Rice Domestication by Reducing Shattering

Changbao Li, *et al.*

Science **311**, 1936 (2006);
DOI: 10.1126/science.1123604

31. S. P. Hazen *et al.*, *Plant Physiol.* **138**, 990 (2005).
32. We thank E. Foss for creation of strain YEF1695, E. Smith for analysis of YEF1695 mapping data, and D. Storton and J. Matese for technical support. Supported by National Institute of General Medical Sciences grant R01 GM046406 and center grant P50 GM071508 (D.B.), NIH grant R37 MH059520 and James S. McDonnell Foundation grant 99-11T (L.K.), and Pew Charitable Trusts award 2000-002558 (S.C.P.). The raw data and software are available at http://genomics-pubs.princeton.edu/SNPscanner/.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1123726/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 to S3
References

12 December 2005; accepted 24 February 2006
Published online 9 March 2006;
10.1126/science.1123726
Include this information when citing this paper.

Rice Domestication by Reducing Shattering

Changbao Li, Ailing Zhou, Tao Sang*

Crop domestication frequently began with the selection of plants that did not naturally shed ripe fruits or seeds. The reduction in grain shattering that led to cereal domestication involved genetic loci of large effect. The molecular basis of this key domestication transition, however, remains unknown. Here we show that human selection of an amino acid substitution in the predicted DNA binding domain encoded by a gene of previously unknown function was primarily responsible for the reduction of grain shattering in rice domestication. The substitution undermined the gene function necessary for the normal development of an abscission layer that controls the separation of a grain from the pedicel.

Cereals, the world's primary food, were domesticated from wild grass species. Because wild grasses naturally shed mature grains, a necessary early step toward cereal domestication was to select plants that could hold on to ripe grains to allow effective field harvest (*1*, *2*) (fig. S1). The selection process might have been mainly unconscious because grains that did not fall as easily had a better chance of being harvested and planted in the following years. Consequently, nonshattering alleles had an increased frequency and eventually replaced the shattering alleles during domestication. The finding that one locus accounted for most phenotypic variance of grain shattering between a cereal crop and its wild progenitor suggested that the domestication process could have been initiated quickly by selection at the locus (*3*–*5*). The molecular genetic basis of the selection, however, has not been characterized.

Rice (*Oryza sativa*) was domesticated from one or both of two closely related species—*O. nivara* and *O. rufipogon*—distributed from southeastern Asia to India (*6*, *7*). Our recent genetic analysis of an F_2 population derived between *O. sativa* ssp. *indica* and the wild annual species *O. nivara* identified three quantitative trait loci (QTL)—*sh3*, *sh4*, and *sh8*—responsible for the reduction of grain shattering in cultivated rice (*5*). Of these QTL, *sh4* explained 69% of phenotypic variance, and the other two explained 6.0% and 3.1% of phenotypic variance. The *sh4* allele of the wild species caused shattering and was dominant.

Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA.

*To whom correspondence should be addressed. E-mail: sang@msu.edu

Two previous QTL studies using crosses between *O. sativa* ssp. *indica* and the wild perennial species *O. rufipogon* detected four and five shattering QTL (*8*, *9*). Both studies identified a QTL at the same location of *sh4* with either the largest or nearly largest phenotypic effect among the detected QTL. Moreover, genetic analyses between *O. sativa* ssp. *japonica* and *O. rufipogon* and two other closely related wild species *O. glumaepetula* and *O. meridionalis* all found that a single dominant allele from each of the three wild species was responsible for grain shattering (*10*, *11*). This locus, named *Sh3*, was mapped to the same chromosomal location as *sh4*.

