



The Cds.71 on *TMS5* May Act as a Mutation Hotspot to Originate a TGMS Trait in Indica Rice Cultivars

Yanning Tan^{1,2*}, Xuewu Sun^{1,2}, Baohua Fang³, Xiabing Sheng⁴, Zheli Li⁵, Zhizhong Sun^{1,2}, Dong Yu^{1,2}, Hai Liu^{1,2}, Ling Liu⁶, Meijuan Duan^{5*} and Dingyang Yuan^{1,2*}

¹ State Key Laboratory of Hybrid Rice, Hunan Hybrid Rice Research Center, Changsha, China, ² State Key Laboratory of Hybrid Rice, Hunan Academy of Agricultural Sciences, Changsha, China, ³ Hunan Rice Research Institute, Hunan Academy of Agricultural Sciences, Changsha, China, ⁴ College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, China, ⁵ College of Agronomy, Hunan Agricultural University, Changsha, China, ⁶ Long Ping Branch, Graduate School of Hunan University, Changsha, China

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Suri Jiao,
BASF, United States
Linglong Liu,
Nanjing Agricultural University, China
Muhammad Qasim Shahid,
South China Agricultural University,
China

*Correspondence:

Yanning Tan
tyncreator@126.com
Meijuan Duan
duanmeijuan@163.com
Dingyang Yuan
yuandingyang@hrrc.ac.cn

Specialty section:

This article was submitted
to Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 05 March 2020

Accepted: 22 July 2020

Published: 07 August 2020

Citation:

Tan Y, Sun X, Fang B, Sheng X, Li Z,
Sun Z, Yu D, Liu H, Liu L, Duan M and
Yuan D (2020) The Cds.71
on *TMS5* May Act as a Mutation
Hotspot to Originate a TGMS Trait in
Indica Rice Cultivars.
Front. Plant Sci. 11:1189.
doi: 10.3389/fpls.2020.01189

The gene *tms5*, which controls thermo-sensitive genic male sterility (TGMS), has been widely used in two-line hybrid rice breeding in China. The *tms5* lines have two sources, namely, AnongS-1 (AnS) and Zhu1S (ZhS) and, interestingly, are commonly subject to an alteration at *cds.71*. However, whether *cds.71* acts as a mutation hotspot is unknown. Herein, another *tms5* mutant named T98S (induced from T98B by irradiation) was used to explore this. First, the gene of *tms(t)* responsible for T98S was fine-mapped on chromosome 2 based on an F₂ group of T98S/R893. In T98S, the candidate gene *TMS5* (LOC_Os02g12290.1) mutated at *cds.71* with a transversion from cytosine (C) to adenine (A), as also observed in AnS and ZhS. Moreover, the entire coding sequence of *TMS5* from T98B converted T98S from sterile to fertile by *Agrobacterium tumefaciens*-mediated transformation, confirming that T98S is controlled by *tms5*. Next, detection on nearly 40,000 single nucleotide polymorphisms (SNPs) on Rice 56K SNP Array revealed T98S was 99.99% similar to T98B but only 72.84% and 77.47% similar to AnS and ZhS, respectively, demonstrating that T98S originated from T98B rather than from existing *tms5* lines. Furthermore, the *cds.70* was found to exist as a T/G haplotype, and it was T rather than G that helped to induce a TGMS trait. The T frequency was 67.52% in indica rice but decreased to 1.75% in japonica rice in 2,644 cultivars tested, which partly explains why *tms5* mutants were mostly found in indica lines. Our findings provide evidence that *cds.71* may act as a mutation hotspot and clues for breeding TGMS lines in a more efficient way.

Keywords: rice (*Oryza sativa* L.), two-line hybrid rice, thermo-sensitive genic male sterility gene 5 (*tms5*), thermo-sensitive genic male sterility (TGMS), mutation hotspot, T98S, unbalanced allele frequency, indica

INTRODUCTION

Improving rice yields has become increasingly urgent due to human population growth, particularly in Asia and Africa. Rice is a crop with strong heterosis of ~ 10% in yield (Li et al., 2013). The heritable trait of male sterility is favored in breeding hybrid rice with strong heterosis and in developing hybrid seeds for commercial use (Yuan, 1966). Hybrid rice bred *via* a three-line system is

based on cytoplasmic-nuclear male sterile lines (CMS lines) in China before the 1990s (Yuan, 1997). The male sterility of a CMS line is retained by a maintainer line and recovered by a restorer line (Yuan, 1997), thereby increasing the labor input in the production of hybrid rice seeds.

To simplify the hybridization procedure, a two-line system was successfully established in the late 1980s that utilizes either photo- or temperature-sensitive genic male sterile lines (P/TGMS lines) (Shi, 1985; Yang et al., 1993; Deng et al., 1999). The P/TGMS lines, because of their control *via* recessive nuclear gene(s), can be recovered by almost any rice cultivar. This greatly improves the hybrid rice breeding efficiency (Huang et al., 2014). In particular, the pollen fertility in P/TGMS lines can be shifted with environmental conditions. This allows them to be pollinated by restoring lines to produce hybrid seeds when sterile and to be self-pollinated for reproduction when fertile. The P/TGMS lines are typically divided into two basic ecotypes, PGMS and TGMS (Yuan, 1990). The PGMS lines in application are derived from a japonica rice cultivar Nongken58S and convert sterile to fertile under a short day-length and low temperature (Shi, 1985; Yuan, 1997). It was found that Nongken58S is determined by several genes, and the dominant one is *pms3* (also named as *p/tms12-1* in indica rice cultivar Peiai 64S) on chromosome 12. The *pms3* gene encodes a long noncoding RNA that may be involved in epigenetic modifications to influence pollen development (Ding J. et al., 2012; Zhou et al., 2012). Such a complicated manipulation on pollen fertility has been greatly improved in TGMS lines. The TGMS lines are restricted solely by temperature and fail to produce pollen grains under high temperature (Chen et al., 1994; Yang et al., 2000; Lee et al., 2005), thus rendering them useful in hybrid production from a safety perspective. Since 2012, the cultivation areas and numbers of TGMS line-based hybrid combinations have exceeded 95% in two-line hybrids, contributing greatly towards food security in China (Zhang et al., 2015).

A practical TGMS line is derived either from Annong S-1 (AnS) or Zhu-1S (ZhS), which are well-known early-season indica rice cultivars in south China. As reported, AnS and ZhS originated independently; AnS was a spontaneous mutant discovered in Annong, and Annong was obtained from a population of a triple cross of Chao40B/H285//6209-3 (Deng et al., 1999). The line ZhS emerged from a F₂ male sterile plant of Ke-fu-2/Xiang-zao-xian 3//02428 (Yang et al., 2000). Interestingly, both AnS and ZhS were found to be controlled by *tms5* (*thermo-sensitive genic male sterile 5*, *tms5*) on chromosome 2 (chr.2) and are commonly subjected to a point altered at cds.71 (Wang et al., 2003; Yang et al., 2007; Sheng et al., 2013; Zhou et al., 2014). In normal fertile rice cultivars, *TMS5* encodes a functional nuclear ribonuclease Z (RNase Z) to process *Ubl40* mRNA into pieces, which is presumed to be a prerequisite for pollen development (Zhou et al., 2014). In contrast, in AnS and ZhS, the change from C to A at cds.71 terminates transcription in advance and accordingly produces a loss-of-function RNase Z to create a TGMS trait (Zhou et al., 2014).

Gene mutation is a common phenomenon in organisms, and the sites on a gene are not equally mutable. The points that mutate at a higher frequency are often called “mutation hotspots” (Coulondre et al., 1978) and are valuable resources to explore the mechanisms underlying gene mutation. Based on the same mechanism underlying both AnS and ZhS, we hypothesize that the cds.71 on *TMS5* act as a mutation hotspot to induce a TGMS trait. To support this, additional *tms5* mutants from diverse genetic backgrounds were screened. We identified a TGMS mutant named T98S that is a ⁶⁰Co- γ irradiation-induced mutant at M₂ from T98B (an elite indica maintainer for three-line hybrid rice in China) (Tan et al., 2018). Observations on morphological specificity revealed that T98S was very similar to T98B in plant shape, yield factors, and flowering habits. Furthermore, T98S was found to be controlled by a recessive nuclear gene, and it presented sterile at 28.0°C and fertile at 22.0°C regardless of the day-length being adjusted to 12.0 h or 13.5 h in the chamber (Tan et al., 2018). Although T98S is quite similar to the TGMS lines AnS and ZhuS in fertility, whether it is controlled by *tms5* remains unknown. Here, using a series of genetic analyses, we confirm that T98S is another *tms5* line with a possible hotspot at cds.71 and explore further why the *tms5* lines were mostly found in indica rice cultivars.

