nature genetics

A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase

Xian-Jun Song[1,2], Wei Huang[1,2], Min Shi[1], Mei-Zhen Zhu[1] & Hong-Xuan Lin[1]

Grain weight is one of the most important components of grain yield and is controlled by quantitative trait loci (QTLs) derived from natural variations in crops. However, the molecular roles of QTLs in the regulation of grain weight have not been fully elucidated. Here, we report the cloning and characterization of *GW2*, a new QTL that controls rice grain width and weight. Our data show that *GW2* encodes a previously unknown RING-type protein with E3 ubiquitin ligase activity, which is known to function in the degradation by the ubiquitin-proteasome pathway. Loss of GW2 function increased cell numbers, resulting in a larger (wider) spikelet hull, and it accelerated the grain milk filling rate, resulting in enhanced grain width, weight and yield. Our results suggest that *GW2* negatively regulates cell division by targeting its substrate(s) to proteasomes for regulated proteolysis. The functional characterization of *GW2* provides insight into the mechanism of seed development and is a potential tool for improving grain yield in crops.

Given the rapid increase in world population, the next century may witness serious global food shortage problems. Consequently, there is a need for an increase in crop grain yield. Many important complex traits in crops, including yield and stress tolerance, are controlled by QTLs derived from natural variations[1,2]. Recent studies have succeeded in isolating and characterizing genes involved in QTLs using map-based cloning techniques[3–10].

Rice (*Oryza sativa* L.), a staple food, is the world's most important cereal crop. Grain weight, number of grains per panicle and number of panicles per plant are the most important components of grain yield. However, thus far, only one QTL for yield in rice has been cloned and functionally characterized: *Gn1a*, a QTL for number of grains per panicle[9]. Grain weight is usually represented by 1,000-grain weight in breeding applications and is determined by grain width, length and thickness[11,12]. Many QTLs for rice grain weight have been mapped in the last decade[13–20]. One of them, *GS3*, a major QTL for grain length and weight and also a minor QTL for grain width and thickness in rice, has been mapped on chromosome 3, and the candidate gene underlying this QTL has been identified[18], although functional analysis has not yet been reported. In the current study, we cloned and characterized a gene underlying a major QTL for rice grain width and weight, elucidating the molecular mechanisms that regulate grain weight and providing critical information for breeding high-yield crops through genetic engineering.

RESULTS

Map-based cloning of the *GW2* QTL

We chose parental varieties that showed highly significant differences in grain size to more easily identify the QTL. We crossed a *Japonica* variety, WY3, with a very large grain (1,000-grain weight, 41.9 g ± 1.3 g) and a high-quality elite *indica* variety, Fengaizhan-1 (FAZ1), with a small grain (1,000-grain weight, 17.9 ± 0.7 g) (**Fig. 1a**) to produce an F_2 population. Using the F_2 population, we mapped several QTLs for grain weight or size, including width, length, thickness and 1,000-grain weight (X.J. Song and H.-X.L., unpublished data). In particular, we mapped a major QTL for grain width, *GW2*, on chromosome 2 (**Fig. 1b**), with the WY3 allele at *GW2* contributing to increased grain width. Previous studies have mapped two QTLs for grain width near the *GW2* region on the short arm of chromosome 2 (refs. 19,20), suggesting they may be the same QTL and that the *GW2* locus may contribute to increased grain width in the various rice varieties. Consequently, we selected *GW2* as the target for map-based cloning.

We performed fine mapping using a BC_2F_2 population and mapped *GW2* between the markers W236 and W239 (**Fig. 1b**). We carried out further high-resolution mapping of *GW2* using 6,013 BC_3F_2 plants and newly developed markers between W236 and W239 (**Fig. 1c**). We localized *GW2* to a high-resolution linkage map by progeny testing of homozygous recombinant plants (BC_3F_4; **Fig. 1d**) and narrowed the *GW2* locus to an 8.2-kb region between markers W024 and W004

[1]National Key Laboratory of Plant Molecular Genetics, Shanghai Institute of Plant Physiology and Ecology, Shanghai Institute for Biological Sciences, The Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China. [2]These authors contributed equally to this work. Correspondence should be addressed to H.-X.L. (hxlin@sibs.ac.cn).

Received 5 December 2006; accepted 26 February 2007; published online 8 April 2007; doi:10.1038/ng2014

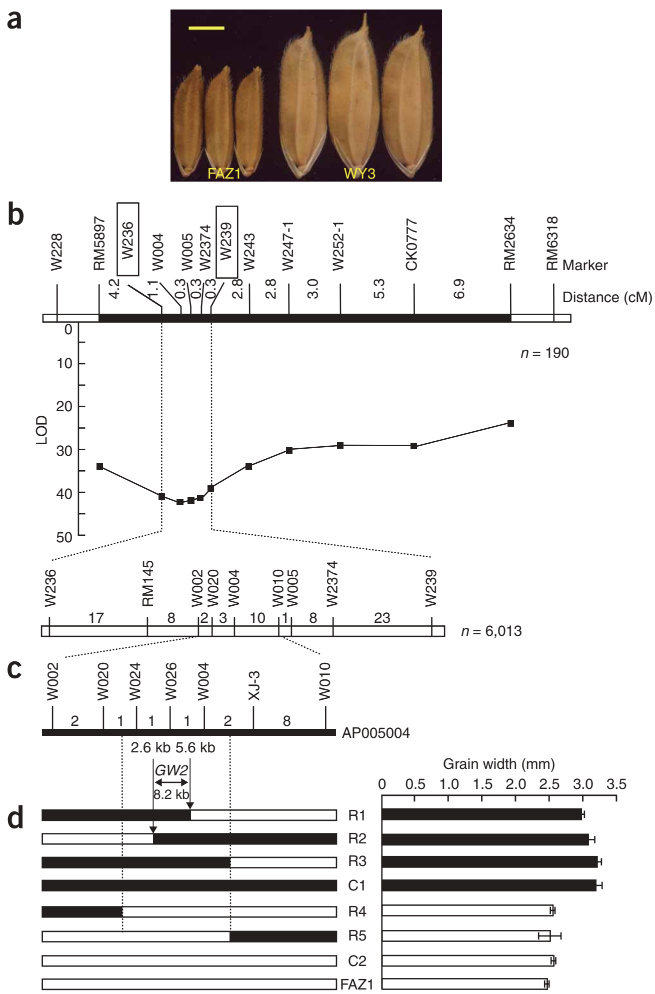


Figure 1 Map-based cloning of *GW2*. (**a**) Grain phenotypes of parents (FAZ1 and WY3). Scale bar, 3 mm. (**b**) Location of *GW2* on rice chromosome 2 in 190 BC_2F_2 plants in which a small region (filled bar) is segregating. (**c**) High-resolution linkage map of the *GW2* region produced with 6,013 BC_3F_2 plants. The number of recombinants between the adjacent markers is indicated above the linkage map. Filled bar shows a part of the PAC clone AP005004. (**d**) Progeny testing of fixed recombinant plants (BC_3F_4) narrowed the *GW2* locus to the region between markers W024 and W004. Grain widths (mean ± s.d.; n = 12 plants) of three recombinant lines (R1–R3) and control 1 (C1; homozygous for WY3 in the target region) were higher than those of the two recombinant lines (R4 and R5), control 2 (C2; homozygous for FAZ1 in the target region) and FAZ1. Filled and open bars represent homozygous chromosomal segments for WY3 and FAZ1, respectively.

