Title

PMS1T, producing phased small-interfering RNAs, regulates photoperiod-sensitive male sterility in rice.

Journal

Proceedings of the National Academy of Sciences of the United States of America, 113(52)

Authors

Fan, Yourong

Yang, Jiangyi

Mathioni, Sandra

et al.

Publication Date

2016-12-27

DOI

10.1073/pnas.1619159114

Peer reviewed

PMS1T, producing phased small-interfering RNAs, regulates photoperiod-sensitive male sterility in rice

Yourong Fan (范优荣)[a], Jiangyi Yang (杨江义)[a,1], Sandra M. Mathioni[b,2], Jinsheng Yu (於金生)[a,3], Jianqiang Shen (沈建强)[a], Xuefei Yang (杨雪菲)[a], Lei Wang (王磊)[a], Qinghua Zhang (张清华)[a], Zhaoxia Cai (蔡朝霞)[a], Caiguo Xu (徐才国)[a], Xianghua Li (李香花)[a], Jinghua Xiao (肖景华)[a], Blake C. Meyers[b,2], and Qifa Zhang (张启发)[a,4]

[a]National Key Laboratory of Crop Genetic Improvement, National Centre of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China; and [b]Department of Plant and Soil Sciences, Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711

Contributed by Qifa Zhang, November 21, 2016 (sent for review October 5, 2016; reviewed by Gynheung An and Yaoguang Liu)

Phased small-interfering RNAs (phasiRNAs) are a special class of small RNAs, which are generated in 21- or 24-nt intervals from transcripts of precursor RNAs. Although phasiRNAs have been found in a range of organisms, their biological functions in plants have yet to be uncovered. Here we show that phasiRNAs generated by the photopheriod-sensetive genic male sterility 1 (*Pms1*) locus were associated with photoperiod-sensitive male sterility (PSMS) in rice, a germplasm that started the two-line hybrid rice breeding. The *Pms1* locus encodes a long-noncoding RNA *PMS1T* that was preferentially expressed in young panicles. *PMS1T* was targeted by miR2118 to produce 21-nt phasiRNAs that preferentially accumulated in the PSMS line under long-day conditions. A single nucleotide polymorphism in *PMS1T* nearby the miR2118 recognition site was critical for fertility change, likely leading to differential accumulation of the phasiRNAs. This result suggested possible roles of phasiRNAs in reproductive development of rice, demonstrating the potential importance of this RNA class as regulators in biological processes.

photoperiod-sensitive male sterility | long-noncoding RNA | phasiRNA

Phased small-interfering RNAs (phasiRNAs) are a special class of small RNAs, which are generated in 21- or 24-nt intervals from transcripts of precursor RNAs (1, 2). PhasiRNAs have been found in a wide range of organisms (1–11). In animal, a class of phasiRNAs known as Zucchini-dependent piwi-interacting RNAs were required for spermatogenesis (3, 4). In some plant species, phasiRNA-generating loci (*PHAS* loci) are found to be distributed in genomic clusters and phasiRNAs preferentially accumulate in reproductive tissues (2, 10). In grass anthers, phasiRNAs are generated from precursor transcripts transcribed from *PHAS* genes or loci, triggered by cleavage of 22-nt miR2118 or miR2275 (1, 2). The 21-nt phasiRNAs are abundant in premeiotic anthers, and 24-nt phasiRNAs in meiotic anthers (10). However, the functions of these reproductive phasiRNAs are yet to be characterized.

The discovery of a photoperiod-sensitive male sterility (PSMS) rice (*Oryza sativa* L.) mutant started the development of two-line hybrid rice (12), which has made tremendous contributions to rice production, especially in China (13). Male fertility in PSMS rice is regulated by photoperiod: it is male-sterile under long-day conditions but fertile under short-day conditions. Studies have shown that PSMS segregates as two Mendelian loci, photopheriod-sensetive genic male sterility 1 (*Pms1*) and *Pms3*, in crosses between the original mutant Nongken 58S (abbreviated as 58S) and many rice varieties (14, 15). Recently, the *Pms3* locus was shown to encode a long-noncoding RNA (lncRNA), *LDMAR* (16, 17), whose abundance is critical for male fertility in long days.

Here we show that *Pms1* functions as a *PHAS* locus, which produces a transcript *PMS1T* targeted and triggered by miR2118 to generate 21-nt phasiRNAs. The abundance of the phasiRNAs was associated with the male fertility of 58S under long-day conditions.

Results

Map-Based Cloning of *Pms1*. It was previously shown that fertility of PSMS rice under long-day conditions was governed by two independent segregating loci, *Pms1* and *Pms3*, such that the pollen was completely sterile only when both loci were homozygous for the alleles from 58S (15). In this study, we developed an F_2 population from a cross between 58S and NIL(MH) [near isogenic line developed by introgressing the *Pms1* locus from Minghui 63 (MH63) to 58S background]. Plants homozygous for the 58S allele were highly sterile and plants homozygous for the MH63 allele were fertile under long-day conditions, whereas spikelet fertility of the heterozygotes varied from completely sterile to fully fertile (Fig. 1 *A*–*C*).

For fine-mapping *Pms1*, more molecular markers were developed surrounding the two markers Fssr and Rssr (now P5 in Fig. 1*D*) based on the previous work (18) to genotype 6,841 individuals from the F_2-mapping population. Genotyping these plants identified 88 recombinants between two flanking markers, Fssr and pj23, covering a 138-kb genomic region in MH63 that corresponded to 300 kb in the Nipponbare reference genome (19). Because of extensive phenotypic overlap between heterozygotes and 58S homozygotes, only the recombinants between MH63 homozygotes and heterozygotes were used for further mapping. Genotyping the recombinants resolved the *Pms1* locus to a 4.0-kb region flanked by

Significance

New discoveries have been continuously made in recent years on the roles of noncoding RNAs in regulating biological processes. Phased small-interfering RNAs (phasiRNAs) may be the newest member discovered in recent years. The photoperiod-sensitive male sterility (PSMS) rice is a very valuable germplasm that started the era of two-line hybrid rice. Here we show that phasiRNAs generated by a long-noncoding RNA *PMS1T* encoded by the *Pms1* locus regulates PSMS in rice. This work provides a case associating the phasiRNAs with a biological trait, especially an agriculturally highly important trait, thus confirming that the phasiRNAs indeed have biological functions.