MATERIALS AND METHODS

Materials

The line T98B is an indica maintainer of wild-type CMS line T98A (Deng et al., 2001). The line T98S is a TGMS mutant identified from T98B through ⁶⁰Co- γ irradiation (Tan et al., 2018). The two indica TGMS lines, AnnongS-1 (ANS) and Zhu1S (ZhS), were reportedly bred from different sources but are commonly controlled by *tms5* (Zhou et al., 2014). These materials were all provided by the Hunan Hybrid Rice Research Center, China.

Mapping the Target Gene for T98S

The bulked-segregant analysis (BSA) method was used to rough map the target gene *tms(t)* responsible for T98S (Michelmore et al., 1991). An F₂ population from T98S crossed with R893 (an indica rice with normal pollen fertility) was established. Twenty-five F₂ sterile/fertile individuals were separately collected to extract DNA using CTAB buffer (McCouch et al., 1988) and establish two bulked DNA samples. Then, 125 polymorphic simple sequence repeat (SSR) markers covering 12 rice chromosomes (<http://www.Gramene.org>) were used to test the genotype difference between TGMS-plants DNA pool and T98S by polyacrylamide gel electrophoresis. Finding associated markers which showed similar genotype between the two samples and determining chromosomal location of *tms(t)*. Later, *tms(t)* was fine mapped following the method of linkage analysis (Qi et al., 2014). In the rough mapping region, developing plenty of polymorphic insert-deletion (InDel) markers and sequence-tagged site markers provided by the Rice SNP-Seek Database (<https://snp-seek.irri.org/>). Subsequently, 1334

F₂ TGMS individuals were used to identify more closely linked markers and fine map *tms(t)* based on the number of the recombinants. Finally, the candidate of *tms(t)* in the mapping region was deduced based on the database of the Rice Genome Annotation Project (RGAP, <http://rice.plantbiology.msu.edu/index.shtml>). Primer information is listed in **Supplemental Table S1**.

Validation of Whether *tms5* Is Responsible for T98S

The full coding sequence (CDS) of *TMS5* (909 bp) cloned from T98B was inserted in the vector pTMS5, which was driven by the maize *ubiquitin* promoter and terminated at the terminator of nopaline synthase gene (*NOS*) from *Agrobacterium tumefaciens*. This construct was then introduced into the callus of T98S induced from mature seeds by *Agrobacterium tumefaciens*-mediated transformation (Holsters et al., 1978). The transgenic plants were detected *via* PCR with specific primers ubi-F/*tms*-R (listed in **Supplemental Table S1**). The tested materials including homozygous T₁ lines, T98S (negative control) and T98B (positive control), were transplanted (with 30 individuals for each line) at one batch in the paddy field at Changsha, China, on May 25th, 2018. The young seedlings undergoing the fertility sensitive phases from IV to VI of young panicle differentiation (July 5th–12th) were treated to heat stress under a daily mean temperature of 29°C. Five panicles of different individuals from each line from July 11th–19th were measured for the ratio of stained ring pollen grains (fertile pollen grains) treated with 3.0% I₂-KI under an optical microscope stain. The values (mean ± sd) were analyzed using one-way analysis of variance (ANOVA) at 0.05 significance level.

Detection of the Genetic Background for T98S

DNA was extracted from the leaves of T98B and T98S following to the CTAB method. The DNA quality was assessed for quantity (≥1.5 µg), concentration (≥30 ng/µl) and purity (ratio of 260 nm/280 nm = 1.8–2.0) with NanoDrop. And then the resultant DNA was visualized on agarose gel for integrity control.

DNA genotyping was performed on Rice 56K SNP Array (Huazhi Biotech Com. LTD). Initially, the DNA was amplified and fragmented. Prior to DNA hybridization, the fragments were purified followed by chip hybridization, scanning and subsequently data quality control. The clean genotyping data (call rate ≥97%) with high resolution and polymorphism was further used to calculate the homozygosity and genetic difference between samples. The loci with identical genotypes for three replicates between two samples are assigned to “identical loci” (IL), while the loci with different genotypes between them are assigned to “polymorphic and homozygous loci” (PHL-1) or “polymorphic and heterozygous loci” (PHL-2). The genetic similarity and genetic difference rate between two samples are calculated with the formulas: genetic similarity = (the number of IL)/(total number of loci) *100%; genetic difference rate = (the number of PHL-1 and PHL-2)/(total number of loci) *100%.

Accordingly, the variation between two pairs of samples (“T98S” and “AnS”; “T98S” and “ZhS”) were respectively evaluated.

In addition, the genetic background of T98S was compared with AnS and ZhS by the same above procedure. The background difference rate equals the ratio of polymorphous SNP numbers to total SNPs. In particular, a region of about 1.0 Mb around *TMS5* (5.39–7.39 M) was studied to identify specific SNPs for T98S.

Allele Frequency Analysis of Cds.70 Among Rice Cultivars

To explain why the original *tms5* mutants tends to be discovered in indica rice rather than other groups, the genotype of *cds.70–72* (from 6,397,411 to 6,397,413 bp on chr.2) was analyzed among rice cultivars obtained from the Rice SNP-Seek Database (<https://snp-seek.irri.org/>) (Wang et al., 2018). The allele frequencies of *cds.70* (T) and *cds.70* (G) were investigated for indica cultivars (including subpopulations of *indx*, *ind1A*, *ind1B*, *ind2*, and *ind3*) and japonica cultivars (including subpopulations of *jap*, *temp*, *subtrop*, and *trop*). Furthermore, the changes in the *TMS5* domains were assessed on whether the genotype of *cds.70–71* was altered from GC to GA, or from TC to TA by CD-Search on the NCBI platform (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

