

Supporting Information

Koide et al. 10.1073/pnas.1711656115

SI Introduction

Genetic factors for hybrid sterility and TRD have been found in the region proximate to the S_1 gene as follows. The factor for the S_1 locus-mediated m TRD was mapped ~40 kb between markers E0506 and E1920 (Fig. S2) (1). In addition, the modifier for the female TRD (f TRD) was located in the region between markers G8008 and RM204 (1) by using the genetic background of T65 (a japonica type of *O. sativa*). The factor for the S_1 locus-mediated f TRD was also mapped into the region between markers RM19357 and RMC6_22028 by using various rice genetic backgrounds (Fig. S2) (2). In addition, Xie et al. (3) mapped the S_1 gene, which causes semisterility in pollen grains and seeds in heterozygotes (S_1^s/S_1^s), to the region between markers 2,170 and 2,199 by using the genetic background of Nipponbare (a japonica type of *O. sativa*). The region for the S_1 locus-mediated m TRD included the regions for the S_1 locus-mediated f TRD and pollen and spikelet sterilities, indicating that the S_1 locus is located in this region, although sex-specificities in the TRD differed. Because our previous study showed that the effect of the S_1 locus-mediated TRD differs slightly depending on the genetic background of the strains used (1, 4), this inconsistency might be due to differences in the genetic backgrounds. We showed that the S_1 locus causes the si TRD in the genetic background of Acc108 (an indica type of *O. sativa* used as a genetic background in the present study).

In addition to the factor for f TRD, Garavito et al. (2) predicted the presence of other factors (S_1A , S_1B , and S_1C). These factors interact with the S_1 locus and determine the viability of female gametes that have the S_1^s allele (2). The regions for the S_1 locus, S_1A and S_1B , were denoted as the “ S_1 regions” (Fig. S2) (5). In the present study, we used “glaberrima sterility complex” (GSC) instead of the “ S_1 regions” (5) to avoid potential confusion with the S_1 locus, although the terms indicate the same chromosomal region.

SI Materials and Methods

Genetic Materials. Pehkuh (an indica type of *O. sativa*, denoted “Acc108”), Taichung 65 (a japonica type of *O. sativa*, denoted “T65”), and three NILs, Acc108 S_1 , T65 wx , and T65 S_1 , were used in the present study. T65 wx contains the $waxy$ (wx) allele at the Wx locus introduced from Kinoshitamochi (6). The Wx locus encodes a granule-bound starch synthase that is required for amylose synthesis; Wx is a functional allele; wx is a nonfunctional allele (7). The Wx gene is expressed in pollen grains, and pollen phenotypes are easily distinguished by I₂-KI staining (see below). Since the Wx locus is tightly linked to the S_1 locus in a coupling phase, the phenotypic difference in pollen grains was used as a marker to detect TRD caused by the S_1 locus (1). Acc108 S_1 and T65 S_1 have the chromosomal segment proximate to the S_1 region, introduced from a strain of *O. glaberrima* (W025 from Guinea). Acc108, T65, and T65 wx contain the S_1^s allele, whereas Acc108 S_1 and T65 S_1 contain the S_1^s allele at the S_1 locus.

Evaluation of Pollen and Seed Fertility. Spikelets at the flowering stage were fixed and stored in 70% ethanol until use. Anthers sampled from a spikelet were squashed in a 0.5% I₂-KI solution on a glass slide. Stained and unstained pollen grains were scored as fertile and sterile, respectively. Sterile pollen grains were further classified as TY or SP abortion types (8). TY-type pollen grains are irregular in shape; SP-type pollen grains are spherical (Fig. 1F). Wx - and wx -type pollen grains were also distinguished by I₂-KI staining: Wx -type pollen grains stain deep blue; wx -type pollen grains stain light brown. Seed fertility was determined by

counting the number of filled and unfilled spikelets of three panicles for each plant. To minimize the effect of the growing environment on seed fertility, the highest value of the seed fertility in the three panicles was used as the seed fertility for an individual plant. The mean fertility for three individual plants per line was used as the seed fertility for the line.

Cytological Observation, Expression Analysis, and in Situ Hybridization.

Spikelets sampled from panicles before the flowering stage were fixed in FAA (1:1:18 formalin–glacial acetic acid–50% ethanol) for 24 h at 4 °C for histological analysis or were fixed with paraformaldehyde (4% wt/vol paraformaldehyde and 1% Triton X-100 in 0.1 M sodium phosphate buffer) for 48 h at 4 °C for in situ hybridization. Ovules were then dehydrated in a graded ethanol–butanol series and embedded in Paraplast Plus (McCormick Scientific). The samples were cut into 8- μ m-thick sections. For histological analyses, sections were stained with Delafield’s hematoxylin (Muto Pure Chemicals) and observed with a light microscope. For in situ hybridization, cDNA fragments of *SSP-gla* and *OgTPRI* were cloned into T-Vector, pMD20 (TaKaRa) and pCR-Blunt II-TOPO (Invitrogen), respectively. For single-target in situ hybridization, digoxigenin-labeled RNA probes of *SSP-gla* and *OgTPRI* were prepared by in vitro transcription using SP6 and T7 RNA polymerase, respectively. In situ hybridization and immunological detection with alkaline phosphatase were performed according to the methods of Kouchi and Hata (9). For double-target in situ hybridization, a biotin-labeled probe of the *SSP* transcript and a digoxigenin-labeled probe of the *OgTPRI* transcript were used. The detection of probe signals and DAPI staining were performed according to the method described in Yoshikawa et al. (10).

For the expression analysis, total RNA of *O. glaberrima* (IRGC103777) was isolated from the plumules, leaf blades, and internodes at the reproductive stage, from roots in the seedling stage, and from anthers and ovules before flowering using TRIzol Reagent (Thermo Fisher Scientific). Total RNA was also isolated from young panicles and leaf blades at the stage before flowering using the RNeasy Plant Mini Kit (QIAGEN). cDNAs were synthesized using the SMARTer RACE cDNA Amplification Kit (Clontech) or the QuantiTect Reverse Transcription Kit (QIAGEN). Expression of the *SSP* gene and *OgTPRI* gene was analyzed by RT-PCR using KOD FX Neo DNA polymerase (Toyobo). The 5′ and 3′ ends of the transcripts of the *SSP* gene were identified by 5′ and 3′ RACE analysis, respectively. For quantifying the expression level of *SSP* and *OgTPRI*, quantitative RT-PCR was conducted using SYBR GreenER qPCR SuperMix (Invitrogen).

Genetic Mapping. To map the S_1 locus in the genetic background of Acc108, we genotyped 223 segregating plants derived from a cross between Acc108 and Acc108 S_1 . Since the S_1 locus induces si TRD, and gametes possessing the S_1^s allele are aborted in the heterozygote (S_1^s/S_1^s), pollen and seed fertility were used to map the S_1 locus. To map the S_1^{mut} , we genotyped 174 segregating plants obtained from a cross between Acc108 S_1M and Acc108. For analyzing the presence of the S_1 locus-mediated si TRD, genotypes of progenies were determined using markers in the GSC. DNA markers used for the mapping are shown in Table S3.

Genomic DNA-Sequencing Analysis. Genomic DNA was isolated from 100 mg of frozen rice leaves using the DNeasy Plant Mini Kit (QIAGEN). Quantified DNA (5 μ g) was used to prepare libraries with the TruSeq Rapid PE Cluster Kit (Illumina). Sequencing was

performed by an Illumina HiSeq2500 DNA sequencer. To identify mutations in the GSC occurring in the mutant Acc108S₇M, we generated a pseudomolecule of the GSC of *O. glaberrima* by concatenating sequences of publicly available BAC clones of CG14 (a variety of *O. glaberrima*) (5) because the GSC of Acc108S₇ was originally introduced from *O. glaberrima*. We then generated a reference genome sequence of the GSC of the WT Acc108S₇ using Perl scripts of the MutMap pipeline version 1.0 prepared by the Iwate Biotechnology Research Center (11). We obtained 14.0 million paired-end short reads from Acc108S₇. These reads were aligned with the Burrows–Wheeler Aligner (BWA) tool to a pseudomolecule of the GSC of *O. glaberrima*. The alignment file was converted to sequence alignment/map (SAM) or binary alignment/map (BAM) tools using SAMtools (www.htslib.org/). To identify reliable SNPs, we applied the filter pipeline Coval (12). We then generated a reference genome sequence of the GSC of WT Acc108S₇ by replacing CG14 nucleotides with Acc108S₇ nucleotides at SNP positions. The aforementioned steps were automatically processed using the MutMap pipeline version 1.0 (11). We then obtained 26.5 million paired-end short reads from the mutant Acc108S₇M. SNPs, and insertions or deletions (indels) were detected by aligning these short reads to a reference genome sequence of the GSC of WT Acc108S₇ using SAMtools in the MutMap pipeline version 1.0 as described above.

In addition, we compared the nucleotide sequences of the GSC of Acc108 and bulked Os-type homozygous plants in the M₂ family to confirm that a mutation occurred on the Og-derived chromosome segment. We obtained 15.3 and 12.4 million paired-end short reads from Acc108 and the bulked sample of six Os-type homozygous plants, respectively. A pseudomolecule of the GSC of Acc108 was generated by aligning short reads of Acc108 to a Nipponbare genome sequence (IRGSP-1.0) and replacing the nucleotide of Nipponbare with those of Acc108 nucleotides at SNP positions. We then aligned short reads obtained from the bulked sample and detected SNPs and indels. All procedures were conducted using MutMap pipeline version 1.0 as described above. DNA sequencing data have been deposited in the DNA Data Bank of Japan (DDBJ) under accession numbers DRA006416–DRA006419.

RNA-Seq Analysis. Total RNA was isolated from anthers at the stage before flowering using TRIzol reagent (Invitrogen). Sequencing libraries were prepared using 1 µg of total RNA according to the protocol in the TruSeq RNA Sample Preparation Kits v2 Guide (Illumina). Sequencing was performed by the KURABO Bio-Medical Department (Kurabo) using Illumina HiSeq. 2000. Short 100-bp paired-end reads were filtered by trimming low-quality [quality value (QV) <19] bases from the 5' and 3' ends of each read. In addition, short reads that had a high percentage (percentage threshold = 30) of low-quality (QV <14) bases were removed. The remaining high-quality reads (23.4 and 19.0 million reads from Acc108S₇ and Acc108S₇M, respectively) were then mapped onto the reference genome sequence of the S₇ locus of Acc108S₇ using Tophat2 software (<https://ccb.jhu.edu/software/tophat/index.shtml>). In addition, RNA-seq data of young panicle and leaf samples from *O. glaberrima* (SRR1174376 and SRR1174378) were obtained from the public database (www.ddbj.nig.ac.jp/) and were used for the analysis. All the aforementioned steps were conducted using the DDBJ Read Annotation Pipeline (13). Gene structure was analyzed using Cufflinks version 2.1.1 (<http://cole-trapnell-lab.github.io/cufflinks/>).

Phylogenetic and Expression Analysis of Trypsin-Like Peptidase Domain-Containing Genes. A total of 17 genes that encode trypsin-like peptidase domain (trypsin 2)-containing protein were annotated in the *O. sativa* cv. Nipponbare genome sequence in the public database Pfam 31.0 (pfam.xfam.org/). Their sequences were downloaded from the database of the Rice Annotation

Project (RAP-DB), and amino acid sequences of peptidase domains were used for the phylogenetic study. The phylogenetic relationship was inferred by the maximum likelihood method using MEGA7 (www.megasoftware.net/) with 1,000 bootstrap replications. Global expression profiles of 14 genes that encode trypsin-like peptidase domain (trypsin 2)-containing protein were compared using the public database RiceXpro (ricexpro.dna.affrc.go.jp/). Gene ontology analysis was conducted using the database Panther version 12.0 (pantherdb.org/).

Genetic Complementation of the Mutant Acc108S₇M. We amplified 7.8- and 4.1-kb DNA fragments containing the entire coding region of the intact *SSP* gene (*SSP-gla*). These fragments also contained 4.6- and 1.0-kb upstream regions, respectively, and a 1.5-kb downstream region of the *SSP* gene. We then subcloned the DNA fragment into a binary vector pGWB1 and produced pGWB1-7.8 kb and pGWB1-4.1 kb. The binary vector was introduced into *Agrobacterium tumefaciens* EHA105. For the transformation, we used two NILs, T65_{wx} and T65S₇, which have the T65 genetic background, because of the high transformation efficiency of T65. The 7.8- and 4.1-kb DNA fragments were introduced into T65_{wx} plants, and the 7.8-kb fragment was also introduced into T65S₇ plants.

