

RESEARCH ARTICLE

Open Access



# Genome-wide identification and transcriptional characterization of DNA methyltransferases conferring temperature-sensitive male sterility in wheat

Dan Li<sup>1†</sup>, Bian-E Feng<sup>1,2†</sup>, Yong-Jie Liu<sup>1</sup>, Jie Gong<sup>1</sup>, Yi-Miao Tang<sup>1</sup>, Li-Ping Zhang<sup>1</sup>, Bin-Shuang Pang<sup>1</sup>, Ren-Wei Sun<sup>1</sup>, Feng-Ting Zhang<sup>1</sup>, Zhao-Bo Chen<sup>1</sup>, Yong-Bo Wang<sup>1</sup>, Xian-Chao Chen<sup>1</sup>, Ai-Ping Wang<sup>2\*</sup>, Chang-Ping Zhao<sup>1\*</sup> and Shi-Qing Gao<sup>1\*</sup> 

## Abstract

**Background:** DNA methyltransferase (DMT) genes contribute to plant stress responses and development by de novo establishment and subsequent maintenance of DNA methylation during replication. The photoperiod and/or temperature-sensitive genic male sterile (P/TGMS) lines play an important role in hybrid seed production of wheat. However, only a few studies have reported on the effect of DMT genes on temperature-sensitive male sterility of wheat. Although DMT genes have been investigated in some plant species, the identification and analysis of DMT genes in wheat (*Triticum aestivum* L.) based on genome-wide levels have not been reported.

**Results:** In this study, a detailed overview of phylogeny of 52 wheat DMT (TaDMT) genes was presented. Homoeolog retention for TaDMT genes was significantly above the average retention rate for whole-wheat genes, indicating the functional importance of many DMT homoeologs. We found that the strikingly high number of TaDMT genes resulted mainly from the significant expansion of the TaDRM subfamily. Intriguingly, all 5 paralogs belonged to the wheat DRM subfamily, and we speculated that tandem duplications might play a crucial role in the TaDRM subfamily expansion. Through the transcriptional analysis of TaDMT genes in a TGMS line BS366 and its hybrids with the other six fertile lines under sterile and fertile conditions, we concluded that *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* might be involved in male sterility in BS366. Furthermore, a correlation analysis showed that *TaMET1-B1* might negatively regulate the expression of *TaRAFTIN1A*, an important gene for pollen development, so we speculated regarding an epigenetic regulatory mechanism underlying the male sterility of BS366 via the interaction between *TaMET1-B1* and *TaRAFTIN1A*.

(Continued on next page)

\* Correspondence: [apdbn2001@163.com](mailto:apdbn2001@163.com); [cp\\_zhao@vip.sohu.com](mailto:cp_zhao@vip.sohu.com); [gshiq@126.com](mailto:gshiq@126.com)

<sup>†</sup>Dan Li and Bian-E Feng contributed equally to this work.

<sup>2</sup>Shanxi Agricultural University, Taigu 030800, China

<sup>1</sup>The Municipal Key Laboratory of the Molecular Genetics of Hybrid Wheat, Beijing Engineering Research Center for Hybrid Wheat, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** Our findings presented a detailed phylogenetic overview of the DMT genes and could provide novel insights into the effects of DMT genes on TGMS wheat.

**Keywords:** *Triticum aestivum* L., Temperature-sensitive genic male sterility (TGMS), Stress response, DNA methyltransferases, Transcript abundance

## Background

DNA methylation is a heritable epigenetic modification widely found in bacteria, animals and plants and is instrumental in transposable element (TE) suppression, gene silencing, genomic imprinting and chromosome inactivation [1, 2]. DNA methylation is catalyzed by DNA methyltransferase (DMT) that introduces a methyl group at the C5 or N4 position of cytosine or the N6 position of adenine. C5-Cytosine DNA methylation (simply named here as DNA methylation) is the most prevalent among all these modifications, especially in plants [3, 4]. DNA methylation occurs in multiple sequence contexts, including CG and CHG, both of which are symmetric, and CHH, which is asymmetric (H = A, T, or C) [5, 6]. Two major methods of modulating DNA methylation have been found, including de novo methylation and maintenance methylation [7, 8]. Maintenance methylation occurs during DNA replication in hemimethylated CG and CHG contexts. De novo methylation occurs at all three unmethylated cytosines [3, 9].