Author contributions: Y.F., J. Yang, and Qifa Zhang designed research; Y.F., J. Yang, S.M.M., J. Yu, J.S., X.Y., L.W., Qinghua Zhang, and Z.C. performed research; C.X., X.L., J.X., and B.C.M. contributed new reagents/analytic tools; Y.F., J. Yang, and Qifa Zhang analyzed data; and Y.F., J. Yang, and Qifa Zhang wrote the paper.

Reviewers: G.A., Kyung Hee University; and Y.L., South China Agricultural University.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo [accession nos. GSE84885 (small RNA and PARE) and KX578835 and KX578836 (PMS1T in 58S and MH63)].

[1]Present address: College of Life Science and Technology, Guangxi University, Nanning 530004, China.

[2]Present address: Donald Danforth Plant Science Center, St. Louis, MO 63132.

[3]Present address: School of Agriculture and Food Science, The Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, Zhejiang A&F University, Linan 311300, China.

[4]To whom correspondence should be addressed. Email: qifazh@mail.hzau.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619159114/-/DCSupplemental.

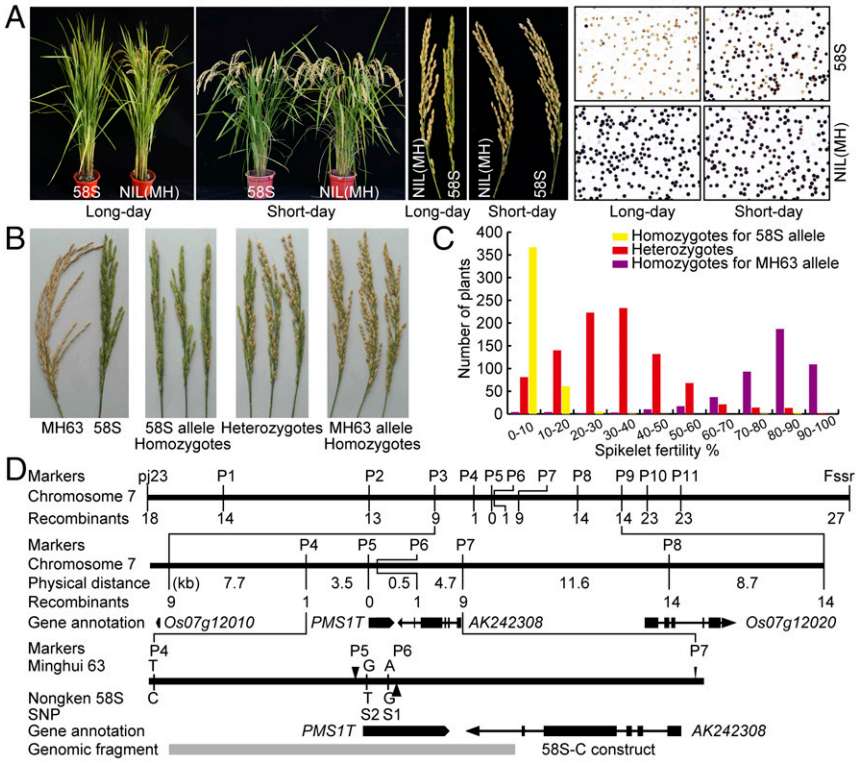

Fig. 1. Genetic mapping of *Pms1*. (*A*) Fertility phenotypes of 58S and NIL(MH) under long- and short-day conditions. Whole plants are shown (*Left*), panicles (*Center*), and pollen (*Right*). (*B* and *C*) Spikelet fertility of plants (*B*) and their frequency distribution (*C*) in the F_2 population from a cross between 58S and NIL(MH) under long-day conditions. The plants were genotyped using marker Fssr. (*D*) Mapping of *Pms1* to a 4.0-kb region on chromosome 7 according to the MH63 genome sequence. The black blocks in the middle represent exons, and the lines between them represent introns, arrows indicate the directions of genes. Sequence polymorphisms between two parents MH63 and 58S are shown below, where vertical lines represent the single nucleotide polymorphisms, and triangles for insertions. The gray block at the bottom indicates the fragment amplified from 58S for preparing the construct 58S-C.

markers P4 and P6, with one recombinant at each side (Fig. 1*D*), with the SSR marker P5 cosegregating with *Pms1*. A sequence comparison of the region bracketed by P4 and P7 between MH63 and 58S revealed two SNPs, S1 and S2, and a 65-bp InDel P6, in addition to P5 (Fig. 1*D*). One recombination event between SNPs S1 and S2 refined the *Pms1* locus to a 3.8-kb region, whereas the S2 SNP with guanine (G) in MH63 and thymine (T) in 58S still cosegregated with *Pms1* (Fig. 1*D* and Fig. S1).

There was no predicted gene within this 3.8-kb region. The nearest predicted genes, LOC_07g12010 and LOC_07g12020, were ~7.7- and 15.0-kb apart, respectively, from P5 according to the MH63 genomic sequence (Fig. 1*D*), and was even longer by the Nipponbare genome sequence. Both genes were annotated as a putative unclassified retrotransposon protein. A full-length cDNA AK242308 (cdna01.dna.affrc.go.jp/cDNA/) on the antisense chain was 1,139-bp apart from P6 (Fig. 1*D*). These data suggested the possibility that *Pms1* encodes a previously unidentified genetic element.

