

Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight

Jianfeng Weng¹, Suhai Gu², Xiangyuan Wan¹, He Gao¹, Tao Guo¹, Ning Su², Cailin Lei², Xin Zhang², Zhijun Cheng², Xiuping Guo², Jiulin Wang², Ling Jiang¹, Huqu Zhai², Jianmin Wan²

¹National Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China; ²Institute of Crop Science and the National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Grain weight is a major determinant of crop grain yield and is controlled by naturally occurring quantitative trait loci (QTLs). We earlier identified a major QTL that controls rice grain width and weight, *GW5*, which was mapped to a recombination hotspot on rice chromosome 5. To gain a better understanding of how *GW5* controls rice grain width, we conducted fine mapping of this locus and uncovered a 1 212-bp deletion associated with the increased grain width in the rice cultivar Asominori, in comparison with the slender grain rice IR24. In addition, genotyping analyses of 46 rice cultivars revealed that this deletion is highly correlated with the grain-width phenotype, suggesting that the *GW5* deletion might have been selected during rice domestication. *GW5* encodes a novel nuclear protein of 144 amino acids that is localized to the nucleus. Furthermore, we show that *GW5* physically interacts with polyubiquitin in a yeast two-hybrid assay. Together, our results suggest that *GW5* represents a major QTL underlying rice width and weight, and that it likely acts in the ubiquitin-proteasome pathway to regulate cell division during seed development. This study provides novel insights into the molecular mechanisms controlling rice grain development and suggests that *GW5* could serve as a potential tool for high-yield breeding of crops.

Keywords: *GW5*, QTL, grain width and weight, domestication, polyubiquitin, rice (*Oryza sativa* L.)

Cell Research (2008) 18:1199-1209. doi: 10.1038/cr.2008.307; published online 18 November 2008

Introduction

Rice is the world's most important staple food crop and is a primary food source for about half of the world's population. As farmland is decreasing and the global population increasing, there is an urgent need to secure grain production [1]. Therefore, increasing grain yield has been the primary objective in current rice-breeding programs.

The shape of rice grain, a complex agronomical trait, plays an important role in determining the rice yield [2, 3]. Currently, more than 22 quantitative trait loci (QTLs) related to rice grain width have been reported in different mapping populations [4-10], and several

genes that regulate rice yield or grain shape have been identified, including *Gn1a* [11], *Ghd7* [12], *GS3* [13] and *GW2* [14]. *Gn1a* encodes a cytokinin oxidase/dehydrogenase required for degrading the phytohormone cytokinin. Reduced expression of *Gn1a* causes cytokinin accumulation in inflorescence meristems and increases the number of reproductive organs, resulting in increased grain yield [11]. *Ghd7* encodes a CCT domain protein. Enhanced expression of *Ghd7* under long-day conditions delays heading and increases plant height and panicle size, resulting in increased rice grain yield [12]. *GS3*, a major QTL for grain length and weight and a minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein [13]. *GW2*, a QTL for rice grain width and weight, encodes a previously unknown RING-type E3 ubiquitin ligase that negatively regulates cell division by targeting its substrate(s) for degradation by the ubiquitin-proteasome pathway [14]. In Arabidopsis, *DA1*, which encodes a predicted ubiquitin receptor, sets the final seed and organ size by restricting the period of

Correspondence: Jianmin Wan

Tel/Fax: +86-10-62186628

E-mail: wanjm@caas.net.cn

Received 3 September 2008; revised 6 October 2008; accepted 6 October 2008; published online 18 November 2008

cell proliferation. The mutant protein encoded by the *dal-1* allele has negative activity toward DA1 and a DA1-related protein, and overexpression of *dal-1* cDNA dramatically increases the seed and organ size of wild-type plants [15]. In addition, plant hormones and their receptors [16-19], as well as the nutritional synthesis pathway [20, 21], have also been implicated in the development of grain in crops and other plant species. For example, the rice *dwarf11* mutant has a defect in a novel cytochrome P450, which shares homology with enzymes involved in brassinosteroid biosynthesis, and bears seeds of reduced length [17]. In maize, the *gln1-4* mutant displays reduced kernel size owing to the loss-of-function in glutamine synthetase isoenzymes [21]. Despite these efforts, the mechanisms that establish the final size of grain or seeds are still poorly understood.

In addition, human civilization has had a major impact on the domestication of crop plants. Regarding the origin of *indica* and *japonica* rice subspecies, one school of thought holds that *indica* and *japonica* rice originated independently from a wild ancestor [22-24]; others favor the notion that the *japonica* were derived from the *indica* [25]. Recent functional genomic studies have helped to excavate a number of domestication genes [26-29]. For example, *sh4* and *qSH1* are essential for effective field harvest because of their ability to reduce grain shattering [29, 30]. In addition, grain or fruit size was important in the evolution of wild species because of the continuous selection for large ones during early domestication. In tomato, *fw2.2* changes the fruit weight by almost 30% and seems to have been responsible for a key transition during domestication [27].

We previously reported a major QTL on chromosome 5, *qGW5*, which is associated with reduced grain width not only in the isogenic Asominori background but also in the recombinant background of Asominori and IR24 under multiple environmental conditions [31]. In this study, we report the fine mapping, cloning and initial characterization of this gene. We also provide evidence that *GW5* is associated with rice domestication. Our work sheds light on the mechanisms of grain development and domestication of rice.

Results

Characterization of the narrow rice line, CSSL28

Grain width is quantitatively inherited; therefore, it is difficult to analyze the mechanism of grain formation using conventional methods. To dissect the loci that control grain width into several single genes, 71 recombinant inbred lines (RILs) were derived from a cross between Asominori and IR24 by single-

seed descent [32]. Sixty-six chromosome segment substitution lines (CSSLs) with largely Asominori background, named CSSL1-66, were produced by nonselectively crossing and backcrossing 19 selected RILs with Asominori to produce the BC₃F₁ generation [33]. One of these CSSLs, CSSL28 (narrow-grain rice), shows a slender-grain phenotype, because it harbors a chromosomal segment between RFLP (restriction fragment length polymorphism) markers C263 and R2289 (Figure 1). This segment derives from IR24 (narrow-grain rice) and is substituted into the Asominori (wide-grain rice) genomic background. CSSL28 shows a 16.4% reduction in grain width of paddy rice, and its 1 000-grain weight is reduced by 18.7% compared with that of Asominori, mainly because of the reduced grain width (Table 1).