Our QTL analysis located *sh4* between simple sequence repeat (SSR) markers RC4-123 and RM280 (*5*), which had a physical distance of about 1360 kb in the *O. sativa* genome (*12*) (Fig. 1A). Because of the large and dominant effect of the *O. nivara* allele, we were able to phenotypically distinguish F_2 individuals that were homozygous recessive (ss) from those that had at least one *O. nivara* allele of *sh4* (ns and nn), regardless of the genotypes at the remaining two QTL of small effect. After evaluating a total of 489 F_2 plants genotyped at the three shattering QTL, we consistently found that plants with the ns and nn genotypes at *sh4* shed all mature grains when hand tapped, whereas plants with the ss genotype at *sh4* did not shed grains or only partially shed mature grains under vigorous hand shaking.

With the reliable phenotyping method available, we grew ~12,000 F_2 seedlings and screened for recombinants between RC4-123 and RM280 (*13*). Plants with the genotype of ss at one marker and ns at the other were selected, and a total of 134 individuals were grown for phenotypic evaluation. By progressively examining SSR and SNP (single-nucleotide polymorphism) markers between RC4-123 and RM280, we finally mapped the mutation responsible for the derivation of nonshattering in cultivated rice to a 1.7-kb region of a gene with a previously unknown function (Fig. 1B and table S1). The gene is predicted to be a transcription factor, and its coding region is physically located between 34,014,305 and 34,012,126 base pairs (bp) on assembly LOC_Os04 g57530 of rice chromosome 4 (The TIGR Rice Genome Annotation Database).

The comparison of the 1.7-kb sequences between the mapping parents revealed seven mutations (Fig. 1C). These include one mutation in the intron: (*a*) a 1-bp substitution; three mutations in the first exon: (*b*) a 15-bp or five–amino acid insertion/deletion, (*c*) a 3-bp or one–amino acid insertion/deletion, and (*d*) a 1-bp or an amino acid substitution; and three mutations 5′ upstream of the start codon: (*e*) a 1-bp substitution at site -55, (*f*) a 3-bp insertion/deletion between sites −343 and −344, and (*g*) an 8-bp insertion/deletion between sites −558 and −559.

To assess the polymorphism and evolutionary direction of these mutations, we sequenced this 1.7-kb region from an additional 14 rice cultivars representing the diversity of *O. sativa* (*14*), 21 accessions of *O. nivara* covering the distributional range of the wild species (*15*), 6 accessions of *O. rufipogon*, and 1 accession of each of the four remaining wild A-genome species (Fig. 1C and table S2). The cultivars were polymorphic for mutation *f*, i.e., some of the cultivars had the same sequence as *O. nivara*. At the remaining six mutation sites, all cultivars shared the same sequences, which were different from those of the *O. nivara* parent.

Surprisingly, three accessions of *O. nivara* had the same sequences as *O. sativa* at these six sites. It was then found that plants grown from these accessions had the nonshattering phenotype. Greenhouse observations indicated that these accessions had additional characteristics of cultivated rice that were not found in *O. nivara*, such as upright tillers, short awns, and/or photoperiod sensitivity. This suggests that the three accessions are weedy rice that has received and fixed the *sh4* allele from cultivars.

The remaining accessions of the wild species with confirmed shattering differed invariably from the cultivars by one mutation, *d*, which was a nucleotide substitution of G for T or an amino acid substitution of

asparagine for lysine in *O. sativa*. At the remaining five mutation sites, sequence polymorphism was found within the wild species (Fig. 1C). That is, some of the wild accessions shared the same sequence with cultivated rice at these sites but had the shattering phenotype. The results thus indicate that the amino acid substitution at site *d* was selected for the development of nonshattering cultivars during rice domestication.

The Blast search of the GenBank for protein sequences identified three predicted genes that are most similar to *sh4*. These include a rice gene (XP_469180) with 32% amino acid sequence identity with *sh4*, and two *Arabidopsis* genes (NP_174416 and NP_181107) with 32% and 29% identity with *sh4*. None of these genes has been functionally characterized. The two *Arabidopsis* genes were predicted to be transcription factors (*16*), and one of them had a cDNA sequence (AAT99796) in the database. The next most similar group of genes was also from rice and *Arabidopsis* but had only 20 to 22% amino acid sequence identity with *sh4*.

