nature genetics

Natural variation in *GS5* plays an important role in regulating grain size and yield in rice

Yibo Li, Chuchuan Fan, Yongzhong Xing, Yunhe Jiang, Lijun Luo, Liang Sun, Di Shao, Chunjue Xu, Xianghua Li, Jinghua Xiao, Yuqing He & Qifa Zhang

Increasing crop yield is one of the most important goals of plant science research. Grain size is a major determinant of grain yield in cereals and is a target trait for both domestication and artificial breeding[1]. We showed that the quantitative trait locus (QTL) *GS5* in rice controls grain size by regulating grain width, filling and weight. *GS5* encodes a putative serine carboxypeptidase and functions as a positive regulator of grain size, such that higher expression of *GS5* is correlated with larger grain size. Sequencing of the promoter region in 51 rice accessions from a wide geographic range identified three haplotypes that seem to be associated with grain width. The results suggest that natural variation in *GS5* contributes to grain size diversity in rice and may be useful in improving yield in rice and, potentially, other crops[2].

In recent years, a number of genes (or QTLs) for yield traits in rice, including tillering[3,4], number of grains per panicle[5–11] and grain weight[12–19], have been isolated using a map-based cloning approach. These genes regulate yield traits and developmental processes by functioning at various stages, in different pathways and through diverse mechanisms[1]. Molecular characterization of genes affecting grain size (such as *GS3*, *GW2* and q*SW5*/*GW5*) indicate that many are negative regulators of grain size, such that the wild-type alleles are associated with small grains, whereas the mutations are associated with large grains.

Using a double haploid (DH) population (92 lines) derived from a cross between Zhenshan 97 and H94 (both *Oryza sativa* L. *ssp. indica*), we detected a QTL, *GS5*, in the interval between two molecular markers RM593 and RM574 on the short arm of chromosome 5 (**Fig. 1**), with the allele from Zhenshan 97 contributing to an increase in yield traits (**Supplementary Table 1**). This interval is consistent with our earlier results[20–23].

To fine map the *GS5* locus, we backcrossed DH27 (a DH line containing the chromosome segment RM593–RM574 from H94 and 55% of the genetic background from Zhenshan 97) three times (BC_3) to Zhenshan 97. Self-pollinating the BC_3F_1 plants heterozygous for this fragment produced two BC_3F_2 populations consisting of 4,373 (population 1) and 5,265 (population 2) individuals, which were planted in the winters of 2005 and 2006 on Hainan Island, China. Ninety-four recombinants between RM593 and RM574 were identified in population 1, and their genotypes at the *GS5* locus were determined by progeny testing in the summer of 2006 in Wuhan (for example, see **Supplementary Table 2**). Using this data, *GS5* was mapped between C35 and RM574 (**Fig. 1b**). Analysis of population 2 detected 15 recombinants between C35 and RM574 (**Fig. 1c**). In total, 10 recombinant plants were found in the two populations in the interval between S2 and RM574. However, progeny testing revealed an inconsistency for three individuals between the marker genotypes determined using C62, which is located in the middle of the *GS5* region, and the genotype deduced by progeny testing (nos. 30, 8396 and 57-5 in **Fig. 1d** and **Supplementary Table 2**). By referencing the previously cloned *qSW5*/*GW5* locus[16,17], located approximately 2 Mb away on the RM574 side, we found that this inconsistency could be explained by the genotype of *qSW5*/*GW5*, which also affects grain size. We discarded these plants from further analysis. Using information from the remaining seven plants, we resolved *GS5* to an 11.6-kb region between RM574 and S2 (**Fig. 1c**).

We next investigated the effects of *GS5* on grain size and filling (see **Fig. 2**). We isolated two near-isogenic lines (NILs) from BC_3F_2 by fixing *qSW5*/*GW5* for the small-grain allele (**Fig. 2a**). Compared to NIL(H94), the grains of NIL(ZS97) were 8.7% wider and 7.0% heavier, leading to a 7.4% increase in grain yield per plant (**Table 1**). No significant differences were detected in other agronomic traits (**Supplementary Table 3**). An analysis of a BC_3F_2 subpopulation showed that the grain width of heterozygotes was almost the same as that for plants homozygous for the Zhenshan 97 allele (**Supplementary Table 4**), suggesting that the gene encoding the wide-grain characteristic is fully dominant.

