Theor Appl Genet (2006) 112: 1164–1171
DOI 10.1007/s00122-006-0218-1

ORIGINAL PAPER

**Chuchuan Fan · Yongzhong Xing · Hailiang Mao
Tingting Lu · Bin Han · Caiguo Xu · Xianghua Li
Qifa Zhang**

GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein

Received: 16 December 2005 / Accepted: 6 January 2006 / Published online: 2 February 2006

Abstract The *GS3* locus located in the pericentromeric region of rice chromosome 3 has been frequently identified as a major QTL for both grain weight (a yield trait) and grain length (a quality trait) in the literature. Near isogenic lines of *GS3* were developed by successive crossing and backcrossing Minghui 63 (large grain) with Chuan 7 (small grain), using Minghui 63 as the recurrent parent. Analysis of a random subpopulation of 201 individuals from the BC_3F_2 progeny confirmed that the *GS3* locus explained 80–90% of the variation for grain weight and length in this population. In addition, this locus was resolved as a minor QTL for grain width and thickness. Using 1,384 individuals with recessive phenotype (large grain) from a total of 5,740 BC_3F_2 plants and 11 molecular markers based on sequence information, *GS3* was mapped to a DNA fragment approximately 7.9 kb in length. A full-length cDNA corresponding to the target region was identified, which provided complete sequence information for the *GS3* candidate. This gene consists of five exons and encodes 232 amino acids with a putative PEBP-like domain, a transmembrane region, a putative TNFR/NGFR family cysteine-rich domain and a VWFC module. Comparative sequencing analysis identified a nonsense mutation, shared among all the large-grain varieties tested in comparison with the small grain varieties, in the second exon of the putative *GS3* gene. This mutation causes a 178-aa truncation in the C-terminus of the predicted protein, suggesting that GS3 may function as a negative regulator for grain size. Cloning of such a gene provided the opportunity for fully characterizing the regulatory mechanism and related processes during grain development.

Communicated by Y. Xue

C. Fan · Y. Xing · H. Mao · C. Xu · X. Li · Q. Zhang (✉)
National Key Laboratory of Crop Genetic Improvement and
National Center of Plant Gene Research (Wuhan),
Huazhong Agricultural University, 430070 Wuhan, China
E-mail: qifazh@mail.hzau.edu.cn
Fax: +86-27-87287092

T. Lu · B. Han
National Center for Gene Research,
Shangai Institutes for Biological Sciences,
Chinese Academy of Sciences, Shanghai 200233, China

Introduction

Grain size is a major determinant of grain weight, one of the three components (number of panicles per plant, number of grains per panicle and grain weight) of grain yield. In breeding applications, grain size is usually evaluated by grain weight, which is positively correlated with several characters including grain length, grain width and grain thickness (Evans 1972; Xu et al. 2002).

Grain size is also a highly important quality trait in rice. Although the preference for rice grain characteristics varies with consumer groups, long and slender grain is generally preferred for *indica* rice by the majority of consumers in China, USA and most Asian countries (Unnevehr et al. 1992; Juliano and Villareal 1993). For example, a length/width ratio of 2.8 is adopted as an enforced threshold for a national quality standard of *indica* rice in China (http://www.knowledgebank.irri.org/regionalSites/china/10_CodesStandards/default.htm). Thus, understanding the genetic and molecular basis of grain size is extremely important for rice improvement programs.

Utilization of molecular markers has greatly facilitated the investigation of the genetic basis of complex quantitative traits. Many QTLs associated with rice grain size were identified in the last decade. Among them, a QTL with major effect on grain size was consistently detected around the centromeric region of chromosome 3 in numerous studies across different genetic backgrounds and environments (Huang et al. 1997; Yu et al. 1997; Redoña and Mackill 1998; Tan et al. 2000; Kubo et al. 2001; Xing et al. 2002; Thomson et al. 2003; Aluko et al. 2004), and was mapped to a region of 93.8-kb in length (Li et al. 2004a, b). Thus this gene, referred to as *GS3* hereafter, can be very useful for

the improvement of both yield and quality in rice breeding programs using molecular marker-assisted selection. Moreover these results also provided a target for molecular cloning and characterization of the grain size gene.