For genetic complementation of the Acc108S₇M mutant, four lines (lines #1 and #4 with the 7.8-kb fragment and lines #5 and #6 with the 4.1-kb fragment) obtained from the independent transgenic event and having normal morphology and high pollen and seed fertilities were selected. The selected lines were crossed with Acc108S₇M and Acc108. To complement the mutation, we developed F₁ hybrids derived from the cross between the transformed plant (T65_{wx} + *SSP-gla*) and Acc108S₇M, instead of transforming *SSP-gla* into Acc108S₇M and crossing resultant transformants (Acc108S₇M + *SSP-gla*) with T65_{wx}. This was because a gene-transformation protocol for strain Acc108 has not been established, and hybrids derived from the former cross are genetically identical to those from the latter cross. Transgenic positive (+) and negative (–; null) segregants in resultant F₁ plants were determined using PCR markers. For analyzing the presence of the S₇ locus-mediated *s*TRD, genotypes of progenies were determined using markers in the S₇ locus.

Distribution Analysis of the *SSP* Gene. The distribution of the *SSP* gene in 112 accessions in nine species in *Oryza* was examined by PCR and genomic sequencing. To determine the presence of *SSP*, we designed PCR primers to amplify a portion of the first exon of the *SSP* gene (Table S3). In addition, to determine the presence of the insertion that contained the *SSP* gene, we designed primers to anneal to the upstream and downstream regions of the insertion. For PCR analysis, genomic DNA was extracted from five accessions of *O. sativa*, 20 of *Oryza rufipogon*, 2 of *O. glaberrima*, 2 of *O. barthii*, 5 of *O. meridionalis*, 3 of *O. officinalis*, 1 of *O. minuta*, and 1 of *O. ridleyi* (Table S2). Among them, six species, *O. sativa*, *O. rufipogon*, *O. glaberrima*, *O. barthii*, *O. longistaminata*, and *O. meridionalis*, share the same genome AA. *O. sativa* and its wild ancestor *O. rufipogon* form the Asian rice gene pool; *O. glaberrima* and *O. barthii* form the African rice gene pool (14). Accessions of *O. officinalis*, *O. minuta*, and *O. ridleyi* were used as outgroups.

For genomic sequencing, we used short-read data from the DDBJ sequence read archive (DRA, www.ddbj.nig.ac.jp/index-e.html). We downloaded short-read archive (SRA) files for 26 accessions of *O. sativa*, 10 of *O. rufipogon*, 16 of *O. glaberrima*, and 20 of *O. barthii* (15, 16). *O. sativa* has been classified into either six groups (I–VI) or five (*indica*, *aus*, *aromatic*, *temperate japonica*, and *tropical japonica*) (17, 18). In addition, 24.8 million short reads were obtained from accession W1508 of *O. longistaminata* using MiSeq (Illumina). The downloaded SRA files were converted to the FASTQ format using the SRA toolkit version 2.4.2. (<https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/>).

The resultant FASTQ files were uploaded to the DDBJ read annotation pipeline (www.ddbj.nig.ac.jp) and analyzed. To determine the presence of the *SSP* gene, we mapped short reads onto the 1,934-bp genomic sequence that includes the entire region of the *SSP* gene, using the software Bowtie2 (bowtie-bio.sourceforge.net) and calculated coverages (the length of nucleotide sequence where short reads were mapped divided by the length of the total nucleotide sequence). The data regarding the presence or absence of the *SSP* gene revealed by these analyses were compiled and arranged on phylogenetic relationships of *Oryza* species (19, 20) in Fig. 5.

SI Results

Histological Analysis. Rice strains Acc108 (S_1^s carrier) and the NIL Acc108 S_1 (S_1^S carrier) had a high level of pollen and seed fertility, whereas seed and pollen of their F_1 hybrids were sterile (Fig. 1 *D* and *F*). Abnormal embryo sacs were observed in female gametophytes in the F_1 hybrid (Fig. 1*E*). Embryo sacs had degenerated by the time a functional sporophyte had formed (Fig. S1 *C–H*), suggesting that the abnormality occurs after meiosis during female gametogenesis. During male gametogenesis, a TY type of abortion was frequently noted in pollen grains in the F_1 hybrid (Fig. 1*F*). Abnormal pollen grains with only one nucleus were observed in the late binucleate stage of pollen development (Fig. S1 *I–N*), suggesting that an abnormality occurs before or during the uninucleate stage of male gametogenesis. These results indicate that seed and pollen sterility in the F_1 hybrids is caused by developmental failure during male and female gametogenesis.

Genetic Mapping of the S_1 Locus in the Acc108 Genetic Background. To analyze the effect of the genetic background of Acc108 on the S_1 locus-mediated TRD, we mapped the S_1 locus by genotyping 223 segregating plants derived from a cross between Acc108 and Acc108 S_1 (Fig. 2*A*). The F_1 hybrids derived from a cross between Acc108 and Acc108 S_1 showed low pollen and seed fertility (31.8% and 54.0%, respectively). In addition, a marked excess of the Og-type allele was observed in the progenies of the F_1 hybrids, indicating the occurrence of *si*TRD in this cross combination, as observed in our previous studies (1, 4). From the 223 segregating plants, we selected three plants that had the heterozygous genotype (Og/Os) at DNA markers E0506 and E1920 and recombination points between DNA markers RM19350 and RM19369. These three plants (one and two plants showing R-1 and R-2 genotypes, respectively) showed low pollen and seed fertility (Fig. 2*A*). Excess of the Og-type allele was also observed (Fig. 2*A*). These results indicated that the factor for *si*TRD is present in the interval between RM19350 and RM19369 where the S_1 locus is located. These results also indicated that the S_1 locus causes the *si*TRD in the genetic background of Acc108.

Heavy-Ion Beam Irradiation and Mutant Screening. A total of 2,478 F_1 seeds were obtained after artificial pollination of Acc108 stigmas with Acc108 S_1 pollen for heavy-ion beam irradiation. A total of 841 and 976 dry seeds were irradiated with carbon ion beams (LET 30 KeV/ μ m) at a dose of 150 Gy during 2012 and 2013, respectively, at the RIKEN RI-beam factory, Wako, Japan (21, 22). In addition, 322 and 339 dry seeds were irradiated with argon ion beams (LET 286 KeV/ μ m) at doses of 7.5 and 10 Gy, respectively, during 2013. The irradiated seeds were sown in the experimental field at the Japan International Research Center for Agricultural Sciences, Tsukuba, Japan, and at the experimental farm at Kyoto University, Kyoto, Japan, during 2012 and 2013, respectively.

The irradiated F_1 plants were expected to produce sterile panicles if there were no mutations at the S_1 locus or at another locus that inhibited the action of the S_1 locus, since all irradiated plants were heterozygous (S_1^S/S_1^s) for the S_1 locus. In contrast, plants with a mutation at the S_1 locus or at another locus that

inhibits action of the S_1 locus were expected to produce a fertile panicle(s). A total of 1,914 panicles were randomly harvested from plants irradiated with carbon ion beams to evaluate their seed fertility during 2012 and 2013. In addition, 839 and 691 panicles were randomly harvested from plants irradiated with argon ion beams at doses of 7.5 and 10 Gy, respectively, during 2013. From 2,478 F_1 hybrids (M_1 generation) (1,817 and 661 plants irradiated with carbon and argon ion beams, respectively), we obtained one plant that had a panicle with >50% seed fertility from the M_1 population irradiated with carbon ion beams (Fig. S7*A*). We used this plant for further analysis.

Genetic Mapping of the Mutation. We used seeds in the panicle of the M_1 plant with >50% seed fertility to develop the M_2 family. In the M_2 family, an excess of Og-type alleles was not observed, but an excess of Os-type alleles was observed (Fig. S7*B*). Among the 37 plants of the M_2 family, 19 were heterozygous (Og/Os) for the marker in the S_1 locus (Fig. S7*B*). We observed seed fertility of these heterozygotes to determine whether the mutation that inhibits the action of the S_1 locus is genetically independent from the S_1 locus. If the mutation was genetically independent from the S_1 locus, one fourth of the heterozygous (Og/Os) plants in the M_2 family would show seed sterility. The results showed that the level of seed fertility in all heterozygous plants was similar to that of homozygous plants (Fig. S7*B*), suggesting that a mutation that affects seed fertility is linked to the S_1 locus.

Then, we examined whether the mutation occurred on the *O. glaberrima*-derived chromosome segment or *O. sativa*-derived chromosome by carrying out crossing experiments using the M_2 family (Fig. S7*C*). We crossed the M_2 plant that is Og-type homozygous for the S_1 locus-linked markers with Acc108. In addition, we randomly selected six M_2 plants that are Os-type homozygous for the S_1 locus-linked markers and crossed them with Acc108 S_1 . If the mutation is on the *O. glaberrima*-derived chromosome segment, the hybrid derived from the former cross will show high seed fertility, while those from the latter cross will show the S_1 locus-mediated hybrid sterility. In contrast, if the mutation is on the *O. sativa*-derived chromosome, the hybrid derived from the former cross will show the S_1 locus-mediated hybrid sterility, while those from the latter cross will show high seed fertility. The results showed that hybrids derived from the Og-type homozygous M_2 plant and Acc108 showed high seed fertility (82.3%), while those from Os-type homozygous M_2 plants and Acc108 S_1 showed low seed fertility (8.8–21.8%) (Fig. S7*C*). These results indicated that a mutation occurred on the *O. glaberrima*-derived chromosome segment. To confirm that a mutation occurred on the *O. glaberrima*-derived chromosome segment, we compared nucleotide sequences of the GSC of Acc108 and bulked Os-type homozygous plants in the M_2 family (DDBJ sequence data nos. DRA006418 and DRA006419). We did not find any differences in the 273-kb region in the GSC between Acc108 and Os-type homozygous plants in the M_2 family. Based on these results, we concluded that the mutation is genetically linked to the S_1 locus and is located on the chromosome segment derived from *O. glaberrima*. Therefore, we self-pollinated an Og-type homozygote from the M_2 plants and designated the resultant line as mutant Acc108 S_1 M.

To examine the detailed location of the mutation on chromosome 6, we then genetically mapped the mutation using 174 segregating plants obtained from a cross between Acc108 S_1 M and Acc108. F_1 hybrids derived from the cross between Acc108 S_1 M and Acc108 showed higher pollen and seed fertility (81.0% and 92.0%, respectively) than those derived from the cross between Acc108 and Acc108 S_1 (Fig. 2*A*). This result indicated the presence of the mutation that affects the S_1 locus-mediated hybrid sterility in Acc108 S_1 M. We evaluated pollen

and seed fertilities of two recombinants (R-4 and R-5) that have a recombination point in the region adjacent to the S_1 locus. R-4 and R-5 have a recombination point between markers E1920 and RM19369 and between markers RM19350 and E0506, respectively (Fig. 2A). As mentioned earlier, we found that the mutation is linked to the S_1 locus and is located on the chromosome segment derived from *O. glaberrima*. Therefore, if the mutation that affects the S_1 locus-mediated hybrid sterility is located outside the region between RM19369 and RM19350, one of these two recombinants does not have the mutation that affects the action of the S_1 locus. Such a plant will show pollen and seed sterility because heterozygosity (S_1^g/S_1^g) at the S_1 locus is effective. However, it turned out that both R-4 and R-5 had high pollen and seed fertility (Fig. 2A). These results indicated that the mutation that affects the S_1 locus-mediated hybrid sterility is not located outside the 273-kb region between markers RM19350 and RM19369 in which the S_1 locus was mapped but is located within the region (Fig. 2A). Based on these results, we

speculated that mutant Acc108 S_1 M has a mutated allele, designated " S_1^{mut} ," at the S_1 locus.

The Presence of the Gene Responsible for the Excess of the Os-Type Allele. An excess of the Os-type allele was observed in the progenies of hybrids derived from the cross between Acc108 S_1 M and Acc108 and in the progenies of R-5 (Fig. 2A and Fig. S7D). In contrast, Mendelian segregation was observed in the progenies of R-4. These results suggested that another gene(s) located in the proximal region (centromere side) of the chromosome 6 is associated with the excess of the Os-type allele in the progenies. To confirm the existence of this gene(s), we evaluated genotypes of the progeny of the recombinant R-6, which was derived from the segregation population and has a recombination point in the proximal region (Fig. 2A). As a result, the excess of Os-type allele was observed, indicating the presence of another gene(s) causing the excess of Os-type allele in a region that differs from the S_1 locus (Fig. 2A). This gene was not investigated further in this study.