Examination of the *sh4* protein, using programs Prosite and PredictNLS, identified a Myb3 DNA binding domain and a nuclear localization signal (Fig. 1D), suggesting that *sh4* is a transcription factor. To test this hypothesis, we fused the gene for a green fluorescent protein (GFP) with *sh4* to make *sh4-GFP*, which was driven by a *Ubi* promoter in the plasmid construct. The construct was introduced into a *japonica* cultivar, Taipei

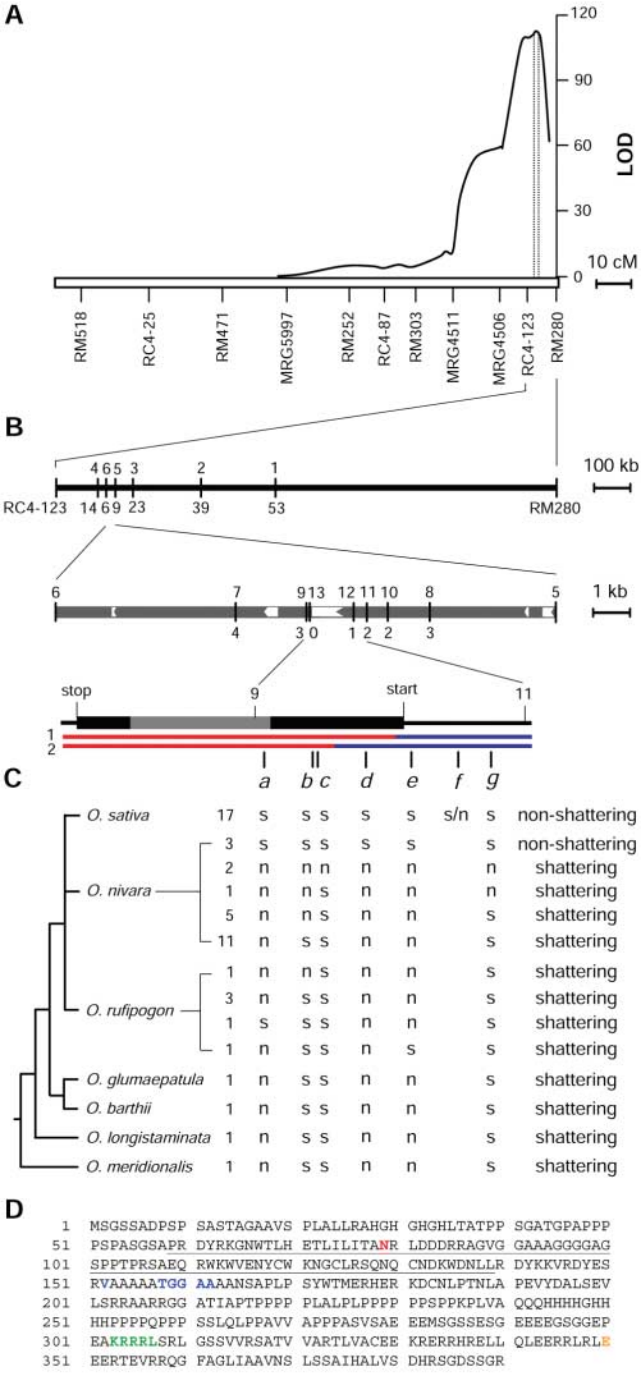

Fig. 1. Molecular cloning of *sh4*. (**A**) Chromosomal location of *sh4* determined by QTL mapping. Dotted lines indicate 1-lod (logarithm of the odds ratio for linkage) supporting interval. (**B**) Fine mapping of *sh4*. Vertical lines indicate SSR and SNP markers. Numbers above lines: markers numbered consecutively according to the order of evaluation; numbers below lines: the number of recombinants left in the chromosomal interval still containing *sh4* after the evaluation of the marker. White horizontal arrows indicate the orientation and size of open reading frames between markers 5 and 6. The mutation responsible for nonshattering was mapped to between markers 9 and 11, in a predicted gene with two exons (black bars) and an intron (gray bar). The start and stop codons of the gene are labeled. Lines below illustrate two constructs made for gene transformation; red and blue represent sequences of *O. nivara* and *O. sativa*, respectively. (**C**) Seven mutations found between the mapping parents are labeled *a* through *g*. Variation at these sites is compared between rice cultivars and wild A-genome species in the phylogenetic context; s and n represent sequences of *O. sativa* and *O. nivara* parents, respectively. The number of accessions of a species with the same combination of sequences is indicated. (**D**) *sh4* protein sequence of *O. sativa*. Mutations between the mapping parents are indicated: red for substitution (mutation *d*; K in *O. nivara*) and blue for insertion/deletion (mutation *b* and *c*; deletion