Using $F_{2:3}$ and recombinant inbred line (RIL) populations derived from a cross between Zhenshan 97 and Minghui 63 (both belong to *Oryza sativa* L. ssp. *indica*), the parents of the most widely cultivated hybrid in China, we detected the effect of the *GS3* locus repeatedly as a major QTL for grain weight (hence yield) when yield and yield component traits were analyzed (Yu et al. 1997; Li et al. 2000; Xing et al. 2002; Hua et al. 2002). This QTL explained approximately 20% of the phenotypic variation of grain weight in all the populations, with the allele from Minghui 63 having a significant positive effect on the trait. When grain size was analyzed as a component of appearance quality, it was shown that *GS3* corresponded to a major locus for grain length (Tan et al. 2000; Xing et al. 2001). This locus alone explained over 55% of the phenotypic variation of grain length, again the allele from Minghui 63 had a large contribution to the increase of grain length. No matter the trait was analyzed as grain weight or grain length, the dominance effect was negative indicating that small grain was dominant over large grain.

In order to fine map the *GS3* locus, Chuan 7, an *indica* cultivar with very small grain, was crossed and successively backcrossed with Minghui 63, which resulted in near isogenic lines (NILs) for *GS3* under the genetic background of Minghui 63. The objectives of this study were (1) to evaluate the effect of the *GS3* locus on grain size including grain length, width, thickness and weight under the near isogenic genetic background; (2) to fine map the *GS3* locus; (3) to identify the candidate of the *GS3* gene and (4) to determine allelic variations at the DNA nucleotide level between the genotypes differing in grain size, which may provide a starting point for functional characterization of the *GS3* gene.

Materials and methods

Genetic materials and field experiments

For developing NILs for the *GS3* locus, a cross was made between two *indica* cultivars, Minghui 63 (large grain) and Chuan 7 (small grain). The resulting F_1 plants were backcrossed with Minghui 63 as the male parent. Plants heterozygous for the target region in the progenies were selected by SSR markers and crossed to Minghui 63. Finally, BC_3F_1 plants were surveyed using 125 SSR markers evenly distributed on the 12 rice chromosomes. One plant (BC_3F_1-19) heterozygous for the target region and containing the least genetic background from Chuan 7 (about 20% of the markers were heterozygous) was selected to produce BC_3F_2.

A total of 5,941 BC_3F_2 plants derived from BC_3F_1-19 were planted in the rice-growing season of 2004 on the experimental farm of Huazhong Agricultural University, Wuhan, China, from which 201 plants were randomly selected for QTL analysis. For the remaining 5,740 BC_3F_2 plants, only those showing recessive phenotype (long grain) were selected for recombinant screening. Twenty BC_3F_3 plants for each of the 201 random BC_3F_2 individuals were planted in the 2005 rice-growing season. Field management followed essentially the normal agricultural practice as described by Fan et al. (2005).

Measurements of grain traits

Harvested paddy rice was air-dried and stored at room temperature for at least 3 months before testing. Fully filled grains were used for measuring grain length, width, thickness and weight. Ten randomly chosen grains from each plant were lined up length-wise along a vernier caliper to measure grain length, and then arranged by breadth to measure grain width. Grain thickness was determined for each grain individually using vernier caliper, and the values were averaged and used as the measurements for the plant. Grain weight was calculated based on 200 grains and converted to 1,000-grain weight for ease of comparison with other studies.

Molecular marker development

SSR markers in the *GS3* locus region were identified from the Gramene database (http://www.gramene.org/). The SSR primers of the RM-series were designed according to Temnykh et al. (2000, 2001), and those of the MRG-series were according to the rice genome sequences of Monsanto Company that were made available by McCouch et al. (2002). Eleven newly developed Indel or cleaved amplified polymorphic sequence (CAPS) markers were listed in Table 1 based on the publicly available rice genome sequences (http://www.rgp.dna.affrc.go.jp; http://www.rise.genomics.org.cn/).

Linkage and QTL analysis

Mapmaker/Exp 3.0 (Lincoln et al. 1992) was used for linkage analysis. The genotypes of the 201 individuals in the random subpopulation at *GS3* locus were determined based on BC_3F_3 progeny test. The Kosambi function was used to calculate genetic distance. QTL analysis was conducted using the program Mapmaker/QTL 1.1 at a threshold of LOD 3.0.

DNA sequence analysis

Four *indica* cultivars, two (Minghui 63 and H94) with long grains and two (Chuan 7 and Zhenshan 97) with short to medium grains (Fig. 1), were sequenced for the target genomic DNA region. DNA fragments that cover

