotropy (Fan et al. 2006; Takano-Kai et al. 2009; Song et al. 2007; Shomura et al. 2008; Weng et al. 2008; Wang et al. 2008). The similar results are also observed in this study. The pleiotropic effects of these genes could cause dilemmas in gene pyramiding since the effects on different seed morphologies regulated by different genes are not always desirable for grain yield breeding. Nevertheless, there are two possible ways to solve the problems. One is to search the most beneficial combination of favorable alleles for improving rice yields through field experiments. Another way is to clarify the signaling pathways and networks that regulate the developmental processes of grains. A clearer understanding of the physiological and molecular mechanisms controlling grain growth and development would accelerate the process of the optimization of gene pyramiding for rice breeding.

Acknowledgments This study was financially supported by the Ministry of Science and Technology of China (Grant No. 2006AA10A102) and the Ministry of Agriculture of China (Grant No. 2009ZX08001-008B), the National Natural Science Foundation of China (Grant No. 31071385), and the Key Subject Construction Program of Zhejiang for Modern Agricultural Biotechnology and Crop Disease Control.

References

Ando T, Yamamoto T, Shimizu T, Ma XF, Shomura A, Takeuchi Y, Lin SY, Yano M (2008) Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. Theor Appl Genet 116:881–890

Bai X, Luo L, Yan W, Kovi MR, Zhan W, Xing Y (2010) Genetic dissection of rice grain shape using a recombinant inbred line population derived from two contrasting parents and fine mapping a pleiotropic quantitative trait locus *qGL7*. BMC Genet 26:11–16

Benfey PN, Mitchell-Olds T (2008) From genotype to phenotype: Systems biology meets natural variation. Science 320:495–497

Colombatti A, Bonaldo P, Doliana R (1993) Type A modules: interacting domains found in several non–fibrillar collagens and in other extracellular matrix proteins. Matrix 13:297–306

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

Fan C, Yu S, Wang C, Xing Y (2009) A causal C-A mutation in the second exon of *GS3* highly associated with rice grain length and validated as a functional marker. Theor Appl Genet 118:465–472

He X, Zhang J (2006) Toward a molecular understanding of pleiotropy. Genetics 173:1885–1891

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by Agrobacterium tumefication and sequence analysis of the boundaries of the T-DNA. Plant J 6:271–282

Huang N, Parco A, Mew T, Magpantay G, McCouch S, Guiderdoni E, Xu J, Subudhi P, Angeles ER, Khush GS (1997) RFLP mapping

of isozymes, RAPD and QTLs for grain shape, brown planthopper resistance in a doubled haploid rice population. Mol Breed 3:105–113

Kubo T, Takano-Kai N, Yoshimura A (2001) RFLP mapping of genes for long kernel and awn on chromosome 3 in rice. Rice Genet Newsl 18:26–28

Lei DY, Xie FM, Xu JL, Chen LY (2008) QTL mapping and epistasis analysis for grain shape and chalkiness degree of rice. Chinese J Rice Sci 22:255–260

Li JM, Thomson M, McCouch SR (2004) Fine mapping of a grain weight quantitative trait locus in the pericentromeric region of rice chromosome 3. Genetics 168:2187–2195

Lin LH, Wu WR (2003) Mapping of QTL underlying grain shape and grain weight in rice. Mol Plant Breed 1:337–342

Lin HX, Min SK, Xiong ZM, Qian HR, ZhuangJY LuJ, Huang N, Zheng KL (1995) RFLP mapping of QTLs for grain shape traits in *indica* rice (*Oryza sativa* L. subsp. *indica*). Scientia Agric Sin 28:1–7

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 USA 107:19579–19584

O'Leary JM, Hamilton JM, Deane CM, Valeyev NV, Sandell LJ, Downing AK (2004) Solution structure and dynamics of a prototypical chordin-like cysteine-rich repeat (von Willebrand Factor type C module) from collagen IIA. J Biol Chem 279:53857–53866

Rabiei B, Valizadeh M, Ghareyazie B, Moghaddam M, Ali AJ (2004) Identification of QTLs for rice grain size and shape of Iranian cultivars using SSR markers. Euphytica 137:325–332

Redona ED, Mackill DJ (1998) Quantitative trait locus analysis for rice panicle and grain characteristics. Theor Appl Genet 96:957–963