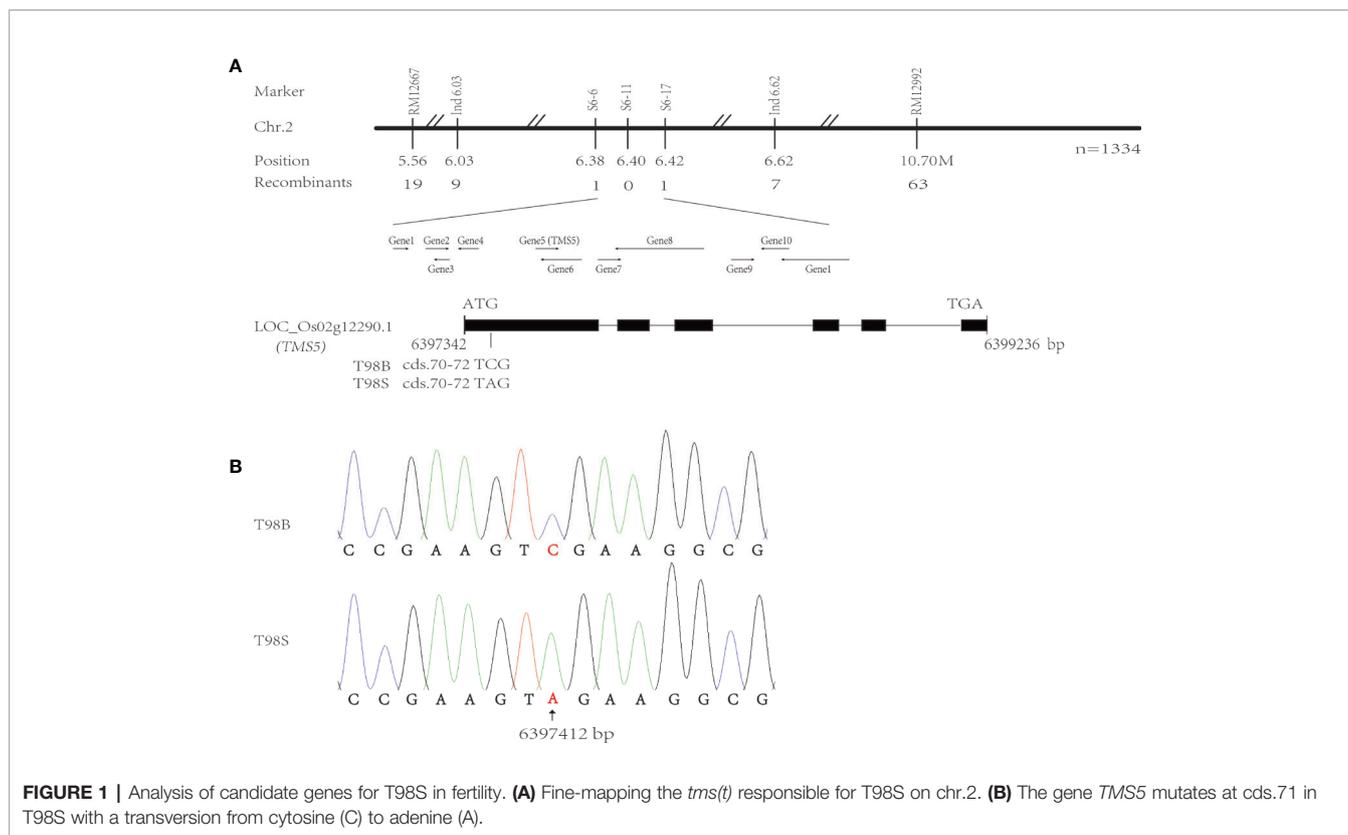
RESULTS

TMS5 Mutated at Cds.71 in T98S

Following the BSA strategy, two F₂ bulked DNA samples from T98S/R996 were used to rough map the *tms(t)* responsible for T98S on chr.2 with the help of the SSR markers RM11267 and RM12992 (**Figure 1A**). We sampled 1,334 F₂ TGMS individuals to fine-map *tms(t)* within 44 Kb between sequence tag site (STS) markers *s6-6* and *s6-17* corresponding to positions 6,387,223–6,429,700 in the referenced Nipponbare (NPB) genome (**Table S1**). This region carried 11 genes, and incidentally, the gene *TMS5* (LOC_Os02g12290.1, labeled as Gene5) controlling the TGMS trait was located on 6,397,342–6,399,236 (**Figure 1A**). The *TMS5* in NPB genome has six exons and five introns and a 909 bp CDS encoding a product of nuclear ribonuclease Z. Through sequencing *TMS5*, we found that only *cds.71* at 6,397,412 mutated with a transversion from cytosine (C) to adenine (A) in T98S (**Figure 1B**). Such an alteration may enable the production of a stop codon (TAG) at *cds.70–72*, terminating transcription. To our surprise, the point of *cds.71* was where the original *tms5* mutants AnS and ZhS had mutated (Zhou et al., 2014), suggesting that *TMS5* is the candidate gene responsible for T98S.

tms5 Is Responsible for T98S in Fertility

A complementary experiment was performed to validate whether *tms5* is responsible for T98S. By *A. tumefaciens*-mediated transformation, a pTMS5 construct overexpressing

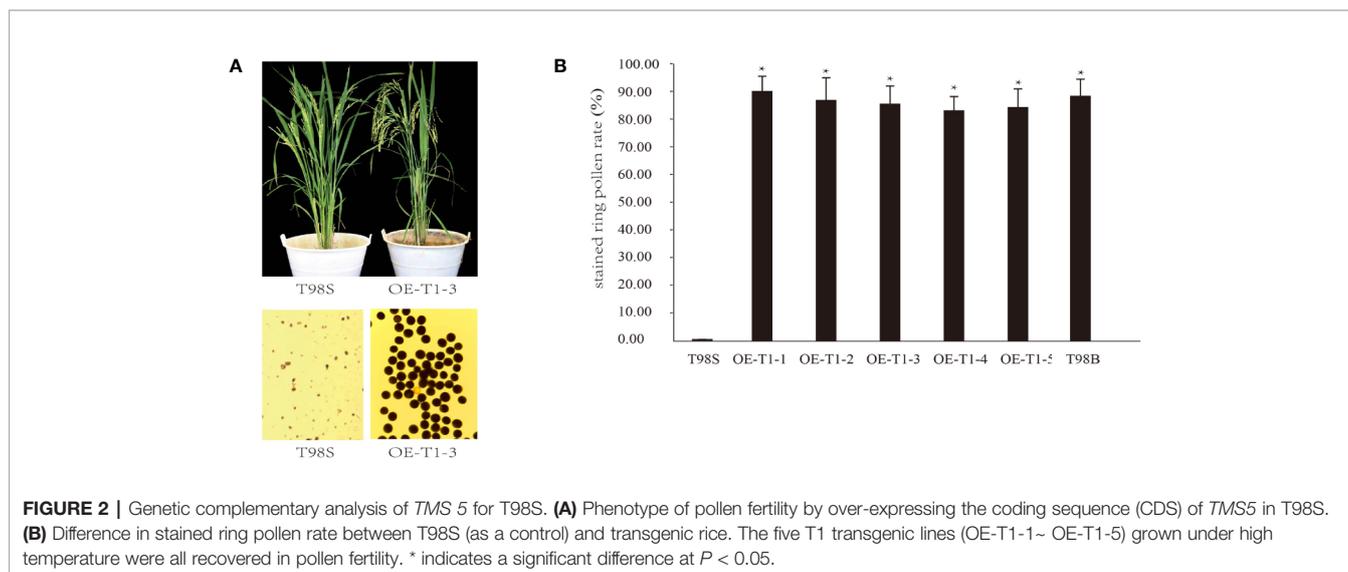


full-length CDS of *TMS5* (cloned from T98B) was introduced into T98S. The 5 T₁ lines including OE-T1-1~ OE-T1-5 were detected positive by PCR (**Figure S1**). They were grown in the paddy field and exposed to a high temperature of ~29°C during a sensitive fertility period. During this period, T98S at heading remained completely sterile with no stained ring pollen grains. However, all transgenic lines that recovered their pollen fertility had an average stained ring pollen rate of 86.53%± 6.32%, similar

to the ones of T98B (89.37% ± 5.27%) (**Figures 2A, B**). These results support that T98S was actually a *tms5* line.

T98S Originated Directly From T98B

We were interested in whether the mutation of T98S at cds.71 on *tms5* originated directly from T98B or was transferred by another *tms5* line. To explore this, the genetic difference between T98S and T98B was assessed on Rice 56K SNP Array using



Nipponbare as a reference. The two samples were tested with three replicates using 39,193 SNPs, achieving a repetition rate of 99.93% for T98S and 99.60% for T98B. The homozygosity rate was found to be 99.15% (38,832/39,165) in T98S and 99.14% (38,701/39,036) in T98B (**Table 1**). Later, a total of 39,015 qualified SNPs covering the whole genome (with a density of ~105 per 1 Mb) were selected to evaluate their similarity. It was estimated that 39,013 SNPs (accounting for 99.99% of total SNPs) showed no polymorphism between T98S and T98B. The remaining two SNPs differed in being homozygous on 6,211,844 chr.2 and heterozygous on 25,891,109 chr.8 (**Figure 3**). This suggested that T98S was a true sibling of T98B.