We transformed a construct (58S-C) containing a 5.6-kb genomic fragment from 58S into NIL(MH) (Fig. 1*D*) . Reciprocally, a construct (MH-C) carrying the homologous fragment from MH63 was introduced into 58S. Spikelet fertility of the transgene-positive plants in the 58S-C T_1 families was significantly lower under long-day conditions than the negative segregants, but was not different under short-day conditions (Fig. 2). Analysis of the MH-C T_1 families showed that the transgene had no effect on fertility under both long- and short-day conditions (Fig. 2*C*). These results confirmed that the transformed genomic fragment indeed contained the *Pms1* locus, and also showed that long-day male sterility by *Pms1* is semidominant rather than completely recessive, as previously assumed (14). We thus designated the allele from 58S as *Pms1* and the one from MH63 as *pms1*.

The Candidate Transcript of *Pms1*. We identified from the sense strand within the transformed fragment a transcript (referred to as *PMS1T*) using 5′ RACE and 3′ RACE (Fig. 1*D*). This transcript had no intron, and the one from 58S was 1,453 nt, which was 65-nt longer than the one from MH63 (1,388 bp). This difference provided an InDel marker P6. The two SNPs, S1 and S2, were both located at the 5′ terminus of *PMS1T* (Fig. 1*D* and Fig. S2).

To investigate the effect of *PMS1T* on fertility, an RNAi construct ("dsi") was transformed into 58S (58S-dsi) and NIL(MH) (MH-dsi), respectively, to knock down the expression of this transcript (see, for example, Fig. S4). Under long days, the spikelet fertility of 58S-dsi transgene-positive plants increased significantly versus negative segregants, but no difference was observed in short days (Fig. 2*C* and Fig. S3*A*). The spikelet fertility of MH-dsi transgenic plants did not change regardless of the day length (Fig. 2*C*). In contrast, overexpression of full-length *PMS1T* from 58S driven by the maize ubiquitin promoter (Ubi:S) in NIL(MH) greatly reduced fertility under long days (Table S1). This result indicates that *PMS1T* is the functional form of *Pms1*.

***Pms1* Encodes a lncRNA.** *PMS1T* was predicted to have three short ORFs (Fig. S2). All had no homology to any known proteins. 58S and MH63 shared the same sequences for ORF1 and ORF3, predicted to encode 44 amino acids and 51 amino acids, respectively. ORF2 was predicted to encode 101 amino acids in MH63 but only 54 amino acids in 58S because of an early stop codon caused by the 65-bp insertion at the P6 marker. To inquire whether *PMS1T* encodes any small peptides or if it is a lncRNA, we modified the construct 58S-C by inserting a guanine residue immediately downstream the ATG start codons of putative ORFs to disrupt the reading frames. These constructs, named 58S-ORF1+G, 58S-ORF2+G, and 58S-ORF3+G, were independently transformed into NIL(MH) (Fig. S4). Under long-day conditions, all transgene-positive individuals showed decreased spikelet fertility, similar to 58S-C (Fig. S3 *B–D* and Table S1). Thus, altering the putative ORFs did not affect the function of 58S *Pms1*, suggesting that the *PMS1T* functions as an RNA rather than protein, regarded as a lncRNA.

The SNP S2 Is Crucial for PSMS. There were three polymorphic sites between 58S and MH63 in the genomic sequences that transcribe *PMS1T*: the 65-bp InDel at P6 and the two SNPs, S2 and

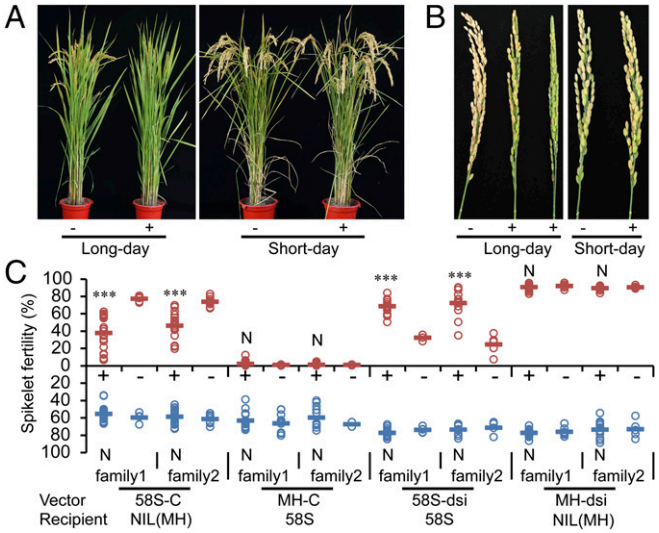

Fig. 2. Fertility of T_1 transgenic plants under long- and short-day conditions. (*A*) Whole plants of transgenic segregants from the T_1 family of 58S-C under long and short days. (*B*) Panicles of transgenic plants of 58S-C. +, positive plants; –, negative plants. (*C*) Spikelet fertility of T_1 transgenic plants from various vectors under long- and short-day conditions. The bars indicate the means. ***$P < 0.001$ significant difference by *t* test between the positive and negative plants in each family; N indicates nonsignificant difference ($P > 0.05$).

S1 (Fig. 1*D*). We compared sequences of 15 different rice lines, including eight PSMS lines bred from 58S, at these sites as well as the SSR marker P5 (Table S2). Lunhui 422 and 1514 were confirmed to contain the fertile allele at the *Pms1* locus by genetic analysis (15), whereas Nongken 58 shared same genotype with 58S (20). Although the P5 marker cosegregated with *Pms1* in the mapping population, it had varied numbers of AT repeats among these lines, thus is an unlikely candidate for PSMS (Table S2). A recombination event between SNP S1 and S2 ruled out the possibility for S1 to be the causal polymorphism (Fig. S1), narrowing the potential causal polymorphism to S2.

No similar sequence was found from other species, except rice using *PMS1T* as the query. The 65-bp inserted sequence, surrounded by sequences similar to those encoding transposon proteins, had high similarity to many loci throughout the genome. We suspected that the 65-bp insertion might be related to the transposon (21).

