Lineage-specific gene acquisition or loss is involved in interspecific hybrid sterility in rice

Yohei Koide[a,b,c,1], Atsushi Ogino[a], Takanori Yoshikawa[d], Yuki Kitashima[a], Nozomi Saito[a], Yoshitaka Kanaoka[a], Kazumitsu Onishi[e], Yoshihiro Yoshitake[b], Takuji Tsukiyama[b], Hiroki Saito[b,f], Masayoshi Teraishi[b], Yoshiyuki Yamagata[g], Aiko Uemura[h], Hiroki Takagi[h], Yoriko Hayashi[i], Tomoko Abe[i], Yoshimichi Fukuta[f], Yutaka Okumoto[b], and Akira Kanazawa[a]

[a]Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan; [b]Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan; [c]Hakubi Center for Advanced Research, Kyoto University, Kyoto 606-8501, Japan; [d]Department of Agricultural Regional Vitalization, Kibi International University, Minami-awaji, Hyogo 656-0484, Japan; [e]Department of Agro-Environmental Science, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080-8555, Japan; [f]Tropical Agriculture Research Front, Japan International Research Center for Agricultural Sciences, Ishigaki, Okinawa 907-0002, Japan; [g]Faculty of Agriculture, Kyushu University, Fukuoka 812-8581, Japan; [h]Department of Genomics and Breeding, Iwate Biotechnology Research Center, Kitakami, Iwate 024-0003, Japan; and [i]Accelerator Applications Research Group, RIKEN, Wako, Saitama 351-0198, Japan

Understanding the genetic basis of reproductive barriers between species has been a central issue in evolutionary biology. The S_1 locus in rice causes hybrid sterility and is a major reproductive barrier between two rice species, Oryza sativa and Oryza glaberrima. The O. glaberrima-derived allele (denoted S_1^g) on the S_1 locus causes preferential abortion of gametes with its allelic alternative (denoted S_1^s) in S_1^g/S_1^s heterozygotes. Here, we used mutagenesis and screening of fertile hybrid plants to isolate a mutant with an allele, S_1^{mut}, which does not confer sterility in the S_1^{mut}/S_1^g and S_1^{mut}/S_1^s hybrids. We found that the causal mutation of the S_1^{mut} allele was a deletion in the peptidase-coding gene (denoted "SSP") in the S_1 locus of O. glaberrima. No orthologous genes of SSP were found in the O. sativa genome. Transformation experiments indicated that the introduction of SSP in carriers of the S_1^s allele did not induce sterility. In S_1^{mut}/S_1^s heterozygotes, the insertion of SSP led to sterility, suggesting that SSP complemented the loss of the functional phenotype of the mutant and that multiple factors are involved in the phenomenon. The polymorphisms caused by the lineage-specific acquisition or loss of the SSP gene were implicated in the generation of hybrid sterility. Our results demonstrated that artificial disruption of a single gene for the reproductive barrier creates a "neutral" allele, which facilitates interspecific hybridization for breeding programs.

rice | reproductive barrier | hybrid sterility | transmission ratio distortion | mutagenesis

Reproductive barriers prevent gene flow and maintain genetic discontinuities between species. One of the central issues in evolutionary biology is how genic incompatibility between species arose and became fixed without causing deleterious interaction within a species. The Bateson–Dobzhansky–Muller (BDM) model indicates that incompatibilities are generally caused by an interaction between genes that have functionally diverged in each population (1). BDM-type genic incompatibilities are caused by several types of genetic differences between species, including nucleotide substitutions, the presence of transposable elements or noncoding repeats, gene transposition, and reciprocal gene loss (2). In addition to these differences, recent genome sequence analyses in animal and plant species revealed widespread gene-presence/absence polymorphisms represented by coding sequences that are present in some genomes but completely absent in others (3, 4). The species-specific genes, also known as "orphan genes," are shown to contribute to species-specific adaptations (5). However, it has not been known whether gene-presence/absence polymorphisms are involved in the reproductive barrier between species.

There are two cultivated rice species, Asian (Oryza sativa) and African (Oryza glaberrima) (6). Although O. glaberrima is rapidly being replaced by O. sativa cultivars that produce higher yields,

O. glaberrima has tolerance to various abiotic and biotic stresses (7, 8) and is therefore a valuable source of useful genes (9). Moreover, hybrids between O. glaberrima and O. sativa are very vigorous (Fig. 1A and Fig. S1) (10), suggesting the potential usefulness of interspecific hybrids for increasing grain production. However, strong hybrid sterility between O. glaberrima and O. sativa (Fig. 1 B and C) has prevented the use of desirable traits of O. glaberrima (9, 11–14). Therefore, elucidating the genetic basis of interspecific hybrid sterility between the two rice species is important not only to understand the evolutionary biology of speciation but also to improve crops for future food production.

In Oryza, several causal genes for hybrid sterility have been cloned (15–19). The presence of "neutral" alleles that do not cause sterility in the heterozygous state at intraspecific hybrid sterility loci in O. sativa has been reported (Table S1). Although neutral alleles have facilitated a breakthrough in hybrid rice production in Asia (18, 20), this type of naturally occurring allele has never been found in interspecific cross-combinations in plants (Table S1). Recently, reverse-genetic approaches through CRISPR/Cas9-mediated gene knockout successfully created neutral alleles of intra- and interspecific hybrid sterility genes in