- Koide Y, et al. (2008) Sex-independent transmission ratio distortion system responsible for reproductive barriers between Asian and African rice species. *New Phytol* 179:888–900.
- Garavito A, et al. (2010) A genetic model for the female sterility barrier between Asian and African cultivated rice species. *Genetics* 185:1425–1440.
- Xie Y, et al. (2017) Interspecific hybrid sterility in rice is mediated by *OgTPR1* at the S_1 locus encoding a peptidase-like protein. *Mol Plant* 10:1137–1140.
- Sano Y (1990) The genic nature of gamete eliminator in rice. *Genetics* 125:183–191.
- Guyot R, et al. (2011) Patterns of sequence divergence and evolution of the S orthologous regions between Asian and African cultivated rice species. *PLoS One* 6:e17726.
- Oka H (1974) Analysis of genes controlling f(1) sterility in rice by the use of isogenic lines. *Genetics* 77:521–534.
- Sano Y (1984) Differential regulation of *waxy* gene expression in rice endosperm. *Theor Appl Genet* 68:467–473.
- Virmani SS, Sun ZX, Mou TM, Jauhar Ali A, Mao CX (2003) *Two-Line Hybrid Rice Breeding Manual* (Int Rice Res Inst, Los Baños, Philippines), p 88.
- Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol Genet* 238:106–119.
- Yoshikawa T, Eiguchi M, Hibara K, Ito J, Nagato Y (2013) Rice *slender leaf 1* gene encodes cellulose synthase-like D4 and is specifically expressed in M-phase cells to regulate cell proliferation. *J Exp Bot* 64:2049–2061.
- Abe A, et al. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30:174–178.
- Kosugi S, et al. (2013) Coval: Improving alignment quality and variant calling accuracy for next-generation sequencing data. *PLoS One* 8:e75402.
- Nagasaki H, et al. (2013) DDBJ read annotation pipeline: A cloud computing-based pipeline for high-throughput analysis of next-generation sequencing data. *DNA Res* 20:383–390.
- Harlan JR, de Wet JMJ (1971) Toward a rational classification of cultivated plants. *Taxon* 2:509–517.
- Xu X, et al. (2011) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotechnol* 30:105–111.
- Wang M, et al. (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat Genet* 46:982–988.
- Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169:1631–1638.
- Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. *Theor Appl Genet* 74:21–30.
- Zhang QJ, et al. (2014) Rapid diversification of five *Oryza* AA genomes associated with rice adaptation. *Proc Natl Acad Sci USA* 111:E4954–E4962.
- Zhu T, et al. (2014) Phylogenetic relationships and genome divergence among the AA- genome species of the genus *Oryza* as revealed by 53 nuclear genes and 16 intergenic regions. *Mol Phylogenet Evol* 70:348–361.
- Kazama Y, et al. (2011) Comprehensive identification of mutations induced by heavy-ion mutagenesis in *Arabidopsis thaliana*. *BMC Plant Biol* 11:161.
- Abe T, Kazama Y, Hirano T (2015) Ion beam breeding and gene discovery for function analyses using mutants. *Nuclear Physics News* 25:30–34.

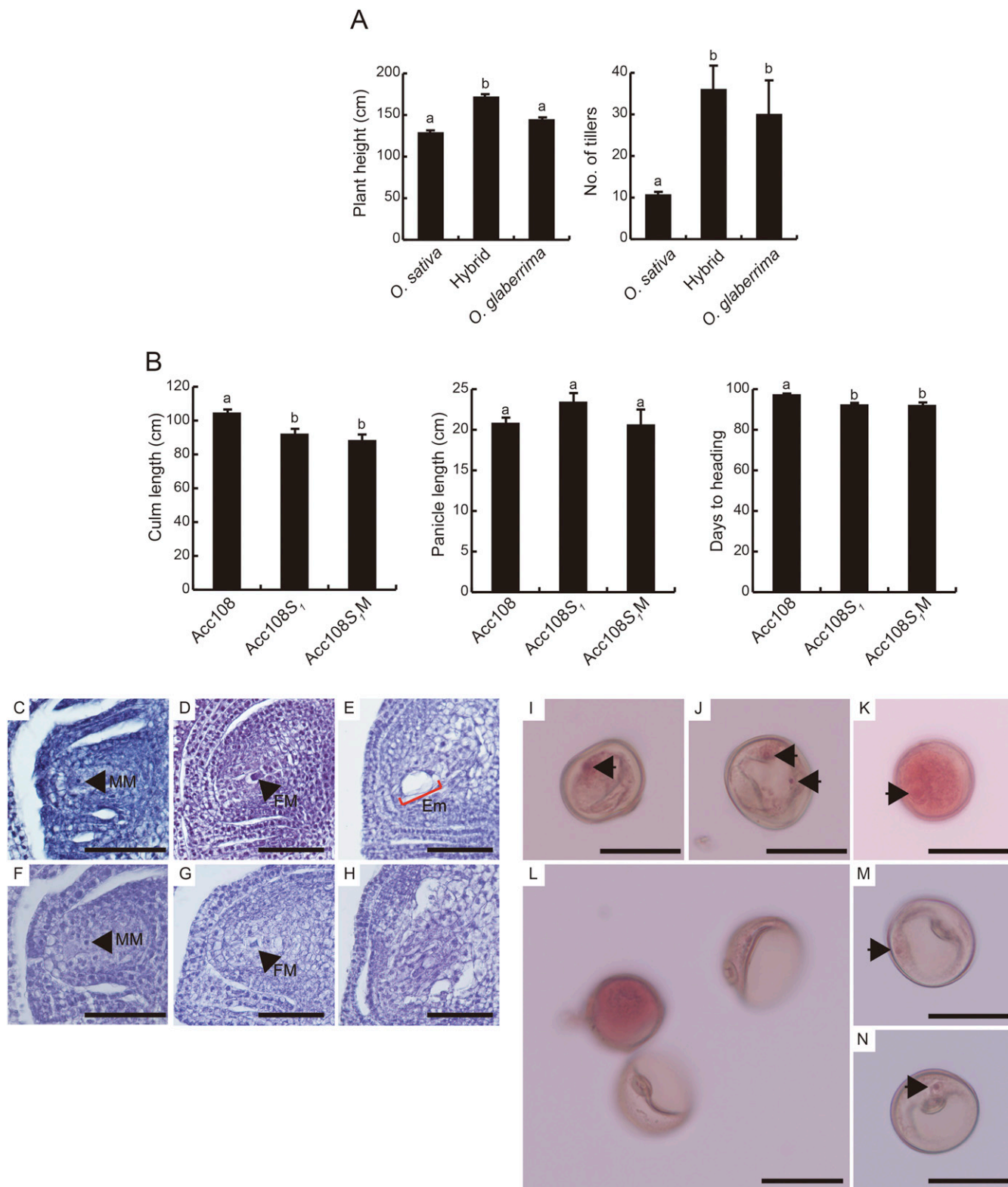


Fig. S1. Phenotype analysis of rice materials. (A) Morphological traits of *O. sativa* (cultivar WAB56-104), *O. glaberrima* (cultivar CG14), and interspecific hybrid. Data are given as means \pm SD ($n = 3$). Means labeled with different letters differed significantly. (B) Morphological traits of Acc108, Acc108S₁, and the mutant Acc108S₁M. Data are given as means \pm SD ($n = 3$). Means labeled with different letters differed significantly. (C–M) Development of embryo sac and pollen in Acc108 and the F₁ hybrid (Acc108 \times Acc108S₁). (C–E) Longitudinal section of the developing embryo sac in Acc108. (Scale bars: 50 μ m.) (C) Megasporocyte. (D) Functional megaspore. (E) Middle-stage embryo sac. (F–H) Longitudinal section of the developing embryo sac in the F₁ hybrid. (Scale bars: 50 μ m.) (F) Megasporocyte. (G) Functional megaspore. (H) Abnormal embryo sac observed in the middle to late stage. (I–K) Developing pollen grains in Acc108. (Scale bars: 25 μ m.) Arrowheads indicate nuclei. (I) Uninucleate stage. (J) Early binucleate stage. (K) Late binucleate stage. (L–N) Developing pollen grains in the F₁ hybrid. (Scale bars: 25 μ m.) Arrowheads indicate nuclei. (L) Normal and abnormal pollen grains at late binucleate stage. (M and N) Abnormal uninucleate pollen grains at late binucleate stage. Em, embryo sac; FM, functional megaspore; MM, megaspore mother cell.

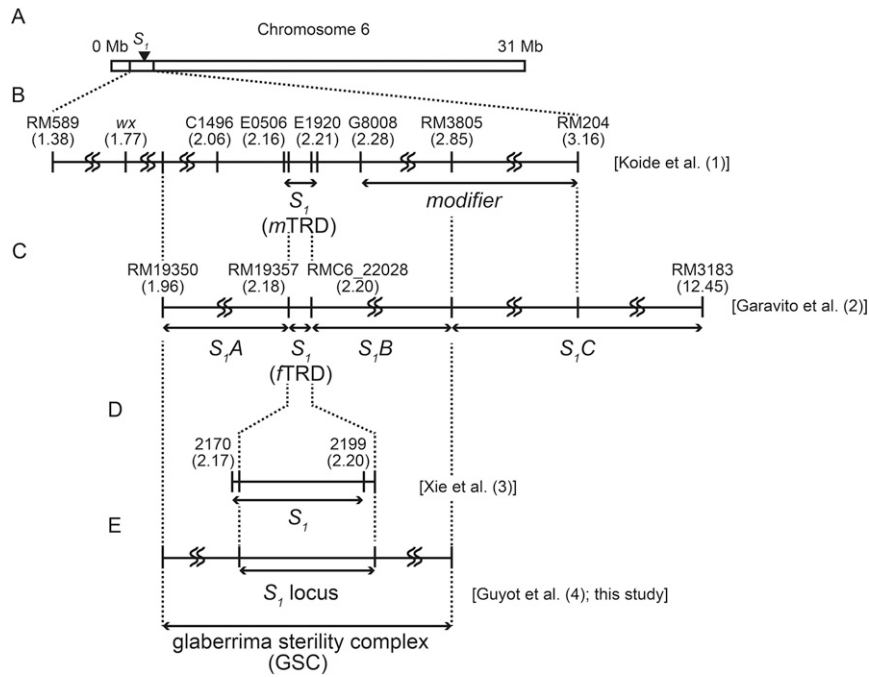


Fig. S2. Graphical summary of previous mapping studies of the S_7 locus. (A) Physical position of the S_7 locus on chromosome 6. (B) Fine mapping of factors for S_7 locus-mediated $mTRD$ (1). Numbers in parentheses indicate physical positions of markers on chromosome 6. The factor for $mTRD$ was mapped between markers E0506 and E1920. In addition, a modifier for $fTRD$ was located in the region between markers G8008 and RM204. (C) Fine mapping of factors for $fTRD$ (2). The factor for $fTRD$ was mapped into the region between markers RM19357 and RMC6_22028. In addition, other factors (S_7A , S_7B , and S_7C) that interact with the S_7 locus and determine the viability of female gametes with allele S_7^g were predicted. (D) The result of mapping of the S_7 locus conducted by Xie et al. (3). The factor for pollen and spikelet fertility was mapped into the region between markers 2,170 and 2,199. (E) Positions of the S_7 locus and GSC equivalent to the S_7 regions used in this and previous (4) studies.

1. Koide Y, et al. (2008) Sex-independent transmission ratio distortion system responsible for reproductive barriers between Asian and African rice species. *New Phytol* 179:888–900.
2. Garavito A, et al. (2010) A genetic model for the female sterility barrier between Asian and African cultivated rice species. *Genetics* 185:1425–1440.
3. Xie Y, et al. (2017) Interspecific hybrid sterility in rice is mediated by *OgTPR1* at the S_7 locus encoding a peptidase-like protein. *Mol Plant* 10:1137–1140.
4. Guyot R, et al. (2011) Patterns of sequence divergence and evolution of the S orthologous regions between Asian and African cultivated rice species. *PLoS One* 6:e17726.