To investigate whether the 65-bp was responsible for fertility change, four constructs were independently transformed into NIL(MH): 58S-C-dP6 was obtained by deleting the 65-bp insertion of 58S-C; Ubi:S CS+P6, Ubi:S P6, and Ubi:S P6+polyA were designed to overexpress parts of *PMS1T* (Fig. S4). Ubi:S P6 contained a 288-bp fragment from 58S-C including the 65-bp sequence, whereas Ubi:S CS+P6 and Ubi:S P6+polyA included the same 288-bp fragment but extended to the 5′ terminus and 3′ terminus, respectively. Under long-day conditions, 58S-C-dP6 transgenic plants showed greatly reduced spikelet fertility (Fig. S3*E*), just like 58S-C. In contrast, there was no significant fertility difference between the transgene-negative and -positive plants of the constructs Ubi:S P6 and Ubi:S P6+polyA (Table S1). These results suggested that the 65-bp insertion in 58S might not be relevant to fertility difference. The T_1-positive transgenic plants of Ubi:S CS+P6, producing a truncated transcript with a 696-bp sequence, showed as much fertility reduction as Ubi:S (Table S1), indicating that there was no important element for fertility change beyond this region, again supporting SNP S2 as the causal polymorphism.

Expression Pattern of *PMS1T*. Sixteen different tissues from vegetative stage to reproductive stage in 58S and NIL(MH) grown under long- and short-day conditions were assayed for the expression pattern of *PMS1T* (Fig. S5). In general, *PMS1T* expression levels in both 58S and NIL(MH) were low in all tissues assayed, and especially so in green tissues, such as leaf and sheath (Fig. S5). It was expressed preferentially in young panicles and florets, with an expression peak when young panicles were ∼2 cm in length, corresponding to pollen mother cell formation stage (P3). This expression pattern corresponded well with previous results that the most critical period for PSMS occurred during the stages from P1 to P3 (22). Moreover, the transcript abundance was lower in 58S under long-day conditions than 58S in short days, and also than NIL(MH) under both long- and short-day conditions at these three stages (Fig. 3*A*).

***PMS1T* Is a Target of miR2118.** Many lncRNAs that are preferentially expressed in developing inflorescences in the grasses are typically targeted by either miR2118 or miR2275 (2, 23). *PMS1T* was predicted to be targeted by miR2118 (plantgrn.noble.org/psRNATarget/). We validated the cleavage site in *PMS1T* in both 58S and NIL(MH) by modified 5′ RLM-RACE (RNA ligase-mediated rapid amplification of cDNA ends), a site that mapped to the expected location between the 10th and 11th nucleotides from the 5′ end of miR2118, in the complementary region (24) (Fig. 3*B*), suggesting that *PMS1T* was a likely target of miR2118. This cleavage site was further confirmed using data of RNA degradome by parallel analysis of RNA ends (PARE) (Fig. 3*C*). The signature sequence of the highest abundance of PARE reads corresponded to the cleavage site, and the signal value from 58S under long days was higher than the others.

Mature miR2118 is 22 nt in length, and expressed exclusively in the immature inflorescence in rice (2, 25). Sequencing of small RNA libraries constructed using young panicle tissues from P2, P3, and P4 stages of 58S and NIL(MH) grown under long- and short-day conditions showed that miR2118 accumulated higher in P3 and P4 stages than in P2 (Fig. S6). The putative target sequence of *PMS1T* by miR2118 was identical between 58S and MH63. To learn whether cleavage triggered by miR2118 was essential for the function of *PMS1T*, we prepared construct 35S:S-CS containing the truncated *PMS1T* starting right from the miR2118 cleavage site, including only half of the target sequence (Fig. S4), and transformed it into NIL(MH). Another construct, 35S:S (having the same rice fragment as Ubi:S but driven by the 35S promoter), which would reduce the spikelet fertility of NIL(MH) under long-day conditions (Table S1), was used for comparison. Without the complete target sequence of miR2118, 35S:S-CS did not reduce spikelet fertility of NIL(MH), unlike the case of 35S:S or Ubi:S (Table S1). Thus, the 5′ region of *PMS1T* including the recognition site of miR2118 was necessary for causing male sterility by *Pms1*.

***PMS1T* Produces 21-nt phasiRNAs Triggered by miR2118.** MiR2118 triggers 21-nt secondary siRNA biogenesis in plants from lncRNAs or protein coding genes (1, 2, 8, 9, 23, 26, 27). Sequencing of the small RNA libraries constructed from young panicles of 58S and NIL(MH) at P2, P3, and P4 stages under long- and short-day conditions revealed small RNAs of various lengths, from 18 nt to 30 nt within the *PMS1T* region, with 21-nt small RNAs being the most abundant, especially in 58S under long days (Table S3).

A cluster of 21-nt small RNAs was localized between the 82nd and 459th nucleotides of *PMS1T*, arranged in 18 phases or registers from both strands and beginning from the cleavage site of miR2118 (Fig. S7*A*). Additionally, a few nonphased small RNAs were also detected. Some of the sense and antisense phasiRNAs correspond to duplexes with 2-nt 3′-overhangs at both termini (Fig. 3*D* and Fig. S7*A*), consistent with processing by DICER-LIKE 4 (DCL4) and RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) (23, 28).

Comparative analysis revealed that the quantities of 21-nt *PMS1T*-phasiRNAs were higher in P3 and P4 stages than P2. Most phasiRNAs were generated from the sense strand and the 4s, 6s, and 12s phasiRNAs were more abundant than others