Table 1 Indel and CAPS markers developed for fine mapping of the *GS3* locus

Marker	Marker type	BAC location	Product size in Nipponbare (bp)	Forward primer (5′–3′)	Reverse primer (5′–3′)	Restriction enzyme
GS06	Indel	OSJNBa0090P23	148	AGCAAAGCTGGAACGAAGAG	TAAATTACGCCGTGTCAACG	
GS09	Indel	OSJNBa0002D18	170	GCAACCAAGTCCACGCTAAT	TAGCCGAAGATCAGCCTCCT	
GS47	CAPS	OSJNBa0030J19	1299	GATTATTGGAGACGGGACGA	GACGGCATGACCACTCTTTT	*Hap*II
GS52	CAPS	OSJNBa0030J19	1400	AGCTTTGGTGTCGTTCTGCT	CCGACTTGGAGAGAATGGAA	*Bgl*I
GS56	CAPS	OSJNBa0030J19	1598	GCTGTGTTGTCCTTTGCTGA	CCAATAAACCCCACTGCAAC	*Bgl*I
GS61	CAPS	OSJNBa0030J19	1598	CTTTACAAAACCGGCGGTAA	TGAAGCGGACCTAGCATTTT	*Bcl*I
GS63	CAPS	OSJNBa0030J19	704	AAGAACGACTACGCGCATCT	CCATCGCTCTCTTTCCTCAG	*Hha*I
GS64	CAPS	OSJNBa0030J19	1154	CAACACCAGCAACGAACAAC	ACGAGGGATTATCAGCCATT	*Eco*RI, *Hap*II
GS65	CAPS	OSJNBa0030J19	772	CGGTATGCCAAGTTGAATGA	TTGCCGCAGTAAACAAGAAG	*Hha*I, *Hap*II
SF18	CAPS	OSJNBa0030J19	1245	CCTTCAGTAAGAGAGATGTG	AGTTGATGGTTTTGTGGGAT	*Bcl*I
SF19	CAPS	OSJNBa0030J19	1224	TCTGCTTGCGGTTATCTGTA	TTAGGTCCCTTTTCTCGTCC	*Sac*I

the target region from these cultivars were amplified with high fidelity using LA-Taq (TakaRa, Dalian, China). The PCR products were cloned into pGEM-T vector (Promega, USA) according to the manufacturer's specification. The T7 and SP6 universal primers and the BigDye Terminator Cycle Sequencing v3.1 (Applied Biosystems, Foster City, CA, USA) were used for sequencing. Sequence contigs were assembled using the computer program SEQUENCHER 4.1.2 (Gene Codes Corporation, Ann Arbor, MI, USA). Sequence alignment was conducted using the computer program Vector NTI 9 (InforMax^{TM} Corporation, USA)

Results

GS3 was a major QTL for grain length and weight and a minor QTL for grain width and thickness

The two parents, Minghui 63 and Chuan 7, showed highly significant differences in all the traits examined (Figs. 1, 2). The distributions of grain weight, grain length, grain width and grain thickness in the random BC_3F_2 subpopulation of 201 individuals are illustrated in Fig. 2. Grain weight showed a bimodal distribution, with 1,000-grain weight 20.5–21.5 g as the boundary. Grain length expressed a discontinuous distribution with a large and clear gap between 8.50 and 9.50 mm (Table 2). Grain length concurred completely with grain weight, such that long grains were heavier than short grains, or vice visa. Whereas, grain width and thickness showed more or less normal distributions. For simplicity, we will describe the grain size trait as grain length so that the large and heavier grains are referred to as long grain, and the opposite type as short grain.

In the BC_3F_3 families from 201 BC_3F_2 plants, grain length displayed three distinct phenotypic groups: 56 families with uniformly long grains, 61 families with uniformly short grains and 84 families with both long and short grains. The ratio of the three groups fit well to the expected ratio (1:2:1) of single locus Mendelian segregation ($\chi^2 = 5.67$, $P > 0.05$). Clearly, the three distinct