(**Fig. 1d**). In this region, we identified only one predicted ORF considered a viable candidate for *GW2*.

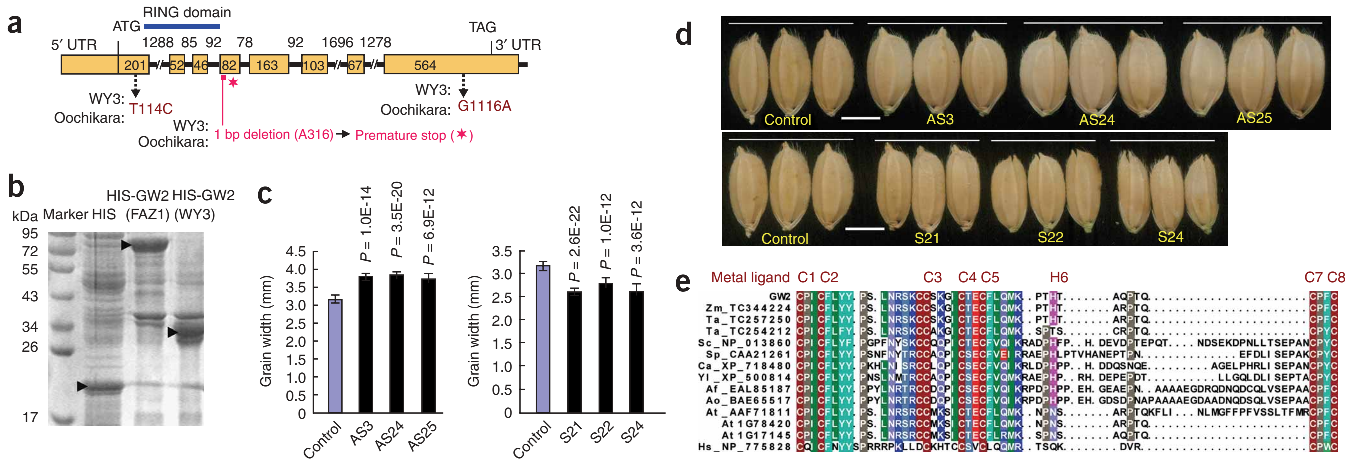
GW2 sequence analysis and natural variation

Through genomic and cDNA sequence comparison of the parent FAZ1 allele of *GW2*, we identified eight exons and seven introns (**Fig. 2a**). The FAZ1 *GW2* allele was predicted to encode a 425-residue polypeptide of ∼47 kDa. Comparison of the nucleotide sequences of the FAZ1 and WY3 alleles of *GW2* uncovered three nucleotide changes, including two nucleotide substitutions in exons 1 and 8 that did not result in any amino acid variations and a 1-bp deletion resulting in a premature stop codon in exon 4 of the WY3 allele (**Fig. 2a**). The premature stop codon led to truncation of 310 amino acid residues; the remaining portion of the protein consisted of a 115-residue polypeptide of ∼13 kDa. This predicted result was confirmed by SDS–PAGE of histidine-tagged GW2 proteins encoded by the FAZ1 and WY3 alleles (**Fig. 2b**). We also analyzed the nucleotide sequence of *GW2* from Oochikara, another rice variety that has a wider grain width, similar to WY3 (**Supplementary Fig. 1** online), and we found a sequence identical to the WY3 allele (**Fig. 2a**). These data indicate that reduction or loss of function of *GW2* results in increased grain width.

We used rice transformation to produce transgenic plants expressing different levels of *GW2*. Because FAZ1 and WY3 were unable to regenerate shoots from the callus, we chose the easily regenerable *Japonica* variety Zhonghua 11, which has a small grain, for the transformation[8]. We generated *35S:: GW2* antisense lines expressing *GW2* cDNA (from FAZ1) in the antisense direction, as *GW2* cDNA (the entire ORF) did not show any significant sequence homology with any other sequences in the rice genome. The transgenic plants with antisense strands of *GW2* and with reduced levels of endogenous expression had a significantly wider grain width than plants containing the vector control (**Fig. 2c,d** and **Supplementary Fig. 2** online). However, we observed reduced grain width in transgenic plants overexpressing *GW2* cDNA under the control of the 35S promoter, which produced high levels of expression (**Fig. 2c,d** and **Supplementary Fig. 2**). These data suggested that the cDNA from the FAZ1 allele represented the coding region of the *GW2* QTL for grain width.

GW2 encodes a RING protein with E3 ubiquitin ligase activity

A search of the National Center for Biotechnology Information (NCBI) database and the Institute for Genomic Research (TIGR) Gene Indices database identified one GW2 homolog in *Zea mays*, two in *Triticum aestivum*, three in *Arabidopsis thaliana*, six in yeast or fungi and one in human, which were all annotated as proteins of unknown function (**Fig. 2e**); we did not find GW2 homologs in rice. Whereas the homologous genes in *Zea mays* and *Triticum aestivum* (TC257250) had high amino acid sequence identities with GW2 (81% and 86.5%, respectively), the one in *Triticum aestivum* (TC254212) had only a 43.5% amino acid sequence identity with GW2. Homologous genes in *A. thaliana* showed a 39%–45% amino acid sequence identity with GW2, and the remaining seven homologous genes in other species had only 11%–25% amino acid sequence identity. None of the homologous genes has been functionally characterized. An analysis of the conserved domain demonstrated that GW2 might share homology with the consensus sequence of the RING-type protein (**Fig. 2e**). The cysteine-rich RING motif was first identified in a protein encoded by the gene *Really Interesting New Gene* (hence the motif name)[21]. RING-type proteins are widely present in animals, plants, yeast, fungi and viruses and are predicted to contain a large diversity of RING domains[22–24]. To date, seven types of RING domain have been identified in *A. thaliana*, including two canonical RING types, RING-HC (C3HC4) and RING-H2 (C3H2C3), and five modified RING domain types, RING-v, RING-C2, RING-D, RING-S/T and RING-G[24–27]. Notably, the GW2 RING-like domain did not correspond to any of the previously described RING domains and is characterized by a cysteine residue at metal ligand position 5 and a histidine residue at metal ligand position 6 (C5HC2) (**Fig. 2e**), rather than the configuration found in the RING-v domain (C4HC3). This feature was also identified in the RING domain of maize, wheat, yeast and fungal homologs (**Fig. 2e**), indicating that GW2 probably represents a new type of RING domain protein in plants. Although