Significance

Hybrid sterility, a major reproductive barrier between species, hinders the transfer of desirable traits from one species to another. We report a forward genetic approach for creating a "neutral" allele of the S_1 locus, a major interspecific hybrid sterility locus in rice. This neutral allele does not induce hybrid sterility in combination with alleles from either Asian or African rice species. The allele carries a deletion in the peptidase-coding gene, SSP, in the S_1 locus. This work provides mechanistic and evolutionary insights into hybrid sterility and demonstrates the feasibility of the approach that allows broader access to desirable traits in distantly related species during crop breeding.
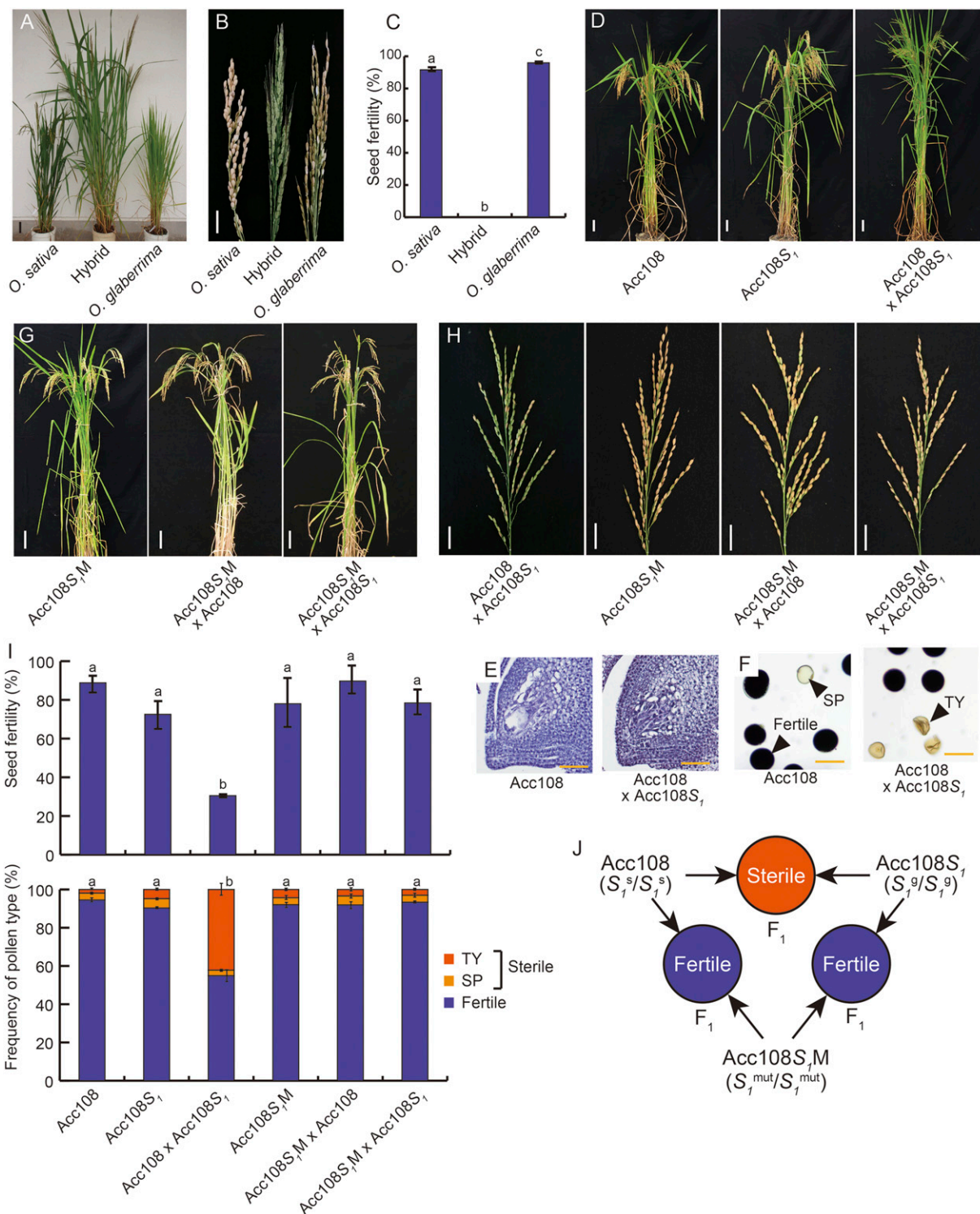
Fig. 1. Effects of lines carrying S_1^s, S_1^g, and the Acc108S_1M on hybrid sterility. (*A*) Hybrid vigor observed in F$_1$ plants derived from a cross between *O. sativa* (cultivar, NERICA10) and *O. glaberrima* (cultivar, CG14). (Scale bar: 10 cm.) (*B*) Panicles of *O. sativa* (cultivar, T65), *O. glaberrima* (cultivar, CG14), and an interspecific hybrid. (Scale bar: 2 cm.) (*C*) Mean ± SD seed fertility of *O. sativa* (cultivar, T65), *O. glaberrima* (cultivar, CG14), and an interspecific hybrid. Means labeled with different letters differed significantly in Tukey's honestly significant difference (HSD) test ($n = 3$ plants). (*D*) Plant type of Acc108, Acc108S_1, and their F$_1$ progeny. (Scale bars: 10 cm.) (*E*) Light micrographs of developing female gametophytes of Acc108 and F$_1$ derived from a cross between Acc108 and Acc108S_1 stained with Delafield's hematoxylin (Muto Pure Chemicals). (Scale bars: 50 μm.) (*F*) I$_2$-KI–stained pollen of Acc108 and of F$_1$ plants derived from a cross between Acc108 and Acc108S_1. TY and SP indicate typical and spherical-type aborted pollen grains, respectively. (Scale bar: 50 μm.) (*G*) Plant types of Acc108S_1M, Acc108S_1M × Acc108, and Acc108S_1M × Acc108S_1. (Scale bars: 10 cm.) (*H*) Panicles of Acc108 × Acc108S_1, Acc108S_1M, Acc108S_1M × Acc108, and Acc108S_1M × Acc108S_1 plants. (Scale bars: 4 cm.) (*I*) Mean ± SD seed fertility (*Upper*) and frequency of pollen types (*Lower*) of Acc108 and Acc108S_1 plants and their F$_1$ progeny. Means labeled with different letters differed significantly in Tukey's HSD test ($n = 3$ plants). (*J*) Interfertility relationships among Acc108, Acc108S_1, and Acc108S_1M. F$_1$ plants derived from a cross between Acc108 and Acc108S_1 are seed and pollen sterile, while F$_1$ plants derived from a cross between Acc108S_1M and either Acc108 or Acc108S_1 are seed and pollen fertile.

rice (21–23). However, such reverse-genetic approaches need detailed information about the causative genes before application; therefore, these approaches are not practical for disrupting reproductive barriers for which the causative genes have not been characterized.