We also compared the NILs for grain filling rate by measuring fresh and dry weight of the grains during grain filling (**Fig. 2c**,**d**). Both fresh and dry weights of NIL(ZS97) were significantly higher ($P < 0.05$) than those of NIL(H94) at 10 d after fertilization, and the differences reached a maximum ~22 d after fertilization, when fresh and dry weights of the grains of NIL(ZS97) were 11.1% and 17.8% higher than NIL(H94), respectively. Thus, the increase in grain weight and yield per plant resulted from increases in both grain width and grain filling rate.

There is only one predicted ORF in this 11.6-kb region; we identified a full-length cDNA of 1822 bp corresponding to the ORF (NCBI accession number AK106800)[24]. Alignment of the cDNA sequence with the genomic sequence of Nipponbare indicated that *GS5* consists of ten exons (**Fig. 1e**) encoding a putative serine carboxypeptidase belonging to the peptidase S10 family (see URLs) and having a

National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan, China. Correspondence should be addressed to Y.H. (yqhe@mail.hzau.edu.cn) or Q.Z. (qifazh@mail.hzau.edu.cn).

Received 28 May 2010; accepted 21 September 2011; published online 23 October 2011; doi:10.1038/ng.977

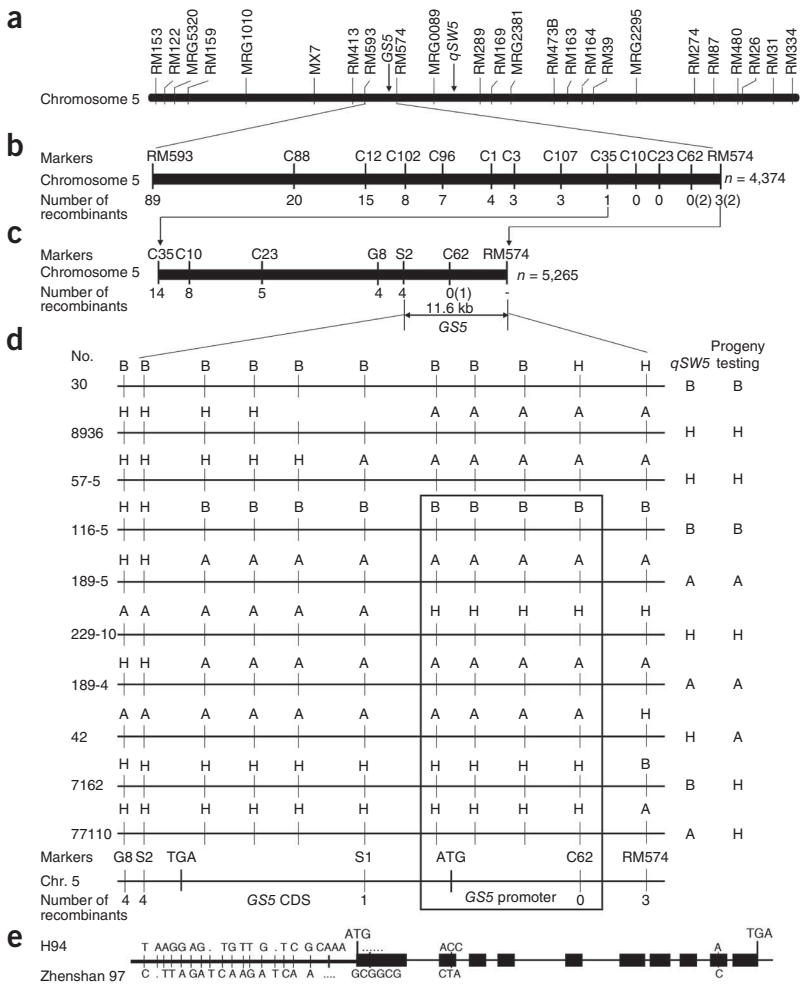

Figure 1 Map-based cloning of *GS5*. (**a**) Locations of *GS5* and *qSW5*/*GW5* in the genetic map. (**b**,**c**) Fine mapping of the *GS5* region using two mapping population, with 4,374 and 5,265 plants, respectively. The thick bar represents the genomic region; numbers underneath the bars indicate the numbers of recombinants between *GS5* and the molecular marker, and numbers in parentheses indicate the numbers of recombinants (for recombinants no. 30, 8936 and 57-5) whose phenotypes were affected by the *qSW5*/*GW5* genotype and that were not used in fine mapping. (**d**) Genotypes of the recombinants assayed by sequencing an 8-kb region between C62 and G8, including the entire coding sequence and 2-kb promoter region (only three well-spaced SNP markers in the promoter region are placed in the map). Each recombinant was phenotyped by progeny testing to deduce the genotype of *GS5* (**Supplementary Table 2**). Genotypes of *qSW5*/*GW5* was determined using two functional markers N1212 and Indel2. A, homozygous for Zhenshan 97 genotype; B, homozygous for H94 genotype; H, heterozygote; No., identification number for each recombinant. (**e**) *GS5* gene structure and natural variations between alleles from Zhenshan 97 and H94.