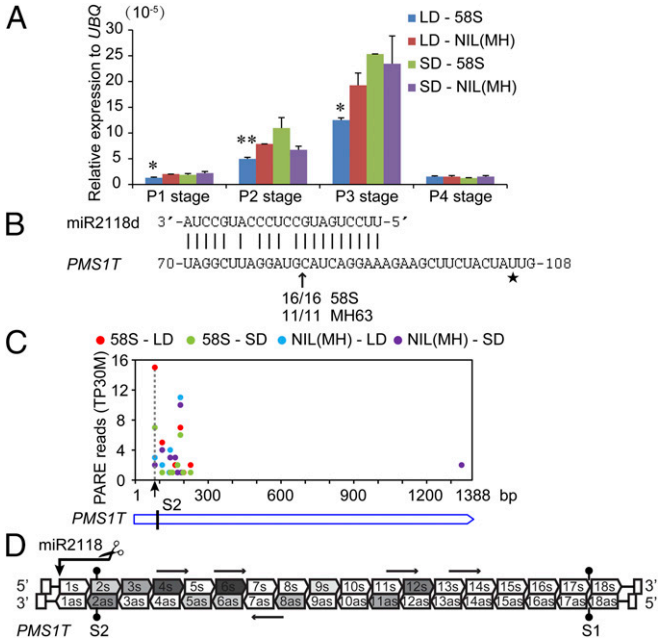

Fig. 3. Targeting of *PMS1T* by miR2118 and production of phasiRNAs. (*A*) Relative abundance of *PMS1T* in 58S and NIL(MH) under long- and short-day conditions at the early stages of panicle development. The descriptions about different stages are as in the legend of Fig. S5. The expression levels are relative to the *UBQ* mRNA. Data are means ± SEM ($n = 3$). Statistically significant difference between LD-58S and LD-NIL(MH) by *t* tests at $*P < 0.05$ and $**P < 0.01$. (*B*) Validation by 5′ RLM-RACE that *PMS1T* is targeted by miR2118 using miR2118d as an example. The arrowhead indicates the cleavage site and sequence frequencies are shown below. Star marks the position of SNP S2. (*C*) PARE results from RNA degradome of *PMS1T* in 58S and NIL(MH) at P3 stage under long- and short-day conditions. Arrow indicates the cleavage sites directed by miR2118. The values of reads in each library are normalized to transcripts per 30 million (TP30M). (*D*) Schematic diagram of the 21-nt phasiRNAs generated from the *PMS1T* transcript in 58S at P3 stage under long-day conditions. The vertical arrow indicates the cleavage site directed by miR2118. Arrow-headed boxes show small RNAs in phase and horizontal arrows indicate ones that are not in phase. Intensity of gray color displays the abundance of 21-nt phasiRNA. as, phasiRNAs from antisense strand; s, phasiRNAs from sense strand.

(Fig. 4*A*). The phasiRNA reads in 58S under long days at P3 stage were the highest followed by 58S under short days, whereas the reads in NIL(MH) were much lower. We noticed a trend of negative relation between the abundance of phasiRNAs and the *PMS1T* transcript (Fig. S8), suggesting that more phasiRNAs and fewer intact *PMS1T* might be related. We also quantified small RNAs from P3 stage in the transgenic plants, which showed that the 21-nt phasiRNA reads of *PMS1T* were much higher in the Ubi:S transgenic-positive plants than in negative ones under long-day conditions (Fig. 4*B*). Conversely, their levels were lower in 58S-dsi transgenic-positive than in negative plants (Fig. 4*C*), despite the total number of 21-nt nonphased small RNA reads in 58S-dsi transgenic positive plants being much larger (Table S3). Examples for detecting the phasiRNAs by RNA blot are illustrated in Fig. 4 *D* and *E*, confirming the presence of the phasiRNAs.

Taken together, all of the results suggested an association between the abundance of the *PMS1T*-phasiRNAs and male sterility under long-day conditions, such that higher accumulation of phasiRNAs result in male sterility.

Discussion

The results obtained in the study can be summarized as the following. The *Pms1* locus encodes a lncRNA, *PMS1T*, and is a 21-*PHAS* gene. MiR2118 triggers the cleavage of *PMS1T* resulting in 21-nt phasiRNAs. These phasiRNAs differentially accumulated in the panicles of 58S compared with NIL(MH), especially under long-day conditions, and the elevated phasiRNAs eventually cause male sterility in rice through yet unknown pathways and genes (Fig. S9).

It is somewhat surprising to find that male sterility conditioned by *Pms1* to be a semidominant trait rather than completely recessive, as we previously expected based on the results of genetic analyses (15, 29). Such semidominance may explain the slow progress in the early days of two-line hybrid rice breeding in China. Knowledge gained from this study may help future two-line hybrid rice breeding using this PSMS system.

In some plant species, phasiRNAs have been found as a class of siRNAs preferentially accumulating in reproductive tissues, particularly anthers, suggesting a possible role in reproductive development (2, 10). The 21-nt phasiRNAs are abundant subepidermally in maize anthers, and 24-nt phasiRNAs in tapetum and meiocytes (10), but the functions of these reproductive phasiRNAs are not known. The rice Agronaute protein, MEIOSIS ARRESTED AT LEPTOTENE 1 (MEL1), binds to 21-nt phasiRNAs generated from over 700 long-intergenic noncoding RNAs (lincRNAs) dispersed on all of the chromosomes. These lincRNAs also contained one or multiple miR2118 recognition sites (1). MEL1 is required for homologous chromosome synapsis in the early meiosis stage and is expressed specifically in germ cells. The *mel1* mutant is male-sterile with vacuolated pollen mother cells (30). Cytological studies by our group showed that the abnormalities in pollen development of 58S under long-day conditions occurred at the pollen mother cell-formation stage when pollen mother cells formed with lots of vacuoles in the large nucleus and the tapetum started to degenerate (16). Thus, there is a lot in common between *MEL1* and *PMS1T* in their roles in regulating pollen development. Nonetheless, *MEL1* encoding an AGO family protein impacts a key component of small RNA generation influencing many *PHAS* clusters, whereas *PMS1T* may impact only a single locus. Studies also showed that loss-of-function of OsDCL4, which is responsible for production of 21-nt phasiRNAs in rice (23), caused abnormal spikelet morphology and also sterility (31). These results also imply that 21-nt phasiRNAs play important roles in rice reproductive development. However, many questions remain to be answered to understand why and how the 21-nt *PMS1T*-phasiRNAs regulate male fertility.