A

```

Acc108S1 ATGTCTGAAGA CAAATCCCAA CCAGTTGGTT GAGACGAGCG CTGAGCTCGC AAGGGACGCG GTGCTCACTG TTATCGTCAT CAAGAACCCT ATCGTCGACT
Acc108SM ATGTCTGAAGA CAAATCCCAA CCAGTTGGTT GAGACGAGCG CTGAGCTCGC AAGGGACGCG GTGCTCACTG TTATCGTCAT CAAGAACCCT ATCGTCGACT

Acc108S1 CTGTCCGATT CAGGGGGGTC GTATTTCGAA CAGCTTTTGT CATCAGAAGC GACGACTGGG GGTCGATCGT CCTCACGGCT CAACACGTTG TTGAGGATCT
Acc108SM CTGTCCGATT CAGGGGGGTC GTATTTCGAA CAGCTTTTGT CATCAGAAGC GACGACTGGG GGTCGATCGT CCTCACGGCT CAACACGTTG TTGAGGATCT

Acc108S1 GCAGCCCGGG GATACACTGC GCGTCCGGAC AATGCTACAA AGCGGCGTCC GAGAGCTATC TGCCTCTGTT CTCTACGAGC ACAGATGGTC AGATGTTGCG
Acc108SM GCAGCCCGGG GATACACTGC GCGTCCGGAC AATGCTACAA AGCGGCGTCC GAGAGCTATC TGCCTCTGTT CTCTACGAGC ACAGATGGTC AGATGTTGCG

Acc108S1 ATTCTGAGTG TTCAGGACT CCGGAATGTG CGGCAGCTTA CCTTTGAACC ATCAGTGCTC AACAGTGACA GTGTAATCGG TGTCGGGTAC GCTAATCCAA
Acc108SM ATTCTGAGTG TTCAGGACT CCGGAATGTG CGGCAGCTTA CCTTTGAACC ATCAGTGCTC AACAGTGACA GTGTAATCGG TGTCGGGTAC GCTAATCCAA

Acc108S1 TTGATCTGTT CCCCAGGATC GAAATGGCGT GTAAGCGTAT TCCTGATTTG TCACCAGGCA GTGTTGTGTC GACTGAATAT GTTGGTTTGT ACCAAGGTGT
Acc108SM TTGATCTGTT CCCCAGGATC GAAATGGCGT GTAAGCGTAT TCCTGATTTG TCACCAGGCA GTGTTGTGTC GACTGAATAT GTTGGTTTGT ACC-----GT

Acc108S1 TGACGTCACC TTTGTTTCATC TTGAAATGGT TGAATGAGA GGCATGTCAG GCATGCCTAT ACTCAACGCA ACAGGGGTCG TTGCCATGCT AATCGAGGGT
Acc108SM TGACGTCACC TTTGTTTCATC TTGAAATGGT TGAATGAGA GGCATGTCAG GCATGCCTAT ACTCAACGCA ACAGGGGTCG TTGCCATGCT AATCGAGGGT

Acc108S1 GAAGGGGCAT ACAGTACGGC AGTTTCGTAT GGTACCCTCT TCCAGGTGCT TAAGGCATAT CTCAGATCC GTCTTCGCCA CCACATGGAT GCTCCGGTAG
Acc108SM GAAGGGGCAT ACAGTACGGC AGTTTCGTAT GGTACCCTCT TCCAGGTGCT TAAGGCATAT CTCAGATCC GTCTTCGCCA CCACATGGAT GCTCCGGTAG

Acc108S1 ACACAATGAC CATGGAGCAG GTCCTTGATG CGCTGCATTG A
Acc108SM ACACAATGAC CATGGAGCAG GTCCTTGATG CGCTGCATTG A

```

B

```

Acc108S1 MSKTNPNQLV ETSAELARDA VLTIVIVIKNR IVDSVRFGRV VFGTAFVIRS DDWGSIVLTA CHVVEDLQPG DTLRVRTMLQ SGVRELSASV LYEHRWSDVA
Acc108SM MSKTNPNQLV ETSAELARDA VLTIVIVIKNR IVDSVRFGRV VFGTAFVIRS DDWGSIVLTA CHVVEDLQPG DTLRVRTMLQ SGVRELSASV LYEHRWSDVA

Acc108S1 ILSVPGLRNV RQLTFEPSVL NSDSVIGVGY ANPIDLFPFI EMACKRIPDL SPGSVVSTEY VGLYQGVVDV FVHLEMVGM R GMSMPILNA TGVVAMLIEG
Acc108SM ILSVPGLRNV RQLTFEPSVL NSDSVIGVGY ANPIDLFPFI EMACKRIPDL SPGSVVSTEY VGLYR*----- GMSMPILNA TGVVAMLIEG

Acc108S1 EGAYSTAVSY GTVFQVLKAY LKIRLRHHMD APVDTMTMEQ VLDALH*
Acc108SM -----

```

C

| Accession | Species | Description | Identities ¹ | Positives ² | Gaps |
|------------|--|--|-------------------------|------------------------|--------------|
| A0A0D3GCN2 | <i>Oryza barthii</i> ³ | Uncharacterized protein | 246/246 (100%) | 246/246 (100%) | |
| K3YMT8 | <i>Setaria italica</i> | Uncharacterized protein | 91/249 (36%) | 132/249 (53%) | 31/249 (12%) |
| K3YJK3 | <i>Setaria italica</i> | Uncharacterized protein | 81/211 (38%) | 114/211 (54%) | 20/211 (9%) |
| K3YLY4 | <i>Setaria italica</i> | Uncharacterized protein | 89/244 (36%) | 128/244 (52%) | 31/244 (12%) |
| B6SLW3 | <i>Zea mays</i> | Putative uncharacterized protein | 73/239 (30%) | 113/239 (47%) | 29/239 (12%) |
| M8C5B7 | <i>Aegilops tauschii</i> | Putative serine protease do-like protein | 62/253 (24%) | 117/253 (46%) | 22/253 (8%) |
| A0A0Q3HHU2 | <i>Brachypodium distachyon</i> | Uncharacterized protein | 73/245 (29%) | 113/245 (46%) | 41/245 (16%) |
| A0A0D3FCT7 | <i>Oryza barthii</i> | Uncharacterized protein | 45/134 (33%) | 66/134 (49%) | 11/134 (8%) |
| Q2QXV9 | <i>Oryza sativa</i> | Expressed protein LOC_Os12g04740 | 60/200 (30%) | 88/200 (44%) | 20/200 (10%) |
| A0A117UT93 | <i>Novosphingobium fuchskuhlense</i> | Peptidase S1 | 57/183 (31%) | 86/183 (46%) | 18/183 (9%) |
| A0A0D9V9R3 | <i>Leersia perrieri</i> | Uncharacterized protein | 29/74 (39%) | 39/74 (52%) | 4/74 (5%) |
| Q2G5N9 | <i>Novosphingobium aromaticivorans</i> | Peptidase S1 and S6, chymotrypsin/Hap | 61/236 (25%) | 107/236 (45%) | 25/236 (10%) |
| B8BM04 | <i>Oryza sativa</i> | Putative uncharacterized protein | 60/201 (29%) | 88/201 (43%) | 21/201 (10%) |
| Q2QXV8 | <i>Oryza sativa</i> | Expressed protein LOC_Os12g04740 | 60/201 (29%) | 88/201 (43%) | 21/201 (10%) |
| A0A0D9V9R4 | <i>Leersia perrieri</i> | Uncharacterized protein | 28/76 (36%) | 40/76 (52%) | 4/76 (5%) |

Result was obtained by using BLASTP 2. 26 by DDBJ. The accessions with E-value less than 1×10^{-3} were listed.

¹ No. of amino acids identical to amino acids of SSP / no. of total amino acids. Parenthesis shows percentage.

² No. of amino acids showing positive score of BLOSUM 62 Matrix against amino acids of SSP / no. of total amino acids. Parenthesis shows percentage.

³ *Oryza barthii* is wild ancestral species of *O. glaberrima*

Fig. S3. (Continued)

D

```

SSP -----M SKTNPQLVLE TSAELARDAV LTIVIVIKNRI VDS-----
OgTPR1 MRKAKRGGGD LPRSFHKNSR AFKNEWISGD LLRNSKQDVW TGLSDGLKSY LSKSVASITL FNGDEILFSC SGIAMEHQFF TKFLTATTLV RALNATTKHH
          * * * * *

SSP -----VRFR GVVFGTAFVI RSDDWGSIVL TAQHVVEDLQ PG----- DTLRVR TMLQSGVREL SASVLYEHR- ----WSDVAI
OgTPR1 DDLKIQRVRLD GTRLYDGYMA EYDLNDNDFSV VEVYSVRDVQ VGFQSALES LPHGEVLAVG RDTSGEIMVK TVELNGDSRV SEDDRDLHCK ISKPWEGGPL
          ** * * * * *

SSP LSVPLGRN-- ----VRQLT FEPS-----
OgTPR1 LSVGGDMVGM NLFFTNRRAI FLPWGTTLNH YLTFVQKKTG LVQSKMKKVH RPEASIGEKS NSHPEVHGDF LNQEQLDLDS MGYPMLPSSM LGAGMILVNS
          *** * * * *

SSP -----VLNSD SVIG-VGYAN PIDLFP---- ----GI EMACKRIPDL
OgTPR1 FEDPFGDIYG EGVWRKFSRR ASILNRNVVA LASFNGEKRF FACTGFFIEW SGSKMILTSA SLVRDSGDEN KIDENLRIKV FLNNQCKEGK LEHCNLHYNI
          * * * * *

SSP SPGSVVSTEY VGLYQGVDTV FVHLEMVGM R GMS----- ----GMP ILNATG----
OgTPR1 ALVSVKYRAL RPLNTSFDCK SSRVAVGRC FNSGTLMAT S GRLVPWTGTL DCQFLARSTC KITKAGIGGP LVNLDGNVIG MNFYDTRIGT PFLWEEICK
          ** * * * *

SSP -----VV AMLIEGEGAY STAVSYGTVF QVLKAYLKIR LRHHMDAP--
OgTPR1 ILASFETKSE SGGDIGNASR ACFWKMPRDV KNKVNRPVVP KPRWCRPEDA ESDDDDKLAF DDIGQLQYSY ILGRKVKLLR LTIPISVPIV EAKSTDEPGV
          * * * * *

SSP -----VDTMTMEQVL DALH-----
OgTPR1 DPFAQRKQKK KRVEKQGNR LENLKAARKV GALPSHIQLA ATSLPITGTK ADLPKKS RKE DLENVAGMGS ATASGGKFDE KLPGEKPPKH PGKHKKFIPIV
          * * * * *

SSP -----
OgTPR1 AEGEGMGNLG KQONNKILMS LLARNSEQLD VCKANTMYKV KKEKRRRDKR EMASRSDK LK PQKKPFKKSS KKKA
  
```

E

```

SSP trypsin-like domain ---GTAFVIR SDDWGSIVLT AQHVVEDLQ GDTLRVRTML QS---GVREL SASVLYEHRW SDVAILSVPG LRNVQRQLTFE
OgTPR1 trypsin-like domain a ILFSCSGIAM EQFFTKFLT TATLVRALNA TTKHHD LKI QVRLDGTRLY DGYMAEYDLD NDFS VVEVYS VRDVQVGFQ
          * * * * *

SSP trypsin-like domain PSVLNSDS-- VIGVGYANPI DLFFGIEMAC KRIPDLSPGS VVSTEYVGLY QGVDTVFVHL EMVGMGRMSG MPILNATGVV
OgTPR1 trypsin-like domain a SALES LPHGE VLA VGRDTS G EIMVKTELVN GDSRVSEDDR DLHCKISKPW EGGPLLSDVG DMVGMN---- LFFTNRRAI-
          * * * * *

SSP trypsin-like domain -GTAFVIRSD DWGSIVLTAQ HVVEDLQPGD TLRVRTMLQS GVRELSASVL YEHRWSDVAI LSVPLGRNVR QLTFFEPSVLN
OgTPR1 trypsin-like domain b NRRNVALASF NGEKRRFFACT GFPIEWSGSK MILTSASLVR DSGDENKIDE NLRIKVF LNN QCKEGLKHC NLHYNIALVS
          * * * * *

SSP trypsin-like domain SDSVIGVGYA NPIDLFPGIE MACKRIPDLS PGSVVSTEYV GLYQGVDTVF VHLEMVGMRG MSG-MPILNA TGVV-----
OgTPR1 trypsin-like domain b VKYRALRPLN TSFDCKSSRV VAVGRCFNSG TLMATSGRLV PWTGTLDCQF LARSTCKITK AGIGGPLVNL DGNVIGMNFY
          * * * * *
  