Fig. 1 Grains of six cultivars used in this work

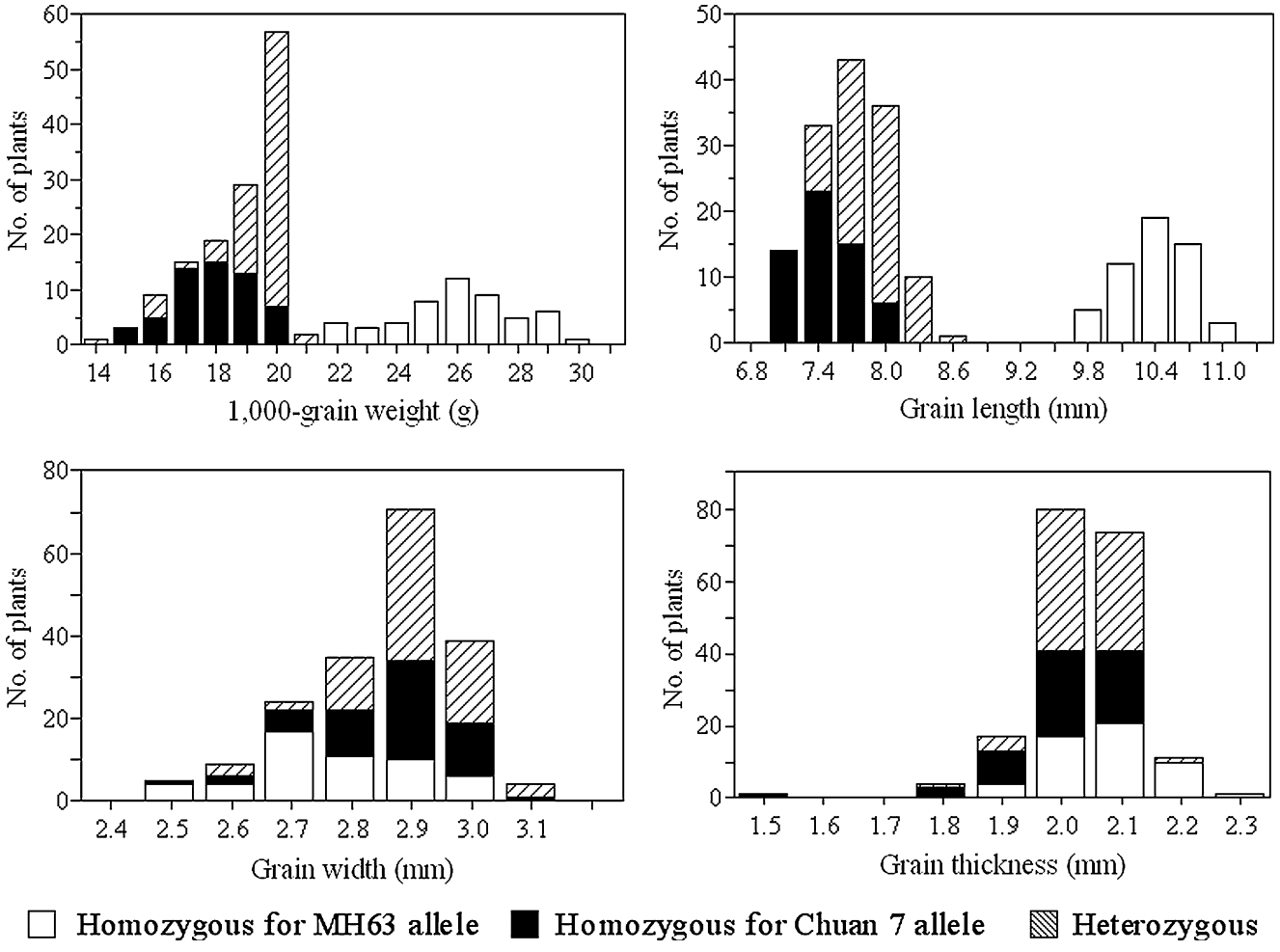

Fig. 2 Frequency distribution of 1,000-grain weight, grain length, grain width and grain thickness in the random BC_3F_2 subpopulation using Minghui 63 as the recurrent parent and Chuan 7 as the donor parent. Three genotype classes of *GS3* assessed by progeny tests are as indicated

Table 2 Descriptive statistics of the traits for parents and the random subpopulation derived from BC_3F_1-19 using Minghui 63 as the recurrent parent and Chuan 7 as the donor parent

Trait	Parent (mean ± SD)		MM		CC		MC	
	Minghui 63	Chuan 7	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
1,000-grain weight (g)	28.6 ± 0.6	12.5 ± 0.4	25.6 ± 2.0	21.5–29.8	17.5 ± 1.3	14.2–20.0	19.0 ± 1.2	14.0–20.5
Grain length (mm)	9.91 ± 0.099	6.30 ± 0.089	10.25 ± 0.29	9.64–10.73	7.32 ± 0.26	6.86–7.84	7.72 ± 0.25	7.24–8.50
Grain width (mm)	2.80 ± 0.03	2.48 ± 0.02	2.72 ± 0.13	2.43–2.96	2.82 ± 0.12	2.45–3.06	2.85 ± 0.09	2.56–3.04
Grain thickness (mm)	2.13 ± 0.06	1.68 ± 0.03	2.03 ± 0.09	1.81–2.21	1.95 ± 0.10	1.45–2.10	1.99 ± 0.06	1.79–2.13

Three genotype classes of *GS3*: *MM* homozygous for Minghui 63 allele, *CC* homozygous for Chuan 7 allele, *MC* heterozygous

phenotypic classes corresponded to the three genotypes of the BC_3F_2 individuals at the *GS3* locus: homozygote for the Minghui 63 allele (long grain), homozygote for the Chuan 7 alleles (short grain), and heterozygote. Using the three phenotypic classes as a marker, *GS3* was directly mapped into a 1-cM region delimitated by an Indel marker GS09 and SSR marker MRG5881 (Fig. 3).