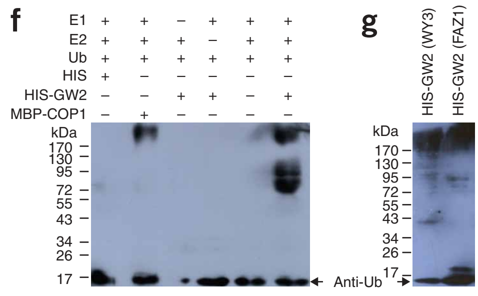


Figure 2 Structure and molecular characterization of *GW2*. (**a**) *GW2* structure and mutation sites, including nucleotide substitutions and deletions, in WY3 and Oochikara. (**b**) SDS-PAGE of GW2 proteins of FAZ1 and WY3. Arrows indicate the HIS-GW2 protein of FAZ1 or WY3 and HIS tag alone. (**c**,**d**) Grain width (**c**) and grain shape (**d**) in transgenic plants compared with control grains. The control plants were a transgenic line carrying an empty vector. AS3, AS24 and AS25 are transgenic lines carrying antisense *GW2* cDNA from FAZ1, and S21, S22 and S24 are transgenic lines carrying sense *GW2* cDNA from FAZ1. Data are mean ± s.d. ($n = \sim 9$ to ~ 20) in **c**. A Student's *t*-test was used to generate the *P* values in **c**. Scale bar in **d** represents 3 mm. (**e**) Amino acid alignments of the GW2 RING domain with 13 RING-like domains from other species. Conserved metal ligand positions are indicated by numbered cysteine (C) or histidine (H) residues. Zm, *Zea mays*; Ta, *Triticum aestivum*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe*; Ca, *Candida albicans*; Yl, *Yarrowia lipolytica*; Af, *Aspergillus fumigatus*; Ao, *Aspergillus oryzae*; At, *Arabidopsis thaliana*; Hs, human. (**f**) E3 ubiquitin ligase activity of GW2. HIS-GW2 (from FAZ1) and MBP-COP1 fusion proteins were assayed for E3 activity in the presence of E1, E2 (UbcH5b) and ubiquitin (Ub). *A. thaliana* E3 ubiquitin ligase COP1 was used as a positive control and HIS as negative control. Ubiquitinated proteins were detected by protein blot analysis using an antibody to ubiquitin. (**g**) E3 ubiquitin ligase activity of WY3 GW2 (lane 1).

RING-like domains of a homologous gene in *Triticum aestivum* (TC254212) and three homologous genes in *A. thaliana* encode proteins that share high amino acid sequence identity with GW2, they might not be considered RING domain proteins, as they lacked one metal ligand amino acid (a histidine residue). Similarly, the RING-like domain of the human homolog, which lacks the same amino acid, also cannot be categorized as a RING domain.

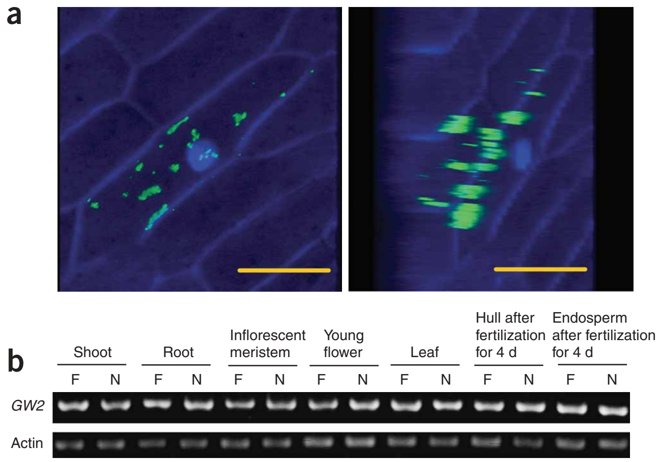
Several studies have reported that RING-type proteins function as E3 ubiquitin ligases in *vitro*, targeting proteins for degradation[24,28–31]. We expressed affinity-purified GW2 protein as a fusion with a six-histidine tag (HIS-GW2) in *Escherichia coli* to investigate whether GW2 also had E3 activity (**Fig. 2f**). *A. thaliana* E3 ubiquitin ligase constitutive photomorphogenic 1 (COP1) was used as a positive control[31]. In the presence of ubiquitin, E1 and E2 (UbcH5b), HIS-GW2 (from FAZ1) or maltose-binding protein (MBP)-COP1 could carry out self-ubiquitination, whereas in the absence of any of the E1, E2 or E3 enzymes, we did not detect any clear protein ubiquitination (**Fig. 2f**). These results demonstrate that GW2 possesses intrinsic E3 ligase activity. Although the WY3 GW2 protein was truncated by 310 amino acids, it possessed an intact RING domain (**Fig. 2a**), and we expected it to have E3 activity, which we confirmed by ubiquitination assays (**Fig. 2g**).

Subcellular localization and tissue localization of *GW2*

To investigate the subcellular localization of GW2, we constructed a green fluorescent protein (GFP)-GW2 fusion whose expression was driven by the CaMV 35S promoter. Transient expression in onion epidermal cells showed that GFP-GW2 localized to the cytoplasm (**Fig. 3a**). We examined temporal and spatial expression patterns of *GW2* by RT-PCR and by analysis of transgenic rice plants expressing the *GW2* promoter–*GFP* transgene. RT-PCR data in both FAZ1 and NIL(*GW2*) showed that *GW2* mRNA was expressed constitutively in shoots and roots of seedlings, inflorescent meristems, young flowers, leaves and spikelet hulls and endosperms 4 d after fertilization (**Fig. 3b**). We did not observe any differences in *GW2* expression between FAZ1 and NIL(*GW2*), suggesting that the sequence change in the coding region accounted for the functional variation in the two alleles (FAZ1 and WY3). The *GW2* promoter–GFP expression analysis showed that GFP was strongly expressed in roots (**Fig. 3c**), leaves (**Fig. 3d**) and floral organs including stamens, pistils and hulls (**Fig. 3e**). Thus, our *GW2* transcript and *GW2* promoter-GFP expression analysis demonstrates that *GW2* is expressed constitutively in various tissues and organs.

GW2 increases grain width and weight and affects other traits

To investigate the effects of *GW2* on the rice grain, we bred a nearly isogenic line, NIL(*GW2*), on the FAZ1 genetic background, containing a very small *GW2* region (1.4 cM, **Fig. 1b**) between markers W236 and W005 from WY3. We observed a substantial increase in grain width (+26.2%) in NIL(*GW2*) but only a slight increase grain thickness (+10.5%) and grain length (+6.6%) (**Fig. 4a,b**) compared with the FAZ1 isogenic control. We also observed a significant increase (+49.8%) in 1,000-grain weight in NIL(*GW2*) (**Fig. 4b**). These results indicated that the increase grain weight in NIL(*GW2*) was primarily due to increased grain width, followed by grain thickness and length. The data indicated that the *GW2* allele from WY3 with the natural mutation predominantly increases grain width and weight but increases grain thickness and length only slightly.