T98S Can Be Distinguished From AnS and ZhS

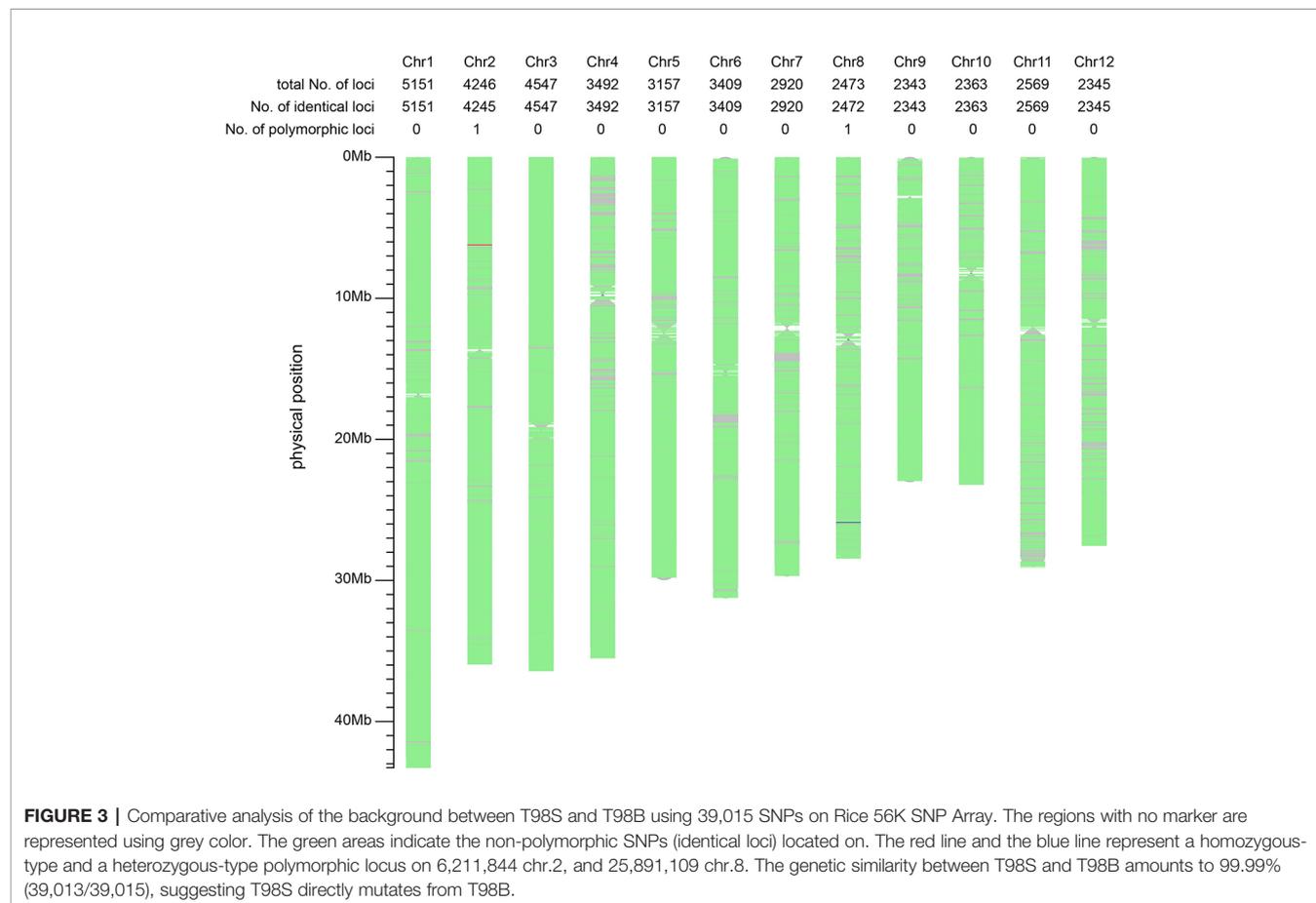
In addition, we compared T98S with two original *tms5* lines, namely AnS and ZhuS, using 39,165 SNPs. The difference rate was 27.16% (10,638/39,165) between T98S and AnS, and 22.53%

(8,825/39,165) between T98S and ZhS (**Figure 4, Table 2**). The polymorphous SNPs covered 12 chromosomes. T98S differed greatly from AnS on chr.12 (36.45%), chr.11 (32.27%), and chr.6 (30.33%), and on chr.11 (39.25%), chr.9 (37.10%), and chr.12 (34.97%) for ZhS. On chr.2, where *TMS5* is located, T98S could distinguish AnS and ZhS at a difference rate of 25.80% and 34.44%, respectively (**Figure 4, Table 2**). These data indicate that T98S has a different genetic background from AnS and ZhuS.

We focused on the area surrounding *tms5* to identify the specific SNPs for T98S. The number of SNPs differing between T98S and AnS (or ZhS) was one in the upstream area (5.39–6.39 M) and 15 in the downstream area (6.39–7.39 M), and they were concentrically distributed on the region from 7.03 M to 7.38 M (**Figure 4, Table 3**). Although AnS and ZhS could be characterized by 13 SNPs located from 5.79 Mb to 6.84 Mb through genome resequencing (Yan et al., 2019), there were no polymorphous SNPs to distinguish them through microarray analysis.

TABLE 1 | Analysis of repetition rate and homozygous degree for T98B and T98S.

Line	TotalSNP	Repeatable SNPs	UnrepeatableSNPs	Repetition rate	Homozygous SNPs	HeterozygousSNPs	Homozygousdegree
T98B	39,163	39,036	157	99.60%	38,701	335	99.14%
T98S	39,193	39,165	28	99.93%	38,832	333	99.15%



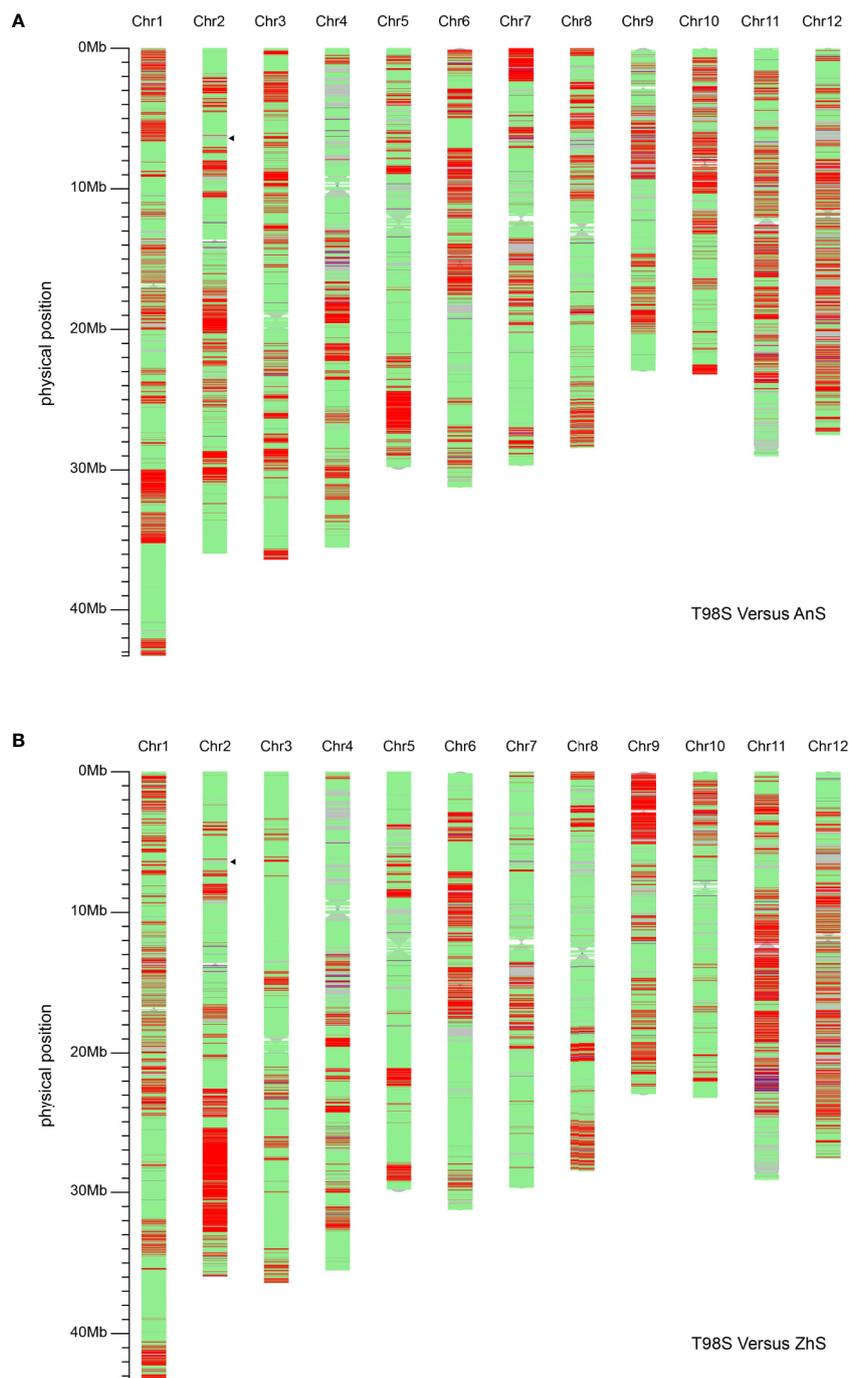


FIGURE 4 | Analysis of the background of T98S compared with AnS **(A)** and ZhuS **(B)** using 39,165 single nucleotide polymorphisms (SNPs) on Rice 56K SNP Array. The green areas indicate the non-polymorphic SNPs located on. The red areas cover homozygous-type polymorphic SNPs, and the blue areas cover heterozygous-type polymorphic SNPs. The position that the black triangle points toward locates *TMS5*.