PF00450 consensus domain (http://pfam.sanger.ac.uk/) ($E = 5 \times 10^{-193}$). Using PROSITE analysis (see URLs), we identified an active site (ISGESYAG) characteristic of the serine carboxypeptidase. A signal peptide was predicted in the N terminus using SignalP 3.0 (see URLs) (**Supplementary Fig. 1**).

We compared genomic sequences corresponding to the ORF and the promoter regions of *GS5* between varieties with narrow and wide grains. Two varieties with narrow grains (H94 and Minghui 63; see images in **Fig. 2a**) had identical sequences, which were very different from the sequence of the wide-grain variety Zhenshan 97. The coding sequences of H94 and Minghui 63 are 1,440 bp in length, encoding a predicted polypeptide of 480 amino acids, whereas the coding sequence of Zhenshan 97 is 1,446 bp, encoding a polypeptide of 482 amino acids. Six bases are inserted 10–15 bp downstream of the translation start site in Zhenshan 97, relative to H94 and Minghui 63, resulting in an in-frame increase of 2 amino acids in the predicted signal peptide (**Fig. 1e** and **Supplementary Fig. 1**). There were also four nucleotide differences in the downstream sequence between the two varietal groups, resulting in substitutions of three amino acids (**Fig. 1e** and **Supplementary Fig. 1**). A comparison of the promoter sequences revealed 18 polymorphisms, including substitutions, deletions and insertions, in one group relative to the other in the 2-kb region upstream of the translation start site (**Fig. 1e** and **Supplementary Fig. 2**).

We assayed the temporal and spatial expression patterns of *GS5* in NIL(ZS97) and NIL(H94) using quantitative real-time PCR (qRT-PCR) with total RNA from 14 tissues (**Supplementary Fig. 3**). The levels of *GS5* transcript varied drastically among the tissues. In particular, the transcript was much more abundant in NIL(ZS97) than in NIL(H94) in the palea/lemma at 2, 4 and 5 d before heading, and in the endosperm at 10 d after fertilization. Such expression differences corresponded well with the critical stages for grain width and grain filling.

Two hypotheses emerged from the comparisons. First, it is possible that the differences in expression levels of *GS5* are attributable to polymorphisms in the promoter, leading to grain width variation.