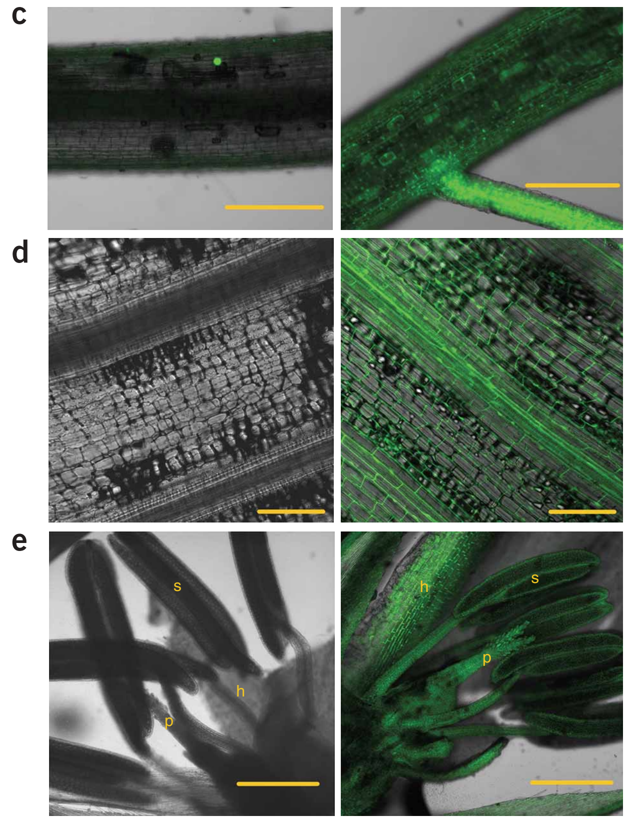


Figure 3 Cellular properties and expression pattern of *GW2*. (**a**) *GW2* was localized to the cytoplasm. The DAPI was used to stain the nucleus of onion epidermal cells. Left: three-dimensional reconstruction of individual image stacks. Right: three-dimensional image rotated 60° along its *y*-axis. Scale bar, 100 μm. (**b**) Expression of *GW2* analyzed using RT-PCR in various organs. F, FAZ1; N, NIL(*GW2*). Actin was used as a control. (**c**–**e**) *GW2* promoter–GFP expression pattern in transgenic rice plants. Left, empty vector plants; right, *GW2* promoter–GFP plants. There was strong GFP expression in roots (**c**), leaves (**d**) and floral organs, including stamens, pistils and hulls (**e**). s, stamen; p, pistil; h, hull. Scale bars: 300 μm (**c**), 100 μm (**d**) and 500 μm (**e**).

To test whether *GW2* affects grain yield, we compared the grain yields of FAZ1 and NIL(*GW2*). We found that the grain yield per plant of NIL(*GW2*) increased by 19.7% ($P < 0.05$), although the number of grains per main panicle was 29.9% lower in NIL(*GW2*) than in FAZ1 (**Fig. 4c**). Additionally, the plant type of NIL(*GW2*) was similar to FAZ1, an elite variety (**Fig. 4c**). Although the grain yield per plant measured in plants grown in paddies under normal cultivation conditions (25 plants) showed that NIL(*GW2*) has the potential to increase grain yield, we still need to carefully evaluate the potential of NIL(*GW2*) for higher grain yield in plots with randomized blocks in

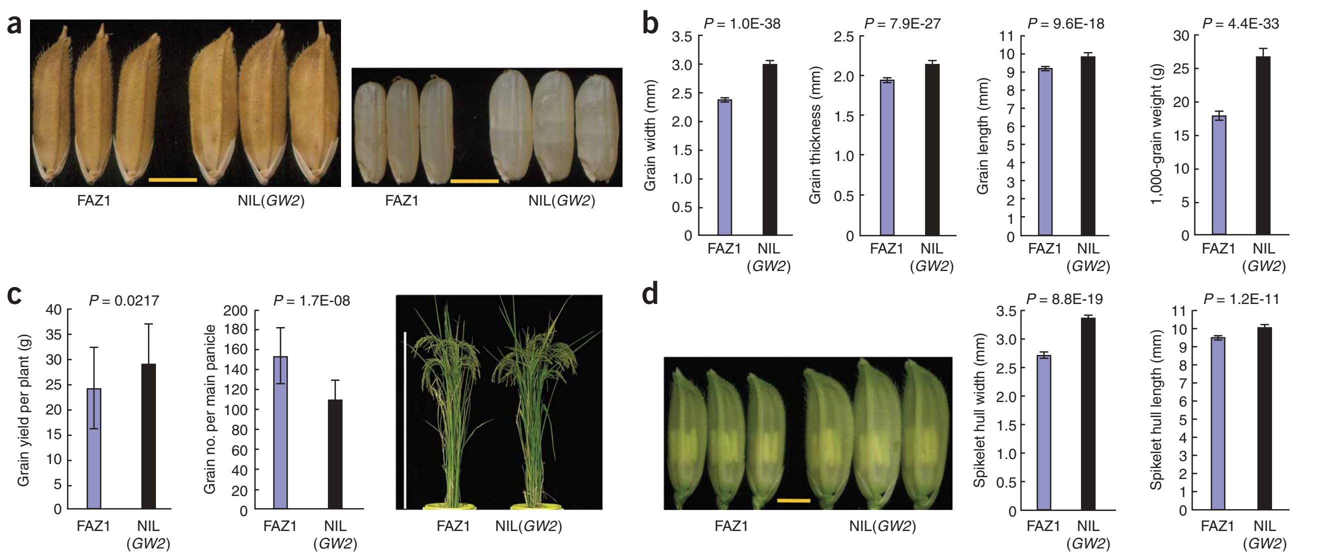
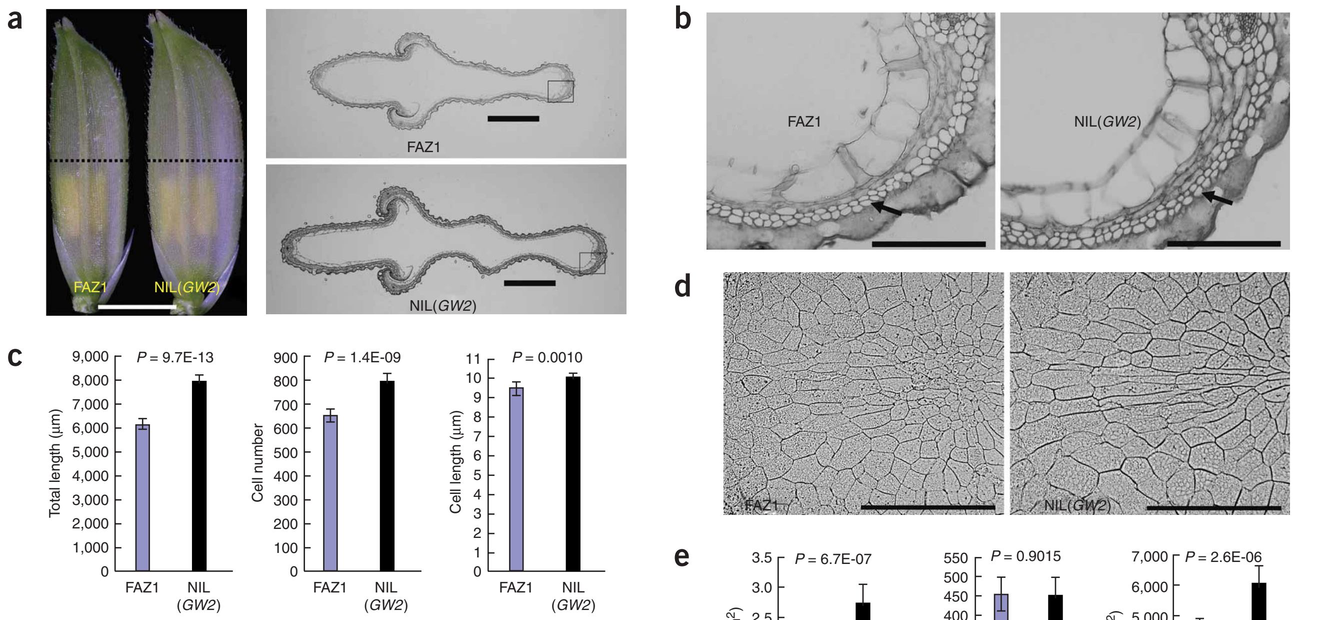