QTL analysis indicated that the interval between GS09 and MRG5881 had effects simultaneously on grain weight, grain length, grain width and grain thickness (Table 3) contributing 83.4, 95.6, 19.8 and 12.1% of the phenotypic variation to these traits, respectively. The allele from Minghui 63 contributed to the increase of grain weight, grain length and grain thickness, but to the

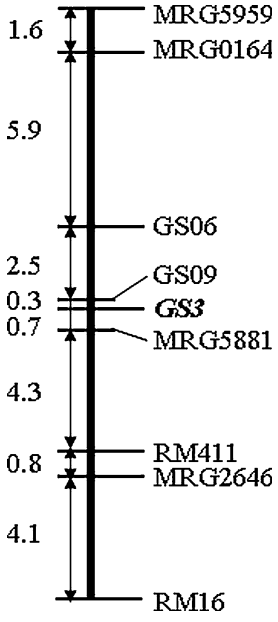

Fig. 3 Location of the *GS3* locus on the molecular linkage map of chromosome 3

decrease of grain width. Moreover, the QTL also showed different modes of gene actions on the traits, such that negative partial dominance was observed for grain weight and grain length, overdominance was detected for grain width, while dominance effect was not significant for grain thickness.

Delimitation of *GS3* to a 7.9-kb DNA fragment

To further narrow down the *GS3* containing genomic region, 1,384 plants with recessive phenotype (long grain, 9.7 mm or longer) from the BC_3F_2 population of 5,740 individuals were selected for recombinant screening. All the 1,384 selected plants were genotyped using an SSR marker MRG5881 and Indel marker GS09, which identified a total of 55 recombinants on both sides. These plants were further confirmed by progeny test for their recessive phenotype.

To obtain more markers in the GS09-MRG5881 interval, nine polymorphic CAPS markers between Minghui 63 and Chuan 7 were developed according to the sequence information of OSJNBa0030J19 (a BAC clone of cv. Nipponbare) and the contig Ctg009226 (from 93-11) encompassing the *GS3* locus. Using these CAPS markers to screen the 55 recombinants, five recombination events were resolved between GS47 and *GS3*, four identified between GS52 and *GS3*, four between GS56 and *GS3*, three between GS61 and *GS3*, and two between GS65 and *GS3* (Fig. 4a). In particular, the assay revealed one recombination event between GS63 and *GS3* and two recombination events between SF19 and *GS3*. In addition, GS64 and SF18 were found to co-segregate with the *GS3* locus (Fig. 4a). Therefore, the genomic region containing the *GS3* locus was narrowed down to the DNA fragment bounded by GS63 and SF19 (Fig. 4a), approximately 7.9-kb in length.

Table 3 Effects of the QTL (in the interval GS09-MRG5881) on grain shape and weight detected in a random subpopulation derived from BC_3F_1-19 using Minghui 63 as the recurrent parent and Chuan 7 as the donor parent

Traits	LOD	A[a]	D[b]	Var. %[c]
1,000-grain weight (g)	72.8	−4.08[d]	−2.52[d]	83.4
Grain length (mm)	129.2	−1.47[d]	−1.06[d]	95.6
Grain width (mm)	8.9	0.05[d]	0.08[d]	19.8
Grain thickness (mm)	5.3	−0.04[d]	0.004[e]	12.1

[a]Additive effect
[b]Dominance effect
[c]Percentage of total phenotypic variance explained by the QTL
[d]Significant at $P < 0.0001$ by t test
[e]Not significant at $P < 0.05$

GS3 encoded a putative transmembrane protein

A full-length complimentary DNA (cDNA), osigcea013f09t3, from the plumule of an *indica* cultivar Guangluai 4 was identified that matched well with the region between positions 1.6 and 7.3 kb of the 7.9-kb fragment (Fig. 4b). Allowing for the regulatory regions on both ends, this is considered as the only candidate gene for *GS3*. Thus the organization of putative *GS3* gene could be well defined by comparing the sequences of osigcea013f09t3 with the genomic DNA sequence of Chuan 7 generated in the study.