in *O. nivara*). The predicted Myb3 DNA binding domain is underlined. The predicted nuclear localization signal is colored green. The *sh4-GFP* construct for subcellular localization encodes a recombined protein beginning from the sh4 N terminus to the amino acid E (colored orange) followed by GFP. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

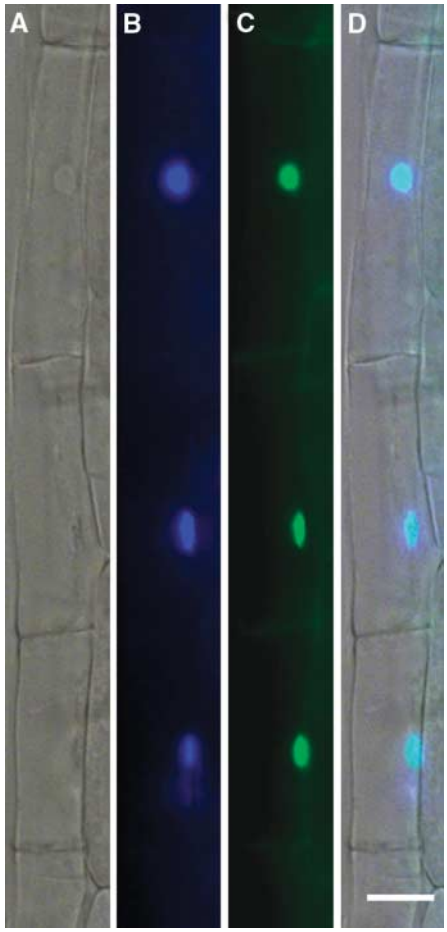

Fig. 2. Subcellular localization of sh4. Roots of rice cultivar Taipei 309 transformed with *Ubi::sh4-GFP* were stained with 4′-6-diamidino-2-phenylindole (DAPI) and observed under various conditions. (**A**) A differential interference contrast image of epidermal cells. (**B**) The same cells showing the DAPI-stained nuclei. (**C**) The same cells showing the nuclear localization of sh4-GFP fluorescence. (**D**) The merged image. Bar, 10 μm.

309, a rice strain tested as suitable for gene transformation (*17*). The nuclear localization of GFP-tagged *sh4* was determined (Fig. 2). This result supports the bioinformatic prediction that *sh4* is a transcription factor.

Reverse transcription–polymerase chain reaction (RT-PCR) detected the expression of *sh4* at the flower and pedicel junction, where mature grains separate from the mother plant (Fig. 3A). Gene expression was not detected in the remaining parts of flowers or pedicels or in the leaves. We amplified, using RT-PCR, the entire coding region of *sh4* cDNA from both mapping parents. The comparison of the cDNA sequences showed that the intron was spliced from the same position as predicted by rice genome annotation.

We conducted real-time PCR to compare the relative levels of *sh4* expression at various stages of flower and seed development in both mapping parents (Fig. 3B). Although there was a trend of increased gene expression as seeds matured, a substantial increase began 12 days after pollination. The expression in *O. sativa* continued to increase on day 18, while the measurement for *O. nivara* was no longer possible due to shattering.