Figure 4 Phenotypic characterization of grains in FAZ1 and NIL(*GW2*). (**a**) Grains (left) and brown rice grains (right) of FAZ1 and NIL(*GW2*). Scale bar, 3 mm. (**b**) Comparisons of grain width, thickness, length and 1,000-grain weight in FAZ1 and NIL(*GW2*) ($n = 25$ plants). (**c**) Comparison of grain yield per plant, grain number per main panicle and plant type in FAZ1 and NIL(*GW2*) plants ($n = 25$ plants). Scale bar, 1 m. (**d**) Spikelets and comparison of spikelet hull width and length in FAZ1 and NIL(*GW2*) just before heading ($n = 12$ plants). All phenotypic data were measured from the plants, which were grown with a distance of 15 × 15 cm in paddies under normal cultivation conditions. Scale bar, 3 mm. All data are given as mean ± s.d. A Student's *t*-test was used to generate the *P* values in **b**–**d**.



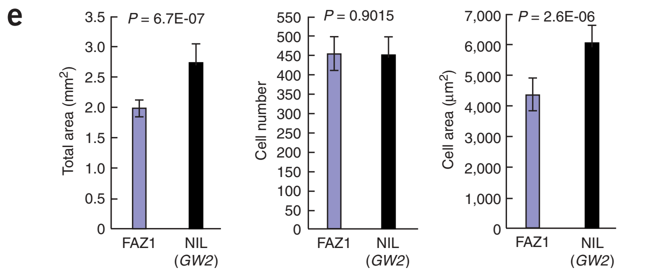


Figure 5 Histological analyses of spikelet hulls just before heading and endosperm at maturity in FAZ1 and NIL(*GW2*). (**a**) Cross-section of spikelet hull. Left: spikelets (scale bar, 3 mm). Right: cross-section of spikelet hull (scale bar, 500 μm). Dotted line indicates position of cross-section. (**b**) Magnified view of spikelet hull cross-section boxed in **a**. Arrow shows the outer parenchymal cell layer. Scale bar, 100 μm. (**c**) Comparison of total length, cell number and mean cell length in the outer parenchymal cell layers of spikelet hulls of FAZ1 and NIL(*GW2*) ($n = 10$ spikelets). (**d**) Portion between the dorsal side and the central point of the endosperm cross-section. Scale bar, 500 μm. (**e**) Comparison of total area, cell number and mean cell area in the endosperm cross-section of FAZ1 and NIL(*GW2*) ($n = 10$ endosperms). All data are given as mean ± s.d. A Student's *t*-test was used to generate the *P* values in **c** and **e**.

paddies. Nevertheless, our present results indicate that *GW2* is a useful locus for high-yield crop breeding.

Full rice grains are formed by milk filling the spikelet hull. Therefore, we measured the width and the length of the spikelet hull just before heading. The spikelet hull of NIL(*GW2*) was markedly wider (23.4%) than the that of FAZ1 (**Fig. 4d**). However, the length of the spikelet hull in NIL(*GW2*) was only 6.7% longer than in FAZ1. These results demonstrated that the increase in the width of the spikelet hull produced by the WY3 allele at *GW2* resulted in rice grains of greater width and weight.

Because *GW2* is expressed constitutively in various tissues or organs, we examined other traits in FAZ1 and NIL(*GW2*) to test whether *GW2* has pleiotropic effects (**Supplementary Fig. 3** online). Plant height and flag leaf width did not differ between FAZ1 and NIL(*GW2*). However, we observed increases in panicle number per plant and days to heading in NIL(*GW2*) compared with FAZ1. We observed a decrease in main panicle length in NIL(*GW2*). These results indicated that *GW2* has pleiotropic effects, at least on the panicle number per plant, days to heading and main panicle length, in addition to on the grain numbers per main panicle (**Fig. 4c**). The presence of these pleiotropic effects implies that *GW2* may be involved in the development of other tissues or organs, beyond its influence on grain size. The *A. thaliana* E3 ubiquitin ligase COP1, which has a RING-type domain (C3HC4), also has pleiotropic effects (the *cop1* mutant has a pleiotropic phenotype), suggesting that COP1 may function in all photoreceptor signaling pathways[32,33].

A genetic analysis of *A. thaliana* seed mass has demonstrated that both maternal and nonmaternal QTLs affect seed mass, implicating both maternal and zygotic genomes in processes that determine seed size[34]. To test whether *GW2* acts maternally or zygotically, we compared grain width, thickness, length and weight in reciprocal crosses between FAZ1 and NIL(*GW2*) (**Supplementary Fig. 4** online). These traits did not differ between the reciprocal crosses FAZ1/NIL(*GW2*) and NIL(*GW2*)/FAZ1, suggesting that *GW2* acts nonmaternally on grain (seed) size. Similarly, there is no maternal effect on seed size in *A. thaliana iku* class mutants[35].

GW2 increases number of cells and milk filling rate

Given that the NIL(*GW2*) spikelet hull was wider than that of FAZ1 before fertilization, we compared cross-sections of the central part of the spikelet hull in NIL(*GW2*) and FAZ1 to investigate the origins of the observed size differences (**Fig. 5a,b**). The outer parenchyma cell layer of NIL(*GW2*) was longer (by 29.6%) and contained substantially more cells (22.4% more) than that of FAZ1 (**Fig. 5c**), with only a 5.8% increase in cell length (**Fig. 5c**). These data demonstrate that the increased width of the NIL(*GW2*) spikelet hull results mainly from an increase in cell number, but not in cell size, suggesting that *GW2* may be involved in regulation of cell division.

We compared cross-sections of mature grains from NIL(*GW2*) and FAZ1 (**Fig. 5d,e**) to investigate differences in endosperm cells. Although endosperm cells of NIL(*GW2*) were larger than those of FAZ1, there was no significant difference in endosperm cell number, suggesting that the increase in NIL(*GW2*) endosperm size (**Fig. 4a**) resulted mainly from cell expansion, not from an increase in cell number.