The *GS3* gene (GenBank accession number: DQ355996) had five exons with a transcript length of 956 bp encoding 232 amino acids (aa) (Fig. 4b). Analysis of the predicted sequence of GS3 protein revealed several known regions and domains (http://www.ebi.ac.uk/InterProScan/) (Figs. 4c, 5). There is a PEBP (phosphatidylethanolamine-binding protein)-like domain of 54 aa at the N-terminus. A transmembrane region is located at aa 97–117. The region of aa 116–155 contains a putative TNFR (tumor necrosis factor receptor)/NGFR (nerve growth factor receptor) family cysteine-rich domain, in which about 40 residues contain six conserved cysteines all involved in intrachain disulfide bonds (http://www.kr.expasy.org/cgi-bin/nicedoc.pl?PDOC00561). The C-terminal cysteine-rich region is similar to the von Willebrand factor type C (VWFC) modules, which are typically 60–80 aa in length and are mainly defined by a consensus sequence of ten cysteines (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00093), with a conserved glycine and an aromatic residue between the first pair of cysteine residues, as well as highly conserved motifs $C_2XXC_3XC_4$ and $C_8C_9XXC_{10}$ ('X' indicates any aa) located in the middle and at the C-terminal end of the repeat (O'Leary et al. 2004).

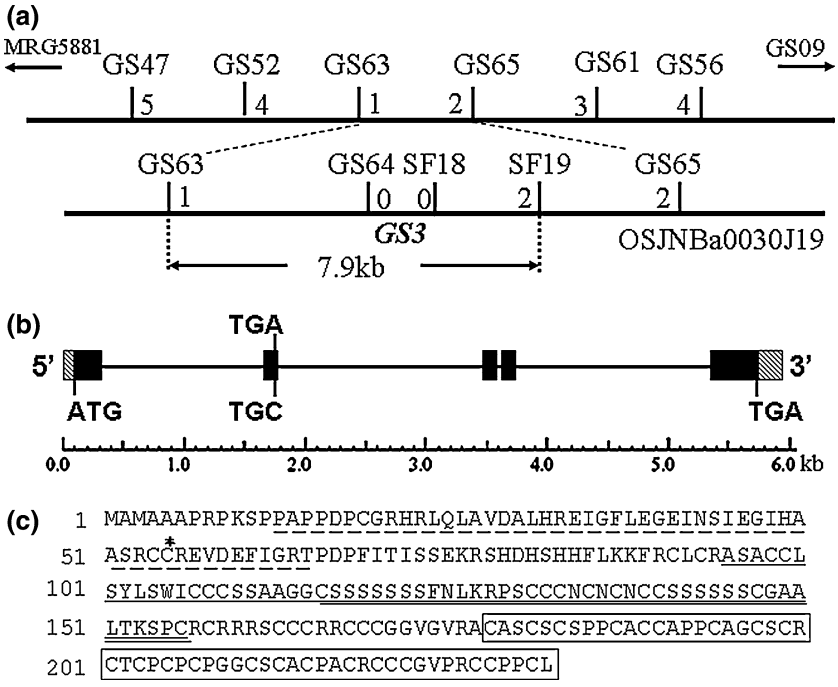

Fig. 4 Maps of the region encompassing the *GS3* locus. **a** The high-resolution linkage map of the *GS3* region generated using 1,384 BC_3F_2 plants with long grains. The numbers between molecular markers indicate the numbers of recombination events detected between the *GS3* locus and respective markers. OSJNBa0030J19 is a BAC clone of cv. Nipponbare encompassing the *GS3* locus. **b** Organization of the *GS3* gene. The positions of coding regions (*black boxes*), 5′ and 3′ UTR (*hatched boxes*), translation start codon (ATG), translation stop codon (TGA), one common single nucleotide mutation in the second exon between the two grain-length groups are indicated, in which a substitution of C (small grain group) by A (large grain group) in the second exon results in an early stop codon in large grain group. **c** Predicted amino acid sequence of the GS3 protein. The position of the amino acid change in large grain group (cysteine to stop codon) is indicated by an *asterisk*. The PEBP-like domain is indicated by *dashed underline*, the transmembrane region by *single solid underline*, the TNFR/NGFR family cysteine-rich domain by *double underline*, and the VWFC module is *boxed*

The long grain phenotype resulted from a premature termination

To identify the mutation that caused the variation of grain length, comparative sequence analysis of the homologous DNA fragment in the 7.9-kb region was conducted using six cultivars including three with long grains: Minghui 63 (9.91 ± 0.10 mm, mean ± SD), H94 (11.16 ± 0.09 mm) and 93–11 (10.02 ± 0.10 mm), and three with short to medium grains: Chuan 7 (6.30 ± 0.09 mm), Zhenshan 97 (8.08 ± 0.07 mm) and Nipponbare (7.12 ± 0.05 mm) (Fig. 1) (The sequences of Nipponbare and 93-11 were from the clone OSJNBa0030J19 and contig Ctg009226, respectively, http://www.ncbi.nlm.nih.gov/).