A major question thus emerged concerning the targets of the *PMS1T*-phasiRNAs. The transcripts of *TAS* genes (tasiRNA-generating loci) are noncoding, and tasiRNAs are triggered and generated by one or two hits of miRNAs (7). *TAS1*- and *TAS2*-tasiRNAs target *PPR* (pentatricopeptide repeat) genes, as well as several genes of unknown function; *TAS3*-tasiRNAs target genes from ARF family; a group of *MYB* (myeloblastosis) genes are the targets of *TAS4*-tasiRNAs (32). A large number of *NB-LRR* (nucleotide-binding site and leucine-rich repeat) genes were discovered to be *PHAS* genes in *Medicago* that generated phasiRNAs targeting *NB-LRR* transcripts in *cis* or *trans* (8). In soybean, the phasiRNAs from a noncoding *PHAS* locus accumulating preferentially in anthers are predicted to target transposable elements (9). The *PMS1T*-phasiRNAs are also predicted to target many transcripts, but they did not belong to a special gene family or group. To identify whether one or more or the sum of the phasiRNAs regulate PSMS remains a challenge for future studies (Fig. S9).

Our study showed SNP S2 to be the causal polymorphism of *PMS1T*, and crucial for PSMS. It was located in the 4th nucleotide of the 2s phasiRNA and the 16th nucleotide of the 2as phasiRNA (Fig. S7*B*), 24 nt downstream of the cleavage site directed by miR2118 (Fig. 3*C*). SNPs outside of miRNA target sites could influence miRNA-mediated gene regulation (33–35). There are three speculated possibilities for the mechanism by which SNP S2 of *PMS1T* regulates PSMS. First, SNP S2 may influence the cleavage efficiency directed by miR2118. We analyzed the alteration of RNA

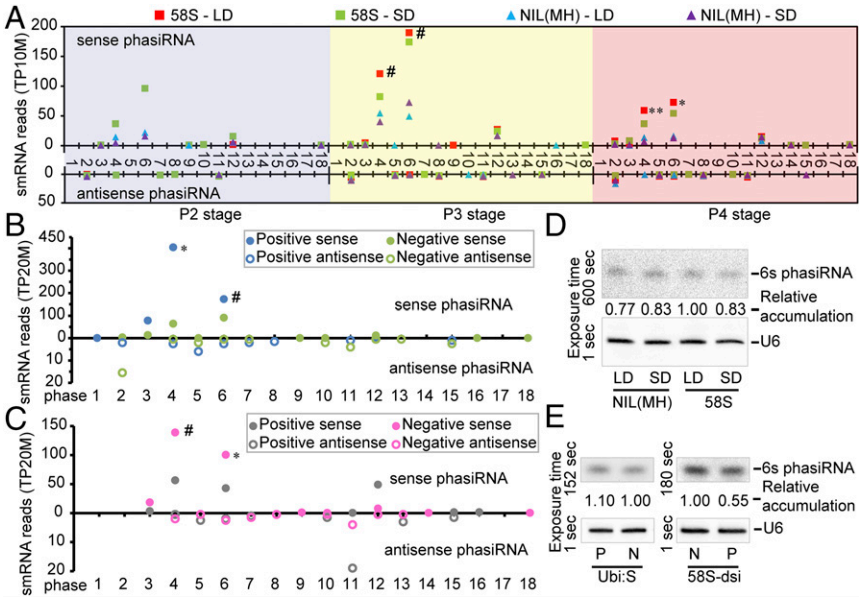

Fig. 4. Expression levels of *PMS1T*-phasiRNAs in various genotypes. (*A*) Abundance of 21-nt *PMS1T*-phasiRNAs from young panicles of 58S and NIL(MH) at P2, P3, and P4 stages under long- and short-day conditions. The smRNA reads are normalized to the whole library with transcripts per 10 million (TP10M) and presented as the mean from two biological replicates. Differences were detected between 58S-LD and NIL(MH)-LD by *t* test at $^{**}P < 0.01$, $^{*}P < 0.05$, or $^{\#}P < 0.1$. (*B* and *C*) Abundance of 21-nt phasiRNAs from young panicles of transgenic T_1 plants from Ubi:S (*B*) and 58S-dsi (*C*) under long-day conditions at P3 stage. Differences were detected by *t* test at $^{*}P < 0.05$ or $^{\#}P < 0.1$, respectively. The smRNA data collecting and handing were as in *A*. (*D* and *E*) RNA blot analysis of 21-nt 6s phasiRNA in transgenic plants. Ten-microgram small RNAs per sample from young panicles at P3 stage were loaded on the gel. The samples in *E* are collected under long-day conditions. The blots were stripped off and rehybridized with U6 probe. The exposure time of the membranes are listed on the left. LD, long days; N, negative plants; P, positive plants; SD, short days.

secondary structure caused by *PMS1T* sequence variation (rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi), predicting a large difference between 58S and MH63 (Fig. S10). Because the 65-bp deletion did not affect male fertility, we used the sequence corresponding to construct 58S-C-dP6 for structure comparison. Contrastive analysis indicated that the S2 SNP brought about a large change in the RNA secondary structure, especially with respect to the miR2118 target region (Fig. S10). This different spatial structure may change the cleavage efficiency directed by miR2118, resulting in more phasiRNAs in 58S at long-day conditions (Fig. S9). Second, SNP S2 may cause differences in target selection of 2s and 2as phasiRNAs between 58S and MH63 under long days, such that one is more effective than the other in targeting the downstream genes for male sterility. Third, SNP S2 may influence the day-length–dependent preferential binding of some unknown RNA binding protein to the *PMS1T*, affecting the processing of phasiRNAs. However, more studies are needed to provide evidence to support these speculations.

Pms1 and *Pms3* both confer PSMS in rice functioning in the same plant. The two share some common characters: they both encode lncRNAs that do not produce proteins; functional mutations are both SNPs; small RNAs can be detected in both transcripts (16, 36). However, there are important differences between them: sterility is of incomplete dominance for *Pms1*, whereas recessive for *pms3*; sufficient amount of the *LDMAR* transcript is required for fertility in long days, whereas abundance of *PMS1T*-phasiRNAs is the main cause for fertility reduction; data obtained at present suggest that RNA-directed DNA methylation is involved in *pms3* (17), but phasiRNAs are produced in *Pms1*. How these two loci function in the same plant to cause male sterility remains to be characterized in future studies.

