A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice

Jihua Ding\textsuperscript{a}, Qing Lu\textsuperscript{a}, Yidan Ouyang\textsuperscript{b}, Hailiang Mao\textsuperscript{a}, Pingbo Zhang\textsuperscript{a}, Jialing Yao\textsuperscript{b}, Caiguo Xu\textsuperscript{a}, Xianghua Li\textsuperscript{a}, Jinghua Xiao\textsuperscript{a}, and Qifa Zhang\textsuperscript{a,1}

\textsuperscript{a}National Key Laboratory of Crop Genetic Improvement and National Centre of Plant Gene Research, and \textsuperscript{b}College of Life Science and Technology, Huazhong Agriculture University, Wuhan 430070, China

Contributed by Qifa Zhang, December 31, 2011 (sent for review August 31, 2011)

Hybrid rice has greatly contributed to the global increase of rice productivity. A major component that facilitated the development of hybrids was a mutant showing photoperiod-sensitive male sterility (PSMS) with its fertility regulated by day length. Transcriptome studies have shown that large portions of the eukaryotic genomic sequences are transcribed to long noncoding RNAs (lncRNAs). However, the potential roles for only a few lncRNAs have been brought to light at present. Thus, great efforts have to be invested to understand the biological functions of lncRNAs. Here we show that a lncRNA of 1,236 bases in length, referred to as long-day-specific male-fertility-associated RNA (LDMAR), regulates PSMS in rice. We found that sufficient amount of the LDMAR transcript is required for normal pollen development of plants grown under long-day conditions. A spontaneous mutation causing a single nucleotide polymorphism (SNP) between the wild-type and mutant altered the secondary structure of LDMAR. This change brought about increased methylation in the putative promoter region of LDMAR, which reduced the transcription of LDMAR specifically under long-day conditions, resulting in premature programmed cell death (PCD) in developing anthers, thus causing PSMS. Thus, a lncRNA could directly exert a major effect on a trait like a structure gene, and a SNP could alter the function of a lncRNA similar to amino acid substitution in structural genes. Molecular elucidating gene, and a SNP could alter the function of a lncRNA similar to amino acid substitution in structural genes. Molecular elucidating epigenetics has important implications for understanding molecular mechanisms of photoperiod regulation of many biological processes and also for developing male sterile germplasms for hybrid crop breeding.

Photoperiod regulates a large array of processes underlying growth and development of plants (1). It has been known for centuries that flowering of plants is regulated by day length (2). In recent years, many genes for photoperiod-controlled flowering have been identified in a number of plants, leading to characterization of pathways in flowering regulation (3–9). However, photoperiod regulation of other developmental processes has been studied in far less detail than flowering. It has been reported in several plants that male fertility is also sensitive to photoperiod; a critical day length is required for normal pollen development (10–15). However, nothing is known about the molecular control of the photoperiod regulation of male fertility, despite the tremendous progress in recent years in genetic and molecular understanding of male reproductive development in many plant species (16–22).

Hybrid rice has greatly contributed to the global increase of rice productivity (23, 24). A spontaneous mutant exhibiting photoperiod-sensitive male sterility (PSMS) was found in 1973 in a japonica rice (Oryza sativa ssp. japonica) variety Nongken 58 (referred to as 58N) in Hubei Province, China (12). Intensive studies of the mutant (58S) established that its pollen fertility is regulated by day length; it is completely sterile when grown under long-day conditions, whereas pollen fertility varies when it is grown under short-day conditions (SI Appendix, Fig. S1); and the critical stage for photoperiod induction comes between secondary branch differentiation and microsporogenesis during panicle development, suggesting that photoperiod regulation of PSMS is independent of photoperiod flowering (25–27). For hybrid rice breeding, PSMS rice can be used to propagate itself under short days, and to produce hybrid seeds by interplanting it with normal fertile lines under long-day conditions. Moreover, PSMS rice has a broad spectrum of restoration; all of the normal rice varieties tested can restore the fertility of F1 hybrids (28). Two-line hybrids developed using this PSMS germplasm have occupied millions of hectares of rice fields mostly in China for more than a decade (29). PSMS germplasms have also been found and explored in several other crops (15, 30–32).

Transcriptome studies have shown that large portions of the eukaryotic genomic sequences are transcribed to long noncoding RNAs (lncRNAs). However, the potential roles for only a few lncRNAs have been brought to light at present (33–39).