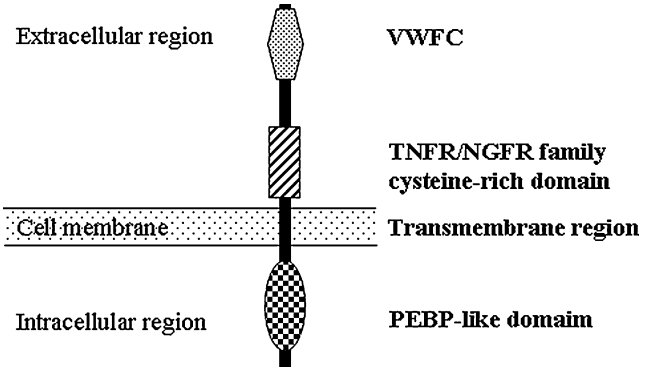

Fig. 5 Organization of the predicted GS3 protein indicating the localization of the various conserved domains and region. The PEBP-like domain is at the N-terminus and the VWFC module at the C-terminus

Although many nucleotide changes were observed among the six cultivars in the 7.9-kb region, there was only one common single nucleotide mutation detected at the second exon of the *GS3* gene between these two different grain-length groups, which changed a cysteine codon (TGC) in the small-grain group to a termination codon (TGA) in the large-grain group (Fig. 4b). This premature termination resulted in a 178-aa truncation in the C-terminus of the predicted protein in the large-grain group, which eliminated part of the PEBP-like domain and all the other three conserved domains. Such mutation is clearly in agreement with the recessive nature of the long grain phenotype, indicating that long grains resulted from the loss of the function of the protein otherwise producing short grains.

Discussion

The most significant finding of the present work was the delimitation of the *GS3* locus to a DNA fragment of approximately 7.9 kb in length. Identification of the full-length cDNA in the target region provided a candidate gene of *GS3*, which has five exons and encodes 232 aa containing a putative PEBP-like domain, a transmembrane region, a putative TNFR/NGFR family cysteine-rich domain and a VWFC module. Comparative sequencing analysis indicated that all the large-grain varieties tested share the same nonsense mutation in the second exon of the *GS3* gene that causes a 178-aa truncation in the C-terminus of the predicted protein, which deleted part of the PEBP-like domain and all the other three domains. These findings suggest that *GS3* may function as a negative regulator to prevent the growth of the grain size. Cloning of such a gene has provided the opportunity for fully characterizing the regulatory mechanism and related processes during grain development.

In this connection, it is interesting to note that the putative VWFC module is also found in the protein of a fruit shape gene, *OVATE* in tomato or *OVATE* homologues in other plant species, which determines the conversion of fruit from round to pear-shape which is a recessive phenotype (Liu et al. 2002). More interestingly, sequence comparison revealed that all the varieties with the pear-shaped fruit had a premature stop-codon in the *OVATE* gene. The similarity in morphological change and DNA sequence deletion between the rice grain and tomato fruit strongly suggests that the putative VWFC module may have a role in regulating the fruit/grain shape by negatively affecting the growth. It is known that the VWFC module, also referred to as Chordin-like cysteine-rich (CR) repeats, is present in a growing number of extracellular matrix proteins, and binds to members of the transforming growth factor-β (TGF-β) superfamily (Abreu et al. 2002). It has been proposed that the general function of VWFC is to regulate growth factor signaling by disrupting the receptor binding sites of TGF-β superfamily in the extracellular matrix, thus preventing activation of the TGF-β receptor (O'Leary et al. 2004). Such inhibitory activity of VWFC on growth factor signaling is clearly consistent with the mechanism of negative regulation in the development of grain size/fruit shape hypothesized for the *GS3* and *OVATE* genes.

Grain size has played important roles in the evolution of cereal crops. The grains of wild relatives are usually small and round in shape, which is often favored under natural selection because it is ideal for high fecundity and advantageous for dispersal by natural vectors. Domestication has greatly increased the diversity of grain shape and size together with many other changes, as consequences of physiologic response and adaptation to diverse natural environments and human needs. Interestingly, a QTL corresponding to *GS3* associated with grain length or grain weight in rice has been identified in at least five different interspecific crosses: V20A (*indica*) × *Oryza rufipogon* (Xiao et al. 1998), BG90-2 (*indica*) × *Oryza glumaepatula* (Brondani et al. 2002), Jefferson (tropical *japonica*) × *Oryza rufipogon* (Thomson et al. 2003), Caiapo (*indica*) × *Oryza glaberrima* (Aluko et al. 2004) and V20A (*indica*) × *Oryza glaberrima* (Li et al. 2004a, b). In all the cases, the wild parent possessed a dominant allele for small grains, suggesting that grain-size gene *GS3* has been conserved in different species of *Oryza* from a common ancestor during the course of evolution, and may be the site of a key allelic change that occurred during domestication. The similarities in both the wild-type gene structure containing the VWFC module and the cause of mutation by premature stop-codons between fruit shape in tomato (Liu et al. 2002) and grain size in rice (this study) suggest that orthologous genes and similar related regulatory processes for this type of traits may be conserved across a broad range of taxa ranging from monocot to dicot species.