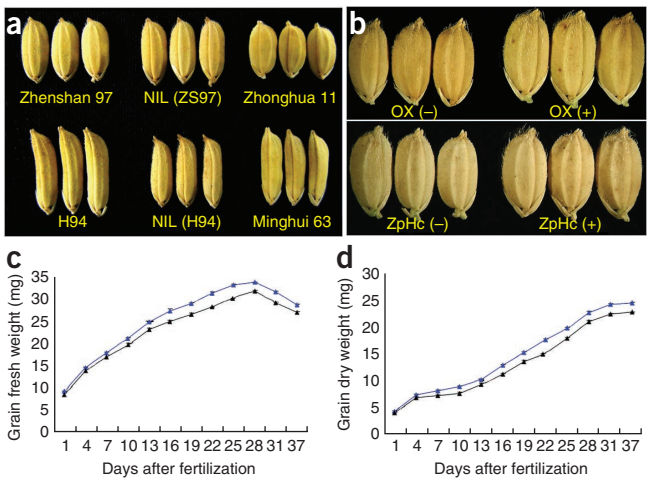

Figure 2 Effects of *GS5* on grain size and filling. (**a**) Grains of Zhenshan 97, H94, NIL(ZS97), NIL(H94), Zhonghua 11 and Minghui 63. (**b**) Grains of the transformants. OX (+) indicates grains from T_1 plants expressing the coding sequence of *GS5* from H94 driven by the 35S promoter; ZpHc (+) indicates grains from T_1 plants expressing the coding sequence of *GS5* driven by the promoter from Zhenshan 97. OX (–) and ZpHc (–) are the corresponding negative segregants. (**c**,**d**) Time-course of grain weight ($n = 90$ grains for each point). Blue line, NIL(ZS97); black line, NIL(H94).

Table 1 Grain width and weight of NILs and transgenic plants

Genotype	No. of plants	Grain width (mm)[a,b]	1,000-grain weight (g)[b,c]	Grain weight per plant (g)[b]
NIL(ZS97)[d]	55	3.12 ± 0.011	24.38 ± 0.13	19.85 ± 0.31
NIL(H94)[d]	77	2.87 ± 0.012	22.79 ± 0.17	18.49 ± 0.32
P^1		8.8×10^{-6}	5.7×10^{-6}	0.003
OX (+)[e]	18	3.70 ± 0.015	25.86 ± 0.21	
OX (–)[f]	19	3.42 ± 0.014	22.37 ± 0.22	
P^2		8.19×10^{-16}	3.26×10^{-13}	
ZpHc (+)[e]	18	3.61 ± 0.029	24.23 ± 0.37	
ZpHc (–)[f]	12	3.40 ± 0.044	21.49 ± 0.55	
P^2		9.30×10^{-4}	4.80×10^{-4}	

P^1 and P^2 are *P* values produced by the Duncan test and two-tailed *t*-test, respectively. [a]At least 30 grains of each plant were measured for grain width. [b]All data are given as means ± s.e.m. [c]At least 200 grains of each plant were measured for 1,000-grain weight. [d]Data for the NILs are based on a field experiment using a randomized complete block design with three replications. [e](+) indicates transgene-positive T_0 plants. [f](–) indicates transgene-negative T_0 plants.

Alternatively, it may be that the grain size differences are attributable to coding variation. We used two approaches to test these hypotheses. First, we sequenced an ~8-kb fragment containing the entire coding region and a 2-kb fragment 5′ (upstream) of the translation start site from the seven recombinant plants that we used in fine mapping (**Fig. 1d**). Genetic polymorphisms in the promoter region, not the coding region, corresponded well with the phenotypes of the progeny test, which is consistent with the hypothesis that the effect on grain size of *GS5* is due to variation in the promoter region.

In the second approach, two transformation constructs were prepared. The first, OX, contained the cDNA of *GS5* from H94 (narrow grain) driven by the 35S promoter; in the second construct, ZpHc, the 2-kb promoter fragment of *GS5* from Zhenshan 97 was fused with the cDNA from H94. Because of the difficulty in regenerating plants from the calli of both Zhenshan 97 and H94, an *Oryza sativa* L. *ssp. japonica* variety, Zhonghua 11 (with a grain even wider than Zhenshan 97; see **Fig. 2a**), was used for transformation.

A total of 37 transgenic plants (T_0) were generated using OX of which 18 were transgene positive, whereas the other 19 were negative. We observed an increase in grain size and weight in the transgene-positive plants, compared to the transgene-negative plants (**Fig. 2b** and **Table 1**), with no effects on plant morphology, plastochron or flowering time (**Supplementary Fig. 4**). Co-segregation tests of T_1 progeny from two T_0 plants showed that grain width is significantly ($P < 0.01$) correlated with the expression level of *GS5* (**Supplementary Table 5**); plants with higher expression levels produced wider grains.