Here we show that a locus regulating PSMS in rice encodes a lncRNA. A SNP between 58N and 58S in this lncRNA caused epigenetic modifications, which reduced expression of this lncRNA, resulting in male sterility under long-day conditions.

Results

Mapping of the \textit{pms3} Locus to a 12-kb Region. It was determined previously that \textit{pms3} on chromosome 12 was the original mutation that changed 58N to 58S (40–44). Using an F2 population of 7,000 plants from a cross between 58S and DH80, a doubled haploid line derived from a cross between 58S and 1514 (a normal japonica variety), which were examined for fertility under natural field long-day conditions in Wuhan, the \textit{pms3} locus was delimited to a 28.4-kb DNA fragment between two molecular markers LJ25 and LK40 (45) (Fig. 1B). We sequenced this 28.4-kb region and found seven polymorphic sites between 58N and DH80 and a SNP resulting from substituting guanine (G) in 58N with cytosine (C) in 58S (Fig. 1B). We developed four additional molecular markers, which resolved \textit{pms3} to a 12-kb region between M1 and M4 (Fig. 1A).

We assayed the SNP genotypes of 52 rice accessions representing diverse germplasms, including both wild and cultivated species. All these accessions had the G nucleotide at this position, indicating that this C substitution was probably unique in 58S (SI Appendix, Table S1).

Although there were three predicted genes in the 28.4-kb region (LOC\textsubscript{12}g36010, LOC\textsubscript{12}g36020, and LOC\textsubscript{12}g36030), the targeted 12-kb region contained only part of LOC\textsubscript{12}g36030 (Fig. 1C). LOC\textsubscript{12}g36030 encodes three alternatively spliced
transcripts supported by full-length cDNAs, AK071577, AK100176, and AK111270 (http://cdna01.dna.affrc.go.jp/cDNA/), each of which had two exons. The ORF was in the first exon that was shared among all three transcripts, whereas the second exon was specific to each transcript.

**Complexity of Transcripts in the **$pms3$** Region.** We prepared four constructs containing overlapping genomic DNA fragments covering the 12-kb region by digesting clone 117H5 from a BAC library of 58N constructed in our laboratory (Fig. 1D) and subcloning into the vector pCAMBIA1301 (SI Appendix, Fig. S2A). We transformed the constructs into 58S. We also overexpressed the three alternative transcripts (AK071577, AK100176, and AK111270) in both 58N and 58S (SI Appendix, Fig. S2A). At least 15 independent T₀ plants were produced for each of the construct/genotype combinations. Surprisingly, no statistically significant difference was detected in spikelet fertility between the transgene-positive and -negative plants in the T₀ and/or T₁ progenies under long-day conditions in any of the construct/genotype combinations.

Because the targeted 12-kb region contained only part of LOC_12g36030 (Fig. 1D and E), we speculated that there might be other transcripts from this region that regulate PSMS. To check this possibility, we designed 24 pairs of oligonucleotide primers to cover the genomic region ~3 kb in length surrounding the SNP site (SI Appendix, Fig. S3A). Three unique transcripts were identified from the two strands (Fig. 1E and SI Appendix, Fig. S3B-D), which we named Transcript-1 to -3. RT-PCR showed that none of the three transcripts had an intron. Transcript-1 and -2 were transcribed from the sense strand relative to AK111270. Transcript-1 was 1,236 bases in length, and its 3′ end had an 110-base overlap with the 3′ end of AK111270. The 1,268-base sequence of Transcript-2 was located upstream of Transcript-1 and overlapped with the second exon of AK111270 in its entire length. Transcript-3, 1,088 bases in length, was transcribed from the complementary strand, overlapping with Transcript-1 by 688 bases. Thus, this region had a complex transcription pattern producing multiple partial overlapping transcripts. This complexity may also explain why transformation with the genomic fragments could not recover male fertility.

**Expression Patterns of the Transcripts.** Using a strand-specific RT-PCR method, we assayed the expression profiles of these transcripts in 58N and 58S using RNA from 15 tissues spanning the entire life cycle of the rice plant (Fig. 2A). Transcript-1 was detected in almost all of the tissues, but clearly higher in young panicles. Transcript-2 was constitutively expressed and slightly higher in young panicles. Transcript-3 was specifically expressed in young panicles from pollen mother cell-formation stage to pollen mitosis stage and hardly detectable in vegetative organs. The levels of all three transcripts were extremely low in all of the tissues assayed.