```

F

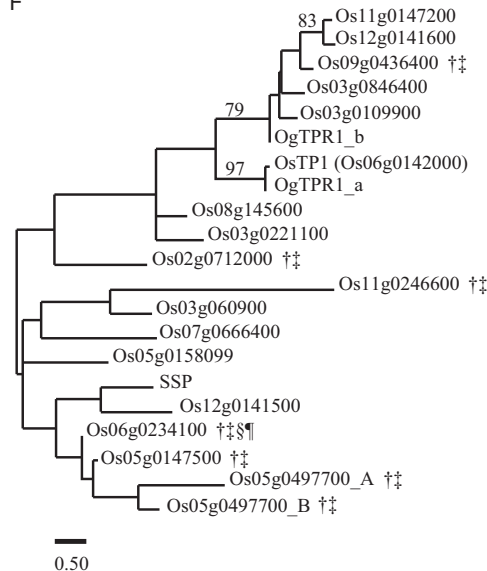


Fig. S3. (Continued)

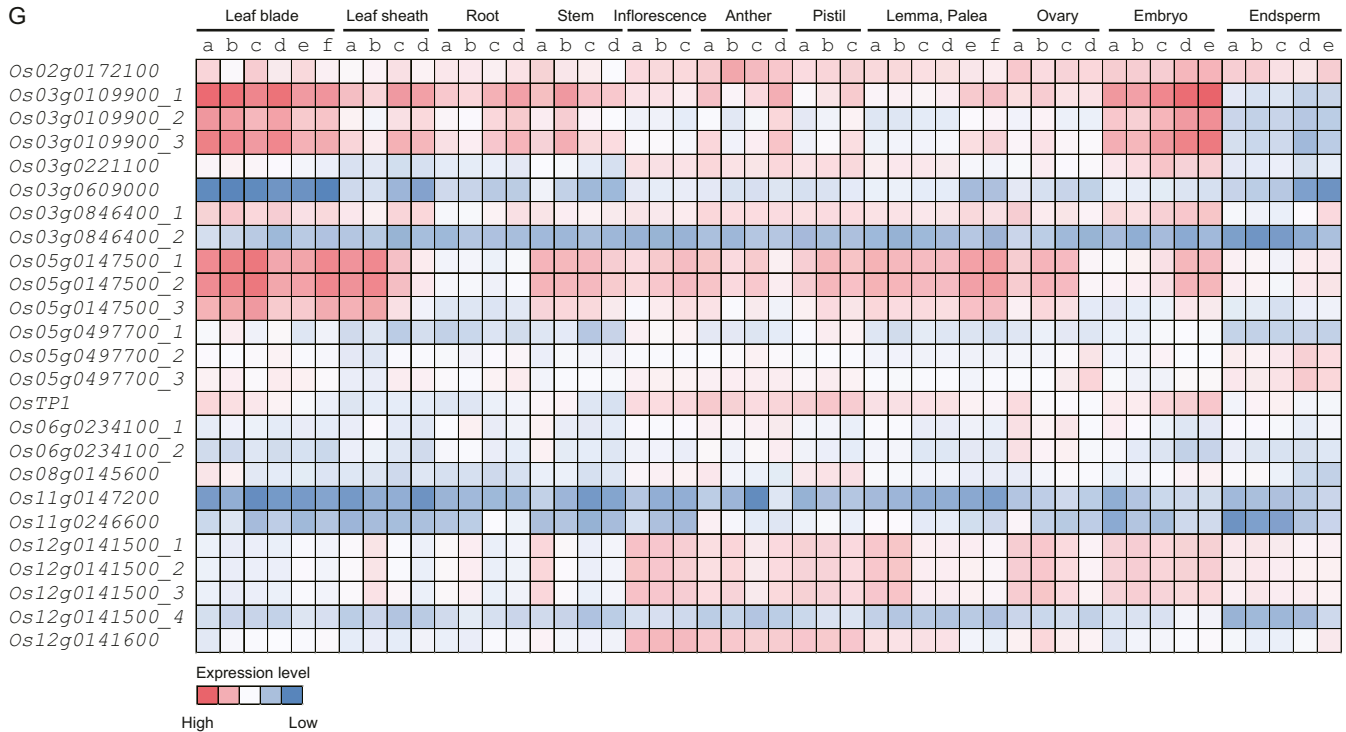


Fig. S3. Nucleotide and amino acid sequences analysis of *SSP* and *OgTPR1* genes. (A) Comparison of nucleotide sequences of *SSP* between *Acc108S*₁ and *Acc108S*_M. The position of the 5-bp deletion in *SSP* is shown by red letters. (B) Protein sequences encoded by *SSP*. Differences between *Acc108S*₁ and *Acc108S*_M are shown in red letters. Asterisks mark the stop codons. The position of the trypsin-like peptidase domain is indicated by bars below sequences. A catalytic triad conserved in serine proteases is shown by boxes. (C) Proteins homologous to *SSP*. The homology was analyzed by BLASTP search. (D and E) Alignment of *SSP* and *OgTPR1* proteins. (D) Comparison of the full-length amino acid sequences of the *SSP* and *OgTPR1* proteins. (E) Comparison of the amino acid sequences of trypsin-like peptidase domains in the *SSP* and *OgTPR1* proteins. The *OgTPR1* protein has two trypsin-like peptidase domains, a and b. Asterisks indicate the positions of conserved residues. (F) Phylogenetic relationship among trypsin-like peptidase domains encoded by *SSP*, *OgTPR1*/*OsTP1*, and the other 16 genes from *O. sativa*. *OsTP1* (Os06g0142000) is an allelic counterpart of *OgTPR1*. The trypsin-like peptidase domain-containing genes from *O. sativa* were identified in a domain-name search of the public database Pfam 31.0 (pfam.xfam.org). Among these genes, 17 genes (including *OsTP1*) annotated in the *O. sativa* reference genome (cv. Nipponbare) were used for the phylogenetic analysis. *OgTPR1* and *Os05g0497700* contained two trypsin-like peptidase domains. Protein sequences of the trypsin-like peptidase domain in each gene were aligned by the software MEGA7. If a gene encodes more than two trypsin-like peptidase domains, the amino acid sequence of each domain shown by the gene name with an alphabet was used for the alignment. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Symbols for gene ontology terms: †, protein folding; ‡, proteolysis; §, intracellular signal transduction; ¶, apoptotic process. (G) Expression profile of 14 genes coding the trypsin-like peptidase domain-containing protein in *O. sativa* (cv. Nipponbare). The trypsin-like peptidase domain-containing protein from *O. sativa* (cv. Nipponbare) was identified in a domain-name search in the public database Pfam 31.0 (pfam.xfam.org). Among these genes, 14 genes that have global expression profile data in the public database RiceXpro (ricexpro.dna.affrc.go.jp) were used. *OsTP1* is an allelic counterpart of the *OgTPR1* gene. Expression profiles for *SSP* and *OgTPR1* are not available because they are not annotated in the *O. sativa* genome. Transcript variants of each gene are shown by the gene name followed by an underscore and a number. The expression levels are shown in the color chart. Average values for the normalized (\log_2) Cy3 signal intensity in three replications were used. These values ranged from -9.306 (low expression) to 2.755 (high expression). Leaf blade a, vegetative stage at 12:00; b, vegetative stage at 0:00; c, reproductive stage at 12:00; d, reproductive stage at 0:00; e, ripening stage at 12:00; f, ripening stage at 0:00; Leaf sheath a, vegetative stage at 12:00; b, vegetative stage at 0:00; c, reproductive stage at 12:00; d, reproductive stage at 0:00; Root a, vegetative stage at 12:00; b, vegetative stage at 0:00; c, reproductive stage at 12:00; d, reproductive stage at 0:00; Stem a, reproductive stage at 12:00; b, reproductive stage at 0:00; c, ripening stage at 12:00; d, ripening stage at 0:00; Inflorescence a, 0.6–1.0 mm; b, 3.0–4.0 mm; c, 5.0–10.0 mm; Anther a, 0.3–0.6 mm; b, 0.7–1.0 mm; c, 1.2–1.5 mm; d, 1.6–2.0 mm; Pistil a, pistil from 5–10 cm inflorescence; b, pistil from 10–14 cm inflorescence; c, pistil from 14–18 cm inflorescence; Lemma, Palea a, lemma from a 1.5–2.0 mm floret; b, palea from a 1.5–2.0 mm floret; c, lemma from 4.0–5.0 mm floret; d, palea from a 4.0–5.0 mm floret; e, lemma from a >7.0 mm floret; f, palea from a >7.0 mm floret; Ovary a, ovary at 1 d after flowering; b, ovary at 3 d after flowering; c, ovary at 5 d after flowering; d, ovary at 7 d after flowering; Embryo a, embryo at 7 d after flowering; b, embryo at 10 d after flowering; c, embryo at 14 d after flowering; d, embryo at 28 d after flowering; e, embryo at 42 d after flowering; Endosperm a, endosperm at 7 d after flowering; b, endosperm at 10 d after flowering; c, endosperm at 14 d after flowering; d, endosperm at 28 d after flowering; e, endosperm at 42 d after flowering. 0:00 and 12:00 indicate midnight and noon, respectively.