GW5 is associated with a 1 212-bp deletion in the Asominori cultivar

For genetic analysis and isolation of the dominant narrow-grain QTL, named *GW5*, the F₂ population was constructed from a cross between Asominori and CSSL28, and the QTL was dissected into a single gene [31]. By means of linkage analysis using the genotype data of both the *GW5* gene and simple sequence repeat (SSR) markers, the *GW5* gene was mapped to an interval between markers RM3328 and RMw513 in 805 homozygotes (Figure 2A). Thus, the *GW5* gene was mapped to a genomic region 2.7 cM in length, located 2.3 cM from RM3328 and 0.37 cM from RMw513. In this region, three bacterial artificial chromosome (BAC) contigs were found – OJ1725_E07, OJ1097_A12 and B1007D10 (Figure 2B) – according to the Nipponbare genome (<http://www.gramene.org>).

For high-resolution mapping of *GW5*, five new cleaved-amplified polymorphic sequence (CAPS)

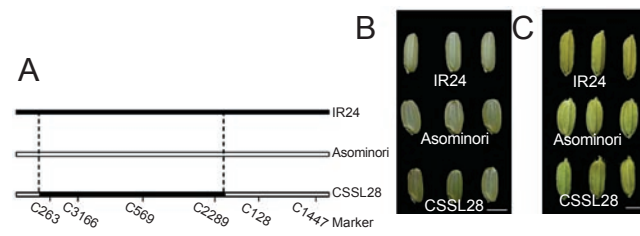


Figure 1 *qGW5* controls grain width and weight in CSSL28. (A) Genotype of the narrow grain line, CSSL28 (a segment substitution between the RFLP markers C263 and R2289 on chromosome 5). The black bar indicates the fragment from IR24 in the Asominori genomic background. (B) Phenotype of the paddy rice grain, IR24 (top), Asominori (middle) and CSSL28 (bottom). (C) Phenotype of the brown rice grain. Scale bar: 3 mm.

Table 1 Phenotypic analysis of Asominori and CSSL28

Traits		Asominori (%)	CSSL28 (%)	<i>P</i> value
Thousand-grain weight (g)	paddy rice	27.30±0.3 (100)	22.20±0.5 (81.3)	3.00E-07
Grain width (mm)	paddy rice	3.36±0.1 (100)	2.81±0.1 (83.6)	7.20E-16
	brown rice	2.90±0.1 (100)	2.43±0.1 (83.8)	6.10E-12
Grain thickness (mm)	paddy rice	2.15±0.1 (100)	1.97±0.1 (91.6)	1.10E-06
	brown rice	1.98±0.1 (100)	1.75±0.1 (89.7)	2.20E-06
Grain length (mm)	paddy rice	7.78±0.1 (100)	8.26±0.2(106.4)	1.00E-09
	brown rice	5.48±0.1 (100)	5.67±0.1(103.5)	2.50E-05

markers were developed on one of the BACs (OJ1097_A12). The *GW5* locus was pinpointed into an interval between the CAPS markers *Cw5* and *Cw6* in 2 180 wide-grain homozygous individuals from the BC₄F₂ population (Figure 2C and Table 2). Within this region, the 21-kb fragment of Asominori and the corresponding 22.2-kb fragments of CSSL28 and IR24 were sequenced. Further comparison of these sequences showed that the Asominori genome harbored a 1 212-bp deletion

(Figure 2D). Meanwhile, two Indel (insertion or deletion) markers – Indel1, including 775 bp of the 1.2 kb fragment, and Indel2, containing the aforementioned 1.2-kb fragment and totaling 1 897 bp in length – were designed (Table 3) and used to confirm the presence of *GW5*. Genetic analysis revealed that these two Indel markers co-segregated with *GW5* in 2 180 wide-grain progenies (Figure 2C).

On the basis of the available sequence annotation

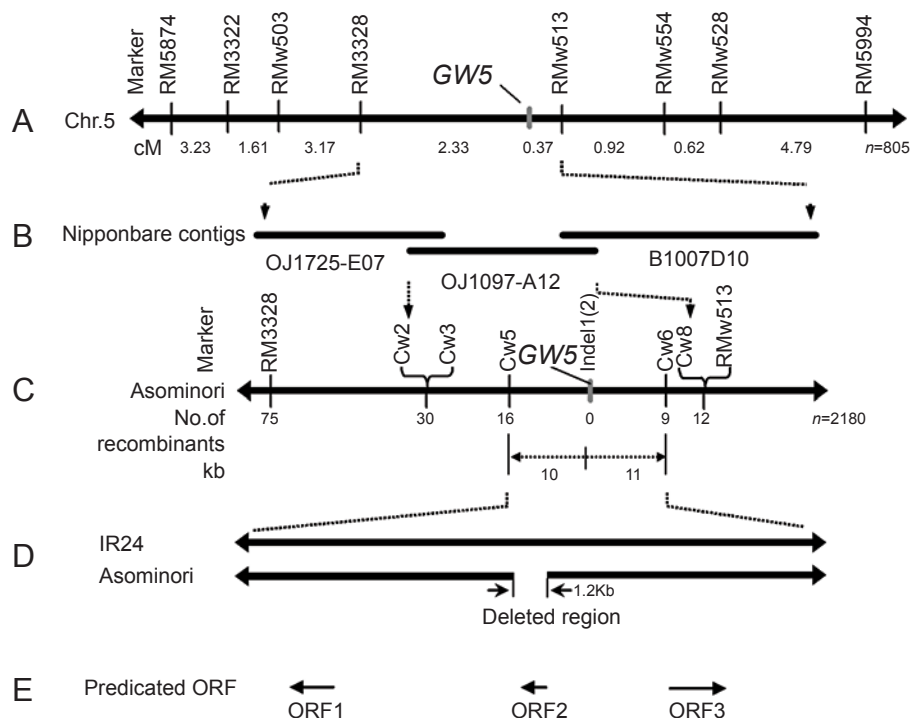


Figure 2 Fine mapping of *GW5*. This map was constructed on the basis of publicly available rice sequences. The CAPS and Indel markers developed in this work are indicated. **(A)** *GW5* was mapped to a region between markers RM3328 and RMW513 in 805 recessive individuals. **(B)** Three BAC contigs, OJ1725_E07, OJ1097_A12 and B1007D10, cover the *GW5* locus in the Nipponbare genomic sequence. **(C)** *GW5* was narrowed down to a 21-kb genomic DNA region between CAPS markers *Cw5* and *Cw6*, and it co-segregated with the Indel1 marker in 2 180 recessive individuals. **(D)** Compared with IR24, 1.2-kb genomic DNA is deleted in Asominori. **(E)** Three ORFs are predicted in the candidate region harboring *GW5* (<http://softberry.com>). The predicted ORF2 is located in the deleted region in Asominori.