We measured the strength of flower and grain attachment to the pedicel at the corresponding developmental stages. Flowers and grains were pulled away at the interface where a mature grain separates from the pedicel, and the force required was measured. For the first 9 days after pollination, the force was not significantly different between the developmental stages in either species (Fig. 3C). The force began to decrease in both species from day 12. The decline continued at a much faster rate in *O. nivara* than in *O. sativa*. On day 18, shattering prevailed in *O. nivara*, which left few grains to measure. In *O. sativa*, the force measured on day 18 was about half of that required at the earlier stages; it then decreased at a rather slow pace but did not reach the level permitting grain shattering in *O. nivara*.

We conducted rice transformation to confirm the gene function and to test the role of the amino acid substitution. We made two constructs that had the *O. sativa* promoter and recombined coding regions between the mapping parents. The two constructs differed only at the mutation site *d*. Construct 1 contained *O. nivara* sequence from the 3′ nontranslated region to the inclusion of mutation site *d* and *O. sativa* sequence from mutation site *e* to the 5′ regulatory region. Construct 2 contained *O. nivara* sequence from 3′ to the inclusion of mutation site *c* and *O. sativa* sequence from mutation site *d* to the 5′ regulatory region (Fig. 1B). The plasmids were introduced into Taipei 309.

The expression of the introduced constructs in the transgenic plants was verified by RT-PCR identification of the 18-bp deletion of the *O. nivara* sequence at mutation sites *b* and *c*. Transformants expressing construct 1 showed significantly reduced strength of grain attachment to pedicel (Student's *t* test, $P = 0.003$), whereas no significant difference was found between transformants expressing construct 2 and the control (*t* test, $P = 0.5$) (Fig. 3D). The results thus support the finding made from the genetic mapping and sequence comparison that the amino acid substitution at site *d* was primarily responsible for the reduction of grain shattering in rice domestication.

Programmed organ detachment, such as the falling of old leaves, withered floral parts, and ripe fruits, is fundamental to plant function and adaptation and is regulated by an abscission zone at the juncture of the organ and the main body of the plant. The molecular genetic control of the abscission zone development, however, is poorly understood. The study of dicotyledonous plants such as bean, tomato, and *Arabidopsis* showed that an abscission zone encompassed several layers of small, densely cytoplasmic cells. In response to environmental and hormonal signals, the activation of abscission is coupled with cell expansion and secretion of hydrolytic enzymes that break the middle lamella between cell layers in the abscission zone (*18*).

In monocotyledons, including grasses, little is known about the development and function of abscission zones. Genes regulating the developmental processes have not been identified. Here we found that the abscission zone between a rice grain and the pedicel consists of mostly one layer of small, thin-walled cells. *O. nivara* has a complete layer of abscission cells between the grain and the pedicel, which is seen in a longitudinal section as continuous lines of abscission cells between the vascular bundle and the epidermis (Fig. 4A). *O. sativa*, however, has an incomplete abscission layer. In the longitudinal section, the line of abscission cells is discontinuous and completely absent near the vascular bundle, where they are replaced by thicker-walled cells similar to adjacent pedicel cells (Fig. 4, B and C). For both species, these anatomical features were seen in young flowers (flowers ~15 days before opening were examined) and remained similar in mature grains.

Because *japonica* cultivars are generally harder to thresh than *indica* cultivars (*19*),