An Indica-Preferred TGMS Trait May Have Resulted From an Unbalanced Allele Frequency of Cds.70

Rice cultivars in China and Asia are divided into two major *O. sativa* subspecies: indica and japonica. Less is known about why all original

tms5 mutants were found in the indica background. Hence, we focused on comparing the difference in the genotype at *cds.70–72* between the two subspecies in Rice SNP-Seek Database. In this database, 2,644 rice cultivars were defined as indica or japonica. Of these, 1,789 were indica (consisting subpopulations indx, ind1A,

TABLE 2 | Background analysis of T98S compared with AnS and ZhuS.

Chromosome	Total number	T98S versus ANS			T98S versus ZhS		
		Non-polymorphism SNP number	Polymorphous SNP number	Difference rate	Non-polymorphism SNP number	Polymorphous SNP number	Difference rate
1	5,167	3,775	1,392	26.94%	3,888	1,279	24.75%
2	4,260	3,161	1,099	25.80%	2,793	1,467	34.44%
3	4,563	3,386	1,177	25.79%	4,143	420	9.20%
4	3,505	2,680	825	23.54%	2,925	580	16.55%
5	3,170	2,404	766	24.16%	2,714	456	14.38%
6	3,419	2,382	1,037	30.33%	2,609	810	23.69%
7	2,930	2,225	705	24.06%	2,619	311	10.61%
8	2,481	1,857	624	25.15%	2,000	481	19.39%
9	2,353	1,711	642	27.28%	1,480	873	37.10%
10	2,371	1,695	676	28.51%	2,063	308	12.99%
11	2,581	1,748	833	32.27%	1,568	1,013	39.25%
12	2,365	1,503	862	36.45%	1,538	827	34.97%
Sum	39,165	28,527	10,638		30,340	8,825	
Average				27.16%			22.53%

TABLE 3 | The single nucleotide polymorphisms (SNPs) within 1.0 Mbp up- and down-stream around TMS5 among three thermo-sensitive genic male sterility (TGMS) lines referenced as Nipponbare.

SNP ID	Position*	T98S	AnS	ZhS
AX-155464611	6,211,844	C/C	T/T	T/T
AX-165095954	7,032,362	T/T	C/C	C/C
AX-165089120	7,044,610	C/C	T/T	T/T
AX-115808897	7,072,374	A/A	G/G	G/G
AX-154242679	7,119,202	G/G	A/A	A/A
AX-154944191	7,132,042	A/A	C/C	C/C
AX-165092511	7,238,480	A/A	G/G	G/G
AX-115797125	7,248,666	T/T	C/C	C/C
AX-165103200	7,282,914	C/C	T/T	T/T
AX-154410183	7,289,955	C/C	T/T	T/T
AX-155524733	7,299,368	A/A	G/G	G/G
AX-165103287	7,324,667	T/T	C/C	C/C
AX-115755253	7,357,661	T/T	C/C	C/C
AX-165095546	7,359,341	C/C	T/T	T/T
AX-165094219	7,366,395	G/G	A/A	A/A
AX-154929815	7,376,951	G/G	A/A	A/A

*The position is referenced as the Nipponbare genome.

ind1B, ind2, and ind3) and 855 were japonica (consisting subpopulations jap, temp, subtemp, and trop) (Table S2). We found that cds.70 existed as a G/T haplotype, while no variation was detected on cds.71 and cds.72 (Figure 5; Table S2). Conserved domain analysis by CD-search showed that the effects on RNase Z differed between the two alleles of cds.70 combined with cds.71 when it mutates from C to A. It revealed that cds.70 (T) along with cds.71 (A) generated a stop codon (TAG) at cds.70–72 to lose a conserved domain inducing a TGMS trait (Figure 6). In contrast, cds.70 (G) only led to a replacement of alanine (Ala) with glutamate (Glu) on the 24th amino acid residue, and therefore conserved domain Z_MBL-fold (amino acid residues 29–231) of RNase remained unaffected and so did the normal pollen development (Figure 6). Interestingly, in the 2,588 cultivars with a homozygous state at cds.71 (1,365 for GG and 1,223 for TT), we noticed an unbalanced allele frequency between the two subspecies. In indica rice, cds.70 (T) reached a higher mean frequency of 67.52%, with 66.67% for indx, 62.68% for ind1A, 63.90% for ind1B, 70.53% for ind2, and 70.53% for ind3 (Table 4, Figure 5B). On the contrary, the allele frequency decreased dramatically in the japonica cultivars,

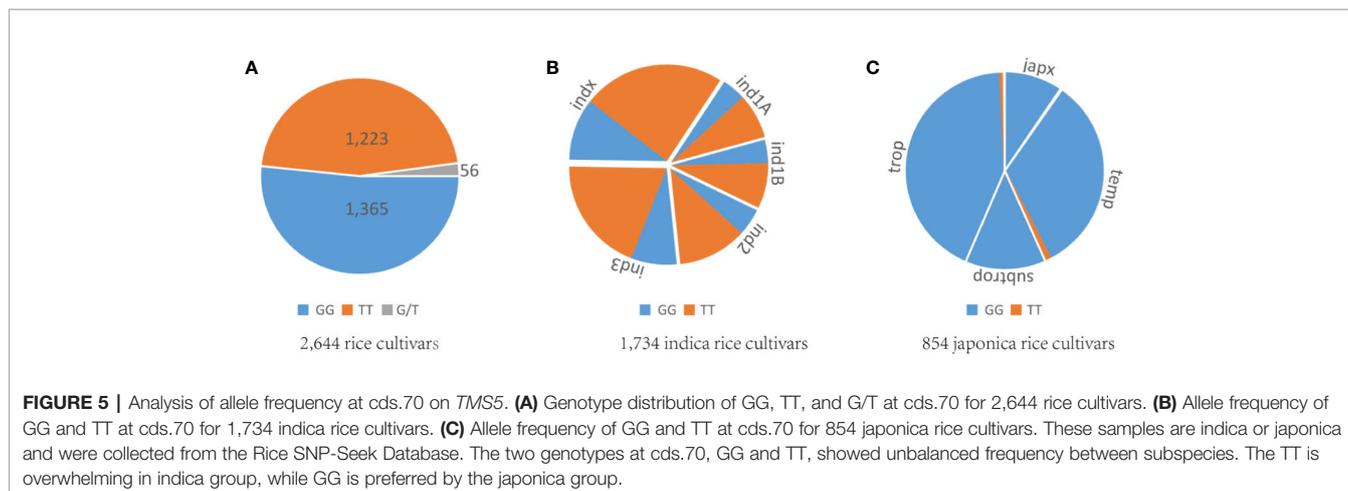
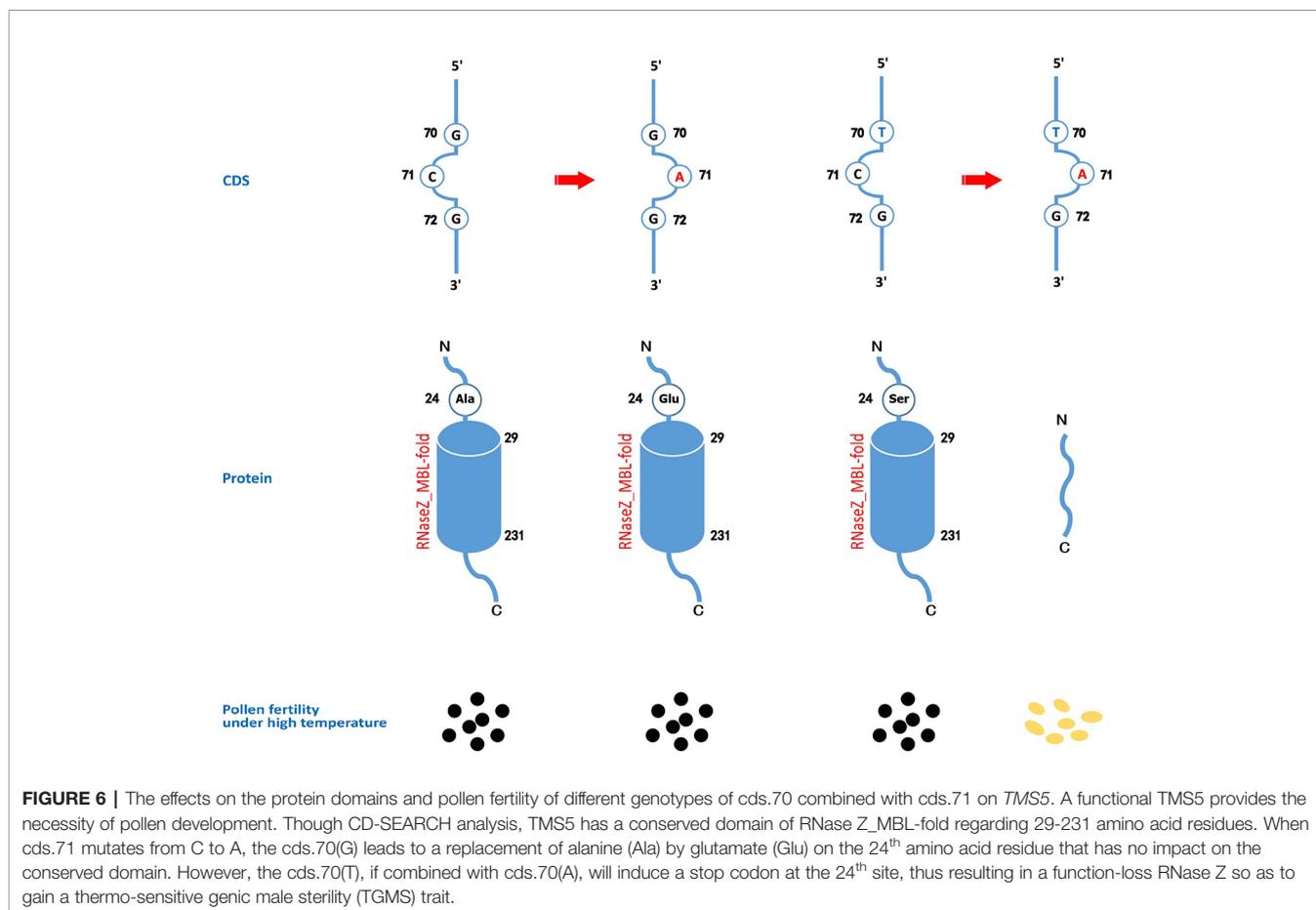