Because NIL(*GW2*) has a larger endosperm and a heavier grain, we investigated the grain milk filling rate in NIL(*GW2*) and FAZ1. There

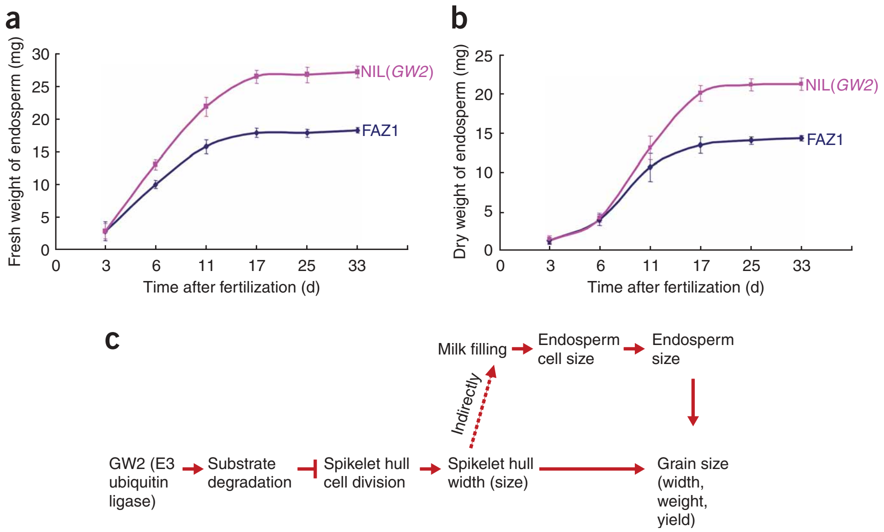


Figure 6 Characterization of grain milk filling in FAZ1 and NIL(*GW2*) and proposed model for the role of GW2. (**a**) Time-course of endosperm fresh weight increase. (**b**) Time-course of endosperm dry weight increase. Data are mean ± s.d. ($n = \sim 10$ to ~ 15 plants) in **a** and **b**. (**c**) Proposed model for the role of GW2 in regulation of grain width (size) and weight. GW2 recruits the targeted substrate for degradation, thereby inhibiting cell division, and then influences spikelet hull size, subsequently indirectly influencing milk filling rate, endosperm cell size, endosperm size and, ultimately, grain size (width, weight and yield).

were no differences in either endosperm fresh weight or dry weight 3 d after fertilization (**Fig. 6a,b** and **Supplementary Fig. 5** online). On day 6 after fertilization, NIL(*GW2*) endosperm fresh weight was slightly higher than that of FAZ1, although there was no difference in dry weight. NIL(*GW2*) endosperm fresh weight and dry weight were significantly higher than FAZ1 starting 11 d after fertilization, with those differences almost reaching a maximum 17 d after fertilization. NIL(*GW2*) increased endosperm fresh weight and dry weight by 47.8% and 49.7%, respectively, on day 17 after fertilization. These results indicated that the larger endosperm (or the larger cell size of endosperm) and heavier grain in NIL(*GW2*) resulted from a faster rate of accumulation of dry matter, suggesting that the WY3 allele of *GW2* may increase the rate of dry matter accumulation.

Because increased grain size may have a negative effect on rice grain quality[36], we observed the six grain quality traits in NIL(*GW2*) and FAZ1 to test whether the WY3 *GW2* allele, which results in a large grain size, also influences grain quality. A grain's cooking and eating quality traits are determined mainly by amylose content and gel consistency[37]. Amylose content and gel consistency did not differ between NIL(*GW2*) and FAZ1 (**Supplementary Fig. 6** online), indicating that the WY3-*GW2* allele does not influence the grain's cooking and eating quality traits. There were significant but small difference in milling quality traits, brown rice percentage and milled rice percentage between FAZ1 and NIL(*GW2*) (**Supplementary Fig. 6**). However, an appearance quality trait, chalky rice grain percentage in NIL(*GW2*), was significantly increased, whereas head milled grain appearance in NIL(*GW2*) was not as poor as in FAZ1 (**Supplementary Fig. 6**). On the other hand, protein content, a nutritional quality trait, did not differ between NIL(*GW2*) and FAZ1 (**Supplementary Fig. 6**). These results suggest that the WY3 *GW2* allele could increase grain size and yield with little influence on appearance and no reduction in cooking or eating quality and thus could be useful in breeding. Whether this is associated with an accelerated milk filling rate mediated by the WY3 *GW2* allele remains to be determined in future studies.

DISCUSSION

In the current study, we have successfully cloned a QTL, *GW2*, that increases grain size and yield in rice. Our data show that GW2 is a new RING-type protein with intrinsic E3 ubiquitin ligase activity (**Fig. 2e,f**) that localizes to the cytoplasm and is constitutively expressed in various tissues (**Fig. 3**). Reduced expression of *GW2* increases grain size (mainly grain width), resulting in enhanced grain weight, whereas overexpression decreases grain size and weight (**Fig. 2c,d**). The naturally occurring WY3 allele of *GW2*, which encodes a truncated version of the protein with a 310–amino acid deletion (**Fig. 2a,b**), increases the number of cells of the spikelet hull, resulting in a wider spikelet hull (**Fig. 5a–c**), and subsequently accelerates the grain milk filling rate (**Fig. 6a,b**), resulting in increased grain width, weight and yield (**Fig. 4a–c**).

Many RING-type proteins function as E3 ubiquitin-protein ligases, targeting proteins for ubiquitin-dependent degradation by the 26S proteasome[28–31,38,39]. Such RING proteins are involved in the regulation of numerous cellular processes, including transcription, signal transduction, recombination and cell cycle progression. A previous study has shown that the RING-type (C3H2C3) protein BIG BROTHER, which has E3 ligase activity, acts as a central negative regulator of *A. thaliana* floral organ size, most likely by marking cellular proteins for degradation[39]. In contrast, we found that GW2, a new RING-type protein, has E3 ubiquitin ligase activity and alters the number of cells in the spikelet hull, suggesting that GW2 E3 ligase functions as a regulator of cell division through ubiquitin-mediated proteolysis. However, the mechanism of cell cycle regulation mediated by GW2 remains to be elucidated.

In the majority of RING-type proteins, the N terminus contains the RING domain that binds E2, and the remainder contains other protein–protein interaction domains that may function as the substrate-binding domain of the E3 ligase[24]. For example, COP1 has both a RING-type domain in the N terminus and a WD-40 repeat domain in the C terminus that can bind several protein targets, including HY5, thereby recruiting an E2 and targeting HY5 and other substrates for ubiquitination and degradation by the proteasome[40]. In addition, the WY3 variant of GW2 has an intact RING domain, thereby retaining E3 ubiquitin ligase activity, but is truncated by 310 amino acids that might contain the substrate-binding domain. The absence of a substrate-binding domain suggests that WY3 *GW2* is a null allele. The coincidence of the *GW2* null allele and the increased number of cells in the spikelet hull suggest that WY3 GW2 protein does not interact with the substrate(s) involved in cell division and does not target them for ubiquitination and subsequent degradation. These data suggest that GW2 E3 ligase is a new negative regulator of cell division, targeting its substrate(s) to proteasomes for regulated proteolysis. We examined GW2 for the presence of other known domains but did not find any known domains in GW2. Identification and characterization of both the substrate-binding domain of GW2 and the GW2 target substrate will be challenging but worth pursuing. *BRCA1*, a well-known gene that encodes a breast- and ovarian-specific tumor suppressor protein with a RING domain (C3HC4), is expressed in numerous tissues, including breast and ovary[41]. Many subsequent studies have found that *BRCA1* not only has an E3 ubiquitin ligase function but also has a transcriptional activation function[42]. Such findings prompted us to speculate whether GW2 may also be involved in transcriptional regulation, aside from its E3 ubiquitin ligase function. Future studies will need to address this question.