RNA in situ hybridization showed that, similar to RT-PCR results, Transcript-1 was detectable in leaf and sheath and produced intensive signals in microspore mother cells and vacuolated pollen cells (SI Appendix, Fig. S4 A–E). Transcript-2 was detectable in all tissues with signals more intense than Transcript-1 (SI Appendix, Fig. S4 F–J). The signal of Transcript-3 was only detectable in anthers at vacuolated pollen cell stage (SI Appendix, Fig. S4 K–O).

Comparative analysis in young panicles (Fig. 2B) showed that the expression level of Transcript-1 in 58N under long-day conditions was significantly higher than the other three treatments: 58N under short-day and 58S under both long-day and short-day conditions in all four developmental stages. The levels of Transcript-1 in 58S were nearly the same under long-day and short-day conditions. The expression of Transcript-2 in 58N under long days was also significantly higher than the other three treatments in the two early stages, but the difference diminished at later stages. The expression of Transcript-3 increased gradually with panicle development but there was no significant difference among the four treatments at any stage.

We analyzed the diurnal expression pattern of Transcript-1 in 58N and 58S in leaves at pistil/stamen primordium differentiation stage at 4-h intervals for a total of 48 h under both long-day and short-day conditions (Fig. 2C). Generally, the expression of Transcript-1 was much higher under long days than short days in
both 58N and 58S. There seemed no identifiable rhythm expression pattern. Unlike in panicles, however, the abundance of this transcript in leaves was similar in 58N and 58S, under both long-day and short-day conditions.

Validation of Transcript-1 as the Candidate for pms3. We took Transcript-1 as the best candidate and overexpressed it in 58S with the construct Actin:NT1 (osfl-actin1-58N-Transcript-1) containing the full-length sequence of Transcript-1 from 58N driven by the rice β-actin1 promoter (SI Appendix, Fig. S2 A and B), which produced 26 independent T0 plants. Transgene-positive T0 plants showed a significant increase in spikelet fertility under natural long days compared with the negative plants (Table 1), without obvious morphological difference (SI Appendix, Table S3). The results were confirmed by further examination of three T1 families randomly selected from T0 plants with single-copy transgene, which showed perfect cosegregation between the transgene and pollen fertility. All positive transgenic plants produced normal pollen grains and were fertile, whereas negative plants produced smaller anthers with collapsed pollen (Fig. 3). We thus named Transcript-1 as pms3.

Pms3 Likely Encodes a lncRNA. We used the ExPASy translate tool to predict the translated protein of Transcript-1, which suggested a short ORF of 73 amino acids, with no homology to any known proteins. To investigate whether this transcript indeed encodes a protein, we changed the predicted start codon ATG to CTG, and overexpressed it (Actin:NT1-mORF, SI Appendix, Fig. S2 A and B) in 58S, resulting in 22 independent T0 plants. Similar to the plants overexpressing Transcript-1, pollen fertility of the transgene-positive plants under natural long days was much higher than that of the negative plants (Table 1). We also prepared a construct overexpressing only the predicted ORF (Actin: NT1-ORF, SI Appendix, Fig. S2 A and B); transformation of this construct in 58S produced 18 independent T0 plants. No fertility recovery was observed in any of the 17 positive plants (Table 1). These results suggested that Transcript-1 may not encode a protein but a lncRNA, which is referred to as long-day fertility-associated RNA (LDMAR).

Functional Importance of the Sequences Surrounding the SNP Site. The SNP between 58N and 58S was located at the 789-bp site of LDMAR for fertility recovery. By contrast, the construct deleting the SNP-did not affect the function of the transcript in 58S, and the full-length sequence of LDMAR from 58N (−200 bp to +300 bp of LDMAR) from leaves of plants grown under LD conditions and young panicles of plants under SD conditions, for 58N and 58S.

Functional Importance of the Sequences Surrounding the SNP Site. The SNP between 58N and 58S was located at the 789-bp site of LDMAR (SI Appendix, Fig. S2B). To investigate how the SNP influenced the function of LDMAR, we prepared three constructs, Actin:ST1 containing the full-length sequence of LDMAR from 58S, Actin:NT1 containing the SNP site was deleted (SI Appendix, Table S3). These constructs were individually transformed into 58S. Examination of transgenic plants (Table 1) showed that Actin:ST1 could recover the spikelet fertility under long days just like LDMAR from 58N, suggesting that sufficient quantity of LDMAR could recover the fertility in 58S despite the presence of the SNP. The 110-bp deletion in the 5′ end did not affect the function of the transcript in fertility recovery. By contrast, the construct deleting the SNP-containing 52-bp sequence failed to recover fertility, indicating that the 52-bp sequence was in a region critical for the function of LDMAR for fertility recovery.