TABLE 4 | Differences in allele frequency of *cds.70* on *TMS5* between indica rice and japonica rice.

Subspecies	Subpopulation	Total number	GG		TT		G/T	
			Number	Frequency	Number	Frequency	Number	Frequency
indica		1,789	526	29.40%	1208	67.52%	55	3.07%
	indx	594	159	26.77%	410	69.02%	25	4.21%
	ind1A	204	68	33.33%	126	61.76%	10	4.90%
	ind1B	196	68	34.69%	122	62.24%	6	3.06%
	ind2	280	74	26.43%	201	71.79%	5	1.79%
	ind3	469	125	26.65%	335	71.43%	9	1.92%
japonica		855	839	98.13%	15	1.75%	1	0.12%
	japx	81	79	97.53%	2	2.47%	0	0.00%
	subtrop	112	112	100.00%	0	0.00%	0	0.00%
	trop	368	363	98.64%	5	1.36%	0	0.00%
	temp	288	279	96.88%	8	2.78%	1	0.35%



with an average level of just 1.75%, with 2.41% for japx, 2.78% for temp, 0.00% for subtrop, and 1.34% for trop (Table 4, Figure 5C). Therefore, the biased allele frequency of *cds.70* would have led to an indica-preferred TGMS trait.

DISCUSSION

In the last three decades, nearly 20 original P/TGMS mutants have been identified (Fan and Zhang, 2018). They display diverse

sensitivity against day-length/temperature. The japonica PGMS mutant Nongken58S can be converted from fertility to sterility when the day-length is longer than 14 h (Zhang et al., 1987), while another japonica mutant *csa* (for carbon starved anther) showed male sterile under short-day conditions with a critical photoperiod of 12.5–13.0 h (Zhang et al., 2012). Furthermore, the Nongken58S derivatives were also noticed to be influenced by temperature in fertility alteration (Zhang et al., 1992), indicating the existence of a compensating mechanism between critical day-length and temperature. The original TGMS mutants were

almost discovered in indica rice. The ones including 5460S (Wang et al., 1995), Hengnong S-1 (Qi et al., 2014), AnS-1 (Chen et al., 1994), ZhS (Sheng et al., 2013), XianS (Peng et al., 2010), and Mian 9S (Huang et al., 2008) became sterile when the temperature was beyond 24°C, while J207S and go543S had completely abortive pollen grains below 31°C (Jia et al., 2001; Liu et al., 2010).

The diverse phenotypes were regulated *via* different genetic mechanisms. The genes *PMS3* and *PMS1T*, two major loci responsible for Nongken58S, both encode long noncoding RNA (lncRNA). The decreased level of *PMS3* together with the increased level of *PMS1* in transcription leads to male sterility under long-day conditions in Nongken58S (Ding J. H. et al., 2012; Fan et al., 2016). A single-locus *TMS5* encoding RNase Z controls the fertility of AnS. In a normally fertile line, *TMS5* itself is insensitive to temperature, but several targets of RNase Z such as ubiquitin-60S ribosomal protein L40 family (*Ub_{L40}*) members *Ub_{L40}1*, *Ub_{L40}2*, and *Ub_{L40}4* dramatically accumulated their mRNA levels in AnS under high temperature.

It is generally accepted that AnS and ZhS are the two original donors for the TGMS lines in practice in China. Through map-based cloning, Zhou et al. (2014) and Sheng et al. (2013) demonstrated that ZhS is another *tms5* mutant allele to AnS, but surprisingly, they are commonly subjected to the critical point of *cds.71* altered from C to A. Because *tms5* gene is very easy for breeders to transfer from one rice cultivar to another, we speculated if AnS and ZhS were derived from a common donor or differed in sources. To explore this, Yan et al. (2019) investigated their genetic differences by genome resequencing and sought 13 polymorphous SNPs around *tms5* to distinguish the AnS-derivatives from the ZhS-derivatives.

From such observations we hypothesized the *cds.71* may act as a mutation hotspot to induce a TGMS trait. The most convincing evidence was that many original *tms5* mutants were found. In fact, several mutants like XianS, Q523S, Q524S, N28S, G421S, and Q527S were also revealed to be controlled by *tms5* (Peng et al., 2006; Peng et al., 2010; Zhou et al., 2014). However, there were no details to verify their true identity, as it is possible for a fake mutant to be unknowingly selected in the paddy due to biological confounding. One example is a well-known TGMS line named “Guangzhan63S”. According to Yang’s description, Guangzhan63S was selected by a crossing between N422S (a Nongken58S derivative) and Guangzhan63 (Yang et al., 2002). However, it was finally proven to be a *tms5* line by molecular analysis (Xu et al., 2011; Zhou et al., 2014). So, in this sense, it was important for us to verify the true identity of the TGMS mutant T98S. Previous works found T98S featured a rapid stability and showed quite similar morphological traits as T98B (Tan et al., 2018). Our work herein demonstrated that T98S is genetically a real sibling of T98B (Figures 3 and 4, Table 2) and another *tms5* mutant with a C-to-A transversion at *cds.71* (Figures 1 and 2). Coincidentally, we also reported another TGMS mutant “BM2-23” (also irradiation-induced from T98B) that was found to be allele to T98S (Tan et al., 2018).

The *cds.71* may tend to mutate and that led us to question why the original *tms5* mutants cannot be found in japonica rice.

The TGMS trait is noticeable and the period for selection lasts for approximately three months, and so it is impossible to go unobserved. This biased disappearance may be related to the effects of genotype of *cds.70-72*. We noticed a T/G haplotype existed at *cds.70* and the allele frequency was opposite between indica rice and japonica rice (Figure 5). Therefore, the T rather than G at *cds.70* resulted in that TGMS trait (Figure 6). The typical Chinese early-season indica cultivars (e.g., XIANGZAOXIAN7, ZAOXIAN14, and ZAOXIAN240) in the “ind1A” subpopulation always have T at *cds.70* in Rice SNP-Seek Database, and it may partly be the reason why AnS, ZhS, and T98S were found in early-season indica rice; whereas in japonica rice, only 1.75% cultivars carry a T at *cds.70* (Figure 5), so it would decrease the probability to originate a *tms5* mutant. Particularly, the subspecies-based preference provides an alternative for developing TGMS lines in the breeding. For example, mutagenesis with chemical or physical treatment would be an effective way to obtain an indica *tms5* line (Tan et al., 2018). For japonica rice, the *tms5* lines are unlikely to be induced *via* mutagenesis, and the possible way will be transferring *via* an indica *tms5* line (Deng, 2005) or gene editing on *TMS5* (Zhou et al., 2016). We also did not find a C-to-A transversion at *cds.71* in fertile rice, and we believe the likely cause is a certain specific locus (like *cds.70(G)*) that inhibits its ability to mutate.