To date, more than 10 loci for hybrid sterility between *O. sativa* and *O. glaberrima* have been reported (24–26). Among them, the *HYBRID STERILITY 1* (S_1) locus on chromosome 6 has frequently been detected, suggesting that the S_1 locus is the major cause of the sterility barrier (11, 13, 14). The hybrid between *O. sativa* and near-isogenic lines (NILs) of *O. sativa* containing a segment of chromosome 6 from *O. glaberrima* showed partial sterility in pollen and seeds (11). This phenomenon was explained by Sano in 1990 as the interaction between two genetically defined alleles (12), the S_1^g allele [formerly the S_1 allele of Sano (12), renamed here to avoid potential confusion with the name of the locus] identified in the chromosomal segment derived from *O. glaberrima* and its alternative allele, the S_1^s allele [formerly the S_1^a allele of Sano (12)], in *O. sativa* at the S_1 locus. The S_1^g allele acts as a "gamete eliminator," and both male and female gametes possessing the S_1^s allele are aborted only in the heterozygote (S_1^g/S_1^s). As a result, the S_1^g allele is transmitted more frequently than the S_1^s allele through both male and female gametes, causing a sex-independent transmission ratio distortion (*si*TRD) in later generations (27). Previous genetic-mapping studies suggest that multiple components affecting hybrid sterility are located in a region on chromosome 6 from *O. glaberrima* [in this article the region is denoted as the "glaberrima sterility complex" (GSC), equivalent to the "S_1 regions" of Garavito et al. (13)], and a factor causing hybrid sterility is located in the 27.8-kb interval (13) (the S_1 locus) between two DNA markers, RM19357 and RMC6_22028, in the GSC (for more detail, see *SI Introduction* and Fig. S2). A recent attempt at a CRISPR/Cas9-mediated knockout of *OgTPR1*, a gene coding a protein containing trypsin-like peptidase and ribosome biogenesis regulatory domains located in the S_1 locus, resulted in an increase in the level of pollen and seed fertility in the heterozygote (S_1^g/S_1^s) (22). Although this study showed that *OgTPR1* is necessary for the S_1 locus-mediated *si*TRD and suggested that a nucleotide substitution causes a functional change between *OgTPR1* and its allelic counterpart *OsTP1*, whether *OgTPR1* is sufficient for inducing this phenomenon has not been examined. Thus, the overall mechanism of the S_1 locus-mediated *si*TRD system has remained unknown. Interestingly, previous studies showed that there are large insertions in the S_1 locus in the *O. glaberrima* genome (the S_1 locus is 27.8 kb in *O. sativa*, while it is 50.3 kb in *O. glaberrima*). In these insertions, five genes that are not present in *O. sativa* genome are predicted (13). These results raised the possibility of the involvement of species-specific genes in the S_1 locus-mediated *si*TRD.

Here, we describe a forward genetic approach to screen a mutant that does not induce sterility in a hybrid with either the S_1^g or S_1^s allele. Using this mutant, we found that a gene coding for a serine/trypsin-like peptidase domain-containing protein, which is not *OgTPR1*, is also an essential gene for inducing hybrid sterility in heterozygotes (S_1^g/S_1^s). The causative gene is located on the *O. glaberrima*-specific insertion, and no orthologous genes were found in the *O. sativa* genome. We demonstrate that the presence of the causal gene alone is insufficient for inducing S_1 locus-mediated *si*TRD, but its presence with another factor(s) is sufficient. This study provides evidence that the gene-presence/absence polymorphism is involved in the reproductive barrier between species. This study also demonstrated that a neutral allele for a hybrid sterility locus can be created through a forward genetic approach, which potentially enables crop breeding using distantly related species between which even uncharacterized hybrid sterility exists.

Results

Histological Analysis and Genetic Mapping of the S_1 Locus.
Rice strains Acc108 (S_1^s carrier) and the NIL Acc108S_1 (S_1^g 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*). In female gametogenesis, abnormal embryo sacs were observed in the F$_1$ hybrid (Fig. 1*E* and Fig. S1). During male gametogenesis, a typical (TY) type of abortion was frequently noted in pollen grains in the F$_1$ hybrid (Fig. 1*F* and Fig. S1; for abortion types of pollen grains, see *SI Materials and Methods*). These results indicate that seed and pollen sterility in the F$_1$ hybrids is caused by developmental failure during male and female gametogenesis (for details, see *SI Results*).

By using the segregating population derived from the cross between Acc108 and Acc108S_1, the gene causing hybrid sterility and *si*TRD [excess of the *O. glaberrima*-type (Og-type) allele] was mapped onto the 273-kb interval between two markers (RM19350 and RM19369) where the S_1 locus was located (Fig. 2*A* and *SI Results*).

Mutant Acc108S_1M Has a Deletion in the S_1 Locus-Specific Peptidase Domain-Containing Protein Gene.
The mutant Acc108S_1M, which does not induce sterility in the hybrid with either the S_1^g or the S_1^s carrier, was obtained through mutagenesis using heavy-ion beam irradiation applied to the seeds derived from the cross between Acc108 and Acc108S_1 (for details, see *SI Results*). Acc108S_1M had a high level of pollen and seed fertility in hybrids with both Acc108 and Acc108S_1 (Fig. 1 *G–J*), whereas quantitative values for other morphological traits were similar to those of Acc108S_1 (Fig. S1*B*).

Genetic experiments suggest that the mutant phenotype is controlled by a single locus (*SI Results*). The crossing experiments and genetic mapping revealed that the mutation was located in the 273-kb region on the chromosome of *O. glaberrima* that contains the S_1 locus (Fig. 2*A* and *SI Results*). Therefore, we speculated that the mutant Acc108S_1M has a mutated allele, designated "S_1^{mut}," at the S_1 locus and that the S_1^{mut} allele acts as the neutral allele in the S_1 locus-mediated *si*TRD system. In addition to the S_1^{mut} allele, the gene responsible for the excess of the Os-type allele was also mapped in a region different from that of the S_1 locus (Fig. 2*A* and *SI Results*). This gene was not investigated further in this study.

Comparison of the 273-kb sequence in the Acc108S_1M and Acc108S_1 revealed that the Acc108S_1M has a 5-bp deletion in the S_1 locus (Fig. 2*B*), with no other mutations observed in the 273-kb region. The gene structure determined by RNA-sequencing (RNA-seq) analysis in the S_1 locus indicated that the deletion had occurred in the gene coding for a serine/trypsin-like peptidase domain-containing protein (Fig. 2*B*). The gene was not found in the corresponding chromosomal region of *O. sativa*. In addition, no homologous genes with >50% amino acid sequence identity were detected in the *O. sativa* genome in a BLASTP search (Fig. S3). A serine proteinase-specific catalytic triad was present in the protein sequence of this gene (Fig. S3). Therefore, we tentatively designated this gene as the S_1 locus-specific peptidase domain-containing protein (SSP). The SSP of Acc108S_1 (denoted "SSP-gla") encodes 246 amino acids, whereas that of the Acc108S_1M (denoted "SSP-mut") has a 5-bp deletion in the second exon, causing a frame-shift mutation and thereby generating a premature stop codon in a peptidase domain (Fig. S3). These results suggested that the 5-bp deletion in the SSP gene is the cause of the S_1^{mut} allele in Acc108S_1M.