In plants, DNA methylation is controlled by four sub-families of DMT genes: methyltransferase 1 (MET1, the homologs of mammalian DNMT1), chromomethylase (CMT, plant-specific), domains rearranged methyltransferase (DRM, the homologs of mammalian DNMT3) and DNA methyltransferase 2 (DNMT2) [10–12]. MET1 genes played a pivotal role in maintaining methylation in the CG context. *Arabidopsis thaliana* harbored four MET1 homologs, but only the loss-of-function of *MET1* bore phenotypic consequences. Ablation of *MET1* resulted in nearly complete loss of CG methylation in *Arabidopsis thaliana* [13–15]. In rice (*Oryza sativa* L.), there were two MET1 genes (*OsMET1A* and *OsMET1B*), but only *OsMET1B* (*OsMET1-2*) plays an essential part in maintaining CG methylation during normal growth and development. A null mutation of *OsMET1B* led to genome-wide hypomethylation [16]; however, knockout of *OsMET1A* (*OsMET1-1*) did not show alteration of DNA methylation at assessed marker loci [17]. The maintenance of major CHG methylation is accomplished by CMT3 [18]. There were three CMT genes (*AtCMT1*, *AtCMT2*, and *AtCMT3*) in *Arabidopsis thaliana*, and *AtCMT3* is the main CHG methyltransferase. In a mutant of *AtCMT3*, strong depletion of CHG methylation was observed [19]. In addition, previous research showed that CMT2 could partly partake in the maintenance of

CHH methylation via cross-talk with histone modifications in plants [11, 12]. In all three contexts, DRMs are primarily responsible for catalyzing de novo DNA methylation, which is achieved by the RNA-directed DNA methylation (RdDM) pathway in plants. RdDM involves two specific RNA polymerases, RNA Pol IV and Pol V, as well as 24 nt small interfering RNAs (siRNAs) [20–23]. DRM2 is required for de novo DNA methylation of all three contexts and for the maintenance of non-CG methylation [24, 25]. After the targeted destruction of *OsDRM2*, the CG and non-CG methylation levels were reduced, and de novo methylation mediated by the RdDM process was deficient [26]. Based on sequence homology, DNMT2 was originally considered a DNA methyltransferase. However, robust DNA methyltransferase activity could not be observed using DNMT2 preparations, and the considerable variability in the target DNA sequences showed that DNMT2 enzymes were actually tRNA transferases [27–29].

Comprehensive studies have showed that plant DNA methylation was predominantly associated with growth and development [30, 31], secondary metabolism [32, 33] and various stress responses, including salt stress [34, 35], drought stress [36, 37], and cold stress [38, 39]. Since DNA methylation is involved in regulating a wide range of biological processes, DMT genes have been identified and analyzed in several plant species, such as *Arabidopsis thaliana* [40, 41], rice [42, 43], maize [44, 45], and wheat [46]. However, to the best of our knowledge, no investigator has focused on the identification and analysis of DMT genes in wheat at the whole-genome level.

Wheat (*Triticum aestivum* L.) is one of the most important staple crops worldwide, contributing a significant amount of proteins and calories to the global human population [47, 48]. Given a growing world population coupled with increasingly challenging cultivation conditions, facilitating breeding and large-scale adoption of hybrid wheat are of great importance to the food supply [49, 50]. The two-line system based on photoperiod and/or temperature-sensitive genic male sterile (P/TGMS) lines plays an important role in hybrid seed production of wheat [51]. The male sterility of the P/TGMS line is heritable but regulated by an appropriate temperature or photoperiod. In the TGMS line of wheat, pollen viability is dramatically decreased to

complete sterility at a certain low temperature occurring during reproductive development [52, 53]. Temperature stress may lead to complete sterility and severe losses in grain yield during reproductive development mainly because the reproductive stage is especially sensitive to environmental stresses [54]. Specifically, male reproductive organs are known to be more sensitive to temperature stress than female organs in higher plants [55]. Plants have developed diverse mechanisms to survive extreme environmental conditions. Changes in DNA methylation have been shown to play a crucial part in plant stress responses [6]. Additionally, some recent reports have revealed that DNA methylation levels are affected by DMTs and regulate the expression of genes involved in cytoplasmic male sterility in rice, maize, and wheat [56–58]. However, few studies have been reported that explore the roles of DMTs in TGMS wheat during the reproductive stage.