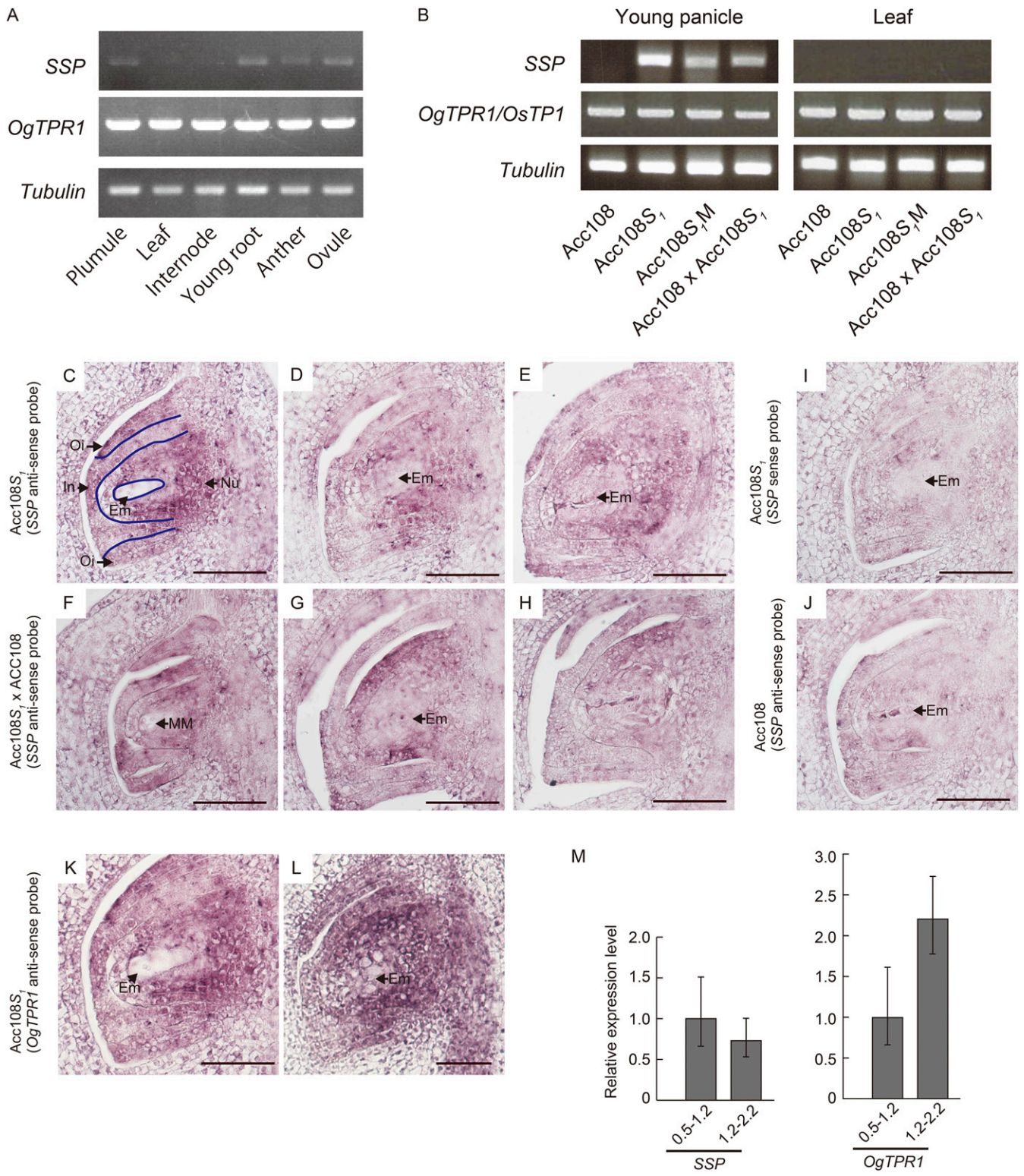


Fig. S4. Expression of *SSP* and *OgTPR1* genes. (A) Expression of *SSP* and *OgTPR1* in plumules, leaves, internodes, young roots, anthers, and ovules in strain IRGC103777 of *O. glaberrima*, which has the S_7^9 allele at the S_7 locus. (B) Expression of *SSP* and *OgTPR1/OsTP1* in young panicles and leaves. Acc108 has the *OsTP1* gene, and Acc108 S_7 and Acc108 S_7 M have the *OgTPR1* gene. The tubulin alpha-1 chain gene (Os07g0574800) was used as an internal control. (C–L) In situ hybridization analysis of *SSP* and *OgTPR1* transcripts during ovule development. (C–E) Longitudinal sections of developing ovule of Acc108 S_7 hybridized with the *SSP* antisense probe. (C) Functional sporophyte stage. (D) Early-stage embryo sac. (E) Early-middle-stage embryo sac. (F–H) Longitudinal sections of developing ovule in F₁ hybrid (Acc108 S_7 × Acc108) hybridized with the *SSP* antisense probe. (F) Megaspore mother cell stage. (G) Normally developing embryo sac in the early-middle stage. (H) Abnormal embryo sac in early-middle stage. (I and J) Negative controls. (I) Longitudinal section of a functional sporophyte-stage ovule in Acc108 S_7 hybridized with the *SSP* sense probe. (J) Longitudinal section of a functional sporophyte-stage ovule in Acc108 hybridized with the *SSP* antisense probe. (K and L) Longitudinal sections of a developing ovule of Acc108 S_7 hybridized with the *OgTPR1* antisense probe. (K) Functional sporophyte stage. (L) Early-middle-stage embryo sac. Em, embryo sac; In, inner integument; MM, megaspore mother cell; Nu, nucellus; Oi, outer integument. (Scale bars: 50 μ m.) (M) Expression levels of *SSP* and *OgTPR1* in two different stages of developing anthers. Sampled anthers were 0.5–1.2 or 1.2–2.2 mm long. The ubiquitin gene (Os01g0328400) was used as an internal control.

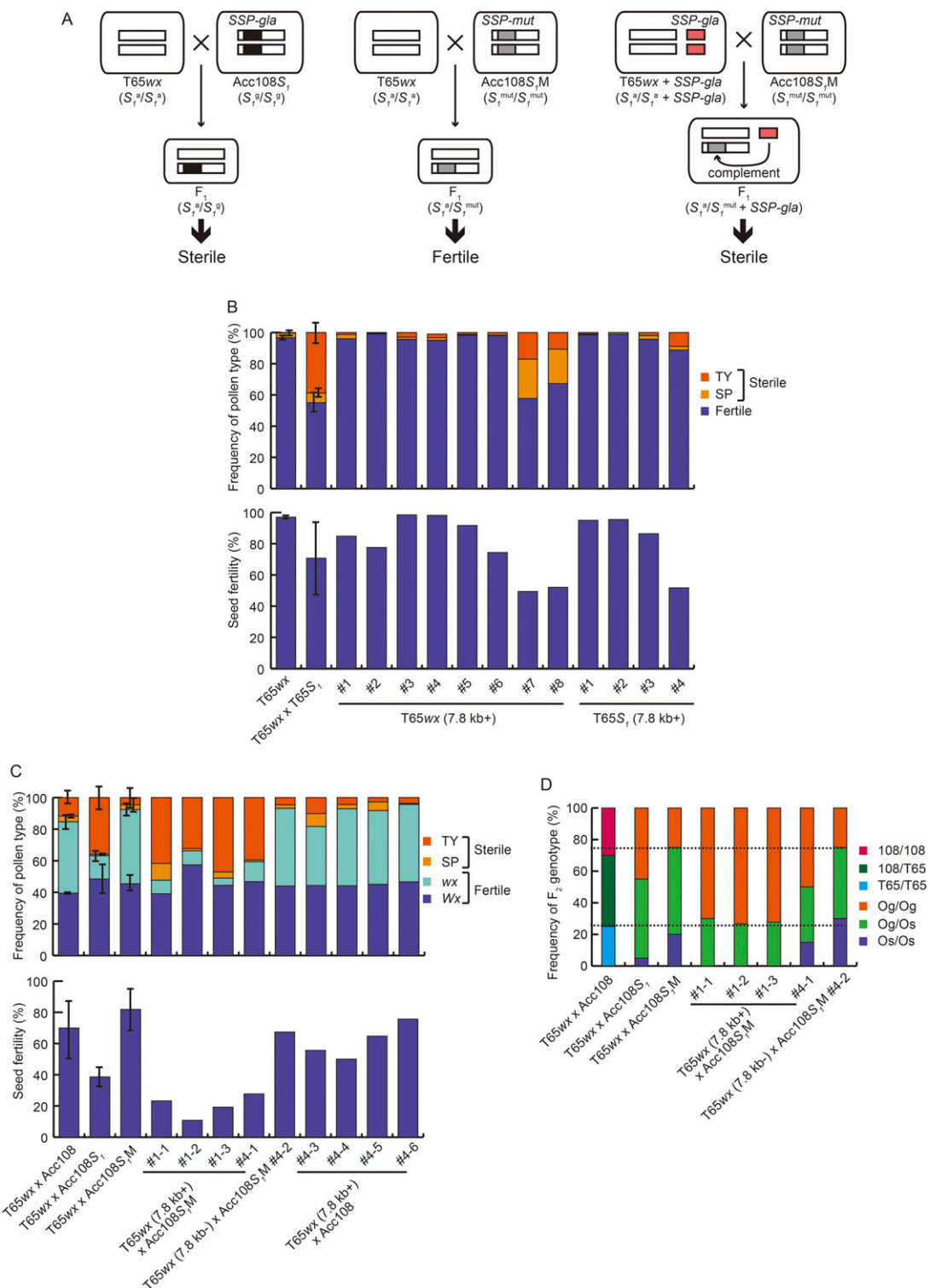


Fig. 55. Transgenic experiments. (A) The scheme of the complementation test. (Left) A cross between T65wx and Acc108S₁ produces sterile F₁ hybrids. (Middle) A cross between T65wx and the mutant Acc108S₁M produces fertile F₁ hybrids. (Right) When the intact *SSP-gla* gene (*SSP-gla*) complements the phenotype of Acc108S₁M, a cross between T65wx carrying the *SSP-gla* transgene and Acc108S₁M produces sterile F₁ hybrids. (B–D) Transgenic experiments with the 7.8-kb fragment. (B) Mean ± SD frequency of pollen types and seed fertility of T65wx, T65S₁, and 12 independent transgenic plants (T65wx, 7.8 kb+; T65S₁, 7.8 kb+). TY and SP indicate typical- and spherical-type aborted pollen grains, respectively. *n* = 3 plants for T65wx and T65wx × T65S₁; *n* = 1 plant for transgenic plants. (C) Mean ± SD frequency of pollen types and seed fertility of T65wx × Acc108, T65wx × Acc108S₁, T65wx × Acc108S₁M, and transgenic plants. For transgenic plants, three genotypes, T65wx (7.8 kb+) × Acc108S₁M, T65wx (7.8 kb-) × Acc108S₁M, and T65wx (7.8 kb+) × Acc108, were used to assess pollen and seed fertility. A plus sign (+) indicates a transgenic positive; a minus sign (-) indicates a null segregant. The frequency of wx-type fertile pollen grains decreased, whereas the frequency of TY-type aborted pollen grains increased in transgenic positive segregants, indicating the presence of the S₁ locus-mediated mTRD. *n* = 3 plants for T65wx × Acc108, T65wx × Acc108S₁, and T65wx × Acc108S₁M; *n* = 1 plant for transgenic plants. (D) Mean ± SD frequency of F₂ genotypes obtained from the self-pollination of plants of each genotype. T65wx × Acc108, T65wx × Acc108S₁, T65wx × Acc108S₁M, T65wx (7.8 kb+) × Acc108S₁M, and T65wx (7.8 kb-) × Acc108S₁M were used. Genotypes were determined using markers in the S₁ locus. Og/Og, *Oryza glaberrima*-type homozygotes; Og/Os, heterozygotes for *O. glaberrima* and *O. sativa* types; Os/Os, *O. sativa*-type homozygotes; T65/T65, T65wx-type homozygotes; 108/T65, heterozygotes for Acc108 and T65wx types; 108/108, Acc108-type homozygotes. Dotted lines mark 25% and 75%.

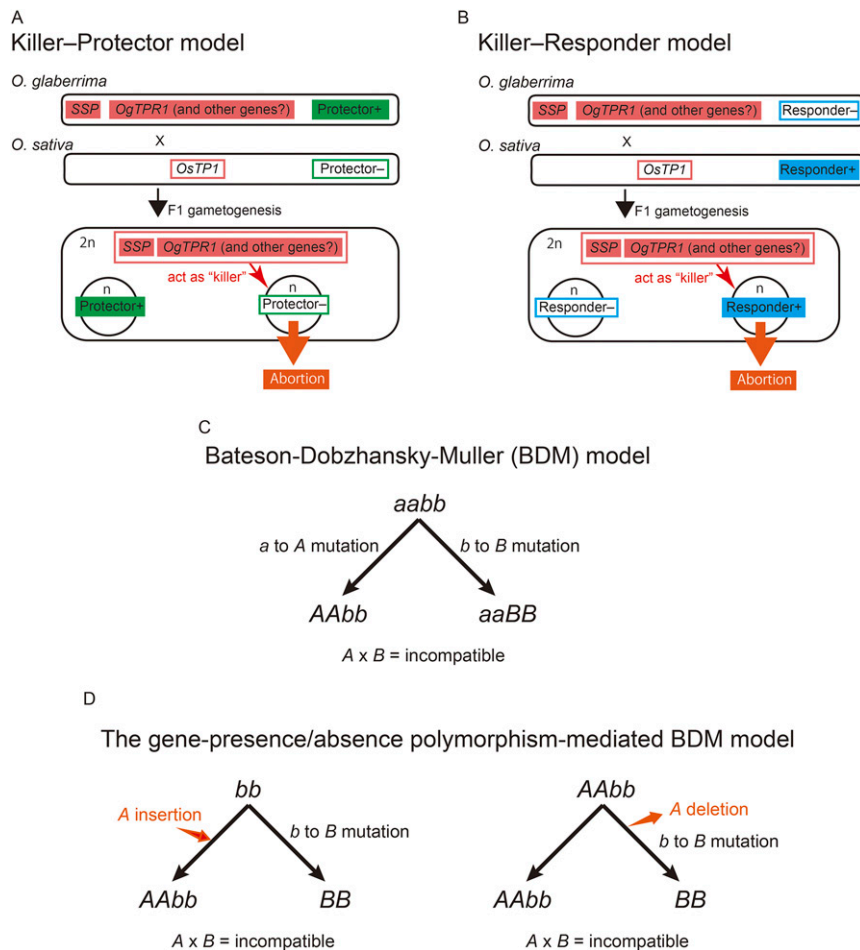


Fig. 56. Hypothetical genetic and evolutionary models for the S_1 locus-mediated $sTRD$ based on the killer–protector and the killer–responder models proposed in other hybrid sterility systems. (A) Killer–protector model. *O. glaberrima* (S_1^9 allele carrier) has *SSP*, *OgTPR1*, and other genes for a functional killer and protector (Protector+) genes, whereas *O. sativa* (S_1^3 allele carrier) has nonfunctional killer, *OsTP1*, and protector (Protector–) genes. In F_1 gametogenesis, *SSP* and *OgTPR1* (and other genes for a functional killer) act together, and gametes without Protector+ preferentially abort. The notations $2n$ and n indicate that *SSP* and *OgTPR1* (and other genes for a functional killer) act in a sporophytic manner ($2n$) and that Protector+ acts in a gametophytic (n) manner. (B) Killer–responder model. *O. glaberrima* (S_1^9 allele carrier) has *SSP*, *OgTPR1*, and other genes for a functional killer and nonfunctional responder (Responder–) genes, whereas *O. sativa* (S_1^3 allele carrier) has a nonfunctional killer and *OsTP1* and functional responder (Responder+) genes. In F_1 gametogenesis, *SSP* and *OgTPR1* (and other genes for a functional killer) act together, and gametes with Responder+ preferentially abort. (C and D) Models for the evolution of genic incompatibility. (C) The BDM model. The genotype of an ancestral population is *aabb*. This population splits into two independent lineages. The $a \rightarrow A$ mutation arises in one lineage, while the $b \rightarrow B$ mutation arises in the other. The coexistence of *A* and *B* causes incompatibility in hybrids. (D) The gene-presence/absence polymorphism-mediated BDM model. (Left) The insertion of the *A* gene in one lineage causes hybrid incompatibility with the *B* allele that arose from the *b* allele and is fixed in the other lineage. (Right) The *B* allele that arose from the *b* allele after the deletion of the *A* gene in one lineage causes incompatibility with *A* gene in the other lineage.