Table 2 Recombination events between *GW5* and molecular markers

Plant No.	Markers closely linked to <i>GW5</i>						
	Cw2	Cw3	Cw5	Cw15	Cw6	Cw8	RMw513
AA1	H	H	A	A	A	A	A
AA14	H	H	H	A	A	A	A
AA152	H	H	H	A	A	A	A
AA165	A	A	A	A	A	H	H
AA166	H	H	H	A	A	A	A
AA185	A	A	A	A	H	H	H
AA213	H	H	H	A	A	A	A
AA270	B	B	A	A	A	A	A
AA302	H	H	A	A	A	A	A
AA313	A	A	A	A	A	H	H
AA363	H	H	H	A	A	A	A
AA395	H	H	A	A	A	A	A
AA401	H	H	H	A	A	A	A
AA461	H	H	A	A	A	A	A
AA485	H	H	A	A	A	A	A
AA569	A	A	A	A	H	H	H
AA570	H	H	A	A	A	A	A
AA598	H	H	A	A	A	A	A
AA653	A	A	A	A	B	B	B
AA690	H	H	H	A	A	A	A
J134	H	H	H	A	A	A	A
J215	H	H	H	A	A	A	A
J442	H	H	H	A	A	A	A
J461	H	H	A	A	A	A	A
J505	H	H	H	A	A	A	A
J585	A	A	A	A	H	H	H
J716	A	A	A	A	H	H	H
J727	A	A	A	A	H	H	H
J830	H	H	A	A	A	A	A
J865	H	H	H	A	A	A	A
J965	A	A	A	A	H	H	H
J991	H	H	H	A	A	A	A
J1085	H	H	H	A	A	A	A
J1119	H	H	A	A	A	A	A
J1122	A	A	A	A	A	H	H
J1140	H	H	A	A	A	A	A
J1180	H	H	H	A	A	A	A
J1231	H	H	A	A	A	A	A
J1254	A	A	A	A	A	A	A
J1314	A	A	A	A	H	H	H
J1316	H	H	H	A	A	A	A

A, B and H denote the marker genotypes for Asominori, CSSL28 and F₁ of CSSL18×Asominori, respectively. AA1, AA14, AA152, ... J1316 represent the No. of 41 F₂ individuals in which recombination occurred between the *GW5* gene and the molecular marker.

(<http://www.softberry.com>; <http://www.ncbi.nlm.nih.gov/BLAST>), two open reading frames (ORFs) were predicted in the 21.0-kb target region of Asominori, and three ORFs were predicted in the corresponding region of CSSL28. However, no sequence differences were found between these lines for ORF1 and ORF3. ORF1 encodes a protein showing high similarity to ubiquitin-protease-like protein, whereas ORF3 encodes an unknown protein containing a calmodulin-binding motif. Interestingly, we found that the 22.2-kb region of CSSL28 contains a third ORF between ORF1 and ORF3, designated ORF2, which is located in the deleted region of Asominori (Figure 2E). Motif scan analysis showed that the predicted product of ORF2 harbors a nuclear localization signal (NLS) and an arginine-rich domain (http://hits.isb-sib.ch/cgi-bin/motif_scan). These results suggest that ORF2, which is present in CSSL28 but absent in Asominori, likely corresponds to *GW5*.

The deletion prevails in wide-grain rice cultivars and defines a domestication-related rice gene

To further confirm the identity of *GW5*, we randomly selected 46 rice lines, including Asominori, IR24 and 44 other cultivars, to determine the co-segregation relationship between the absence of the candidate *GW5* gene and the wide-grain phenotype (Table 4). The 46 rice lines were divided into two groups: group I contained 23 narrow-grain varieties (grain width ranging from 2.40 to 2.85 mm) and group II consisted of 23 wide-grain ones (3.30–3.92 mm) (Table 4). Segregation analysis of the Indel1 marker showed that the 775-bp target fragment was observed only in group I cultivars, but not in group II cultivars (Figure 3A and 3B). Similarly, PCR analysis with the Indel2 marker detected a 1 897-bp fragment in group I cultivars, but only a 697-bp product in group II cultivars (Figure 3C and 3D). These results suggested that the 1.2-kb fragment containing the *GW5* gene was absent in all 23 wide-grain cultivars examined here but was present in the narrow-grain ones. The strict correlation between the grain-width phenotype and deletion of the 1.2-kb genomic DNA strongly supports the notion that the candidate ORF2 located on the deleted 1.2-kb fragment represents *GW5*.

To address how this deletion in *GW5* prevails in rice domestication, we further analyzed the 46 cultivars. Group I contained 21 *indica* species (grain width ranged from 2.40 to 2.84 mm) and two slender *japonica* varieties (one 2.74 mm wide, the other 2.85 mm wide), whereas most *japonica* (3.30–3.92 mm) fell within group II, with the exception of two wide *indica* lines (3.46 and 3.48 mm). PCR analysis showed that *GW5* was deleted in all group II lines, including both *japonica* and

Table 3 A list of primers used in this work

Marker name	Marker type	BAC location	Size(bp)	Forward primer	Reverse primer	Motif length or restriction enzyme
RM3322	SSR	AC087553	172	CTTCTCCACCCATGCCAC	CCTGCAACGAACACCCAC	(ct)14
RM3328	SSR	B1140B01	119	TCTACTCGTGCAGCTTGTGC	GAAGCAGCAAGAAAGGGATG	(ct)14
RM5874	SSR	AC10427	208	GAAAAGATCCTGGCTCGTTG	GCATCATCGCCAGAGCTC	(atg)8
RM5994	SSR	AC135914	105	CAGGCACTGCAGCATGTC	AGGCTCGGCGAGATATTCTC	(ccg)8
RMw-503	SSR	OSJNBa 0029B02	125	TCAGGCTGCAACCTCAAATGG	CCGACGACGACCCCGACAA	(ag)11
RMw-513	SSR	B1007D10	171	GTATTTGTTTGTGCGATTC	TAGGACCATAGATGTGAGTTA	(ct)10
RMw-528	SSR	P0473H02	168	GATGGAGCAAATGGAAGAAT	GATTTCCAAGTAGCCCCTA	(tct)8
RMw-554	SSR	B1140B01	499	TACGTCGAAATCGTAGC	AAACCCTTTGTAGTCCC	(ta)13
Cw-2	CAPS	OJ1097_A12	516	GAACAGCAGGACGAAGA	GCATGTAAATTGCGAGAC	HindIII
Cw-3	CAPS	OJ1097_A12	588	GGAATGGAGGAAGTAGAAAG	AACAGACTGCTCCCAAAA	HpyCH4III
Cw-5	CAPS	OJ1097_A12	382	GGAGTAACTGTTATCGTGCT	CTCCCATTACCTCTACTTCAA	BsaH I
Cw-6	CAPS	OJ1097_A12	665	ATGAGGAGCGGACCACAGA	GCATGATTCCACCCACCAG	BfaI
Cw-8	CAPS	OJ1097_A12	427	AACTAATCCC GCATGTCC	AAGTCACCCAAATCCACGA	ApoI
Indel1	Indel	OJ1097_A12	775	AGTACGACCATGATGTTTCCC	GACCTAACCCATCTCATTCCA	
Indel2	Indel	OJ1097_A12	1897	GCGTCGTCAGAGGTAGA	GTGGGATAGGATGAAACC	
Marker01			140	AGACGGAGGAGGAGGAACGG GCGGCCAGTG	GAATATCTTCCCAGATCCAG GACGAGG	
Marker02			200	AGACGGAGGAGGAGGAACGG GCGGCCAGTG	TCGATCCTATTTTCGAGCTG TTTGGGTAG	