Thirty transgenic plants were obtained using ZpHc; transgene-positive plants also showed an increase in grain width and grain weight (**Fig. 2b** and **Table 1**). Again, the increase in grain size was significantly correlated with the expression level of *GS5* ($P < 0.001$; **Supplementary Table 6**). This is quite similar to the high (though negative) correlation between transcript abundance and fruit mass resulting from a dosage effect of *fw2.2* reported in tomato[25].

Taken together, results from the sequencing, qRT-PCR and transformation studies suggest that polymorphisms in the *GS5* promoter region may be the cause of the *GS5* effect on grain width. In addition, we obtained a mutant (3C-00077) from a T-DNA insertion library using an *O. sativa* L *ssp. japonica* variety, Dongjin (see URLs) [26,27], in which the T-DNA was inserted into the fourth intron of the gene (**Supplementary Fig. 5a**). To eliminate possible variations in the genetic background arising in tissue culture of this insertion line, we crossed and backcrossed this line with the wild-type parent to produce BC_1F_2 progeny. As expected, progeny carrying the *gs5* mutation produced smaller grains than wild type, and the grain size co-segregated with the T-DNA insertion and the lower expression level of *GS5* (**Supplementary Fig. 5b–d**).

We next sequenced the *GS5* promoter regions (~2 kb, including the 6-base deletion in the first exon) and measured grain width in a total of 51 rice accessions, including 35 cultivated varieties of *O. sativa*, mostly from China, and 16 of the wild rice *Oryza rufipogon*, originating from a wide geographic range across Asia (**Supplementary Table 7**). The sequences of the cultivated varieties could be divided into three haplotypes: H94 type (narrow grain), Zhenshan 97 type (medium grain) and Zhonghua 11 type (wide grain (**Supplementary Table 8**)). Accessions within each group had exactly the same sequences. There were 18 polymorphic sites between H94 and Zhenshan 97; Zhonghua 11 differed from H94 and Zhenshan 97 by 22 and 26 sites, respectively (**Supplementary Fig. 2**). Six wild-rice narrow-grain accessions of disparate geographic origins had exactly the same sequence as H94. One wild-rice accession (IRGC 103823) from China differed by two nucleotides from Zhenshan 97. The remaining 9 accessions had little similarity to any of the three types.

We examined cross-sections of the central parts of the palea/lemma of the spikelet between NIL(ZS97) and NIL(H94) (**Fig. 3a–c**). The inner parenchyma cell layer of NIL(ZS97) contained a substantially greater number of cells than did NIL(H94) (**Fig. 3d**). In addition, the cell size of palea in NIL(ZS97) was also larger than in NIL(H94) (**Fig. 3e**). These results suggest that *GS5* positively regulates grain size by increasing cell number and also, to some extent, cell size, leading to enhanced latitudinal growth in the grain.

Cell division in eukaryotic organisms is controlled by a highly conserved basic cell cycle machinery[28–32]. To investigate the relationship of *GS5* with genes regulating the plant cell cycle, we searched the Rice Genome Browser database (see URLs) by querying sequences of all previously published core cell cycle regulators with a cutoff *E*-value of 5×10^{-5}. This search identified 25 genes, including 14 putatively involved in the G_1/S and 11 in the G_2/M phase (**Supplementary Table 9**). We analyzed expression of these 25 genes in plants carrying the *gs5* mutation and in plants overexpressing *GS5* relative to the wild type. The transcript levels of the five putative G_1/S-phase genes, *CDKA1*, *CAK1*, *CAK1A*, *CYCT1* and *H1*, were greatly elevated in plants overexpressing *GS5* as compared to plants not carrying the transgene (**Fig. 4a**). In contrast, the expression of these five genes was significantly reduced in the *gs5* mutant, relative to wild type (**Fig. 4b**). Thus, *GS5* functions putatively as a positive modulator upstream of cell cycle genes, and its overexpression may result in an increase in cell numbers by promoting mitotic division. This feature of *GS5* is in contrast to those of some