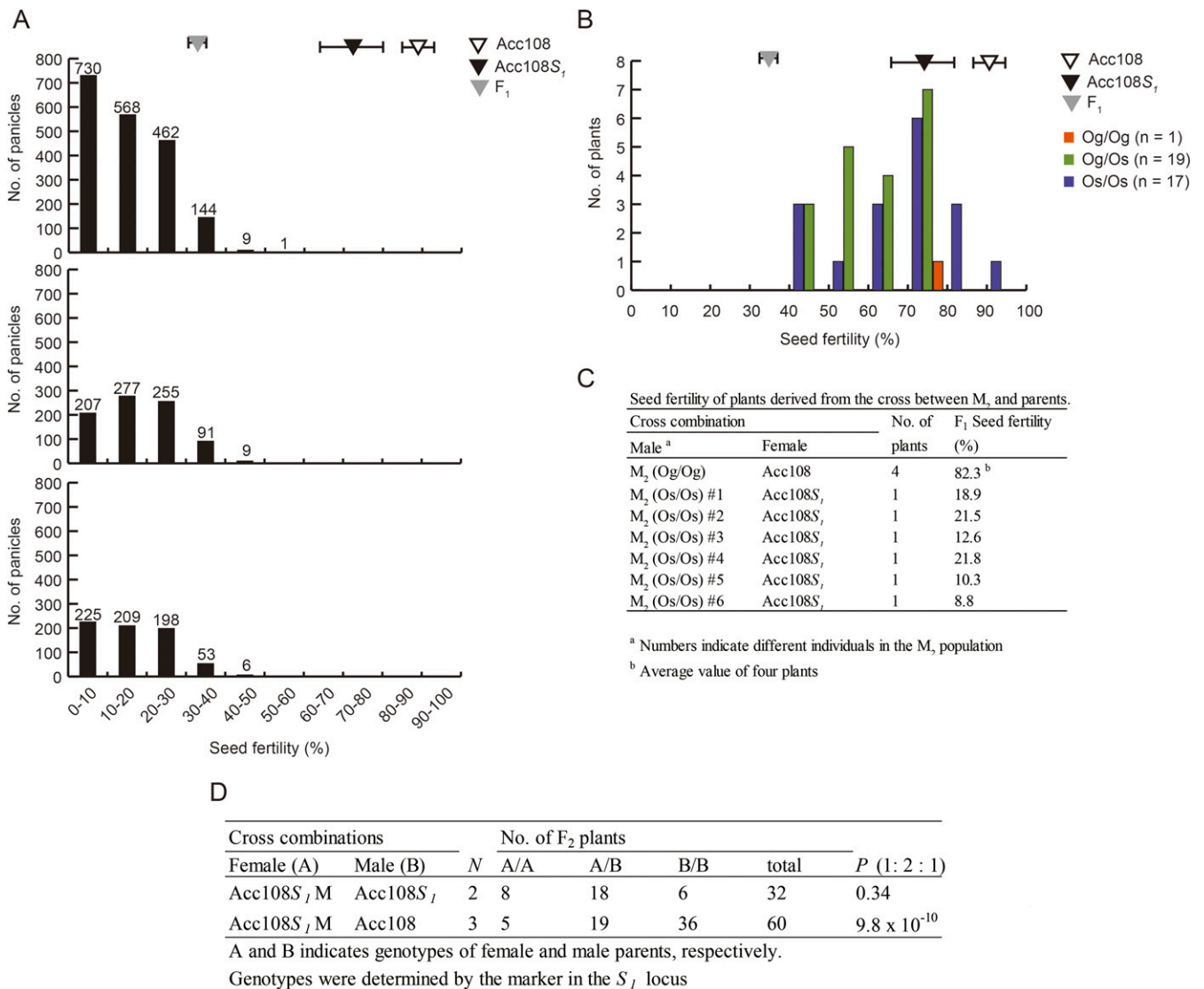


Fig. S7. Screening and characterization of the mutant. (A–C) Seed fertilities of M₁, M₂, and plants derived from the cross between M₂ and their parental rice varieties. (A) Frequency distribution of seed fertility of panicles in M₁ plants. M₁ plants were obtained by irradiation of carbon ions (LET 30 KeV/μm) at a dose of 150 Gy (Top) and argon ions at doses of 7.5 Gy (Middle) and 10 Gy (Bottom). The number of panicles is shown above the bars. Average levels of seed fertility of Acc108, Acc108S₁, and their F₁ progeny are indicated by arrowheads, with whiskers showing SDs. (B) Frequency distribution of seed fertility of plants in the M₂ population. The M₂ population was obtained from the M₁ panicle with high seed fertility level (>50%). Plants were classified based on the genotype of the marker in the S₁ locus. Og/Og, Og/Os, and Os/Os indicate *O. glaberrima*-type homozygotes and heterozygotes and *O. sativa*-type homozygotes, respectively. One *O. glaberrima* homozygote, 19 *O. glaberrima* heterozygotes, and 17 *O. sativa* homozygotes were obtained. Mean seed fertilities of Acc108, Acc108S₁, and their hybrids are indicated by arrowheads, with whiskers showing SDs. (C) Seed fertility of plants derived from the cross between M₂ and parents. (D) Segregation of F₂ plants derived from the cross between the mutant Acc108S₁M and two lines, Acc108 and Acc108S₁.

Table S1. Hybrid sterility loci on which the neutral allele were reported

| Locus | Species | References |
|--------------------|---|--|
| <i>Ge</i> | <i>Lycopersicon esculentum</i> and <i>Lycopersicon pimpinellifolium</i> * | Rick (1, 2) |
| <i>S5, f6</i> | <i>O. sativa</i> | Ikehashi and Araki (3); Wang et al. (4) |
| <i>S6</i> | <i>O. sativa</i> and <i>O. rufipogon</i> † | Koide et al. (5) |
| <i>S9</i> | <i>O. sativa</i> | Wan et al. (6) |
| <i>S15</i> | <i>O. sativa</i> | Wan et al. (6) |
| <i>S24, Sb, f5</i> | <i>O. sativa</i> | Wang et al. (7); Shi et al. (8); Zhao et al. (9) |
| <i>S31</i> | <i>O. sativa</i> | Li et al. (10) |
| <i>S32</i> | <i>O. sativa</i> | Li et al. (10) |
| <i>Sa</i> | <i>O. sativa</i> | Wang et al. (11) |
| <i>Sd</i> | <i>O. sativa</i> | Liu et al. (12) |
| <i>Se</i> | <i>O. sativa</i> | Liu et al. (12) |

**L. esculentum* and *L. pimpinellifolium* are the cultivated tomato and the closely related wild tomato, respectively. Reproductive isolation was not found between these species (13), indicating that they belong to the same gene pool.

†These two species belong to the same gene pool.

- Rick CM (1966) Abortion of male and female gametes in the tomato determined by allelic interaction. *Genetics* 53:85–96.
- Rick CM (1971) The tomato *ge* locus: Linkage relations and geographic distribution of alleles. *Genetics* 67:75–85.
- Ikehashi H, Araki H (1988) Multiple alleles controlling F1 sterility in remote crosses of rice (*Oryza sativa* L.). *Jpn J Breed* 38:283–291.
- Wang GW, He YQ, Xu CG, Zhang Q (2005) Identification and confirmation of three neutral alleles conferring wide compatibility in inter-subspecific hybrids of rice (*Oryza sativa* L.) using near-isogenic lines. *Theor Appl Genet* 111:702–710.
- Koide Y, et al. (2008) The evolution of sex-independent transmission ratio distortion involving multiple allelic interactions at a single locus in rice. *Genetics* 180:409–420.
- Wan J, Yamaguchi Y, Kato H, Ikehashi H (1996) Two new loci for hybrid sterility in cultivated rice (*Oryza sativa* L.). *Theor Appl Genet* 92:183–190.
- Wang GW, He YQ, Xu CG, Zhang Q (2006) Fine mapping of *f5-Du*, a gene conferring wide-compatibility for pollen fertility in inter-subspecific hybrids of rice (*Oryza sativa* L.). *Theor Appl Genet* 112:382–387.
- Shi LG, et al. (2009) Identifying neutral allele *Sb* at pollen-sterility loci in cultivated rice with *Oryza rufipogon* origin. *Chin Sci Bull* 54:3813.
- Zhao ZG, et al. (2011) Molecular analysis of an additional case of hybrid sterility in rice (*Oryza sativa* L.). *Planta* 233:485–494.
- Li D, et al. (2007) Fine mapping of *S32(t)*, a new gene causing hybrid embryo sac sterility in a Chinese landrace rice (*Oryza sativa* L.). *Theor Appl Genet* 114:515–524.
- Wang Y, et al. (2010) Fine mapping of a gene causing hybrid pollen sterility between Yunnan weedy rice and cultivated rice (*Oryza sativa* L.) and phylogenetic analysis of Yunnan weedy rice. *Planta* 231:559–570.
- Liu B, et al. (2011) Identification of neutral genes at pollen sterility loci *Sd* and *Se* of cultivated rice (*Oryza sativa*) with wild rice (*O. rufipogon*) origin. *Genet Mol Res* 10:3435–3445.
- Miller JC, Tanksley SD (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *Theor Appl Genet* 80:437–448.

Table S2. Distribution of the SSP gene among *Oryza* species

| Species | Accession name | Subspecies/type | Allele* | PCR [†] | | Coverage [‡] | Presence of SSP |
|---------------------|----------------|------------------------|----------------------------------|------------------|-------|-----------------------|-----------------|
| | | | | SSP | 47-55 | | |
| <i>O. sativa</i> | Acc27590 | IND | <i>S₇^a</i> | – | + | Nd | – |
| | Acc27593 | IND | <i>S₇^a</i> | – | + | Nd | – |
| | Acc414 | IND | <i>S₇^a</i> | – | + | Nd | – |
| | Kasalath | AUS | <i>S₇^a</i> | – | + | Nd | – |
| | A5 | TEJ | <i>S₇^a</i> | – | + | Nd | – |
| | IRGC12793 | ARO [§] | — | Nd | Nd | 16.1 | – |
| | IRGC31856 | ARO [§] | — | Nd | Nd | 16.9 | – |
| | IRGC38994 | ARO [§] | — | Nd | Nd | 16.4 | – |
| | IRGC9060 | ARO [§] | — | Nd | Nd | 14.5 | – |
| | IRGC9062 | ARO [§] | — | Nd | Nd | 15.0 | – |
| | RA4952 | ARO [§] | — | Nd | Nd | 16.6 | – |
| | IRGC12883 | AUS [§] | — | Nd | Nd | 16.8 | – |
| | IRGC45975 | AUS [§] | — | Nd | Nd | 16.1 | – |
| | IRGC6307 | AUS [§] | — | Nd | Nd | 17.2 | – |
| | IRGC8555 | AUS [§] | — | Nd | Nd | 17.2 | – |
| | IRGC6513 | III [§] | — | Nd | Nd | 17.0 | – |
| | IRGC25901 | IND [§] | — | Nd | Nd | 7.8 | – |
| | IRGC27762 | IND [§] | — | Nd | Nd | 17.0 | – |
| | IRGC30416 | IND [§] | — | Nd | Nd | 17.1 | – |
| | IRGC8231 | IND [§] | — | Nd | Nd | 17.2 | – |
| | IRGC9148 | IND [§] | — | Nd | Nd | 16.3 | – |
| | IRGC9177 | IND [§] | — | Nd | Nd | 15.6 | – |
| | IRGC60542 | IV [§] | — | Nd | Nd | 16.7 | – |
| | IRGC2540 | TEJ [§] | — | Nd | Nd | 17.0 | – |
| | IRGC27630 | TEJ [§] | — | Nd | Nd | 15.5 | – |
| | IRGC32399 | TEJ [§] | — | Nd | Nd | 17.4 | – |
| | IRGC55471 | TEJ [§] | — | Nd | Nd | 17.4 | – |
| | IRGC11010 | TEJ [§] | — | Nd | Nd | 16.9 | – |
| | IRGC38698 | TRJ [§] | — | Nd | Nd | 16.4 | – |
| | IRGC50448 | TRJ [§] | — | Nd | Nd | 17.1 | – |
| | IRGC66756 | TRJ [§] | — | Nd | Nd | 15.5 | – |
| <i>O. rufipogon</i> | W106 | Annual | <i>S₇^a</i> | – | + | Nd | – |
| | W107 | Annual | <i>S₇^a</i> | – | + | Nd | – |
| | W1551 | Annual | <i>S₇^a</i> | – | + | Nd | – |
| | W130 | Intermediate | <i>S₇^a</i> | – | + | Nd | – |
| | W149 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W1681 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W2005 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W2007 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W172 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W1943 | Perennial | — | – | + | Nd | – |
| | W1944 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W1945 | Perennial | — | – | + | Nd | – |
| | W1952 | Perennial | <i>S₇^a</i> | – | + | Nd | – |
| | W1718 | Weedy | — | – | + | Nd | – |
| | W1714 | Weedy | <i>S₇^a</i> | – | + | Nd | – |
| | W120 | Perennial | — | – | + | Nd | – |
| | W154 | Perennial | — | – | + | Nd | – |
| | W2099 | OrII [¶] | — | – | + | Nd | – |
| | W1970 | OrI [¶] | — | – | + | Nd | – |
| | W1236 | OrII [¶] | — | – | + | Nd | – |
| | IRGC105327 | Annual [§] | — | Nd | Nd | 16.6 | – |
| | IRGC106105 | Annual [§] | — | Nd | Nd | 17.0 | – |
| | IRGC106154 | Annual [§] | — | Nd | Nd | 16.4 | – |
| | IRGC80470 | Annual [§] | — | Nd | Nd | 16.7 | – |
| | IRGC89215 | Annual [§] | — | Nd | Nd | 17.2 | – |
| | IRGC105958 | Perennial [§] | — | Nd | Nd | 16.1 | – |
| | IRGC105960 | Perennial [§] | — | Nd | Nd | 17.1 | – |
| | P46 | Perennial [§] | — | Nd | Nd | 17.0 | – |
| | VOC4 | Perennial [§] | — | Nd | Nd | 17.2 | – |
| | Yuan3 | Perennial [§] | — | Nd | Nd | 15.5 | – |