The results also have important implications for rice genetic improvement. It has been repeatedly demonstrated that *GS3* can simultaneously increase grain-length which improves the appearance quality of (*indica*) rice, and grain-weight which increases grain yield. Surprisingly, however, this gene has not been intensively utilized in many rice breeding programs especially in China, as indicated by the high frequency of the medium grain cultivars. In fact, the female parents for most of the widely cultivated *indica* hybrids are of medium grain, resulting in medium grain hybrids. This is likely the result of historical distributions of the rice germplasms such that the long grain allele occurred frequently in South Asia but rarely in China. Thus incorporating the long-grain allele into the medium grain cultivars, especially in Chinese hybrid rice breeding programs to make both parents carry the long-grain allele, should have the potential to greatly improve both yield and quality of the cultivars and hybrids.

Acknowledgements This work was supported in part by grants from the National Program on the Development of Basic Research, the National Special Key Project of Functional Genomics and Biochips, and the National Natural Science Foundation of China.

References

Abreu JG, Coffinier C, Larraín J, Oelgeschläger M, Robertis EMD (2002) Chordin-like CR domains and the regulation of evolutionarily conserved extracellular signaling systems. Gene (Amst) 287:39–47

Aluko G, Martinez C, Tohme J, Castano C, Bergman C, Oard HJ (2004) QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. Theor Appl Genet 109:630–639

Brondani C, Rangel PHN, Brondani RPV, Ferreira ME (2002) QTL mapping and introgression of yield-related traits from *Oryza glumaepatula* to cultivated rice (*Oryza sativa*) using microsatellite markers. Theor Appl Genet 104:1192–1203

Evans LT (1972) Storage capacity as a limitation on grain yield. In: Banos Los (ed) Rice breeding. International Rice Research Institute, Manila, pp 499–511

Fan CC, Yu XQ, Xing YZ, Xu CG, Luo LJ, Zhang Q (2005) The main effects, epistatic effects and environmental interactions of QTLs on cooking and eating quality of rice in a doubled haploid line population. Theor Appl Genet 110:1445–14452

Hua JP, Xing YZ, Xu CG, Sun XL, Yu SB, Zhang Q (2002) Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. Genetics 162:1885–1895

Huang N, Parco A, Mew T, Magpantay G, McCouch S, Guiderdoni E, Xu JC, 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

Juliano BO, Villareal CP (1993) Grain quality evaluation of world rices. International Rice Research Institute, Manila

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

Li JX, Yu SB, Xu CG, Tan YF, Gao YJ, Li XH, Zhang Q (2000) Analyzing quantitative trait loci for yield using a vegetatively replicated F_2 population from a cross between the parents of an elite rice hybrid. Theor Appl Genet 101:248–254

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

Li J, Xiao J, Grandillo S, Jiang L, Wan Y, Deng Q, Yuan L, McCouch SR (2004b) QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. Genome 47:697–704

Lincoln S, Daly M, Lander E (1992) Constructing genetics maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, Whitehead Institute, Cambridge

Liu JP, Eck JV, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L (2002) Development and mapping of 2,240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9:199–207

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

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

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

Temnykh S, Park WD, Ayres N, Cartihour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Temnykh S, Declerck G, Luashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441–1452

Thomson M, Tai T, McClung A, Xai XH, Hinga M, Lobos K, Xu Y, Martinez P, McCouch S (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

Unnevehr LJ, Duff B, Juliano BO (1992) Consumer demand for rice grain quality. International Rice Research Institute, Manila, and International Development Research Center, Ottawa

Xiao JH, Li JM, Grandillo S, Ahn SN, Yuan LP, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. Genetics 150:899–909

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 43:721–726

Xing YZ, Tan YF, Hua JP, Sun XL, Xu CG, Zhang Q (2002) Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. Theor Appl Genet 105:248–257

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

Yu SB, Li JX, Xu CG, Tan YF, Gao YJ, Li XH, Zhang Q, Saghai Maroof MA (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc Natl Acad Sci USA 94:9226–9231