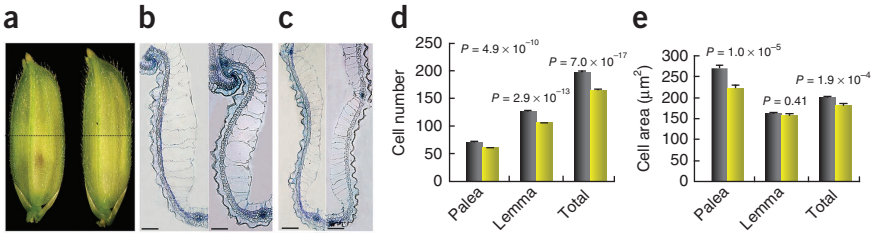

Figure 3 The effect of *GS5* on cell number and size in lemma/palea. (**a**) Spikelets of NIL(ZS97) (left) and NIL(H94) (right) 4 d before heading. (**b**,**c**) Cross-sections of palea (**b**) and lemma (**c**) cut horizontally at the middle of the spikelets shown in **a**. Scale bars, 200 μm for both **b** and **c**. (**d**,**e**) Comparisons of cell number (**d**) and cell size (**e**) between NIL(ZS97) and NIL(H94) in the cross-sections of the inner parenchyma cell layer of spikelets. All *P* values are based on two-tailed *t*-tests. Black bars, NIL(ZS97); yellow bars, NIL(H94). Error bars, s.e.m.

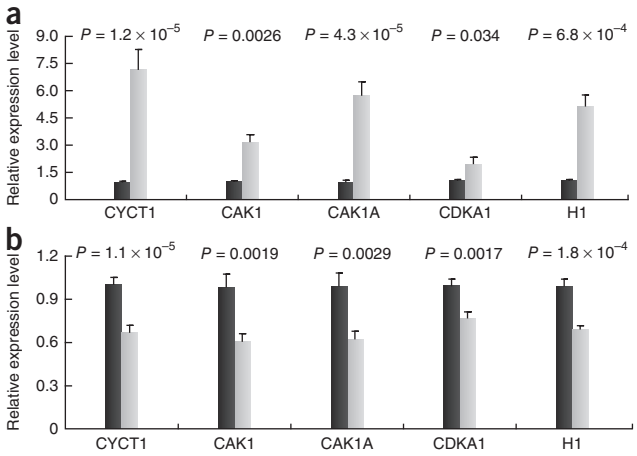

Figure 4 Regulation by *GS5* of the expression of genes involved in the cell cycle. (**a**) Transcript levels of genes associated with cell cycle regulation in *GS5* overexpressor OX(+) relative to negative segregants OX(−). Black bars, OX(−); light bars, OX(+). (**b**) Transcript levels of genes associated with cell cycle regulation in the *gs5* mutant, relative to wild type. Black bars, wild type; gray bars, mutant. Expression levels were determined by qRT-PCR using 6- to 8-cm young panicles from at least five plants, in at least three biological samples and three replicates. Error bars, s.e.m. All *P* values are based on two-tailed *t*-tests.

previously reported genes affecting grain size in rice[12–17] and fruit size in tomato[33,34] that function as negative regulators of the grain or fruit size. Elevated expression of *GS5* appears to be associated with a larger grain, whereas loss of function or reduced expression of these previously reported genes is associated with larger grain or fruit.

URLs. Peptidase S10 family, http://merops.sanger.ac.uk/cgi-bin/famsum?family=S10; *O. sativa* L *ssp. japonica* variety Dongjin, http://www.postech.ac.kr/life/pfg/risd/; Rice Genome Browser, http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/; PROSITE database of protein domains, families and functional sites, http://prosite.expasy.org/; SignalP 3.0, http://www.cbs.dtu.dk/services/SignalP-3.0/; Rice T-DNA Insertion Sequence Database (RISD), http://www.postech.ac.kr/life/pfg/risd/; ImageJ, http://rsb.info.nih.gov/ij/.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Accession codes. GenBank: cDNA of *GS5* in H94, JN256055; cDNA of *GS5* in Zhenshan 97, JN256056; genomic DNA of *GS5* in H94, JN256057; genomic DNA of *GS5* in Zhenshan 97, JN256058.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank J. Hu, H. Du, W. Dong and G. Gao for their help and suggestions, and G. An (Kyung Hee University, Crop Biotech Institute, Republic of Korea) for the mutant line. This work was supported by grants from the National 863 Project, the National Program on the Development of Basic Research, the National Program on R&D of Transgenic Plants, the National Natural Science Foundation of China and the Bill & Melinda Gates Foundation.