To better understand the dynamic evolution of the DMT gene in wheat and to facilitate further research on this important gene family, all members of DMT genes in wheat were identified, and their phylogenetic relationships and gene structures were investigated. We found that the wheat DRM subfamily was significantly larger than expected and hypothesize that high homoeolog retention in the DRM subfamily contributed to the expansion of wheat DMT genes. Moreover, the transcript abundance of DMT genes under sterile and fertile conditions was analyzed. Notably, we found that most TaDMT genes were significantly upregulated under sterile conditions compared to fertile conditions. After further expressional analysis, we concluded that *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* were more likely associated with male sterility in the TGMS line BS366, particularly the interaction between *TaMET1-B1* and *TaRAFTINIA*. The results of this study present a detailed overview of the phylogeny of DMT genes and provide novel insights into the functional roles of DMT genes in the male sterility of TGMS wheat.

## Results

### Identification and phylogenetic analysis of DMT genes

To analyze the function of DMT genes, we performed genome-wide identification and phylogenetic analysis of DMT genes in wheat. A total of 52 DMT genes were identified on the basis of the functional annotation (Pfam domains) in the IWGSC archive v.1.0 [59]. We named all TaDMT genes according to their subfamily association and subgenome location (Supplemental Table S1). *Arabidopsis* and rice had 10 and 7 DMTs, respectively (Fig. 1). Wheat had the largest number of DMTs across the three species; this was partly the result of the hexaploid nature of wheat.

However, even after correcting for ploidy level, the number of DMT genes in wheat was significantly higher than that in rice (c. 2.5-fold higher;  $\chi^2$  test,  $P=0.02$ ). The increase in the number of TaDMT genes was mainly a result of the gene counts in the TaDRM subfamily, which was significantly larger than expected ( $\chi^2$  test,  $P=0.01$ ). The numbers of genes in the remaining subfamilies were not significantly different from the expected 3:1 ratio.