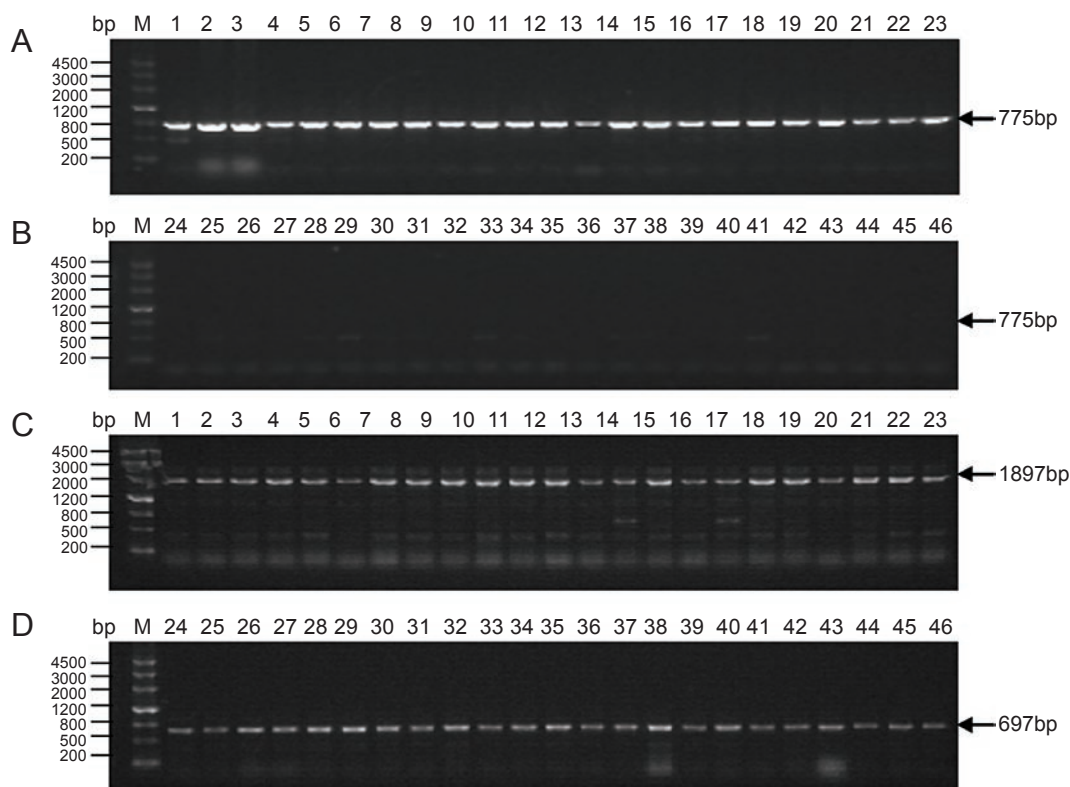


Figure 3 Genotyping analyses in 46 rice cultivars. **(A)** PCR products of the Indel1 marker in 23 narrow-grain rice lines; **(B)** PCR products of the Indel1 marker in 23 wide-grain rice varieties; **(C)** PCR products of the Indel2 marker in 23 narrow-grain rice lines; and **(D)** PCR products of the Indel2 marker in 23 wide-grain rice varieties. The DNA sample number is listed in Table 3.

Table 4 The 46 rice cultivars analyzed in this work and their places of origin

DNA No.	Cultivar name	Grain width (mm)	Grain length (mm)	Place of origin	Indica/ Japonica
1	IR24	2.45	8.95	Philippines	indica
2	CB1128	2.4	8.6	America	indica
3	Lemont	2.61	9.95	America	indica
4	Kasalath	2.71	7.42	Japan	indica
5	Xianxiaozhan	2.24	9.33	Guangdong, China	indica
6	Texianzhan 25	2.51	8.85	Guangdong, China	indica
7	Yuexianzhan	2.63	8.54	Guangdong, China	indica
8	Gui 99	2.67	9.53	Guangxi, China	indica
9	Minghui 63	2.84	9.86	Fujian, China	indica
10	Feigai B	2.7	10	Sichuan, China	indica
11	Zhongjian 100	2.57	9.64	Zhejiang, China	indica
12	Zhongyouzao 81	2.58	10.19	Zhejiang, China	indica
13	Zhouyou 903	2.58	9.85	Zhejiang, China	indica
14	Zhongsu 2	2.65	9.29	Zhejiang, China	indica
15	9311	2.81	9.17	Jiangsu, China	indica
16	Nanjing 16	2.59	9.45	Jiangsu, China	indica
17	Xieqingzao B	2.69	9.43	Anhui, China	indica
18	Pa64s	2.73	9.84	Hunan, China	indica
19	Xiangzhaoxian 31	2.48	9.46	Hunan, China	indica
20	Ce 64-7	2.52	9.29	Hunan, China	indica
21	Xiangwanxian 9	2.64	8.88	Hunan, China	indica
22	H305-84	2.74	10.32	Hungary	Japonica
23	KAROLINA	2.85	10.45	Hungary	Japonica
24	Asominori	3.36	7.78	Japan	Japonica
25	Nipponbare	3.31	7.47	Japan	Japonica
26	Teqing	3.31	7.57	Guangdong, China	Japonica
27	Tesanai 2	3.3	7.6	Guangdong, China	Japonica
28	Xiushui 110	3.35	7.04	Zhejiang, China	Japonica
29	Yongjing 18	3.92	7.1	Zhejiang, China	Japonica
30	Xiushui 63	3.41	7.63	Zhejiang, China	Japonica
31	Wuyujing 3	3.52	7.37	Jiangsu, China	Japonica
32	Lianjing 3	3.35	6.92	Jiangsu, China	Japonica
33	Yanjing 5	3.46	7.2	Jiangsu, China	Japonica
34	Zaofeng 9	3.7	7.22	Jiangsu, China	Japonica
35	Yujing 6	3.33	7.03	Henan, China	Japonica
36	Xiangjing R 109	3.41	7.42	Hunan, China	Japonica
37	Jijing 66	3.36	7.32	Hebei, China	Japonica
38	Liaojing 294	3.54	7.34	Liaoning, China	Japonica
39	Liao 294	3.37	7.06	Liaoning, China	Japonica
40	Tongyu 124	3.47	7.52	Jilin, China	Japonica
41	Songjing 6	3.77	7.34	Heilongjiang, China	Japonica
42	Hejiang 19	3.76	6.92	Heilongjiang, China	Japonica
43	Fushiguang	3.59	8.01	Heilongjiang, China	Japonica
44	Kendao 8	3.41	7.6	Heilongjiang, China	Japonica
45	Zhenshan 97B	3.46	8.68	Jiangxi, China	indica
46	Sanlicun	3.48	10.25	Jiangsu, China	indica