AUTHOR CONTRIBUTIONS

Y.L. conducted most of the experiments, including fine mapping, gene cloning, genetic transformation, expression analysis, mutant analysis, histological analysis and other functional analysis; C.F., Y.X. and L.L. conducted the QTL primary mapping analysis and developed the NILs; Y.J. and L.S. carried out part of the cell division and expression analysis; D.S., C.X., X.L. and J.X. participated in the promoter sequencing; Y.H. and Q.Z. designed and supervised the study; and Y.L. and Q.Z. analyzed the data and wrote the paper. All of the authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Xing, Y. & Zhang, Q. Genetic and molecular bases of rice yield. *Annu. Rev. Plant Biol.* **61**, 421–442 (2010).
2. Zhang, Q. Strategies for developing green super rice. *Proc. Natl. Acad. Sci. USA* **104**, 16402–16409 (2007).
3. Li, X. *et al.* Control of tillering in rice. *Nature* **422**, 618–621 (2003).
4. Takeda, T. *et al.* The *OsTB1* gene negatively regulates lateral branching in rice. *Plant J.* **33**, 513–520 (2003).
5. Ashikari, M. *et al.* Cytokinin oxidase regulates rice grain production. *Science* **309**, 741–745 (2005).
6. Xue, W. *et al.* Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* **40**, 761–767 (2008).
7. Li, S. *et al. Short panicle1* encodes a putative PTR family transporter and determines rice panicle size. *Plant J.* **58**, 592–605 (2009).
8. Jiao, Y. *et al.* Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* **42**, 541–544 (2010).
9. Miura, K. *et al. OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* **42**, 545–549 (2010).
10. Wei, X. *et al. DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol.* **153**, 1747–1758 (2010).
11. Yan, W.H. *et al.* A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in Rice. *Mol. Plant* **4**, 319–330 (2011).
12. Fan, C. *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.* **112**, 1164–1171 (2006).
13. Takano-Kai, N. *et al.* Evolutionary history of *GS3*, a gene conferring grain size in rice. *Genetics* **182**, 1323–1334 (2009).
14. Mao, H. *et al.* Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. USA* **107**, 19579–19584 (2010).
15. Song, X., Huang, W., Shi, M., Zhu, M. & Lin, H.A. QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630 (2007).
16. Shomura, A. *et al.* Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* **40**, 1023–1028 (2008).
17. Weng, J. *et al.* Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res.* **18**, 1199–1209 (2008).
18. Wang, E. *et al.* Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat. Genet.* **40**, 1370–1374 (2008).
19. Wu, C. *et al.* Brassinosteroids regulate grain filling in rice. *Plant Cell* **20**, 2130–2145 (2008).
20. Li, J. *et al.* 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 (2000).
21. Tan, Y. *et al.* Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. *Theor. Appl. Genet.* **101**, 823–829 (2000).
22. Xing, Y. *et al.* 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 (2002).
23. Yu, S. *et al.* Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* **94**, 9226–9231 (1997).
24. Kikuchi, S. *et al.* Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* **301**, 376–379 (2003).
25. Liu, J., Cong, B. & Tanksley, S. Generation and analysis of an artificial gene dosage series in tomato to study the mechanisms by which the cloned quantitative trait locus *fw2.2* controls fruit size. *Plant Physiol.* **132**, 292–299 (2003).
26. Jeong, D. *et al.* Generation of flanking sequence-tag database for activation-tagging lines in *japonica* rice. *Plant J.* **45**, 123–132 (2006).
27. Jeon, J. *et al.* T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* **22**, 561–570 (2000).
28. Francis, D. The plant cell cycle-15 years on. *New Phytol.* **174**, 261–278 (2007).
29. Inzé, D. Green light for the cell cycle. *EMBO J.* **24**, 657–662 (2005).
30. Inzé, D. & De Veylder, L. Cell cycle regulation in plant development. *Annu. Rev. Genet.* **40**, 77–105 (2006).
31. De Veylder, L., Beeckman, T. & Inzé, D. The ins and outs of the plant cell cycle. *Nat. Rev. Mol. Cell Biol.* **8**, 655–665 (2007).
32. Dewitte, W. & Murray, J. The plant cell cycle. *Annu. Rev. Plant Biol.* **54**, 235–264 (2003).
33. Frary, A. *et al. fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85–88 (2000).
34. Cong, B., Liu, J. & Tanksley, S. Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. *Proc. Natl. Acad. Sci. USA* **99**, 13606–13611 (2002).