The SSP gene was located ~7 kb from the *OgTPR1* gene, a gene recently reported to be involved in S_1 locus-mediated *si*TRD (22). Although both these genes encode a protein containing trypsin-like peptidase domains, there is less than 12.5% identity between the domains encoded by these genes (Fig. S3).
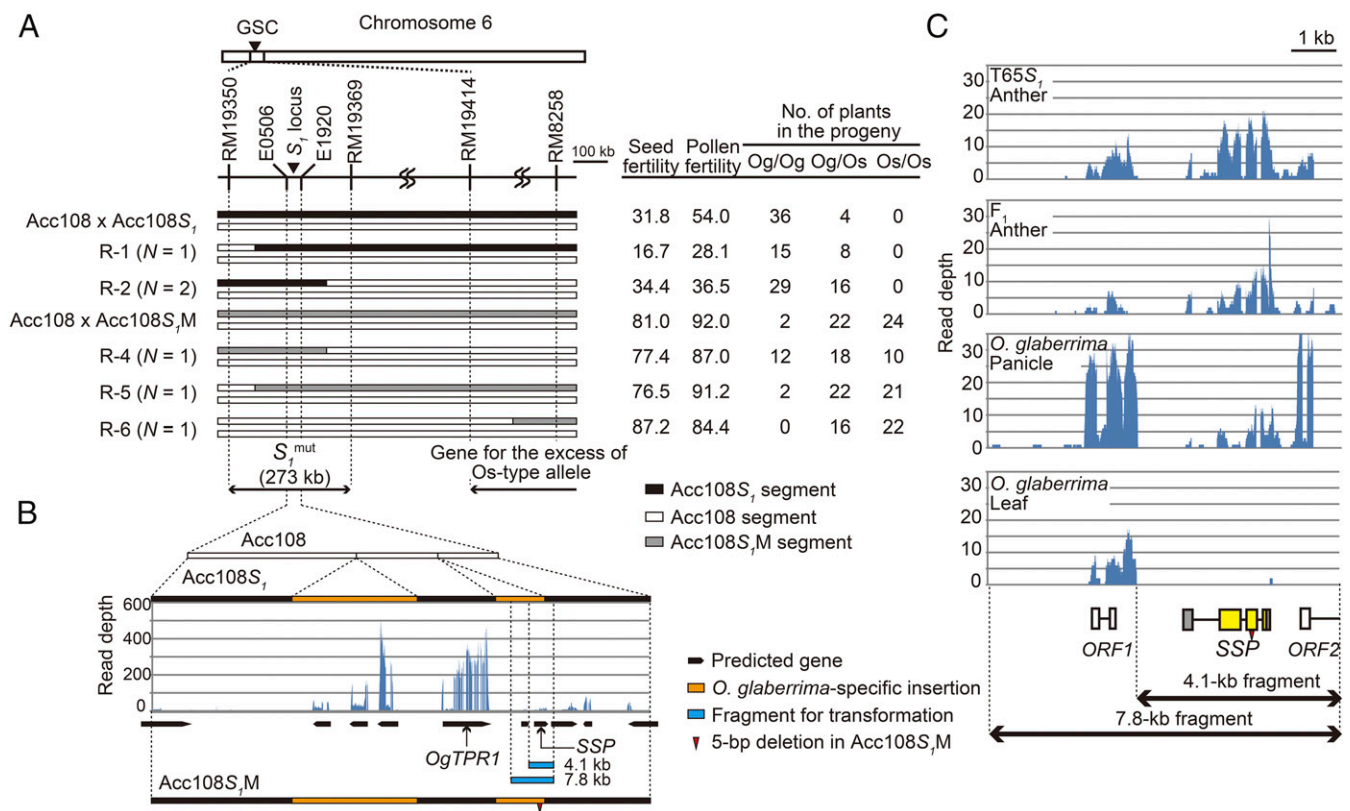
Fig. 2. Map-based cloning of the S_1^{mut} allele in the Acc108S_1M mutant. (*A*, *Left*) Substitution mapping of S_1^{mut}. S_1^{mut} was mapped onto the region between markers RM19350 and RM19369, where the S_1 locus was mapped. The genotypes of plants used for the mapping are shown graphically. Black, white, and gray boxes indicate chromosomal segments from *O. glaberrima*, *O. sativa*, and the Acc108S_1M mutant, respectively. (*Right*) The table shows the means for seed and pollen fertility in each genotype and the number of *O. glaberrima*-type homozygotes (Og/Og) and heterozygotes (Og/Os) and *O. sativa*-type homozygotes (Os/Os) in progenies. (*B*) Genomic structures of the S_1 locus in Acc108 and the Acc108S_1 and Acc108S_1M mutants. Sequence analysis revealed that Acc108S_1 and Acc108S_1M possess insertions specific to *O. glaberrima* in this region (orange boxes). Ten genes were predicted in the S_1 locus in Acc108S_1 based on the RNA-seq analysis using young anthers of T65S_1. The number of mapped reads in each nucleotide is shown. The Acc108S_1M had a 5-bp deletion in the *O. glaberrima*-specific insertion in the S_1 locus (red downward-pointing triangle). The positions of the *SSP* and *OgTPR1* are shown by arrows. Two WT genomic fragments (4.1 and 7.8 kb; pale blue boxes) were used for transgenic experiments. (*C*) RNA-seq analysis to determine the gene structure surrounding the 5-bp deletion in the Acc108S_1M mutant. Short reads obtained from RNA samples of young anthers in T65S_1 and T65wx × T65S_1 were mapped to the 7.8-kb genomic sequence surrounding the 5-bp deletion. In addition, RNA-seq data of young panicle and leaf samples from *O. glaberrima* were obtained from the public database and used (*SI Materials and Methods*). The number of mapped reads in each nucleotide is shown. Three predicted genes in the region are depicted by a box with the gene name below.

In *O. sativa*, 17 genes have been reported to encode trypsin-like peptidase domain-containing protein. A phylogenetic analysis indicated that trypsin-like peptidase domains encoded by *SSP* and *OgTPR1* belong to different clades, which is consistent with the low level of sequence similarity between them (Fig. S3). The expression profile of genes encoding trypsin-like peptidase suggested that they are not preferentially expressed in any specific tissues or stages (Fig. S3). The gene ontology analysis showed that some of these genes (Fig. S3) are related to protein folding, proteolysis, intracellular signal transductions, and apoptotic processes.