Nonetheless, the phasiRNAs of *PMS1T* provided an incidence associating the phasiRNAs with a biological trait, especially an agriculturally highly important trait that started the industry of two-line hybrid rice. Such an association confirmed that the phasiRNAs indeed have biological functions that expand our understanding of PSMS, and offer soundly based opportunities for artificially generating male-sterile lines for rice breeding, which may also be useful for other crops.

Materials and Methods

Plant Materials and Field Growth. For natural field evaluation, all plants were grown and investigated under natural field conditions in the experimental farm of Huazhong Agricultural University (114.36°E, 30.48°N), Wuhan, China. For natural long-day conditions, seeds were sown in mid-April to mid-May to place the flowering time of 58S before September 3rd, ensuring the day-length at the young panicle-developing stage to be longer than 14 h. For a natural short day, seeds were sown in mid-to-late June to place the flowering time between September 8th and 20th, so that the day-length of the young panicle-developing stage to be shorter than 13.5 h (22). Plants with the heading time later than September 20th were not included because low temperature may cause sterility in the field.

Fine-Mapping of *Pms1*. The mapping population was the BC_5F_2 generated from the cross between MH63 and 58S with five successive backcrosses using 58S as the female and recurrent parent. Further backcrosses were made to BC_8 from which NIL(MH) (MH63 homozygote in near-isogenic lines) was obtained for subsequent molecular analysis. Molecular markers for mapping are listed in Table S4.

Full-Length cDNA Determination and Modified 5′ RLM-RACE. Young panicles were collected from NIL(MH) and 58S plants growing in the field and stored in liquid nitrogen. Total RNA was isolated following the instruction of RNA extraction kit (TRIzol reagent, Invitrogen).

To synthesize the first-strand cDNA, 1 μg total RNA was reverse-transcribed using the protocol provided by the SMART RACE cDNA Amplification Kit (Clontech), with a final volume of 100 μL. For 5′-RACE, two rounds of PCR amplification were conducted according to the protocol provided with the Advantage 2 PCR kit (Clontech), using two nested adapter primers (UPM and NUPM) in the kit and gene-specific primers (RACE5-1 paired with UPM, RACE5-2 paired with NUPM). The 3′-RACE was essentially the same except that primer RACE3-1 was used to pair with UPM in the first round of reaction, and RACE3-2 with NUPM in the second round of reaction.

The modified 5′ RLM-RACE, which was adapted to validate the cleavage site of miRNA on mRNA, was performed using the FirstChoice RLM RACE Kit (Ambion) by modifying the manufacturer's instructions without enzymatic pretreatment, as described previously (24). During the nested PCR, primer RACE5-3 and RACE5-4 were paired with adapter primer 5′ RACE outer primer and 5′ RACE inner primer, respectively, as supplied in the kit.

All above PCR products were gel-purified, cloned (pGEM-T Easy, Promega), and sequenced. The sequences of the primers are listed in Table S4.

Small RNA-seq and PARE Analysis. Following the TruSeq Small RNA Sample Preparation Guide supplied in the TruSeq Small RNA Sample Prep Kit (Illumina), the libraries for Small RNA-seq were constructed and immediately sequenced on an Illumina HiSEq. 2000 instrument. The small RNA data were analyzed as previously described (8). PARE library construction and sequencing followed previously described methods (37).

Small RNA Page-Northern Blotting Analysis. Total RNA extracted from the young panicles at P3 stage with use of TRIzol reagent (Invitrogen) was mixed with one-third volume 25% (wt/vol) PEG8000 and one-third volume 2.5M NaCl to isolate low molecular weight RNAs. Ten micrograms of low molecular weight RNAs was separated by 15% (wt/vol) denaturing TBE-urea polyacrylamide gels (Invitrogen) and transferred to Hybond-N^+ membranes (GE Healthcare) by use of a transblot semidry transfer cell (Bio-Rad). Antisense DNA oligonucleotides of 21-nt complementary to 6s phasiRNA were used as probes (5′-TACACTATCCTTAGATCATCC-3′). The probe was modified with biotin on the 5′ terminus. The membrane was hybridized with hybridization buffer containing 50-pmol/mL–labeled probes. The hybridization was carried out with previously described methods (38). After probe detection, the blots were stripped and reanalyzed with U6 biotin-labeled probe (5′-TGTATCGTTCCAATTTTATCGGATGT-3′), and 1-pmol/mL–labeled probe was added to the hybridization buffer. The biotin-labeled probes were detected using Chemiluminescent Nucleic Acid Detection Module Kit (Thermo Scientific). Hybridized membranes were placed in Molecular imager ChemiDOC XRS+ (Bio-Rad) for exposure. The exposed images were captured and analyzed with Image Lab software to quantify the relative abundance of small RNAs.

Details of the other materials and methods are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Qili Fei, Jixian Zhai, and Mayumi Nakano for assistance with data handling and helpful discussions. This work was supported by National Natural Science Foundation Grant 91540101, the National Key Research and Development Program Grant 2016YFD0100903 of China, and a grant from the Bill & Melinda Gates Foundation.