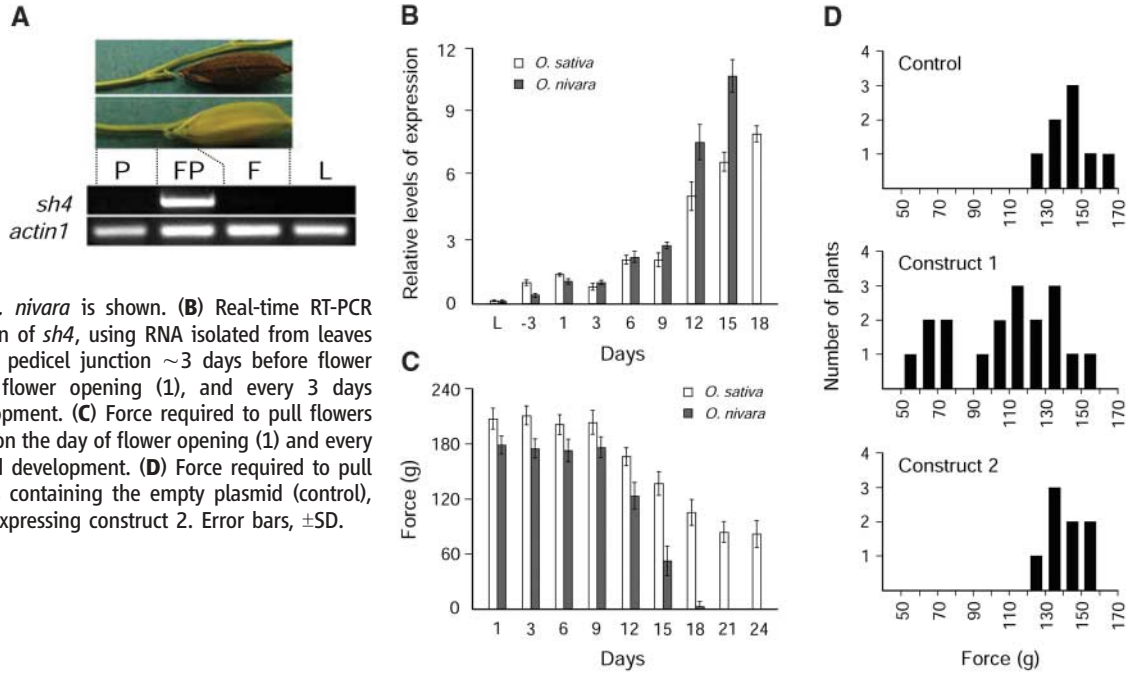

Fig. 3. Expression of *sh4* and flower and grain detachment. (**A**) RT-PCR results, using total RNA isolated from the flower and pedicel junction (FP), from the remaining portions of pedicel (P) and unopened flower (F), and from leaves (L). Above the flower, the separation location of a mature grain from the pedicel of *O. nivara* is shown. (**B**) Real-time RT-PCR estimate of relative expression of *sh4*, using RNA isolated from leaves (L) and the flower/grain and pedicel junction ~3 days before flower opening (−3), the day of flower opening (1), and every 3 days thereafter during seed development. (**C**) Force required to pull flowers or grains away from pedicels on the day of flower opening (1) and every 3 days thereafter during seed development. (**D**) Force required to pull away grains of transformants containing the empty plasmid (control), expressing construct 1, and expressing construct 2. Error bars, ±SD.

Taipei 309 had a stronger grain attachment to the pedicel than the *indica* mapping parent (Fig. 3, C and D). Accordingly, the abscission layer of the *japonica* cultivar showed a higher degree of discontinuity and further retreat from the vascular bundle. The transgenic plants with the strength of grain attachment reduced to less than 100 g had substantially improved abscission layers that were more continuous and extended closer to the vascular bundles (Fig. 4D).

The results indicate that *sh4* plays an important role in the establishment of the abscission layer from the early stage of flower development. The increased expression of *sh4* in the late stage of seed maturation suggests that the gene may also play a role in the activation of the abscission process. One or both of the roles were undermined by the amino acid substitution of asparagine for lysine in cultivated rice.

In the process of rice domestication, human selection was likely to have favored mutations that reduced grain shattering but did not eliminate the formation or function of the abscission layer. In this way, grain loss due to shattering was largely prevented during harvest while a certain level of grain abscission was maintained so that the yield increase was not offset by creating difficulties in threshing. The inverse correlation between the expression level of *sh4* and the strength of grain attachment in *O. sativa* at the late stage of grain maturation seems to suggest that the amino acid substitution did not knock out the gene function in cultivated rice.