The *SSP* and *OgTPR1* Genes Are Coexpressed in Developing Gametophytes. RT-PCR analysis showed that *SSP* is expressed in the plumule, young root, anther, and ovule in *O. glaberrima*, while *OgTPR1* is expressed in various tissues including the anther and ovule (Fig. S4*A*). *SSP* and *OgTPR1* were also expressed in young panicles in Acc108S_1, Acc108 × Acc108S_1 hybrids, and Acc108S_1M, suggesting the coexpression of these two genes in developing gametophytes (Fig. S4*B*).

To investigate the spatiotemporal expression of *SSP* and *OgTPR1* in detail, we performed single-target (Fig. S4 *C–L*) and double-target in situ hybridizations (Fig. 3). In the functional sporophyte stage of female gametogenesis of Acc108S_1, *SSP* transcripts were detected in the tissues surrounding the embryo sac, including the nucellus and the inner and outer integuments (Fig. 3*C* and Fig. S4*C*). In the early to middle embryo sac stages, a weak signal for the *SSP* transcript was also detected in the nucellus (Fig. 3*H* and Fig. S4 *D* and *E*). The same expression pattern of the *SSP* gene was observed during female gametogenesis in Acc108 × Acc108S_1 hybrids (Fig. S4 *F–H*). The *OgTPR1* transcripts were also detected in tissues where the *SSP* transcripts were detected during female gametogenesis (Fig. 3 *D* and *I* and Fig. S4 *K* and *L*). Furthermore, the coexpression of the *SSP* and *OgTPR1* genes in the nucellus and the inner and outer integuments was confirmed by colocalization of the signals in the double-target in situ hybridization analysis (Fig. 3 *E* and *J*).

We could not detect any specific signals in the developing anthers of Acc108S_1 using the *SSP* antisense probe. To confirm the expression of *SSP* and *OgTPR1* during male gametogenesis, we carried out quantitative RT-PCR analysis and confirmed the expression of *SSP* and *OgTPR1* in young anthers in homozygotes for the S_1^g allele in the genetic background of a japonica rice strain, T65 (T65S_1) (Fig. S4*M*). The expression of the *SSP* and *OgTPR1* genes in the young anthers and young panicles was also confirmed by RNA-seq analysis of T65S_1 and RNA-seq
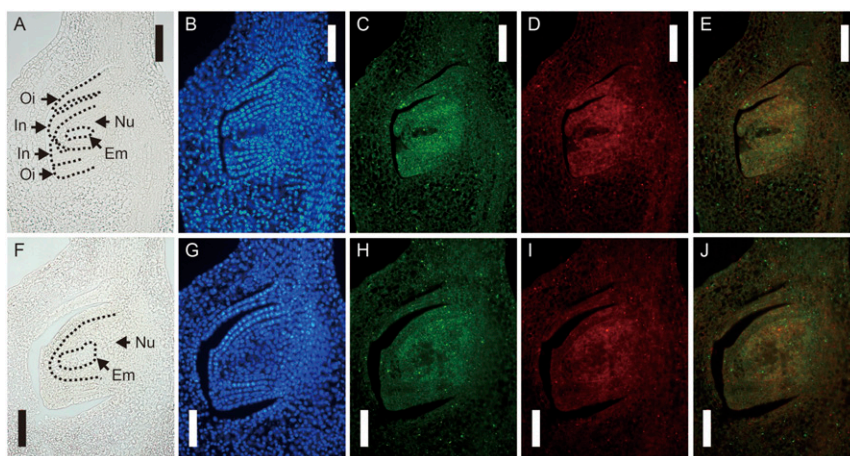
Fig. 3. Double-target in situ hybridization of *SSP* and *OgTPR1* in developing ovules in Acc108S_1. A biotin-labeled probe of the *SSP* transcript and a digoxygenin-labeled probe of the *OgTPR1* transcript were used. (*A–E*) Longitudinal sections of ovules at the functional sporophyte stage. (*F–J*) Longitudinal sections of ovules at the early-middle stage of embryo sac development. Bright-field light micrograph (*A* and *F*), DAPI staining (*B* and *G*), and expression pattern of *SSP* (*C* and *H*) and *OgTPR1* (*D* and *I*) are shown. A merged view of *C* and *D* is shown in *E*; a merged view of *H* and *I* is shown in *J*. Em, embryo sac; In, inner integuments; Nu, nucellus; Oi, outer integuments. (Scale bars: 50 μm.)

data of *O. glaberrima*, in the public database, respectively (Fig. 2 *B* and *C*).

Introduction of the *SSP* Gene Complements the Hybrid Sterility Phenotype. If the *SSP* gene is a necessary and sufficient factor for inducing hybrid sterility, transformation of an S_1^s carrier with the intact *SSP* gene (*SSP-gla*) might induce hybrid sterility. To examine this hypothesis, we used two types of WT genomic fragments (4.1 and 7.8 kb) (Fig. 2) containing *SSP-gla* to transform rice strain T65*wx* (S_1^s carrier). Intriguingly, pollen and seed fertility of the resultant transgenic lines was not significantly lower than that of nontransgenic controls (T65*wx*) (Fig. 4*A* and Fig. S5*B*). This result suggested that *SSP-gla* is not sufficient to induce sterility and that another factor(s) including *OgTPR1* collaboratively acts to induce sterility.

We then conducted a complementation test by crossing these transgenic lines with the Acc108S_1M (S_1^{mut} carrier). If the inability to induce sterility in the Acc108S_1M is caused by the 5-bp deletion in the *SSP* gene, transformation with the WT genomic fragment containing the *SSP-gla* gene will complement the loss-of-function phenotype of the Acc108S_1M (Fig. S5*A*). As a consequence, the F₁ hybrids between the Acc108S_1M and the transgenic lines will show hybrid sterility and *si*TRD identical to that mediated by the S_1 locus, although the *SSP-gla* transgene is present at a different place in the genome. All hybrids harboring the S_1^{mut} allele and transformed with WT genomic fragments indeed showed a high proportion of TY-type aborted pollen, as observed in T65*wx* × T65S_1 (S_1^g/S_1^s) (Fig. 4 *B* and *C* and Fig. S5*C*). In addition, a marked reduction of the proportion of visibly distinguishable *waxy* (*wx*)-type pollen was observed in these hybrids (Fig. 4 *B* and *C* and Fig. S5*C*; for analysis of pollen abortion on the basis of the *wx* phenotype, see *SI Materials and Methods*). Since the *wx* allele is linked in a coupling phase with the S_1^s allele (27), this result indicated the presence of male transmission ratio distortion (*m*TRD) at the S_1 locus in these hybrids. Seed fertility of these hybrids was also lower than that of the hybrids between T65*wx* and Acc108 or Acc108S_1M (Fig. 4*C* and Fig. S5*C*). Furthermore, a marked excess of the Og-type allele was observed in the progenies of the F₁ hybrids between Acc108S_1M and transgenic lines (Fig. 4*D* and Fig. S5*D*), indicating the occurrence of *si*TRD at the S_1 locus. Based on these results, we concluded that the 5-bp deletion in the *SSP* gene caused the functional

change from the S_1^g allele to the S_1^{mut} allele at the S_1 locus and that *SSP-gla* acts with other factor(s) at the S_1 locus to induce sterility.