Note: The values are the means of rice grains.

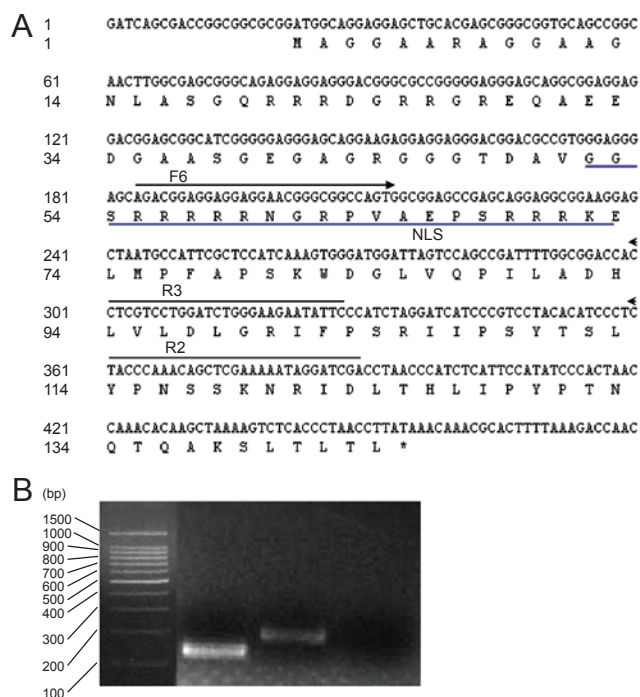


Figure 4 *GW5* expression analyses. **(A)** The sequence of the predicted ORF2. The primers (F6, R2 and R3) used for RT-PCR analysis and the predicted NLS are denoted by arrows and a line, respectively. In addition, amino-acid residues 20-71 formed an arginine-rich region. **(B)** RT-PCR analysis of ORF2 using mRNA derived from the young panicle of CSSL28. Lane 1: molecular marker. Lane 2: a 140-bp RT-PCR product using primers F6 and R3 (marker 1). Lane 3: a 200-bp RT-PCR product using primers F6 and R2 (marker 02). Lane 4: control (no template was added in the PCR reaction).

indica subspecies. We thus conclude that the deletion is highly correlated with the grain-width phenotype among the *japonica* rice in group II. Our results suggest that *GW5* was kept in most *japonica* cultivars during rice domestication and that this gene may underlie the relatively wide- and short-grain phenotype [23].

GW5 is expressed in slender-grain rice

The sequence of 9 311 (a narrow-grain *indica* rice cultivar) in the cognate region is similar to that in the *GW5* region in CSSL28. Analysis of the *indica* rice genome sequence showed that *GW5* represents a single-copy gene without any expressed sequence tags (ESTs) or cDNA support. However, we noticed that expression of the candidate ORF was detected in the tiling microarray database of 9 311 (GenBank Acc: CL971152) [34]. To further confirm the expression of *GW5*, we performed RT-PCR analysis using mRNAs derived from the young panicle of CSSL28 (Figure 4A). We obtained RT-PCR

products with primer combinations of F6 and R2, as well as with F6 and R3 (Figure 4B and Table 3). These PCR products were confirmed by sequencing analyses. Therefore, *GW5* is expressed in CSSL28.

GW5 physically interacts with polyubiquitin

Sequence analysis indicated that *GW5* encodes a novel protein without significant homology to any proteins of known biochemical function. The protein is predicted to contain an NLS and an arginine-rich domain. To test the functionality of the predicted NLS, we fused the coding region of *GW5* to the N-terminus of GFP. Transient expression in onion epidermal cells showed that the *GW5*-GFP fusion protein is exclusively localized to the nucleus (Figure 5).

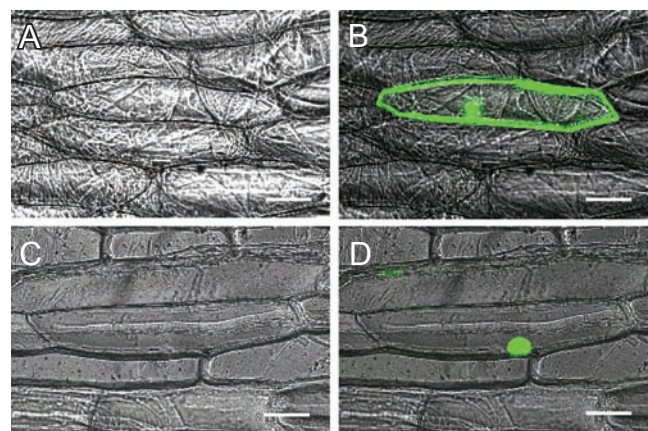


Figure 5 Sub-cellular localization of *GW5*. **(A)** and **(C)** Bright-field images of onion epidermal cells. **(B)** GFP by itself (pCambia 1302 vector) localizes to both cytoplasm and nucleus. **(D)** *GW5*-GFP fusion protein is localized only to the nucleus. Scale bars: 50 μ m.