The slower pace of increase in the level of *sh4* expressed in *O. sativa* than in *O. nivara* during grain maturation might have been a result of selection in the regulatory region of the gene for a finer adjustment of the shattering/threshing balance during rice cultivation. A comparison of the regulatory sequences of *sh4*, the levels of gene expression, and the phenotypic difference among diverse rice cultivars should provide further insights into the genetic basis of agricultural selection continued through the history of rice cultivation.

Genetic analyses of crop domestication, especially the cloning of domestication-related genes, have shed a light on plant development and evolution. Mutations in regulatory genes were found responsible for drastic morphological modifications during maize and tomato domestication (*20*–*23*). Here we show that the substitution of a neutral for a positively charged amino acid in a predicted DNA binding domain led to a physiological transition key to rice domestication. This is consistent with the finding that positively charged amino acids are critical residues on the surface of DNA binding domains (*24*). The cloning of *sh4* opens opportunities for understanding the developmental pathway of programmed cell separation and seed dispersal in monocotyledonous plants and potentially for optimizing the methods of grain harvest.

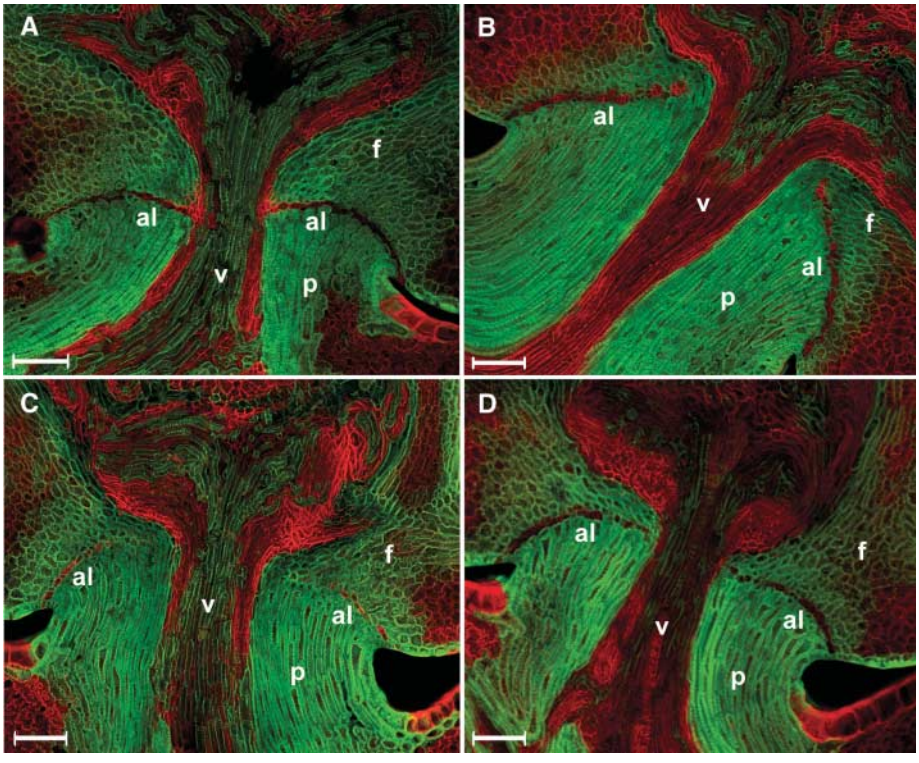

Fig. 4. Fluorescence images of longitudinal section of flower and pedicel junction. (**A**) *O. nivara* mapping parent, with complete abscission layer (al). (**B**) *O. sativa* ssp. *indica* mapping parent, with incomplete abscission layer. (**C**) *O. sativa* ssp. *japonica* Taipei 309, with incomplete abscission layer. (**D**) Transformant of Taipei 309 expressing construct 1, with improved abscission layer. f, flower side; p, pedicel side; v, vascular bundle. Bar, 50 μm.