1. Komiya R, et al. (2014) Rice germline-specific Argonaute MEL1 protein binds to phasiRNAs generated from more than 700 lincRNAs. *Plant J* 78(3):385–397.
2. Johnson C, et al. (2009) Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Res* 19(8):1429–1440.
3. Han BW, Wang W, Li C, Weng Z, Zamore PD (2015) Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* 348(6236):817–821.
4. Mohn F, Handler D, Brennecke J (2015) Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* 348(6236):812–817.
5. Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during *trans-acting* siRNA biogenesis in plants. *Cell* 121(2):207–221.
6. Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006) A two-hit trigger for siRNA biogenesis in plants. *Cell* 127(3):565–577.
7. Chen HM, Li YH, Wu SH (2007) Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis. *Proc Natl Acad Sci USA* 104(9):3318–3323.
8. Zhai J, et al. (2011) MicroRNAs as master regulators of the plant *NB-LRR* defense gene family via the production of phased, *trans-acting* siRNAs. *Genes Dev* 25(23):2540–2553.
9. Arikit S, et al. (2014) An atlas of soybean small RNAs identifies phased siRNAs from hundreds of coding genes. *Plant Cell* 26(12):4584–4601.
10. Zhai J, et al. (2015) Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proc Natl Acad Sci USA* 112(10):3146–3151.
11. Xia R, Ye S, Liu Z, Meyers B, Liu Z (2015) Novel and recently evolved miRNA clusters regulate expansive F-box gene networks through phasiRNAs in wild diploid strawberry. *Plant Physiol* 169(1):594–610.
12. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive recessive male sterile rice (*Oryza sativa* L. subsp. *japonica*). *Zhongguo Nong Ye Ke Xue* 2:44–48.
13. Si H, Liu W, Fu Y, Sun Z, Hu G (2011) Current situation and suggestions for development of two-line hybrid rice in China. *Chin J Rice Sci* 25(5):544–552.
14. Zhang Q, et al. (1994) Using bulked extremes and recessive class to map genes for photoperiod-sensitive genic male sterility in rice. *Proc Natl Acad Sci USA* 91(18):8675–8679.
15. Mei M, Dai X, Xu C, Zhang Q (1999) Mapping and genetic analysis of the genes for photoperiod-sensitive genic male sterility in rice using the original mutant Nongken 58S. *Crop Sci* 39(6):1711–1715.
16. Ding J, et al. (2012) A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci USA* 109(7):2654–2659.
17. Ding J, et al. (2012) RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. *Mol Plant* 5(6):1210–1216.
18. Liu N, et al. (2001) Identification of an 85-kb DNA fragment containing *pms1*, a locus for photoperiod-sensitive genic male sterility in rice. *Mol Genet Genomics* 266(2):271–275.
19. Yu J, et al. (2007) Rapid genome evolution in *Pms1* region of rice revealed by comparative sequence analysis. *Chin Sci Bull* 52(7):912–921.
20. Wang F, Mei M, Xu C, Zhang Q (1997) pms1 genomic region does not cause fertility difference between the photoperiod-sensitive male sterile rice Nongken 58S and normal Nongken 58. *Acta Bot Sin* 39(10):922–925.
21. Kelley D, Rinn J (2012) Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol* 13(11):R107.
22. Yuan S, Zhang Z, Xu C (1988) Studies on the critical stage of fertility change induced by light and its phase development in HPGMR. *Acta Agron Sin* 14(1):7–13.
23. Song X, et al. (2012) Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J* 69(3):462–474.
24. Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297(5589):2053–2056.
25. Song X, et al. (2012) Rice RNA-dependent RNA polymerase 6 acts in small RNA biogenesis and spikelet development. *Plant J* 71(3):378–389.
26. Chen HM, et al. (2010) 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc Natl Acad Sci USA* 107(34):15269–15274.
27. Shivaprasad PV, et al. (2012) A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* 24(3):859–874.
28. Rajeswaran R, Pooggin MM (2012) RDR6-mediated synthesis of complementary RNA is terminated by miRNA stably bound to template RNA. *Nucleic Acids Res* 40(2):594–599.
29. Mei M, et al. (1999) *pms3* is the locus causing the original photoperiod-sensitive male sterility mutation of 'Nongken 58S'. *Sci China C Life Sci* 42(3):316–322.
30. Nonomura K, et al. (2007) A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *Plant Cell* 19(8):2583–2594.
31. Liu B, et al. (2007) *Oryza sativa dicer-like4* reveals a key role for small interfering RNA silencing in plant development. *Plant Cell* 19(9):2705–2718.
32. Howell MD, et al. (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell* 19(3):926–942.
33. Mishra PJ, et al. (2007) A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. *Proc Natl Acad Sci USA* 104(33):13513–13518.
34. Chen AX, et al. (2011) Germline genetic variants disturbing the *Let-7*/LIN28 double-negative feedback loop alter breast cancer susceptibility. *PLoS Genet* 7(9):e1002259.
35. Li J, Reichel M, Millar AA (2014) Determinants beyond both complementarity and cleavage govern microR159 efficacy in *Arabidopsis*. *PLoS Genet* 10(3):e1004232.
36. Zhou H, et al. (2012) Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. *Cell Res* 22(4):649–660.
37. Zhai J, Arikit S, Simon SA, Kingham BF, Meyers BC (2014) Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. *Methods* 67(1):84–90.
38. Huang Q, Mao Z, Li S, Hu J, Zhu Y (2014) A non-radioactive method for small RNA detection by northern blotting. *Rice (N Y)* 7(1):26.
39. Peng K, Zhang H, Zhang Q (1998) BAC library constructed to the rice cultivar "Minghui 63" for cloning gene of agronomic importance. *Acta Bot Sin* 40(12):1108–1114.
40. Yuan B, Shen X, Li X, Xu C, Wang S (2007) Mitogen-activated protein kinase OsMPK6 negatively regulates rice disease resistance to bacterial pathogens. *Planta* 226(4):953–960.
41. Zhou Y, et al. (2009) Over-expression of aspartate aminotransferase genes in rice resulted in altered nitrogen metabolism and increased amino acid content in seeds. *Theor Appl Genet* 118(7):1381–1390.
42. Qiu D, et al. (2007) OsWRKY13 mediates rice disease resistance by regulating defense-related genes in salicylate- and jasmonate-dependent signaling. *Mol Plant Microbe Interact* 20(5):492–499.
43. Lin YJ, Zhang Q (2005) Optimising the tissue culture conditions for high efficiency transformation of indica rice. *Plant Cell Rep* 23(8):540–547.
44. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):e45.