SSP Is Not Present in the Asian Rice Gene Pool. A BLASTP search showed that there is no homologous gene for *SSP* in the genus *Oryza* other than in *Oryza barthii*, a wild ancestral species of *O. glaberrima* (Fig. S3), suggesting that the *SSP* gene is present only in limited species in this genus. To comprehensively analyze the distribution of the *SSP* gene, we examined 112 rice accessions in nine species of the genus *Oryza* (Fig. 5 and Table S2). The *SSP* gene was detected in *O. glaberrima*, *O. barthii*, and *Oryza meridionalis* but not in the Asian rice gene pool and other distantly related *Oryza* species.

Discussion

The present study showed that the neutral allele S_1^{mut} at the S_1 locus was obtained through the forward genetic approach via mutagenesis using heavy-ion beam irradiation. Such a neutral allele will greatly promote the transfer of elite genes from different species in rice. The cause of the S_1^{mut} was the 5-bp deletion in the *SSP* gene (*SSP-mut*). These findings showed that the intact *SSP* gene (*SSP-gla*) is essential for the S_1 locus-mediated *si*TRD system. However, the introduction of the *SSP-gla* gene into the S_1^s carrier did not induce hybrid sterility or *si*TRD. These results suggested that the ability to induce sterility and *si*TRD is due to the *SSP-gla* gene and other factor(s): the S_1 locus consists of multiple genes, and the epistatic interaction between them induces hybrid sterility and the S_1 locus-mediated *si*TRD. Recently, Xie et al. (22) suggested the involvement of another gene, *OgTPR1*, in the S_1 locus-mediated *si*TRD. However, it was still unclear whether the presence of *OgTPR1* is sufficient for the S_1 locus-mediated *si*TRD. Our results showed that mutant Acc108S_1M, which does not have any mutations in *OgTPR1*, produces fertile hybrids when crossed with Acc108 (Fig. 1 *G–I*). These results indicated that the intact *OgTPR1* alone does not induce the *si*TRD mediated by the S_1 locus.

In well-characterized examples of hybrid sterility involving *Segregation distorter* (*SD*) in *Drosophila*, the *t*-haplotype in mouse, and the *Sa* locus in *O. sativa*, selective dysfunction of gametes is induced by the interaction between two factors, a "distorter" and a "responder" (15, 28, 29). In this distorter–responder system, a distorter and a responder are linked in a repulsion phase, and the distorter acts to raise its transmission by inducing dysfunction of
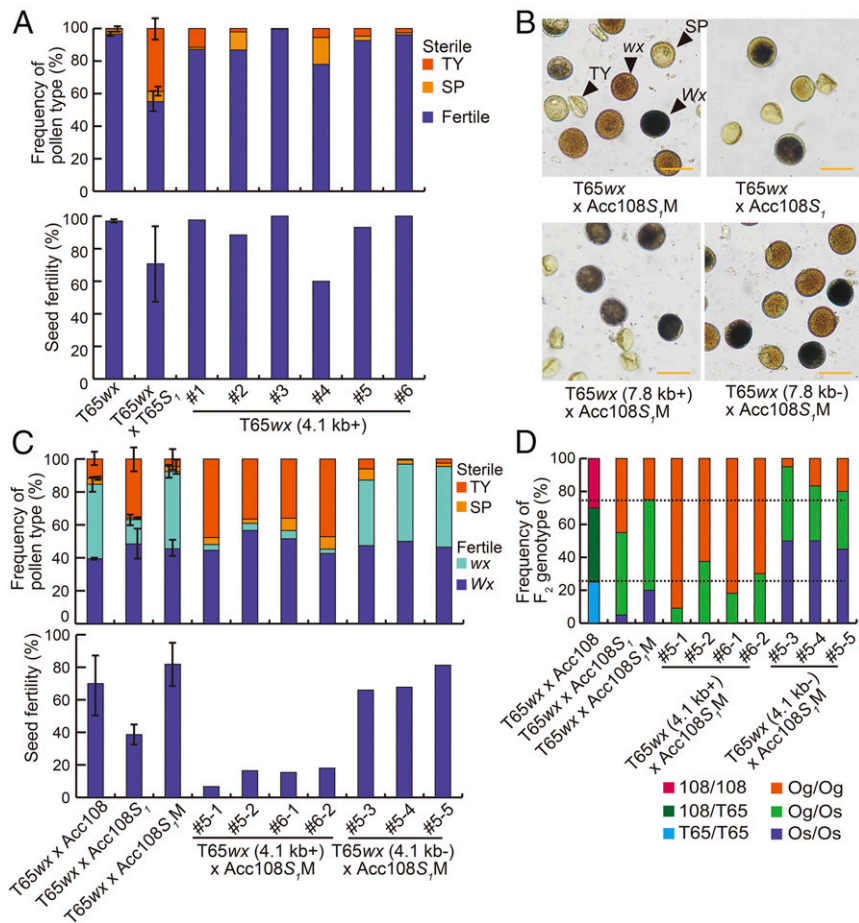
Fig. 4. Transgenic experiments. (*A*) Mean ± SD frequency of pollen types and seed fertility of T65*wx*, T65*S₁*, and six independent transgenic plants (T65*wx*, 4.1 kb+). TY and SP indicate typical- and spherical-type aborted pollen grains, respectively. *n* = 3 plants for T65*wx* and T65*wx* × T65*S₁*; *n* = 1 plant for transgenic plants. (*B*) Pollen grains of T65*wx* × Acc108*S₁*M, T65*wx* × Acc108, T65*wx* (7.8 kb+) × Acc108*S₁*M, and T65*wx* (7.8 kb−) × Acc108*S₁*M plants. *Wx* and *wx* indicate *Wx*- and *wx*-type fertile pollen grains with a functional and nonfunctional allele, respectively. (Scale bar: 50 μm.) (*C*) Mean ± SD frequency of pollen types and seed fertility of T65*wx* × Acc108, T65*wx* × Acc108*S₁*, T65*wx* × Acc108*S₁*M, and transgenic plants. For transgenic plants, two genotypes, T65*wx* (4.1 kb+) × Acc108*S₁*M and T65*wx* (4.1 kb−) × Acc108*S₁*M, were used to assess pollen and seed fertility. A plus sign (+) indicates a transgenic positive; a minus sign (−) indicates a null segregant. *n* = 3 plants for T65*wx* × Acc108, T65*wx* × Acc108*S₁*, and T65*wx* × Acc108*S₁*M; *n* = 1 plant for transgenic plants. (*D*) Mean ± SD frequency of F₂ genotypes obtained from the self-pollination of plants of each genotype. T65*wx* × Acc108, T65*wx* × Acc108*S₁*, T65*wx* × Acc108*S₁*M, T65*wx* (4.1 kb+) × Acc108*S₁*M, and T65*wx* (4.1 kb−) × Acc108*S₁*M were used. 108/108, Acc108-type homozygotes; 108/T65, heterozygotes for Acc108 and T65*wx* types; T65/T65, T65*wx*-type homozygotes; Og/Og, *O. glaberrima*-type homozygotes; Og/Os, heterozygotes for *O. glaberrima* and *O. sativa* types; Os/Os, *O. sativa*-type homozygotes. Dotted lines mark 25% and 75%.