Rogers SO, Bendch AJ (1988) Extraction of DNA from plant tissues. Plant Mol Biol Man 6:1–10

Shi ZY, Wang J, Wan XS, Shen GZ, Wang XQ, Zhang JL (2007) Overexpression of rice *OsAGO7* gene induces upward curling of the leaf blade that enhanced erect–leaf habit. Planta 226:99–108

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40:1023–1028

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

Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B, Bustamante C, Yoshimura A, Doi K, McCouch S (2009) Evolutionary history of *GS3*, a gene conferring grain size in rice. Genetics 182:1323–1334

Tan YF, Xing YZ, Li JX, Yu SB, Xu CG, Zhang Q (2000) Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. Theor Appl Genet 101:823–829

Thomson MJ, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. Theor Appl Genet 107:479–493

Tian F, Li DJ, Fu Q, Zhu ZF, Fu YC, Wang XK, Sun CQ (2005) Construction of introgression lines carrying wild rice (O*ryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. Theor Appl Genet 112:570–580

Wan XY, Wan JM, Jiang L, Wang JK, Zhai HQ, Weng JF, Wang HL, Lei CL, Wang JL, Zhang X, Cheng ZJ, Guo XP (2006) QTL analysis for rice grain length and fine mapping of an identified QTL with stable and major effects. Theor Appl Genet 112:1258–1270

Wan XY, Weng JF, Zhai HQ, Wang JK, Lei CL, Liu XL, Guo T, Jiang L, Su N, Wan JM (2008) Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele *gw-5* in a recombination hotspot region on the chromosome 5. Genetics 179:2239–2252

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H, Ma H, Zhang G, He Z (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nat Genet 40:1370–1374

Wang C, Chen S, Yu S (2011) Functional markers developed from multiple loci in *GS3* for fine marker-assisted selection of grain length in rice. Theor Appl Genet 122:905–913

Weng JF, Gu SH, Wan XY, Gao H, Guo T, Su N, Lei CL, Zhang X, Cheng ZJ, Guo XP, Wang JL, Jiang L, Zhai HQ, Wan JM (2008) Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. Cell Res 18:1199–1209

Williams GC (1957) Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411

Xie XB, Song MH, Jin FX, Ahn SN, Suh JP, Hwang HG, McCouch SR (2006) Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived from a cross between *Oryza sativa* and *Oryza rufipogon*. Theor Appl Genet 113:885–894

Xing YZ, Tan YF, Xu CG, Hua JP, Sun XL (2001) Mapping quantitative trait loci for grain appearance traits of rice using a recombinant inbred line population. Acta Bot Sin 43:840–845

Xu JL, Xue QZ, Luo LJ, Li ZK (2002) Genetic dissection of grain weight and its related traits in rice (*Oryza sativa* L.). Chinese J Rice Sci 16:6–10

Yamagishi M, Takeuchi Y, Kono I, Yano M (2002) QTL analysis for panicle characteristics in temperate japonica rice. Euphytica 128:219–224

Yan CJ, Liang GH, Chen F, Li X, Yi CD, Tian S, Lu JF, Gu MH (2003) Mapping quantitative trait loci associated with rice grain shape based on an *Indica*/*Japonica* backcross population. Acta Genetica Sin 30:711–716

Yan CJ, Zhou JH, Yan S, Chen F, Yeboah M, Tang SZ, Liang GH, Gu MH (2007) Identification and characterization of a major QTL responsible for erect panicle trait in *japonica* rice (*Oryza sativa* L.). Theor Appl Genet 115:1093–1100

Yan CJ, Yan S, Yang YC, Zeng XH, Fang YW, Zeng SY, Tian CY, Sun YW, Tang SZ, Gu MH (2009) Development of gene-tagged markers for quantitative trait loci underlying rice yield components. Euphytica 169:215–226

Yoon DB, Kang KH, Kim HJ, Ju HG, Kwon SJ, Suh JP, Jeong QY, Ahn SN (2006) Mapping quantitative trait loci for yield components and morphological traits in an advanced backcross population between *Oryza grandiglumis* and the *O. sativa japonica* cultivar Hwaseongbyeo. Theor Appl Genet 112:1052–1062

Zhang JL, Huang Y, Qiu LY, Nickel J, Sebald W (2007) von Willebrand factor type C domain-containing proteins regulate bone morphogenetic protein signaling through different recognition mechanisms. J Biol Chem 282:20002–20014