Different Methylation Patterns in the pms3 Region Between 58N and 58S. To understand why LDMAR was expressed differently in 58N and 58S, we performed bisulfite sequencing of DNA

**Fig. 2.** Expression analysis of transcripts identified in the pms3 region. (A) Expression profile of transcripts in 58N (tissue description in SI Appendix, Table S2). Numbers after gene names represent the number of cycles in RT-PCR amplification. (B) Comparative expression analysis of the transcripts between 58N and 58S under both long-day (LD) and short-day (SD) conditions in young panicles at four developing stages: stage3, secondary branch primordium differentiation; stage4, pollen mother cell formation; and stage6, pollen meiosis. All three transcripts were quantified using a common check. Significant differences detected between LD-58N and LD-58S are indicated by **P < 0.01 or *P < 0.05, respectively. Data are means ± SEM (n = 4). (C) Diurnal expression patterns of Transcript-1 under LD and SD conditions in 58N and 58S. All the black bars indicate the dark period, and the white bars indicate the light period. The numbers on the Top of the bars indicate hours of the day. Data are means ± SEM (n = 4). (D) DNA methylation levels (%) of the 1.5-kb genome region, including 450-bp putative promoter and the entire LDMAR region (1,050 bp), from young panicles of 58N and 58S grown under LD conditions. (E) DNA methylation levels (%) of the 500-bp genome region (~200 bp to +300 bp of LDMAR) from leaves of plants grown under LD conditions and young panicles of plants under SD conditions, for 58N and 58S.
methylation in the 1.5-kb region of this transcript, including the 450-bp putative promoter and the 1,050-bp transcribed region, using DNA from young panicles at pistil/stamen primordium differentiation stage of plants grown under natural long-day conditions. There was a contrasting pattern of DNA methylation between the putative promoter and transcribed region of LDMAR (Fig. 2D). In the promoter region, the methylation level of CG contexts in 58S was much higher than that in 58N, whereas in the transcribed region the methylation level of CG contexts in 58S was much higher than in 58N. Similar difference in methylation patterns between 58S and 58N was observed using leaf tissues of plants grown under long days and panicle tissues of plants under short days (Fig. 2E). Thus, the DNA methylation status was not tissue or day-length specific.

### Table 1. Fertility of plants overexpressing various constructs of Transcript-1 (LDMAR)

<table>
<thead>
<tr>
<th>Transgene</th>
<th>Genotype</th>
<th>No. of plants</th>
<th>Pollen fertility, %</th>
<th>Spikelet fertility, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actin:NT1 (T₀)</td>
<td>Positive</td>
<td>19</td>
<td>48.7 ± 6.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>7</td>
<td>2.9 ± 1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.0000</td>
</tr>
<tr>
<td>Actin:ST1 (T₀)</td>
<td>Positive</td>
<td>14</td>
<td>46.6 ± 5.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>5</td>
<td>3.2 ± 1.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>7.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.0000</td>
</tr>
<tr>
<td>Actin:NT1 (T₁) from three families</td>
<td>Positive</td>
<td>58</td>
<td>83.7 ± 1.03</td>
<td>88.0 ± 1.00</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>26</td>
<td>5.1 ± 0.74</td>
<td>5.6 ± 0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>61.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.0000</td>
</tr>
<tr>
<td>Actin:ST1 (T₁) from three families</td>
<td>Positive</td>
<td>58</td>
<td>80.8 ± 1.08</td>
<td>79.3 ± 1.85</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>24</td>
<td>4.3 ± 0.99</td>
<td>4.7 ± 0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>52.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>35.96</td>
</tr>
<tr>
<td>Actin:NT1-mORF (T₀)</td>
<td>Positive</td>
<td>14</td>
<td>43.6 ± 8.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>8</td>
<td>2.9 ± 1.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.0007</td>
</tr>
<tr>
<td>Actin:NT1-ORF (T₀)</td>
<td>Positive</td>
<td>17</td>
<td>7.4 ± 1.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>1</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.2300</td>
</tr>
<tr>
<td>Actin:NT1-ΔSNP (T₀)</td>
<td>Positive</td>
<td>15</td>
<td>7.6 ± 2.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>2</td>
<td>0.4 ± 0.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.1891</td>
</tr>
<tr>
<td>Actin:NT1-Δ110bp (T₀)</td>
<td>Positive</td>
<td>17</td>
<td>40.7 ± 7.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>4</td>
<td>4.8 ± 4.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P</td>
<td>0.0177</td>
</tr>
</tbody>
</table>

All data are given as mean ± SEM. P values were obtained by t tests between positive and negative plants.