gametes carrying the responder. On the other hand, Yang et al. (18) proposed a "killer–protector" system at the *S5* locus encoded by three tightly linked genes (ORF3–ORF5) in hybrid sterility. In this killer–protector system, a killer and a protector are linked in a coupling phase, and the killer selectively kills the gametes that do not have the protector, thus raising its transmission. In this model,
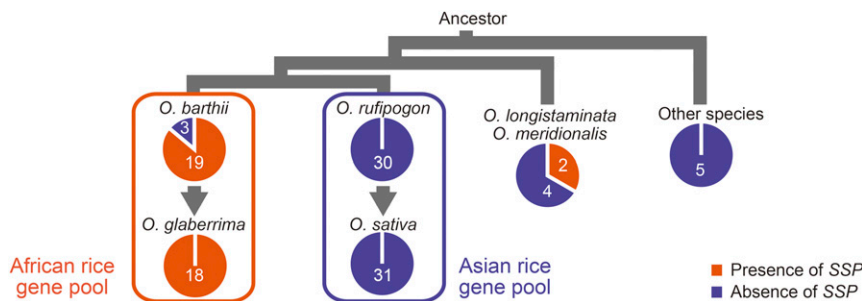

Fig. 5. Distribution of the *SSP* in *Oryza* species. The proportions of the number of strains with and without the *SSP* gene and the number of strains are shown in pie charts. The charts are arranged on phylogenetic relationships of *Oryza* species (for details, see *SI Materials and Methods*). The process of domestication is shown by arrows. *Oryza longistaminata* and *O. meridionalis* have the same AA genome as species in the Asian and African rice gene pools. *Oryza officinalis*, *Oryza minuta*, and *Oryza ridleyi* were used as outgroups.

ORF5, which encodes an aspartic protease, acts like a killer when partnered with ORF4, whereas the functional ORF3 (protector) allele prevents dysfunction of gametes that have this allele. Thus, all these systems of hybrid sterility involve two or more factors: a distorter–responder or a killer–protector to express hybrid sterility.

Are these systems involved in the hybrid sterility mediated by the S_1 locus? The fact that Acc108S_1M and the CRISPR/Cas9-mediated mutant of $OgTRP1$ produced fertile hybrids in crosses between both S_1^g and S_1^s carriers indicates that neither of these mutants has such a killer (distorter) function. Because Acc108S_1M and the CRISPR/Cas9-mediated mutant of $OgTPR1$ possess intact $OgTPR1$ and SSP-gla, respectively, these results suggested that both SSP-gla and $OgTPR1$ are necessary but not sufficient for the killer function. This idea is strongly supported by the lack of sterility induction when the SSP-gla gene is introduced into the S_1^s carrier. Although the SSP-gla transgene did not singly induce sterility, it complemented the phenotype of the Acc108S_1M and induced the S_1 locus-mediated siTRD; the observed frequency of sterility suggests that the SSP gene acts in a sporophytic manner, which is also consistent with its expression in the nucellus (the sporophytic tissue surrounding the developing embryo sac). Considering the results collectively, we propose a model in which at least three genes are involved in this phenomenon (Fig. S6 A and B): The SSP-gla gene, which acts in a sporophytic manner, and the $OgTPR1$ gene cooperate as a killer, and another gene(s) in the S_1 locus for factor(s) corresponding to the responder or protector controls the selective dysfunction of gametes in hybrids. Interestingly, both the SSP-gla and the $OgTPR1$ genes encode peptidase domain-containing proteins, suggesting that genes related to protein degradation are involved in the killer function, not only in the case of the $S5$ locus in intraspecific hybrid sterility but also in interspecific hybrid sterility.

Our expression analysis showed the coexpression of SSP and $OgTPR1$ in developing gametophytes, whereas their expression patterns in other tissues are not identical. These results suggest that SSP and $OgTPR1$ genes may collaboratively act in specific tissues and developmental stages including developing gametophytes. Although there is no experimental support, Xie et al. (22) suggested that $OgTPR1$ was involved either in the activation of certain precursor molecules via the peptidase domains or in the ribosome assembly through the RRS domain. It can be hypothesized that the SSP gene may also be involved in the activation of the precursor together with the $OgTPR1$ or in the activation of the OgTPR1 protein by the peptidase domain. Identification of all the other causative factor(s) in the S_1 locus might provide mechanistic insights regarding the protein interactions involved in the S_1 locus-mediated hybrid sterility.

If hybrid sterility induced by the S_1 locus fits the killer–receptor or killer–protector models, its evolutionary process may be explained by a model similar to the BDM model. Assuming that the killer and receptor arose separately and were fixed in two divergent lineages or that the killer and protector were fixed in one of two divergent lineages, hybrid sterility could evolve without having a deleterious effect in each lineage. One of significant features of our results is that the SSP gene is not found in one of the two diverged lineages (i.e., the Asian rice gene pool). Thus, the present results broaden our understanding of the BDM-type hybrid incompatibility: Hybrid incompatibility can occur as a consequence of a lineage-specific gene acquisition or loss that generated the gene-presence/absence polymorphism between the two lineages (Fig. S6 C and D). Interestingly, in the genus $Oryza$, we did not find any genes homologous to SSP except for three species. If the SSP gene exists only in limited species in the genus $Oryza$, our results lead to a hypothesis that the species-specific gene is involved in a reproductive barrier between species.