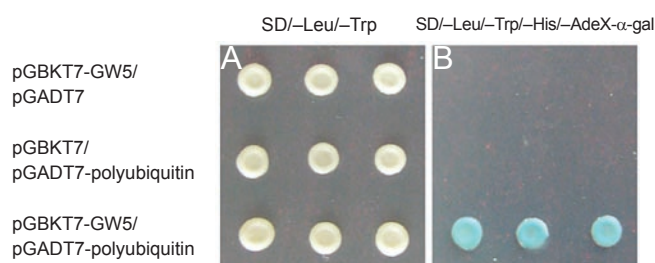


Figure 6 *GW5* interacts with polyubiquitin in yeast two-hybrid assay. **(A)** AH109 yeast cells expressing pGBKT7-*GW5*/pGADT7 (top), pGBKT7/pGADT7-polyubiquitin (middle) and pGBKT7-*GW5*/pGADT7-polyubiquitin (bottom) were selected on synthetic growth medium without Leu and Trp. **(B)** X-gal filter lift assay showing that *GW5* interacts with polyubiquitin.

In an effort to identify the functional partners of GW5, we carried out a yeast two-hybrid screen using full-length GW5 as the bait. A prey library was constructed using mRNAs derived from the young panicle of IR24 before heading, at which time *GW5* expression was detected. We repeatedly found that polyubiquitin protein interacts with GW5. This result was found 14 times in about 200 candidate positive clones on synthetic growth medium without leucine, tryptophan, histidine and adenine. X-gal filter lift assays also detected a clear interaction between GW5 and polyubiquitin (Figure 6B). This result suggests that the GW5 protein may play an important role in regulating the grain shape through involvement with the ubiquitin-proteasome pathway.

Discussion

In this study, we report the fine mapping and cloning of *GW5*, a major QTL controlling grain width and weight in rice. We identified a 1 212-bp deletion between the CAPS markers Cw5 and Cw6 on chromosome 5 in Asominori rice (*japonica*, wide grain) in comparison with IR24 rice (*indica*, slender grain). In this deleted region, we identified an ORF that is expressed in CSSL28. Furthermore, through genotyping analyses of 46 randomly selected rice cultivars, we detected a high correlation between grain shape and *GW5*; that is, loss of *GW5* occurred in the wide-grain lines, including most *japonica* and a few *indica* rice, whereas *GW5* was present in the slender-grain rice, including most *indica* and a few *japonica* rice cultivars (Figure 3 and Table 4). Together, these findings strongly support the notion that deletion of *GW5* is responsible for the change in grain width and weight and that the ORF we identified in this deleted genomic region represents the *GW5* gene.

In addition, it is believed that wild ancestors with small seeds are usually favored by natural selection because small seed size is frequently associated with a higher number of seeds per plant, early maturity and wider geographic distribution [35]. Domestication of wild rice probably started about 9 000 years ago, and *O. sativa* evolved and differentiated into two subspecies, *indica* and *japonica* [23]. As numerous single-nucleotide polymorphisms (SNPs) and Indels were identified between 9 311 (*indica*) and Nipponbare (*japonica*) [36], it was suggested that chromosomal rearrangements might have occurred during the domestication of these two subspecies, which caused significant divergence between the compositions of the two corresponding genomes. In our study, the deletion of *GW5* was detected in most *japonica* and a few *indica* cultivars with the wide-grain phenotype. Our data are in agreement with the hypothesis

that *indica* and *japonica* rice originated independently from a mutual wild ancestor.

Earlier studies have determined that some domestication genes have key regulatory roles in plant morphology and architecture. For example, *TBI*, one of these key regulatory genes, controls the morphological transition from wild species to cultivated maize [26], and the major QTLs *fw2.2* [27], *ovate* [28] and *qSH1* [29] control tomato fruit size, the transition of tomato fruit from round to pear-shaped and seed shattering, respectively. The identification of *GW5* and the earlier reported *GW2* as major QTLs that control rice-grain width lends strong support to the notion that “major morphological changes during evolution and domestication can be attributable to a few major regulatory genes of large effects rather than to many genes, each of which contributes a small effect to the changes” [37].

The *GW5* ORF encodes a predicted protein of 144 amino acids without significant sequence homology to any proteins of known molecular function. It contains a predicted NLS and an arginine-rich domain. Using a GFP fusion strategy, we showed that GW5 protein is exclusively localized to the nucleus. Furthermore, GW5 physically interacts with polyubiquitin, as shown by the yeast two-hybrid assay. These observations suggest that GW5 might be involved in the ubiquitin-proteasome pathway to regulate grain width and weight. It is particularly interesting to note that *GW2*, another QTL for rice-grain width and weight, encodes a RING-type protein with E3 ubiquitin ligase activity [14]. Loss of GW2 causes increased cell numbers, a wider spikelet hull and an accelerated grain milk-filling rate, resulting in enhanced grain width, weight and yield. GW2 may negatively regulate cell division by targeting its substrate(s) for regulated proteolysis by proteasomes. It will be of great interest to determine the functional relationship between GW2 and GW5 in controlling seed development. GW5 may work together with GW2 in targeted degradation of certain substrates that promote cell division and grain growth. GW5 may also work together with the deubiquitinating enzymes and might be involved in reversing ubiquitination by substrate-specific cleavage of ubiquitin moieties [38-40]. Recent studies have also pointed to a critical role of the ubiquitin pathway in grain development in other plant species [14, 15]. For example, loss-of-function of a ubiquitin receptor significantly increased the seed and organ size in *Arabidopsis* [15]. Further studies to determine the molecular function of GW5 will yield more insights into the mechanism of grain development, which could provide a potential tool for improving the crop grain yield.

Materials and Methods

Plant materials

CSSL28, harboring the *GW5* gene, was used to build the BC₄F₂ population by backcrossing with Asominori followed by subsequent self-pollination. When the plants matured, seeds were collected from the primary panicles and dried for 72 h at 50 °C. Paddy rice and brown rice were used for evaluating grain size. A total of 2 180 wide-grain homozygous individuals were collected for fine mapping. These plants were collected in three consecutive years (2004–2006) at three locations (Nanjing, Hainan and Beijing). The widths of 20 rice grains randomly selected from each line were estimated using an electronic digital caliper (Guanglu Measuring Instrument Co. Ltd, China) with a precision of 0.1 mm. The plant materials and markers used in this study are listed in Table 2.

For co-segregation analysis of *GW5* and rice grain shape, we selected Asominori, IR24 and 44 local varieties, including 21 normal *japonica* lines, two *indica* lines with the wide-grain phenotype, 21 normal *indica* lines and two slender *japonica* lines. These rice cultivars were supplied by the National Rice Germplasm Bank (Table 3).