The persistent endosperm forms the bulk of the mature seed in most monocots, including rice, maize and wheat. Consequently, the seed size and weight are attributable mainly to the extent of endosperm growth[43,44]. In the current study, we found that the *GW2* null allele (the naturally occurring WY3 allele) has a wider (larger) spikelet hull owing to increased numbers of cells. The larger spikelet hull allows greater endosperm growth and provides a greater area of contact for endosperm with the seed coat (**Supplementary Fig. 5** and ref. 45), leading to an accelerated milk filling rate (**Fig. 6a,b**) and ultimately to enhanced endosperm size, grain size and grain weight. Therefore, we suggest that the enhanced endosperm size might be an indirect effect of *GW2*. Based on our data, we hypothesize a model to explain the potential role of GW2 in the regulation of grain width (size), weight and yield, shown in **Figure 6c**.

Although a thorough understanding of seed development is important for the improvement of grain yield through genetic manipulation, little is known about the genetic mechanisms that determine final seed size and weight in plants. Thus, our findings provide an initial step toward dissecting the mechanism by which the ubiquitination system contributes to the regulation of seed development and seed yield in plants. As a key regulator of grain (seed) size, the *GW2* gene (and its homologs in other cereals such as maize and wheat) will facilitate breeding efforts to improve grain yield in staple crops.

METHODS

The molecular marker primers and primers for *GW2* molecular analysis are listed in **Supplementary Table 1** online.

Plant materials. A large-grain WY3 *japonica* variety was crossed with Fengaizhan-1 (FAZ1), an elite *indica* small-grain variety. The resultant F_1 plants were selfed to produce F_2 seeds and were backcrossed with FAZ1 plants as the recurrent to produce BC_1F_1 seeds. We selected several plants in which the region around *GW2* was heterozygous, and almost all other regions were homozygous for FAZ1 so as to develop the segregating populations for fine mapping and high-resolution mapping of *GW2* by repetitive backcrossing and marker-assisted selection. From the BC_4F_2 generation, we developed a nearly isogenic line for *GW2*, NIL(*GW2*), with a very small WY3 chromosomal region containing the *GW2* locus in the FAZ1 genetic background.

Fine mapping and high-resolution mapping. A BC_2F_2 population was used for fine mapping of *GW2*, based on rough mapping of the F_2 population. To perform *GW2* QTL analysis, we used the grain width measured in 190 BC_2F_2 plants and 11 molecular markers in a target region containing *GW2*, as described previously[46]. Molecular markers W236 and W239, which flank *GW2*, were used to detect recombinants in 6,013 BC_3F_2 plants. To further determine the location of the recombinations nearest to *GW2*, we developed markers on the basis of the sequence of the bacteriophage P1–derived artificial chromosome (PAC) clone AP005004 and determined genotypes of the recombinants with these markers. The BC_3F_3 progeny derived from recombinant plants were used to screen for homozygous recombination products. We used fixed recombinant plants (BC_3F_4) to measure the grain width and determine the *GW2* genotypes. The candidate *GW2* genes from FAZ1 and WY3 genomic DNA were sequenced and compared.

RNA extraction, cDNA isolation and RT-PCR. Total RNA was extracted from various plant tissues in FAZ1 and NIL(*GW2*) and was converted into first-strand cDNA. The full-length *GW2* cDNA (1,634 bp) was amplified from the first-strand cDNA and sequenced. RT-PCR was carried out to amplify the *GW2* transcripts with 30 PCR cycles, using the first-strand cDNA as a template. Actin was also amplified as the control.

Transgenic analysis. We overexpressed *GW2* using a full-length *GW2* cDNA from FAZ1 that was inserted into the plant binary vector pHB[47], in which transgene expression was under the control of the CaMV 35S promoter. We performed antisense expression of *GW2* using an antisense fragment of the entire *GW2* coding region (1,278 bp) from FAZ1 that was inserted into the binary vector pHB. Both constructs were introduced into *Agrobacterium tumefaciens* strain EHA105 and transferred into a *japonica* variety, Zhonghua 11, as reported previously[48]. The empty pHB vector was also transformed into Zhonghua 11 as a control.

E3 ubiquitin ligase activity assay. We cloned cDNA encoding *GW2* from FAZ1 or WY3 into pET32a (+) (Novagen) and prepared the fusion protein following the manufacturer's protocol. Assays for *in vitro* ubiquitination were carried out as described previously[49], with slight modifications. In brief, 0.25 μg E1, 0.2 μg E2, 5 μg of ubiquitin (Ub) and 0.5 μg purified FAZ1 GW2, purified WY3 GW2 or purified COP1 fusion protein were incubated in a 30-μl reaction mix containing 50 mM Tris-HCl (pH 7.4), 2.5 mM $MgCl_2$, 6.6 mM ATP and 0.5 mM DTT (incubated at 30 °C for 5 h). The reaction was stopped with 1× SDS-PAGE loading buffer (100 °C, 5 min). Samples (15 μl) were analyzed by SDS-PAGE. Polyubiquitinated proteins were detected by protein blotting using an antibody to ubiquitin. Human E1 and E2 (UbcH5b) were purchased from Merck.

Subcellular localization and tissue localization. To investigate the cellular localization of *GW2*, a 35S *GFP-GW2* (from FAZ1) fusion construct was bombarded into onion epidermal cells using a helium biolistic device (Bio-Rad PDS-1000). Onion epidermal cell nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, 5 μg ml^{-1} in PBS). We amplified a 1.5-kb *GW2* promoter region upstream of the ATG start codon from FAZ1 genomic DNA by PCR to make the *GW2* promoter-GFP fusion construct (the binary vector pHB) and generated transgenic plants carrying this construct as described above. Samples were examined with a Zeiss LSM510 confocal laser microscope.

Sample preparation and microscopy. Plant materials were fixed in FAA (50% ethanol, 5% glacial acetic acid and 5% formaldehyde) for 16 h, dehydrated in an ethanol series, and embedded in Paraplast (Sigma). Tissue sections (8 μm thick) were cut with a rotary microtome, mounted and stained with safranine T and fast green. Endosperms were embedded in Epon812 resin, and sections (2 μm) were stained with toluidine blue. Sections were photographed under an Olympus BX51 with a DP70 CCD camera.

Accession codes. GenBank: *GW2*, EF447275.

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

ACKNOWLEDGMENTS
We thank P. Qi and Z.-Z. Piao for technical assistance. We thank S. Luan for critically reading the manuscript. This work was supported by grants from the Ministry of Science and Technology of China and the Shanghai Science and Technology Development Fund to H.-X.L.

AUTHOR CONTRIBUTIONS
H.-X.L. designed the experiments. X.J.S. and W.H. performed most of the experiments. H.-X.L., S.M. and M.Z.Z. performed some of the experiments. H.-X.L. wrote the manuscript.

COMPETING INTERESTS STATEMENT
The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Tanksley, S.D. Mapping polygenes. *Annu. Rev. Genet.* **27**, 205–233 (1993).
2. Yano, M. Genetic and molecular dissection of naturally occurring variations. *Curr. Opin. Plant Biol.* **4**, 130–135 (2001).