The mechanisms underlying gene mutation are extremely complex. A locus in a gene may be conserved or vary in different ways and at different frequencies. For example, the loci at an intron vary more frequently than those at splicing sites in *Arabidopsis thaliana* (Brown, 1996); G/C bases are commonly found to be mutated in RGYW/WRCY on AID (encoding activation-induced cytosine deaminase) (Muramatsu et al., 1999). In patients with breast-ovarian cancer, the gene *BRCA2* has a high frequency of the “AAGA” deletion at *cds.1310-1313* (Laarabi et al., 2017). To date, few mutation hotspots have been reported in rice. Interestingly, regardless of being spontaneous (like AnS), induced by gamma irradiation (like T98S) or genetically recombined by fertile line crossing (like ZhS), the *TMS5* tended to mutate at *cds.71* with a C-to-A transversion. However, we are not sure whether TCG is more easily mutable in other position. Any gene can be known to potentially have hotspots if it induces a noticeable trait such as the one found here in the paddy. Additionally, local DNA sequences were found to strongly influence mutation frequency. In the *lacI* gene of *Escherichia coli*, several spontaneous deletions were caused by DNA repeats with specific termini of palindromic sequences (Farabaugh et al., 1978), and the hotspots with a C-to-T substitution may result from the spontaneous deamination of 5-methylcytosine (Coulondre et al., 1978). Sometimes, the mutation frequency was found to be closely correlated with the level of transcription mediated by regulatory elements (Yoshikawa et al., 2002). For exploring the mechanisms underlying hotspots in higher plants, the *cds.71* on *TMS5* would be an ideal locus to focus on for future fundamental studies.

CONCLUSION

The *tms5*-based two-line hybrid rice has greatly contributed to increasing the cereal supply and is going to be expanded to a larger scale in China and abroad. However, our understanding of *TMS5* is still limited. In this study, we identified another *tms5* mutant T98S and investigated the behavior of the *cds.70–72* on *TMS5*. By detecting the genetic background and using map-based cloning, we demonstrated that T98S originated independently from T98B, and *cds.71* was found to be the mutation point responsible for T98S. Furthermore, the discovery of an unbalanced allele frequency of *cds.70* partly offers an explanation for an indica-preferred TGMS trait. These findings improve our understanding of *tms5* and provide new approaches for developing *tms5* indica lines.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

YT designed the study, performed all experiments, summarized the results, and wrote the manuscript. DY and MD directed and supervised all experiments. XS, BF, XS, and LL participated in the work of map-based cloning. ZS, DY, and HL detected the genetic background. ZL analyzed the allele frequency. All authors contributed to the article and approved the submitted version.

REFERENCES

- Brown, J. (1996). *Arabidopsis* intron mutations and pre-mRNA splicing. *Plant J.* 10 (5), 771–780. doi: 10.1046/j.1365-3113.1996.10050771.x
- Chen, L. B., Zhou, G. Q., and Huang, Y. X. (1994). Effects of temperature and photoperiod on fertility and physiological activities of rice Annon S-1 and Hengnong S-1. *Acta Bot. Sin.* 36, 119–123.
- Coulondre, C., Miller, J., Farabaugh, P., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia Coli*. *Nat.* 274, 775–780. doi: 10.1038/274775a0
- Deng, H. F., Shu, F. B., and Yuan, D. Y. (1999). An overview of research and utilization of Annon S1. *Hybrid Rice* 14 (3), 1–3.
- Deng, X. L., Yuan, D. Y., and Xu, R. F. (2001). A brief introduction of T98A, a high outcrossing and good grain quality CMS line. *Hybrid Rice* 16 (4), 14–14.
- Deng, Q. Y. (2005). Breeding of the PTGMS line Y58S with wide adaptability in rice. *Hybrid Rice* 20 (2), 15–18.
- Ding, J., Shen, J., Mao, H., Xie, W., Li, X., and Zhang, Q. (2012). RNA-directed DNA methylation is involved in regulating photoperiod-sensitive male sterility in rice. *Mol. Plant* 5, 1210–1216. doi: 10.1093/mp/sss095
- Ding, J. H., Lu, Q., Ouyang, Y. D., Mao, H. L., Zhang, P. B., Yao, J. L., et al. (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2654–2659. doi: 10.1073/pnas.1121374109
- Fan, Y., and Zhang, Q. (2018). Genetic and molecular characterization of photoperiod and thermo-sensitive male sterility in rice. *Plant Reprod.* 31, 1–12. doi: 10.1007/s00497-017-0310-5
- Fan, Y., Yang, J., Mathioni, S., Yu, J., Shen, J., Yang, X., et al. (2016). *PMS1T*, producing phased small-interfering RNAs, regulates photoperiod-sensitive

FUNDING

This study was supported by the Agriculture Science and Technology Innovation Program of HHRRC and HAAS, China (Grant No. YB201907 and 2017XC09), the Natural Science Foundation of Hunan, China (Grant No. 2019JJ40206), the Hunan Science and Technology Talents Support Project (Grant No. 2019TJ-Q08) and the National Key Research and Development Program of China (Grant No. 2016YFD0101101).

ACKNOWLEDGMENTS

We would like to thank Dr. Yongfei Zhang of the Research Institute of Agriculture Biotechnology, Hunan, China for advice on designing this study. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.01189/full#supplementary-material>

FIGURE S1 | The PCR amplification of T1 transgenic lines over-expressing the *cds* of *TMS5*. M, DNA marker; 1, T98S(negative control); 2, pTMS5 (positive control); 3-7, transgenic lines referring to OE-T1-1, OE-T1-2, OE-T1-3, OE-T1-4 and OE-T1-5. The samples were detected by PCR using the specific primer ubtms-F/R.

TABLE S1 | The primers used in this study. *The position is referenced as the Nipponbare genome.

- male sterility in rice. *Proc. Natl. Acad. Sci. U. S. A.* 113, 201619159. doi: 10.1073/pnas.1619159114
- Farabaugh, P. J., Schmeissner, U., Hofer, M., and Miller, J. H. (1978). Genetic studies of the lac repressor: vii. on the molecular nature of spontaneous hotspots in the lacI gene of *Escherichia coli*. *J. Mol. Biol.* 126, 847–857. doi: 10.1016/0022-2836(78)90023-2
- Holsters, M., de Waele, D., Depicker, A., Messens, E., van Montagu, M., and Schell, J. (1978). Transfection and transformation of *Agrobacterium tumefaciens*. *Mol. Gen. Genet.* 163 (2), 181–187. doi: 10.1007/bf00267408
- Huang, T., Wang, Z., Hu, Y., Shi, S., Peng, T., Chu, X., et al. (2008). Genetic analysis and primary mapping of *pms4*, a photoperiod-sensitive genic male sterility gene in rice (*Oryza sativa*). *Rice Sci.* 15, 153–156. doi: 10.1016/S16726308(08)60035-9
- Huang, J. Z., Zhi-Guo, E., Zhang, H. L., and Shu, Q. Y. (2014). Workable male sterility systems for hybrid rice: Genetics, biochemistry, molecular biology, and utilization. *Rice* 7, 13. doi: 10.1186/s12284-014-0013-6
- Jia, J., Zhang, D., Li, C., Qu, X., Wang, S., Chamarer, V., et al. (2001). Molecular mapping of the reverse thermo-sensitive genic male-sterile gene (*rtms1*) in rice. *Theor. Appl. Genet.* 103, 607–612. doi: 10.1007/pl00002916
- Khlimongkhon, S., Chakhonkaen, S., Pitngam, K., Dithab, K., Sangarwut, N., Panyawut, N., et al. (2019). Molecular markers and candidate genes for thermo-sensitive genic male sterile in rice. *Rice Sci.* 26, 147–156. doi: 10.1016/j.rsci.2018.08.006
- Laarabi, F., Ratbi, I., Elalaoui, S., Mezzouar, L., Doubaj, Y., Bouguenouch, L., et al. (2017). High frequency of the recurrent c.1310_1313delAAGABRCA2 mutation in the North-East of Morocco and implication for hereditary breast-ovarian cancer prevention and control. *BMC Res. Notes* 10, 188. doi: 10.1186/s13104-017-2511-2