In plants, neutral alleles have frequently been reported for intra- but not for interspecific hybrid sterility loci (Table S1); thus, further knowledge regarding the neutral allele in the latter locus is required. In the present study, we used the S_1 locus as a model system and created an artificial neutral allele by disrupting one of the causative genes of hybrid sterility through forward genetic screening of fertile hybrids. In the case of the S_1 locus, the causal gene, SSP, was not found and annotated in the reference genome of rice (cv. Nipponbare). In general, there are numerous hybrid sterility loci that have not been annotated and characterized in crop gene pools. In such a case, a forward genetic screening is more practical than approaches for which gene identification and characterization are prerequisite (e.g., genome editing) for creating neutral alleles. Thus, the present study demonstrated the feasibility of an approach that allows broader access to desirable traits in distantly related species during crop breeding.

Materials and Methods

Pehkuh (an indica type of O. sativa, denoted "Acc108"), Taichung 65 (a japonica type of O. sativa, denoted "T65") and three NILs (Acc108S_1, T65wx, and T65S_1) were used. T65wx contains the wx allele at the Wx locus. Acc108S_1 and T65S_1 have the chromosomal segment proximate to the S_1 region, introduced from a strain of O. glaberrima (W025 from Guinea). Acc108, T65, and T65wx contain the S_1^s allele, whereas Acc108S_1 and T65S_1 contain the S_1^g allele at the S_1 locus.

Heavy-ion beam irradiation was carried out at the RIKEN RI-beam factory, Wako, Japan. Mutant screening was carried out 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. Detailed procedures of cytological observation, gene mapping, sequencing analysis, expression analysis, in situ hybridizations, transformation analysis, and gene distribution analysis are described in SI Materials and Methods.

1. Coyne JA, Orr HA (2004) Speciation (Sinauer Associates, Sunderland, MA).
2. Maheshwari S, Barbash DA (2011) The genetics of hybrid incompatibilities. Annu Rev Genet 45:331–355.
3. Ding J, et al. (2007) Highly asymmetric rice genomes. BMC Genomics 8:154.
4. Marroni F, Pinosio S, Morgante M (2014) Structural variation and genome complexity: Is dispensable really dispensable? Curr Opin Plant Biol 18:31–36.
5. Wissler L, Gadau J, Simola DF, Helmkampf M, Bornberg-Bauer E (2013) Mechanisms and dynamics of orphan gene emergence in insect genomes. Genome Biol Evol 5:439–455.
6. Oka HI (1988) Origin of Cultivated Rice (Elsevier, Amsterdam).
7. Jones MP, Semon M, Aluko K (1997) Diversity and potential of Oryza glaberrima Steud in upland rice breeding. Breed Sci 47:395–398.
8. Li XM, et al. (2015) Natural alleles of a proteasome α2 subunit gene contribute to thermotolerance and adaptation of African rice. Nat Genet 47:827–833.
9. Jones MP, Dingkuhn M, Aluko GK, Semon M (1997) Interspecific Oryza sativa L. X O. glaberrima Steud. progenies in upland rice improvement. Euphytica 92:237–246.
10. Ghesquière A, Séquier J, Second G, Lorieux M (1997) First steps towards a rational use of African rice, Oryza glaberrima, in rice breeding through a 'contig line' concept. Euphytica 96:31–39.

11. Sano Y, Chu YE, Oka HI (1979) Genetic studies of speciation in cultivated rice, 1. Genic analysis for the F_1 sterility between *O. sativa* L. and *O. glaberrima* Steud. *Jpn J Genet* 54:121–132.
12. Sano Y (1990) The genic nature of gamete eliminator in rice. *Genetics* 125:183–191.
13. Garavito A, et al. (2010) A genetic model for the female sterility barrier between Asian and African cultivated rice species. *Genetics* 185:1425–1440.
14. Zhou J, et al. (2010) Genetic dissection of a chromosomal region conferring hybrid sterility using multi-donors from *Oryza glaberrima*. *Euphytica* 175:395–407.
15. Long Y, et al. (2008) Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. *Proc Natl Acad Sci USA* 105:18871–18876.
16. Mizuta Y, Harushima Y, Kurata N (2010) Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proc Natl Acad Sci USA* 107:20417–20422.
17. Yamagata Y, et al. (2010) Mitochondrial gene in the nuclear genome induces reproductive barrier in rice. *Proc Natl Acad Sci USA* 107:1494–1499.
18. Yang J, et al. (2012) A killer-protector system regulates both hybrid sterility and segregation distortion in rice. *Science* 337:1336–1340.
19. Kubo T, Takashi T, Ashikari M, Yoshimura A, Kurata N (2016) Two tightly linked genes at the *hsa1* locus cause both F1 and F2 hybrid sterility in rice. *Mol Plant* 9:221–232.
20. 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.
21. Shen R, et al. (2017) Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat Commun* 8:1310.
22. Xie Y, et al. (2017) Interspecific hybrid sterility in rice is mediated by *OgTPR1* at the *S1* locus encoding a peptidase-like protein. *Mol Plant* 10:1137–1140.
23. Xie Y, et al. (2017) Suppression or knockout of *SaF/SaM* overcomes the Sa-mediated hybrid male sterility in rice. *J Integr Plant Biol* 59:669–679.
24. Doi K, Yasui H, Yoshimura A (2008) Genetic variation in rice. *Curr Opin Plant Biol* 11:144–148.
25. Li J, et al. (2008) Identification of four genes for stable hybrid sterility and an epistatic QTL from a cross between *Oryza sativa* and *Oryza glaberrima*. *Euphytica* 164:699–708.
26. Shen Y, et al. (2015) Fine mapping of *S37*, a locus responsible for pollen and embryo sac sterility in hybrids between *Oryza sativa* L. and *O. glaberrima* Steud. *Plant Cell Rep* 34:1885–1897.
27. 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.
28. Kusano A, Staber C, Chan HY, Ganetzky B (2003) Closing the (*Ran*)GAP on segregation distortion in *Drosophila*. *BioEssays* 25:108–115.
29. Bauer H, Willert J, Koschorz B, Herrmann BG (2005) The *t* complex-encoded GTPase-activating protein Tagap1 acts as a transmission ratio distorter in mice. *Nat Genet* 37:969–973.