References and Notes

1. J. R. Harlan, *Crops & Man* (American Society of Agronomy, Madison, WI, 1975).
2. J. F. Hancock, *Plant Evolution and The Origin of Crop Species* (CABI, Cambridge, MA, ed. 2, 2004).
3. V. Poncet *et al.*, *Theor. Appl. Genet.* **100**, 147 (2000).
4. A. H. Paterson, *New Phytol.* **154**, 591 (2002).
5. C.-B. Li, A.-L. Zhou, T. Sang, *New Phytol.* **170**, 185 (2006).
6. H. I. Oka, *Origin of Cultivated Rice* (Japan Scientific Society Press, Tokyo, 1988).
7. S. D. Sharma, S. Tripathy, J. Biswal, in *Rice Breeding and Genetics: Research Priorities and Challenges*, J. S. Nanda, Ed. (Science Publications, Enfield, NH, 2000), pp. 349–369.
8. L. Xiong, K. Liu, K. Dai, C. Xu, Q. Zhang, *Theor. Appl. Genet.* **98**, 243 (1999).
9. H. W. Cai, H. Morishima, *Theor. Appl. Genet.* **100**, 840 (2000).
10. Sobrizal, K. Ikeda, P. L. Sanchez, A. Yoshimura, *Rice Genet. Newslett.* **16**, 74 (1999).
11. Y. S. Nagai *et al.*, *Rice Genet. Newslett.* **19**, 74 (2002).
12. Q. Feng *et al.*, *Nature* **420**, 316 (2002).
13. To screen a large number of seedlings for recombinants, we developed a method of rapid DNA isolation for plants with tough leaves unsuitable for high-throughput protocols (*25*).
14. A. J. Garris, T. H. Tai, J. Coburn, S. Kresovich, S. McCouch, *Genetics* **169**, 1631 (2005).
15. D. A. Vaughan, *The Wild Relatives of Rice: A Genetic Resources Handbook* (International Rice Research Institute, Philippines, 1994).
16. J. L. Riechmann *et al.*, *Science* **290**, 2105 (2000).
17. D. Choi, Y. Lee, H.-T. Cho, H. Kende, *Plant Cell* **15**, 1386 (2003).
18. S. E. Patterson, *Plant Physiol.* **126**, 494 (2001).
19. T. T. Chang, in *Rice: Origin, History, Technology, and Production*, C. W. Smith, R. H. Dilday, Eds. (Wiley, New Jersey, 2003), pp. 3–25.
20. J. Doebley, A. Stec, L. Hubbard, *Nature* **386**, 485 (1997).
21. A. Frary *et al.*, *Science* **289**, 85 (2000).
22. J. Liu, J. Van Eck, B. Cong, T. D. Tanksley, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13302 (2002).
23. H. Wang *et al.*, *Nature* **436**, 714 (2005).
24. K. Yamasaki *et al.*, *Plant Cell* **17**, 944 (2005).
25. Materials and methods are available as supporting material on *Science* Online.
26. Data have been deposited into GenBank with accession numbers DQ383371 to DQ383414. We thank D. Choi for helping set up rice transformation; S. Owens for suggesting and assisting with confocal microscopy; T. Briggeman for photographing; M. Yano for providing plasmids; J.-P. Hu and J.-L. Fan for helping with the subcellular localization; F. Ewers, N. Gibson, M. Grillo, and W.-X. Zhu for discussion and comments on the manuscript; and S. Ge, B.-R. Lu, and the International Rice Research Institute for providing DNA samples and plant material. The research was supported by the National Science Foundation (USA) and the Rackham Research Endowment Fund.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1123604/DC1
Materials and Methods
Fig. S1
Tables S1 and S2
References

8 December 2005; accepted 28 February 2006
Published online 9 March 2006;
10.1126/science.1123604
Include this information when citing this paper.