DNA preparation, PCR protocol and DNA marker analysis

DNA was extracted from fresh leaves of BC₄F₂ individuals and 46 rice lines using the methods described earlier [41]. PCR was performed using a procedure described earlier [42], with minor modifications. PCR products were separated on 8% non-denaturing polyacrylamide gels and detected using the silver staining method [43].

Primer sequences, map position and amplified length of new SSR, CAPS and Indel markers used for fine mapping of *GW5* are listed in Table 3. New CAPS markers were developed by comparing the original length with CAPS length, using SNP2CAPS software [44]. PCR products of CAPS and Indel markers were separated on 1–2% agarose gels.

Primary mapping of the *GW5* gene

A total of 805 BC₄F₂ plants were genotyped using eight SSR markers to construct a small-scale linkage map. Among these molecular markers, four were new SSRs developed in our laboratory. The 805 BC₄F₂ plants were classified according to their measured phenotypic performance as one of two types: slender-grain rice, such as IR24 or CSSL28, or wide-grain rice, such as Asominori. On the basis of the segregation of grain width within the 805 BC₄F₂ families, the BC₄F₂ population was further partitioned into three groups: slender plants (genotype *GW5GW5*), slender-wide plants with segregation (genotype *GW5gw5*) and wide-grain individuals without segregation (genotype *gw5gw5*). To combine the molecular marker data with the *GW5* gene, small-scale genetic mapping of the *GW5* gene was performed using the program Mapmaker/Exp 3.0.

High-resolution mapping

For high-resolution mapping of the *GW5* gene, the bulked-extreme and recessive-class approach was used to calculate the recombination frequencies (*c*) between the *GW5* gene, the five newly developed CAPS and the two Indel markers in the 2 180 homozygous wide-grain BC₄F₂ plants. Thus, $c=(N1+N2/2)/N$,

where *N* is the total number of wide-grain plants surveyed, *N1* is the number of wide-grain individuals with the banding pattern of the slender-grain parent and *N2* is the number of wide-grain plants with heterozygous banding patterns [45].

mRNA isolation and RT-PCR

Total RNA was extracted from the young spike rice IR24 (narrow-grain line) about 5 days before heading, using an RNA prep pure Plant Kit (Tiangen Biotech Co., Ltd). The mRNA was then further purified using PolyAtract[®] mRNA isolation systems (Promega Corporation), according to the manufacturer's instructions. First-strand cDNA was synthesized using a PrimeScript[™] RT-PCR Kit (TaKaRa Biotechnology Co., Ltd). RT-PCR was carried out using the first-strand cDNA as the template with the primers F6 (5'-AGA CGG AGG AGG AAC GGG CGG CCA GTG-3') and either R2 (5'-TCG ATC CTA TTT TTC GAG CTG TTT GGG TAG-3') or R3 (5'-GAA TAT TCT TCC CAG ATC CAG GAC GAG G-3'). The PCR products were cloned into the pMD18-T vector and sequenced at the Invitrogen Sequencing Facility.

Protein sub-cellular localization

To investigate the cellular localization of *GW5*, the full-length cDNA of *GW5* was cloned into the pCAMBIA1302 vector at the *NcoI* and *SpeI* sites. The *GFP-GW5* fusion gene was driven by the 35S promoter. The control vector pCAMBIA1302 and the *GFP-GW5* fusion gene construct were bombarded into onion epidermal cells using a helium ballistics device (Bio-Rad PDS-1000). The samples were examined with a Zeiss LSM510 confocal laser microscope.

Yeast two-hybrid screen

The full-length *GW5* coding region was introduced into the pGBKT7 vector at the *EcoRI* and *SalI* sites. A BD Matchmaker library construction and screening kit (Clontech Laboratories, Inc.) was used for yeast two-hybrid assays. All protocols were performed according to the manufacturer's user manual. The yeast strain Y187 transformed with pGBKT7-GW5 was mated with the yeast strain AH109 transformed with a rice young spike AD fusion library and the resulting progeny were selected on SD/-Leu/-Trp/-His and SD/-Leu/-Trp/-His/-Ade/X-gal plates.

All selected pGADT7 clones were sequenced at the Invitrogen Sequencing Facility to ensure that the prey proteins were in-frame fusions with the GAL4 AD domain and that the sequences were accurate.

Accession codes

Gene bank: 9311, Exon-trapped, CL971152.

Gene bank: IR24, *GW5*, DQ991205.

Acknowledgments

This work was supported by grants from the National Basic Research Program of China (grant no. 2006CB101706) and the Hi-Tech Research and Development Program of China (863 Program) (grant no. 2006AA100101). We are grateful to Professor A Yoshimura, Kyushu University, Japan, for kindly

providing us with the CSSL populations. We thank Dr Hai Yang Wang at the Boyce Thompson Institute of Cornell University for critically reading this paper. We also thank Dr Xing Wang Deng at the Department of Molecular, Cellular and Developmental Biology of Yale University for sharing the tiling array data of the 9311 cultivar.

Note

During the preparation of our paper, we noticed a report of a newly identified QTL, *qSW5* (a QTL for seed width on chromosome 5), the deletion of which resulted in a significant increase in seed size owing to an increase in cell number in the outer glume of the rice flower. *qSW5* was mapped to the same genomic location as *GW5*. An 11.2-kb genomic fragment from the Kasalath cultivar, covering the *GW5* region, restored the slender-grain phenotype in Nipponbare [46]. These results further confirm that *qSW5* corresponds to *GW5*.