- Lee, D., Chen, L., and Suh, H. (2005). Genetic characterization and fine mapping of a novel thermo-sensitive genic male-sterile gene *tms6* in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 111, 1271–1277. doi: 10.1007/s00122-005-0044-x
- Li, J., Ni, J., Chen, Q., and Li, X. (2013). Development and prospect of hybrid rice seed industry in China. *Seed* 32, 56–20.
- Liu, X., Li, X., Zhang, X., and Wang, S. (2010). Genetic analysis and mapping of a thermosensitive genic male sterility gene, *tms6(t)*, in rice (*Oryza sativa* L.). *Genome* 53, 119–124. doi: 10.1139/G09-092
- McCouch, S., Kochert, G., Yu, Z., Wang, Z., Khush, G., Coffman, W., et al. (1988). Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76, 815–829. doi: 10.1007/bf00273666
- Michelmore, R., Paran, I., and Kesseli, R. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U.S.A.* 88, 9828–9832. doi: 10.2307/2357899
- Muramatsu, M., Sankaranand, V., Anant, S., Sugai, M., Kinoshita, K., Davidson, N., et al. (1999). Specific expression of activation-induced cytosine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274, 18470–18476. doi: 10.1074/jbc.274.26.18470
- Peng, H. F., Qiu, Z. G., Chen, X. H., Wan, B. H., Zhang, G. Q., and Lu, Y. P. (2006). Pollen fertility and cytological observation of a thermosensitive genic male sterile line of non-pollen type XianS in rice (*Oryza sativa* L.). *Acta Ecol. Sin.* 26, 2322–2327.
- Peng, H. F., Chen, X. H., Lu, Y. P., Peng, Y. F., Wan, B. H., Chen, N. D., et al. (2010). Fine mapping of a gene for non-pollen type thermosensitive genic male sterility in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 120, 1013–1020. doi: 10.1007/s00122-009-1229-5
- Qi, Y., Liu, Q., Zhang, L., Mao, B., Yan, D., Jin, Q., et al. (2014). Fine mapping and candidate gene analysis of the novel thermo-sensitive genic male sterility *tms9-1* gene in rice. *Theor. Appl. Genet.* 127, 1173–1182. doi: 10.1007/s00122-014-2289-8
- Sheng, Z. H., Wei, X. J., Shao, G. N., Chen, M. L., Song, J., Tang, S. Q., et al. (2013). Genetic analysis and fine mapping of *tms9a*, a novel thermo sensitive genic male-sterile gene in rice (*Oryza sativa* L.). *Plant Breed.* 132, 159–164. doi: 10.1111/pbr.12024
- Shi, M. S. (1985). The discovery and preliminary studies of the photoperiod-sensitive recessive male sterile rice (*Oryza sativa* L. subsp. japonica). *Sci. Agric. Sin.* 2, 44–48.
- Tan, Y. N., Sun, X. W., Fang, B. H., Yu, D., Sun, Z. Z., Wang, W. P., et al. (2018). Conversion of a rice CMS maintainer into a photo- or thermo-sensitive genic male sterile line. *Mol. Breeding* 38, 56. doi: 10.1007/s11032-018-0805-2
- Wang, B., Xu, W., Wang, J., Wu, W., Zheng, H., and Yang, Z. (1995). Tagging and mapping the thermo-sensitive genic male-sterile gene in rice (*Oryza sativa* L.) with molecular markers. *Theor. Appl. Genet.* 91, 1111–1114. doi: 10.1007/bf00223928
- Wang, Y. G., Xing, Q. H., Deng, Q. Y., Liang, F. S., Yuan, L. P., Weng, M. L., et al. (2003). Fine mapping of the rice thermo-sensitive genic male-sterile gene *tms5*. *Theor. Appl. Genet.* 107 (5), 917–921. doi: 10.1007/s00122-003-1327-8
- Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557 (7703), 43–49. doi: 10.1038/s41586-018-0063-9
- Xu, J., Wang, B., Wu, Y., Du, P., Wang, J., Wang, M., et al. (2011). Fine mapping and candidate gene analysis of *ptgms2-1*, the photoperiod-thermo-sensitive genic male sterile gene in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 122, 365–372. doi: 10.1007/s00122-010-1452-0
- Yan, W., Zhou, H., He, H., and Deng, X. W. (2019). Analysis of the genetic relationship between Annon S-1 And Zhu 1S based on genome-wide polymorphism. *Hybrid Rice* 34 (1), 57–61. doi: 10.16267/j.cnki.1005-3956.20180411.116
- Yang, R. C., Wang, N. Y., Liang, K. J., Chen, S. H., Chen, Q. H., and Chen, Q. J. (1993). Study on indica thermo-sensitive genic male-sterile rice 5460S. *J. Fu Jian Agric. Coll. (Natural Sci. Edition)* 22, 135–140.
- Yang, Y. Z., Tang, P. A., Yang, W. C., Liu, A. M., Chen, Y. Q., Ling, W. B., et al. (2000). Breeding and utilization of TGMS Line Zhu 1 S in Rice. *Hybrid Rice* 15, 6–9. doi: 10.16267/j.cnki.1005-3956.2000.02.003
- Yang, Z. Y., Zhang, G. L., Zhang, C. H., Chen, J. J., Wang, H. Q., Zhang, J. J., et al. (2002). Breeding of fine quality ptgms line guangzhan 63s in medium indica rice. *Hybrid Rice* 17 (4), 4–6. doi: 10.16267/j.cnki.1005-3956.2002.04.002
- Yang, Q. K., Liang, C. Y., Zhuang, W., Li, J., Deng, H. B., Deng, Q. Y., et al. (2007). Characterization and identification of the candidate gene of rice thermo-sensitive genic male sterile gene *tms5* by mapping. *Planta* 225, 321–330. doi: 10.1007/s00425-006-0353-6
- Yoshikawa, K., Okazaki, I. M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H., et al. (2002). AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* 296, 2033–2036. doi: 10.1126/science.1071556
- Yuan, L. P. (1966). A preliminary report on male sterility in rice, *Oryza sativa* L. *Chin. Sci. Bull.* 11, 185–188.
- Yuan, L. P. (1977). Practice and theory of breeding hybrid rice. *Sci. Agric. Sin.* 10, 27–31.
- Yuan, L. P. (1990). Progress of two-line system hybrid rice breeding. *Sci. Agric. Sin.* 23, 1–6.
- Yuan, L. P. (1997). Current status and developing prospects in two-line hybrid rice research in China. *Res. Agric. Modernization* 18, 1–4.
- Zhang, Z., Yuan, S., and Xu, C. (1987). The influence of photoperiod on the fertility changes of hubei photo-sensitive genic male-sterile rice (HPGMR). *Chin. J. Rice Sci.* 1, 137–143.
- Zhang, Z., Zeng, H., Yuan, S., Zhang, D., Wang, B., and Li, Y. (1992). Studies on the response to photoperiod and temperature of photoperiod sensitive genic male sterile rice (PGMR). *J. Huazhong Agric. Univ.* 11, 1–6.
- Zhang, H., Xu, C., He, Y., Zong, J., Yang, X., Si, H., et al. (2012). Mutation in CSA creates a new photoperiod sensitive genic male sterile line applicable for hybrid rice seed production. *Proc. Natl. Acad. Sci. U.S.A.* 110, 76–81. doi: 10.1073/pnas.1213041110
- Zhang, H. L., Chen, X. Y., Huang, J. Z., Z. G., E., Gong, J. Y., Shu, Q. Y., et al. (2015). Identification and transition analysis of photo-/thermo-sensitive genic male sterile genes in two-line hybrid rice in China. *Sci. Agric. Sin.* 48, 1–9. doi: 10.3864/j.issn.0578-1752
- Zhou, H., Liu, Q. J., Li, J., Jiang, D. G., Zhou, L. Y., Wu, P., 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, 649–660. doi: 10.1038/cr.2012.28
- Zhou, H., Zhou, M., Yang, Y. Z., Li, J., Zhu, L. Y., Jiang, D. G., et al. (2014). RNase ²⁵¹ processes *Ubl40* mRNAs and controls thermosensitive genic male sterility in rice. *Nat. Commun.* 5, 4884–4892. doi: 10.1038/ncomms5884
- Zhou, H., He, M., Li, J., Chen, L., Huang, Z. F., Zheng, S. Y., et al. (2016). Development of commercial thermo-sensitive genic male sterile rice accelerates hybrid rice breeding using the CRISPR/Cas9 mediated TMS5 editing system. *Sci. Rep.* 6:37395. doi: 10.1038/srep37395

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Tan, Sun, Fang, Sheng, Li, Sun, Yu, Liu, Liu, Duan and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.