### Fig. 3. The effect of overexpressing Transcript-1 (LDMAR) on male fertility. (A) Expression level of Transcript-1 (Top) in transgene-positive and -negative plants (Bottom) in a T₁ family. Data are means ± SEM (n = 3). (B) Fertility scores of transgene-positive and -negative plants in three T₁ families. There are 58 positive and 26 negative segregants from three T₁ families. Data are means ± SEM. (C) Whole plants of transgene-negative (Left) and -positive (Right) segregants from T₁ family. (D) Anther of transgene-negative (Left) and -positive (Right) plants. (Scale bars, 0.5 mm.) (E) Panicles of transgene-negative (Left, mostly sterile spikelets) and -positive (Right, mostly fertile) plants. (F) Pollen grains of transgene-negative (Top, sterile) and -positive (Bottom, fertile) plants. (Scale bars, 50 μm.)
Recent reports showed that DNA methylation is always correlated with small RNA (46–48). We visualized the RNA secondary structure of LDMAR to find a possible stem-loop structure, which is usually involved in genes encoding micro-RNAs (http://www.tbi.univie.ac.at/RNA/man/RNAfold.html). A 145-base fragment including the SNP site was predicted to form a stem loop, in which the SNP site was located in the stem region and the G–C change of the SNP influenced the stem-loop structure (SI Appendix, Figs. S2B and S5A). Several small RNAs were also predicted, which may be encoded by this region (SI Appendix, Fig. S5B). We tried to verify these predicted small RNAs. Three small RNAs could be detected by a stem-loop qRT-PCR method but not by small RNA gel blot analysis (SI Appendix, Fig. S5D). The average Ct values of the three small RNAs studied were larger than 32 (SI Appendix, Fig. S5D), whereas Ct values of most known miRNAs are less than 25 (49, 50), which might explain why small RNA gel blot analysis could not detect the signal. These results suggest a likely involvement of small RNAs in regulation of PSMS.

Abnormal Degeneration of Tapetum in Anthers of 58S. We examined cross-sections of anthers from 58N and 58S plants grown under natural long-day conditions. No obvious differences were detected between 58N and 58S before sporogenous cell stage. Visible abnormalities in anther development of 58S were observed soon after the formation of sporogenous cells and thereafter, featured by condensed, deformed, and vacuolated cytoplasm of tapetal cells and microspore mother cells and deformed, crumpled, and empty microspores, as well as delayed degeneration of tapetal cells (SI Appendix, Fig. S6 A–L). A terminal deoxynucleotidyl transferase-mediated dUTP nick-end labeling (TUNEL) assay of anthers at different developmental stages showed that programmed cell death (PCD) occurred earlier in 58S than in 58N (SI Appendix, Fig. S7 A–H).

Discussion

On the basis of the results obtained in this study, we proposed a model to explain the genetic and molecular basis of the PSMS, centered on the role of LDMAR (SI Appendix, Fig. S8). A sufficient amount of LDMAR is required for male fertility under long-day conditions. The spontaneous G–C mutation causing a SNP between 58N and 58S altered the RNA secondary structure of this region. This alteration eventually brought about heritable increased methylation in the promoter region of LDMAR, thus reducing the level of transcription specifically under long-day conditions. Such a reduced level of LDMAR results in premature PCD in anther development under long days, thus causing male sterility. However, several questions remain to be answered for complete understanding of PSMS including: (i) the biochemical mechanism by which LDMAR regulates pollen development; (ii) the reason why LDMAR is needed in large quantity in long-days but not in short-days; and (iii) the relation of pms3 to pms1, the other locus for PSMS identified in genetic analysis of the same cross (42–44, 51), in regulating PSMS.

Elucidation of the genetic and molecular bases of PSMS, although still far from completion at this stage, provides important implications for the studies of other photoperiod-regulated processes. It indicates that each of the processes may have a distinct genetic and molecular control at least at the lower level of regulatory hierarchy of photoperiodism. Therefore, the processes have to be investigated individually to understand them. We also speculate that at higher levels, such as the perception of day length, time keeping, and photoreceptors, among other things, these processes should be subjected to the same regulatory machineries as photoperiodism. Thus, the example provided by this study may expand the norm of studies on photoperiodism in plants.