3. Frary, A. *et al.* *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85–88 (2000).
4. Yano, M. *et al.* *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* **12**, 2473–2484 (2000).
5. El-Din El-Assal, S., Alonso-Blanco, C., Peeters, A.J., Raz, V. & Koornneef, M.A. A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2*. *Nat. Genet.* **29**, 435–440 (2001).
6. Liu, J., Eck, J.V., Cong, B. & Tanksley, S.D. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc. Natl. Acad. Sci. USA* **99**, 13302–13306 (2002).
7. Takahashi, Y., Shomura, A., Sasaki, T. & Yano, M. *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. *Proc. Natl. Acad. Sci. USA* **98**, 7922–7927 (2001).
8. Ren, Z.H. *et al.* A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* **37**, 1141–1146 (2005).
9. Ashikari, M. *et al.* Cytokinin oxidase regulates rice grain production. *Science* **309**, 741–745 (2005).
10. Konishi, S. *et al.* An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392–1396 (2006).
11. Evans, L.T. Storage capacity as a limitation on grain yield. in *Rice Breeding* (International Rice Research Institute, Manila, 1972).
12. Xu, J.Y., Xue, Q.Z., Luo, L.J. & Li, Z.K. Genetic dissection of grain weight and its related traits in rice (*Oryza sativa* L.). *Chin J Rice Sci* **16**, 6–10 (2002).
13. Lin, H.X. *et al.* RFLP mapping of QTLs for yield and related characters in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **92**, 920–927 (1996).
14. Huang, N. *et al.* RFLP mapping of isozymes, RAPD, and QTLs for grain shape, brown planthopper resistance in a doubled-haploid rice population. *Mol. Breed.* **3**, 105–113 (1997).
15. Redona, E.D. & Mackill, D.J. Quantitative trait locus analysis for rice panicle and grain characteristics. *Theor. Appl. Genet.* **96**, 957–963 (1998).
16. Thomson, M.J. *et al.* 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 (2003).
17. Li, J., Thomson, M. & McCouch, S.R. Fine mapping of a grain weight quantitative trait locus in the pericentromeric region of rice chromosome 3. *Genetics* **168**, 2187–2195 (2004).
18. 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).
19. Lin, H.X. *et al.* RFLP mapping of QTLs for grain shape traits in *indica* rice (*Oryza sativa* L. subsp. *indica*). *Scientia Agricultura Sinica* **28**, 1–7 (1995).
20. Yoon, D.B. *et al.* Mapping quantitative trait loci for yield components and morphological traits in an advanced backcross population between *Oryza grandiglumis* and the *O. sativa japonica* cultivar Hwaseongbyeo.. *Theor. Appl. Genet.* **112**, 1052–1062 (2006).
21. Freemont, P.S., Hanson, I.M. & Trowsdale, J. A novel cysteine-rich sequence motif. *Cell* **64**, 483–484 (1991).
22. Borden, K.L. & Freemont, P.S. The RING finger domain: a recent example of a sequence-structure family. *Curr. Opin. Struct. Biol.* **6**, 395–401 (1996).
23. Saurin, A.J., Borden, K.L., Boddy, M.N. & Freemont, P.S. Does this have a familiar RING? *Trends Biochem Sci.* **21**, 208–214 (1996).
24. Stone, S.L. *et al.* Functional analysis of the RING-type ubiquitin ligase family of *Arabidopsis*. *Plant Physiol.* **137**, 13–30 (2005).
25. Hewitt, E.W. *et al.* Ubiquitylation of MHC class I by the K3 viral protein signals internalization and TSG101-dependent degradation. *EMBO J.* **21**, 2418–2429 (2002).
26. Dasgupta, A., Ramsey, K.L., Smith, J.S. & Auble, D.T. Sir Antagonist 1 (San1) is a ubiquitin ligase. *J. Biol. Chem.* **279**, 26830–26838 (2004).
27. Albert, T.K. *et al.* Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO J.* **21**, 355–364 (2002).
28. Lorick, K.L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl. Acad. Sci. USA* **96**, 11364–11369 (1999).
29. Joazeiro, C.A. *et al.* The tyrosine kinase negative regulator c-Cbl as a RING-type, E2 dependent ubiquitin-protein ligase. *Science* **286**, 309–312 (1999).
30. Xie, Q. *et al.* SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. *Nature* **419**, 167–170 (2002).
31. Saijo, Y. *et al.* The COP1–SPA1 interaction defines a critical step in phytochrome A-mediated regulation of HY5 activity. *Genes Dev.* **17**, 2642–2647 (2003).
32. Deng, X.W., Caspar, T. & Quail, P.H. *cop1*: A regulatory locus involved in light-controlled development and gene expression in *Arabidopsis*. *Genes Dev.* **5**, 1172–1182 (1991).
33. Moon, J., Parry, G. & Estelle, M. The ubiquitin-proteasome pathway and plant development. *Plant Cell* **16**, 3181–3195 (2004).
34. Alonso-Blanco, C., Blankestijn-De Vries, H., Hanhart, C.J. & Koornneef, M. Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **96**, 4710–4717 (1999).
35. Luo, M., Dennis, E.S., Berge, F., Peacock, W.J. & Chaudhury, A. *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **102**, 17531–17536 (2005).
36. Xie, X. *et al.* Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived from a cross between *Oryza sativa* and *Oryza rufipogon*. *Theor. Appl. Genet.* **113**, 885–894 (2006).
37. Tan, Y.F. *et al.* Genetic bases of appearance quality of rice grains in Shanyou 63, an elite rice hybrid. *Theor. Appl. Genet.* **101**, 823–829 (2000).
38. Dornan, D. *et al.* The ubiquitin ligase COP1 is a critical negative regulator of p53. *Nature* **429**, 86–92 (2004).
39. Disch, S. *et al.* The E3 ubiquitin ligase BIG BROTHER controls *Arabidopsis* organ size in a dosage-dependent manner. *Curr. Biol.* **16**, 272–279 (2006).
40. Osterlund, M.T., Hardtke, C.S., Wei, N. & Deng, X.W. Targeted destabilization of HY5 during light-regulated development of *Arabidopsis*. *Nature* **405**, 462–466 (2000).
41. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* **266**, 66–71 (1994).
42. Venkitaraman, A.R. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
43. Reddy, V.M. & Daynard, T.B. Endosperm characteristics associated with rate of grain filling and kernel size in corn. *Maydica* **28**, 339–355 (1983).
44. Chojecki, A.J.S., Bayliss, M.W. & Gale, M.D. Cell production and DNA accumulation in the wheat endosperm, and their association with grain weight. *Ann. Bot. (Lond.)* **58**, 809–817 (1986).
45. Schruff, M.C. *et al.* The *AUXIN RESPONSE FACTOR 2* gene of *Arabidopsis* links auxin signalling, cell division, and the size of seeds and other organs. *Development* **133**, 251–261 (2005).
46. Lander, E.S. & Botstein, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199 (1989).
47. Mao, J., Zhang, Y.C., Sang, Y., Li, Q.H. & Yang, H.Q. A role for *Arabidopsis* cryptochromes and COP1 in the regulation of stomatal opening. *Proc. Natl. Acad. Sci. USA* **102**, 12270–12275 (2005).
48. 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).
49. Zhang, X., Garreton, V. & Chua, N.H. The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes Dev.* **19**, 1532–1543 (2005).