Table S2. Cont.

| Species | Accession name | Subspecies/type | Allele* | PCR [†] | | Coverage [‡] | Presence of SSP |
|--------------------------|----------------|-----------------|---------|------------------|-------|-----------------------|-----------------|
| | | | | SSP | 47-55 | | |
| <i>O. glaberrima</i> | W025 | | S_1 | + | - | Nd | + |
| | CG14 | | S_1 | + | - | Nd | + |
| | IRGC103469 | | — | Nd | Nd | 97.8 | + |
| | TOG5457 | | — | Nd | Nd | 92.1 | + |
| | TOG5467 | | — | Nd | Nd | 95.6 | + |
| | TOG5923 | | — | Nd | Nd | 96.6 | + |
| | TOG5949 | | — | Nd | Nd | 96.9 | + |
| | TOG7025 | | — | Nd | Nd | 96.6 | + |
| | TOG7102 | | — | Nd | Nd | 97.7 | + |
| | IRGC101049 | | — | Nd | Nd | 97.9 | + |
| | IRGC103472 | | — | Nd | Nd | 96.7 | + |
| | IRGC103520 | | — | Nd | Nd | 97.9 | + |
| | IRGC103632 | | — | Nd | Nd | 97.3 | + |
| | IRGC103937 | | — | Nd | Nd | 95.4 | + |
| | IRGC104206 | | — | Nd | Nd | 89.9 | + |
| | IRGC104574 | | — | Nd | Nd | 95.3 | + |
| IRGC104955 | | — | Nd | Nd | 97.6 | + | |
| W3104 | | — | Nd | Nd | 99.0 | + | |
| <i>O. barthii</i> | W720 | | — | + | - | Nd | + |
| | W1588 | | — | + | - | Nd | + |
| | IRGC100122 | | — | Nd | Nd | 97.7 | + |
| | IRGC100931 | | — | Nd | Nd | 95.6 | + |
| | IRGC100934 | | — | Nd | Nd | 97.6 | + |
| | IRGC103895 | | — | Nd | Nd | 96.4 | + |
| | IRGC104084 | | — | Nd | Nd | 98.0 | + |
| | IRGC104119 | | — | Nd | Nd | 97.4 | + |
| | IRGC105608 | | — | Nd | Nd | 97.5 | + |
| | IRGC106234 | | — | Nd | Nd | 98.2 | + |
| | IRGC103912 | | — | Nd | Nd | 98.0 | + |
| | WAB0028952 | | — | Nd | Nd | 26.6 | - |
| | WAB0028903 | | — | Nd | Nd | 16.9 | - |
| | WAB0028980 | | — | Nd | Nd | 16.3 | - |
| | WAB0028987 | | — | Nd | Nd | 97.9 | + |
| | WAB0028958 | | — | Nd | Nd | 97.9 | + |
| | WAB0028976 | | — | Nd | Nd | 99.8 | + |
| | WAB0028979 | | — | Nd | Nd | 94.0 | + |
| | WAB0028992 | | — | Nd | Nd | 97.3 | + |
| WAB0028938 | | — | Nd | Nd | 97.4 | + | |
| WAB0030151 | | — | Nd | Nd | 97.4 | + | |
| W3106 | | — | Nd | Nd | 99.1 | + | |
| <i>O. longistaminata</i> | W1508 | | — | Nd | Nd | 17.9 | - |
| <i>O. meridionalis</i> | W1625 | | — | - | + | Nd | - |
| | W1635 | | — | - | + | Nd | - |
| | W2069 | | — | + | - | Nd | + |
| | W2079 | | — | - | + | Nd | - |
| | W2112 | | — | + | - | Nd | + |
| <i>O. officinalis</i> | W1131 | | — | - | + | Nd | - |
| | W1302 | | — | - | + | Nd | - |
| | W1814 | | — | - | + | Nd | - |
| <i>O. minuta</i> | W1319 | | — | - | + | Nd | - |
| <i>O. ridleyi</i> | W2033 | | — | - | + | Nd | - |

Nd, not determined.

*Alleles based on Koide et al. (1). The S_7^8 and S_7 alleles are equivalent to S_7^8 and S_7^9 alleles, respectively.

[†]Markers SSP and 47-55 were designed to detect presence (+) and absence (-) of the SSP gene, respectively.

[‡]The length of nucleotide sequence where short reads were mapped divided by the length of the total nucleotide sequence of the SSP gene.

[§]Classification of subspecies/types was based on Xu et al. (2). ARO, AUS, IND, TEJ, and TRJ indicate aromatic, aus, indica, temperate japonica, and tropical japonica, respectively. III and IV indicate group III and IV, respectively, based on Glaszmann's classification (2).

Annual and perennial types of *O. rufipogon* indicate *Oryza nivara* and *O. rufipogon*, respectively, in Xu et al. (3).

[¶]Classification of *O. rufipogon* was based on Huang et al. (4).

1. Koide Y, et al. (2008) Sex-independent transmission ratio distortion system responsible for reproductive barriers between Asian and African rice species. *New Phytol* 179:888–900.
2. Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. *Theor Appl Genet* 74:21–30.
3. Xu X, et al. (2011) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotechnol* 30:105–111.
4. Huang X, et al. (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490:497–501.

Table S3. DNA markers used in this study

| Primer application | Marker name | Forward primer sequence (5'–3') | Reverse primer sequence (5'–3') |
|--------------------------------|--------------------------|---|---|
| Genetic mapping and genotyping | RM19350 | CCAGCTGCTTGTGATCAAGTCG | TGCTAGTGCATGCCAATACTACTG |
| | RM19357 | TGTTTAAACGGTTGATCTCTGC | CCACCCCTGCCTTCTTCAGG |
| | RM19369 | GGTGAGACTACTAAATATGACG | GTAAACCGATGGCTTAGAGTCC |
| | RM19369* | CACCATCCTCTCCATGTTT | CCCTTTCATTTTGGCGTCTA |
| | C6_21788* | GTCGATGATAAAGCAATA | CCCTAGCTAGTTGGTGTGA |
| | RM19414 | GTCAGAACTTCAACCAAGG | GCTGTATAGCTTGTATCTAGGAGTAG |
| Expression analysis | RM8258 | ACCCCTATCTGTGCATCA | CTCGACTAAACTCGAAATCCT |
| | RM3805 | ACACCACCATCAACCTACCAAC | AACTCGAGAGGAAGCAAGG |
| | SSP | ACCTTTGAACCATCAGTGCTCA | TCTCCTTGAGCCCTCAAAACTG |
| | <i>Tubulin</i> | TACAACGGTTGGGTCGCAC | AACTTGGCCACACGGTCCAG |
| | <i>OgTPR1</i> | AGTAGAGCCCTCAAGATGAGTGG | AAACATGGAGGATGTAACCTTAGGA |
| RACE analysis | <i>SSP_qRT-PCR</i> | AGGAGTGTGTGTGACTG | GTGCTACCCGGAGCATCCAT |
| | <i>OgTPR1_qRT-PCR</i> | GCTCAGGCATCGCTATGGAA | TCCACATGGGGCAGACTTTC |
| | <i>Ubiquitin_qRT-PCR</i> | ACGCCTAAGCCTGCTGGTT | ACCACCTCGACCCGCCACTACT |
| | IGS2_6257L | GATCAGACGGCCCTTGACCTCGCCCTCACCCACCCACCC | |
| | IGS2_6305R | GTGGTGGTGGGGCGGAGTCAAGGCCGCTGTGAT | |
| | IGS2_7503L | GTACCAAGGTGTGACGCTCACCTTGTTC | |
| | IGS2_7507R | GGTACAACCAACATATTAGTCGACACAAC | |
| | IGS2_8190R | GAGAACCAAAAATCAGATTGTCACCTGGATAACC | |
| | P9R3347 | TATCCAGGGTTCACATCAGTC | |
| | Transgenic analysis | 13-084_GW | GGGGACAAGTTTGTCAAAAAGCAGGGCTTCTGCAAGACCCGACATACCTCT |
| IGS2_5303_GW | | GGGGACAAGTTTGTCAAAAAGCAGGGCTTCTGCAAGACCCGACATACCTCT | GGGACACATTTGTTTACAAGAAAGCTGGTCTCTGTGTGTCCTTTTGTGC |
| IGS2_8806_M13r | | AGTCCAAGTTCCTCCATGGTA | CAGGAAAACAGCTATGAC |
| IGS2_7260_9576 | | ACCTTTGAACCATCAGTGCTCA | CGTGTAACAAGAAATGGATGCAC |
| <i>Hygromycin_1</i> | | TTTCTGATCGAAAAGTTCGACAGCTCT | GGCAGTTCGGTTTCAGGCAGGCTTGGCAA |
| 13-084seq-primerFw1 | | AATCCGTTTATTTTAACC | |
| Sequence analysis | 13-084seq-primerFw2 | GGCAAATCGAGGGTTAGTTC | |
| | 13-084seq-primerFw3 | ACAAAATCCTTCAATAAAC | |
| | 13-084seq-primerFw4 | TGTGTAGAAAAGTTTGTATG | |
| | 13-84seq-primerFw5 | ACTGCGCAGCGCTCGCCAC | |
| | 13-084seq-primerFw6 | ACGATCTGGCGAGCATCAAG | |
| | 13-084seq-primerFw7 | TTTTATATCTCGGTGACAAG | |
| | 13-084seq-primerFw8 | CTCCTTTTCAAGTATTC AAC | |
| | 13-084seq-primerRv1 | CCATGGGAACTTGGACTTG | |
| | 13-084seq-primerRv2 | CAGGACATCCTTTGGATGG | |
| | 13-084seq-primerRv3 | TATTCAGTCGACCTTCACAC | |
| | 13-084seq-primerRv4 | AACATGCTAAACGTCGGAAG | |
| | 13-084seq-primerRv5 | CTAATTAGGCCGGTAAATCTC | |
| | SSP | ACCTTTGAACCATCAGTGCTCA | TCTCCTTGAGCCCTCAAAACTG |
| | 47-55 | CCATCATTTGGCCTGCTGTTC | TACCATGGGAACTTGGACTTG |

*Markers were based on Garavito et al. (1).

1. Garavito A, et al. (2010) A genetic model for the female sterility barrier between Asian and African cultivated rice species. *Genetics* 185:1425–1440.