Analyses of transcriptomes have revealed surprising complexity of IncRNAs that are often overlapping with, or interspersed in multiple coding and noncoding transcripts (52, 53). They are classified into different types on the basis of their genomic proximity to protein-coding genes, including overlapping, cis-antisense, bidirectional, and intronic ncRNAs (54). In this study, we detected three transcripts (LDMAR, Transcript-2 and Transcript-3) that were transcribed from the same locus: all of them are likely IncRNAs as none of them was predicted to encode a protein. LDMAR and Transcript-2 were transcribed from the same strand and overlapped with each other. Transcript-3 was transcribed from the complementary strand of LDMAR, thus they are mutually cis-antisense. We also showed that all of them had tissue-specific expression patterns, implying that they are involved in specific developmental processes. This was in agreement with recent findings that the expression of many IncRNAs is restricted to particular developmental contexts (54). Besides associating LDMAR with PSMS, the IncRNAs identified here provide additional opportunities to reveal the roles of IncRNAs in plants.

Two-line hybrids enabled by PSMS rice represent a significant advance in rice breeding. The mechanistic understanding of PSMS and the DNA polymorphism identified in the pms3 region may greatly facilitate the development of new PSMS lines in rice breeding programs.

Methods

Plant Materials and Field Growth Condition. 58N and 1514 are japonica rice varieties. 58S is a spontaneous mutant from 58N and DH80 is a doubled haploid line derived from a cross between 58S and 1514. The field growth conditions of rice plants are described in SI Appendix, SI Methods.

Generation of Constructs and Transformation. Genomic fragments were cut from the BAC clone 117HS of 58N and cloned into the vector pCAMBIA1301 (SI Appendix, Fig. S2A). To generate overexpression constructs of LDMAR (SI Appendix, Fig. S2A), the full-length transcribed sequence was amplified from the panicles of 58N and 58S by RT-PCR using primers OSNP1 and OSNP2 (SI Appendix, Table S4). The PCR products were digested with BamHI and PstI and inserted into pcCAMBIA1301A under the control of the rice α-actin1 promoter (55). Other overexpression constructs were generated by similar strategy. All of these constructs were individually transformed into the Agrobacterium strain EHA105. Transformation was conducted according to a published protocol (56).

RNA Quantification. The diurnal expression patterns of the transcripts were carried out using strand-specific RT-PCR, in which 1 μg total RNA was reverse transcribed with SuperScript III reverse transcriptase (Invitrogen) with the mixture of gene-specific primers for sense: SNP-F5, P-R1, and UBQR1; for antisense: pms3-F9 and UBQR1). We used primers SNP-F5 and SNP-R1 for amplifying the transcript of Transcript-1, P-R1 and pms3-F8 for Transcript-2, and pms3-F9 and pms3-R9 for Transcript-3 (SI Appendix, Table S4). Quantitative RT-PCR was performed on an Applied Biosystems 7500 Real-Time PCR system according to the manufacturer's instructions. The measurements were obtained using the relative quantification method (57). The entire life cycle expression profile of these transcripts was carried out using semi-quantitative RT-PCR. PCR products were separated in 2% (wt/vol) agarose gels and normalized against the rice ubiquitin (UBQ, LOC_Os03g13170) gene.

Bisulfite Sequencing. Genomic DNA from young panicles and leaves was bisulfite treated using the EpiTect Bisulfite kit (Qiagen). The candidate regions were amplified using bisulfite primers (SI Appendix, Table S4) designed using MethPrimer software and cloned into the pGEM-T vector. Each fragment was sequenced in at least 30 clones. Sequence analysis was performed by the Web-based tool, Kismeth.

More methods and data collection are provided in SI Appendix, SI Methods. DNA sequences of the two alleles of LDMAR have been deposited in GenBank under accession nos. JQ317784 (Nongken 58) and JQ317785 (Nongken 58).

Acknowledgments. We thank D. Zhou for his suggestions and critically reading the manuscript. This work was supported by B63 Project Grant 2012AA100103 and National Natural Science Foundation of China Grant 3092109.
4. Tsuji H, Taoka K, Shimamoto K (2011) Regulation of
7. Fisher JE (1972) The transformation of stamens to ovaries and of ovaries to in-751
8. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive re-33
11. Fisher JE (1972) The transformation of stamens to ovaries and of ovaries to in-751
12. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive re-33
14. Knox RB, Heslop-Harrison J (1966) Control of pollen fertility through the agency of
25. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive re-33
28. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive re-33
30. Shi M (1985) The discovery and preliminary studies of the photoperiod-sensitive re-33