References

- Rosegrant MW, Cline SA. Global food security: challenges and policies. *Science* 2003; **302**:1917-1919.
- Luo YK, Zhu ZW, Chen N, Duan BW, Zhang LP. Grain types and related quality characteristics of rice in China. *Chin J Rice Sci* 2004; **18**:135-139.
- Xu JY, Xue QZ, Luo LJ, Li ZK. Genetic dissection of grain weight and its related traits in rice (*Oryza sativa* L.). *Chin J Rice Sci* 2002; **16**:6-10.
- Aluko G, Martinez C, Tohme J, et al. QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. *Theor Appl Genet* 2004; **109**:630-639.
- Redoña ED, Mackill DJ. Quantitative trait locus analysis for rice panicle and grain characteristics. *Theor Appl Genet* 1998; **96**:957-963.
- Huang N, Parco A, Mew T, et al. RFLP mapping of isozymes, RAPD and QTLs for grain shape, brown planthopper resistance in a doubled haploid rice population. *Mol Breeding* 1997; **3**:105-113.
- Lin HX, Min SK, Xiong ZM, et al. RFLP mapping of QTLs for grain shape traits in *indica* rice (*Oryza sativa* L. subsp. *Indica*). *Sci Agric Sin* 1995; **28**:1-7.
- Tan YF, Xing YZ, Li JX, et al. Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. *Theor Appl Genet* 2000; **101**:823-829.
- Xu JL, Xue QZ, Luo LJ, Li ZK. Genetic dissection of grain weight and its related traits in rice (*Oryza sativa* L.). *Chin J Rice Sci* 2002; **16**:6-10.
- Li JM, Xiao JH, Grandillo S, et al. QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. Sativa* L.) and African (*O. glaberrimum* S.) rice. *Genome* 2004; **47**:697-704.
- Ashikari M, Sakakibara H, Lin SY, et al. Cytokinin oxidase regulates rice grain production. *Science* 2005; **309**:741-745.
- Xue WY, Xing YZ, Weng XY, et al. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet* 2008; **40**:761-767.
- Fan CC, Xing YZ, Mao HL, et al. *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* 2006; **112**:1164-1171.
- Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet* 2007; **39**:623-630.
- Li YH, Zheng LY, Corke F, Smith C, Bevan MW. Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. *Genes Dev* 2008; **22**:1331-1336.
- Schruff MC, Spielman M, Tiwari S, Adams S, Fenby N, Scott RJ. The *AUXIN RESPONSE FACTOR 2* gene of *Arabidopsis* links auxin signaling, cell division, and the size of seeds and other organs. *Development* 2006; **133**:251-261.
- Tanabe S, Ashikari M, Fujikoka S, et al. A novel cytochrome P450 is implicated in brassinosteroid biosynthesis via the characterization of a rice dwarf mutant, *dwarf11*, with reduced seed length. *Plant Cell* 2005; **17**:776-790.
- Riefler M, Novak O, Strnad M, Schmu T. *Arabidopsis* cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell* 2006; **18**:40-54.
- Ashikari M, Wu JZ, Yano M, Sasaki T, Yoshimura A. Rice gibberellin-insensitive dwarf mutant gene *Dwarf 1* encodes the α -subunit of GTP-binding protein. *Proc Natl Acad Sci USA* 1999; **96**:10284-10289.
- Kim JY, Mahe A, Brangeon J, Prioul JL. Maize vacuolar invertase, *IVR2*, is induced by water stress organ/tissue specificity and diurnal modulation of expression. *Plant Physiol* 2000; **124**:71-84.
- Martin A, Lee J, Kichey T, et al. Two cytosolic glutamine synthetase isoforms of maize are specifically involved in the control of grain production. *Plant Cell* 2006; **18**:3252-3274.
- Kato S, Kosaka H, Hara S. On the affinity of rice varieties as shown by fertility of hybrid plants. *Bull Sci Fac Agric* 1928; **3**:132-147.
- Khush GS. Origin, dispersal, cultivation and variation of rice. *Plant Mol Biol* 1997; **35**:25-34.
- Zhu QH, Ge S. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytol* 2005; **167**:249-265.
- Ting Y. The origin and evolution of cultivated rice in China. *Acta Agron Sin* 1957; **8**:243-260.
- Doebley J, Stec A, Hubbard L. The evolution of apical dominance in maize. *Nature* 1997; **386**:485-488.
- Frery An, Nesbitt TC, Frery Am, et al. *fw2.2*: A quantitative trait locus key to the evolution of tomato fruit size. *Science* 2000; **289**:85-88.
- Liu JP, Eck JV, Cong B, Tanksley SD. A new class of regulatory genes underlying the cause of pear-shape tomato fruit. *Proc Natl Acad Sci USA* 2002; **99**:13302-13306.
- Konishi S, Izawa T, Lin SY, et al. An SNP caused loss of seed shattering during rice domestication. *Science* 2006; **312**:1392-1396.
- Li CB, Zhou AL, Sang T. Rice domestication by reducing shattering. *Science* 2006; **311**:1936-1939.
- Wan XY, Weng JF, Zhai HQ, et al. Fine mapping of the key allele *gw-5* of a grain-width QTL in a recombination hotspot region on rice chromosome 5. *Genetics* 2008; **179**:2239-2252.
- Tsunematsu H, Yoshimura A, Harushima Y, et al. RFLP framework map using recombinant inbred lines in rice. *Breed*

- Sci* 1996; **46**:279-284.
- 33 Kubo T, Nakamura K, Yoshimura A. Development of a series of *Indica* chromosome segment substitution lines in *Japonica* background of rice. *Rice Genet Newslett* 1999; **16**:104-106.
- 34 Li L, Wang XF, Stolt V, *et al.* Genome-wide transcription analyses in rice using tiling microarrays. *Nat Genet* 2006; **38**:124-128.
- 35 Harlan J. *Crops & Man*. Madison, WI: American Society of Agronomy/Crop Science Society of America, 1992.
- 36 Feng Q, Zhang Y, Hao P, *et al.* Sequence and analysis of rice chromosome 4. *Nature* 2002; **420**:316-320.
- 37 Doebley J, Lukens L. Transcriptional regulators and the evolution of plant form. *Plant Cell* 1998; **10**:1075-1082.
- 38 Van der Horst A, De Vries-Smits AMM, Brenkman AB, *et al.* FOXO₄ transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nat Cell Biol* 2006; **8**:1064-1073.
- 39 Zhang D, Zaugg K, Mak T, Elledge S. A role for the deubiquitinating enzyme USP28 in control of the DNA-damage response. *Cell* 2006; **126**:529-542.
- 40 Sridhar VV, Kapoor A, Zhang K, *et al.* Control of DNA methylation and heterochromatic silencing by histone H₂B deubiquitination. *Nature* 2007; **447**:735-738.
- 41 Dellaporta SL, Wood J, Hicks JB. A plant DNA mini preparation: Version II. *Plant Mol Biol Rep* 1983; **1**:19-21.
- 42 Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR. Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theor Appl Genet* 1997; **95**:553-567.
- 43 Sanguinetti CJ, Dias NE, Simpson AJG. Rapid silver staining and recover of PCR products separated on polyacrylamide gels. *Biotechniques* 1994; **17**:915-919.
- 44 Thiel T, Kota R, Grosse I, Stein N, Graner A. SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. *Nucleic Acids Res* 2004; **32**:e5.
- 45 Zhang QF, Shen BZ, Dai XK, Mei MH, Maroof MAS, Li ZB. Using bulked extremes and recessive class to map genes for photoperiod-sensitive genic male sterility in rice. *Proc Natl Acad Sci USA* 1994; **91**:8675-8679.
- 46 Shomura A, Izawa T, Ebana K. Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet* 2008; **40**:1023-1028.