ONLINE METHODS

Primers. All primers used in this study are listed in **Supplementary Tables 9** and **10**.

Field planting and trait measurement. Rice plants were examined under natural field conditions in the Experimental Stations of Huazhong Agricultural University, Wuhan and Hainan, China. The planting density was 16.5 cm between plants in a row, and the rows were 26 cm apart. Field management, including irrigation, fertilizer application and pest control, followed essentially the normal agricultural practice. Harvested rice grains were air-dried and stored at room temperature for at least 3 months before testing. Fully filled grains were used for measuring grain width, length and weight. Ten randomly chosen grains from each plant were lined up width-wise along a vernier caliper to measure grain width and then arranged length-wise to measure grain length. Grain weight was calculated based on 200 grains and converted to 1,000-grain weight.

RNA preparation and reverse transcription. Total RNA was extracted from various plant tissues using a RNA extraction kit (TRIzol reagent, Invitrogen). The first-strand cDNA was synthesized using 2 μg RNA and 200 U M-MLV reverse transcriptase (Promega Kit) in a volume of 20 μl. The full-length *GS5* cDNA (1,443 bp) was amplified from the first-strand cDNA of H94 and confirmed by sequencing.

Constructs and transformation. For preparing the overexpression construct (OX), the full-length *GS5* cDNA of H94 was inserted into the plant binary vector pCAMBIA1301S (Cambia), in which transgene expression was under the control of the CaMV 35S promoter. For preparing the chimeric construct (ZpHc) in which the 2-kb promoter fragment of *GS5* from Zhenshan 97 was fused with the cDNA from H94, *GS5* promoter region was obtained by PCR from Zhenshan 97 and confirmed by sequencing. The promoter fragment was ligated with the cDNA from H94 and then inserted into the plant binary vector pCAMBIA1301 (Cambia). The constructs were introduced into *Agrobacterium tumefaciens* strain EHA105 and transferred into Zhonghua 11 by *Agrobacterium*-mediated transformation as previously described[35], with minor modifications.

Real-time PCR. qRT-PCR was carried out in a total volume of 25 μl containing 2 μl of reverse-transcribed product (see RNA preparation and reverse transcription, above), 0.2 mM gene-specific primers and 12.5 μl *SYBR Premix EX Taq* and 0.5 μl of Rox Reference Dye II (Takara Kit), using an ABI 7500 Real-Time PCR System according to the manufacturer's instructions. The measurements were obtained using the relative quantification method[36]. A rice *Actin1* gene was used as the internal control. Each measurement was determined in at least two biological samples and three replicates for each sample.

Histological analysis. Spikelet hulls, fresh-picked 4 d before heading, were fixed in FAA solution (50% (v/v) ethanol, 5% (v/v) glacial acetic acid and 5% (v/v) formaldehyde) for more than 16 h and then transferred to freshly prepared 70% (v/v) ethanol four times within 24 h to wash out the FAA solution. Hulls were then placed in 70% ethanol for long-term storage. Tissue sections (<8 mm thick) were cut with a rotary microtome, mounted and stained with toluidine blue and fast green. Sections were photographed under a light microscope (Leica DM 4000B) with a DFC480 CCD camera. The number of cells in each sample were counted, and the area of the cells in each sample was obtained using ImageJ. The mean and s.e.m. were calculated from the measurements of more than 20 hulls for each NIL.

35. Hiei, Y., Ohta, S., Komari, T. & Kumashiro, T. Efficient transformation of rice (*Oryza sativa L.*) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**, 271–282 (1994).
36. Livak, K. & Schmittgen, T. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408 (2001).