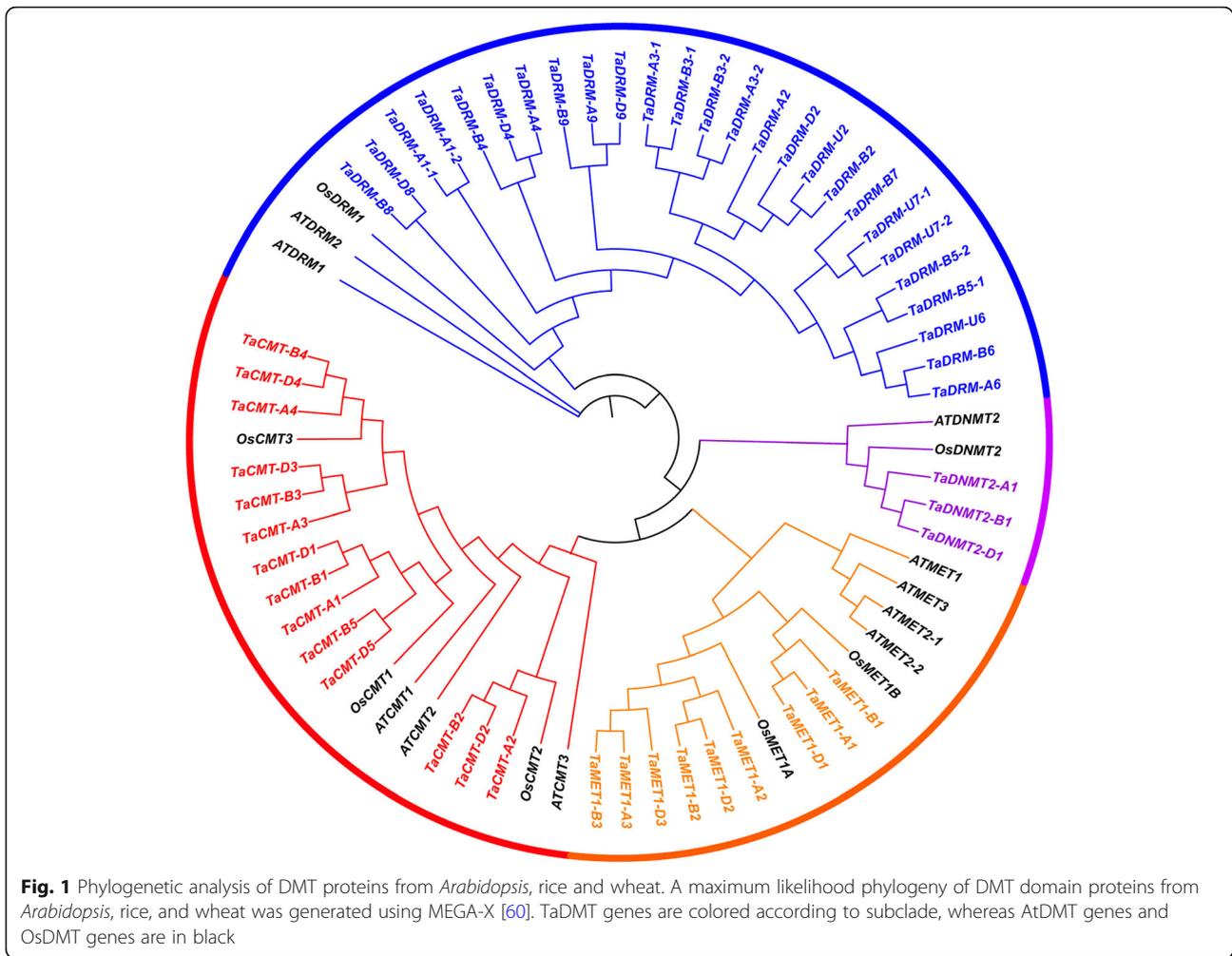
To elucidate the phylogenetic relationship of DMT genes from *Arabidopsis thaliana*, rice (*Oryza sativa* L.), and wheat, a total of 69 DMT protein sequences were used. A maximum likelihood phylogenetic tree of DMT genes showed that TaDMT genes could be naturally grouped into four subfamilies: MET1, CMT, DRM, and DNMT2 (Fig. 1). In all subclades, the DMT phylogeny roughly followed species phylogeny. As the *Arabidopsis* genes were a sister group related to the grass genes, OsDMT genes were more closely related to a triad of wheat homologs such as those in the MET1 and CMT subfamilies (Fig. 1).

### Chromosomal location and homoeologous genes of DMT genes in the wheat genome

DMT genes were unevenly distributed across wheat chromosomes: chr1A, chr1B, and chr1D had no TaDMT gene, and the highest number of genes was located on chr4B (7 genes) (Fig. 2). Interestingly, different members of DMT gene subfamilies were located in chromosomal regions that might represent homoeologous segments resulting from ancestral events [49]. Approximately 69% (12 triads) of the 52 DMT genes identified were present in triads. Additionally, 23% (six pairs) of DMT genes lost one homoeologous gene of the triads in the wheat genome. Thus, the high homoeologous retention rate can partly explain the high number of wheat DMT genes. Intriguingly, there were 5 paralogs in the wheat genome that were all from the DRM subfamily, indicating that tandem duplication might be an important mechanism for subfamily expansion (Fig. 2).

### Conserved motifs and gene structures of DMTs in wheat

Phylogenetic relationships of wheat DMTs showed that the MET1 and CMT subfamilies were more similar and belonged to the same clade (Fig. 3a). Furthermore, the motifs in all members of TaDMT genes were analyzed. A total of 10 different motifs were identified in 52 TaDMT proteins (Fig. 3b). Among the motifs, motifs 1, 3, 8, and 9 were found in all TaCMT and TaMET1 proteins, hence, these two subfamily proteins were clustered together in a small branch. TaDNMT2 was the smallest subfamily in the wheat phylogenetic tree and contained only 3 genes. The TaDNMT2 proteins included motifs 3 and 8. Most TaDRM genes were composed of 8 motifs,



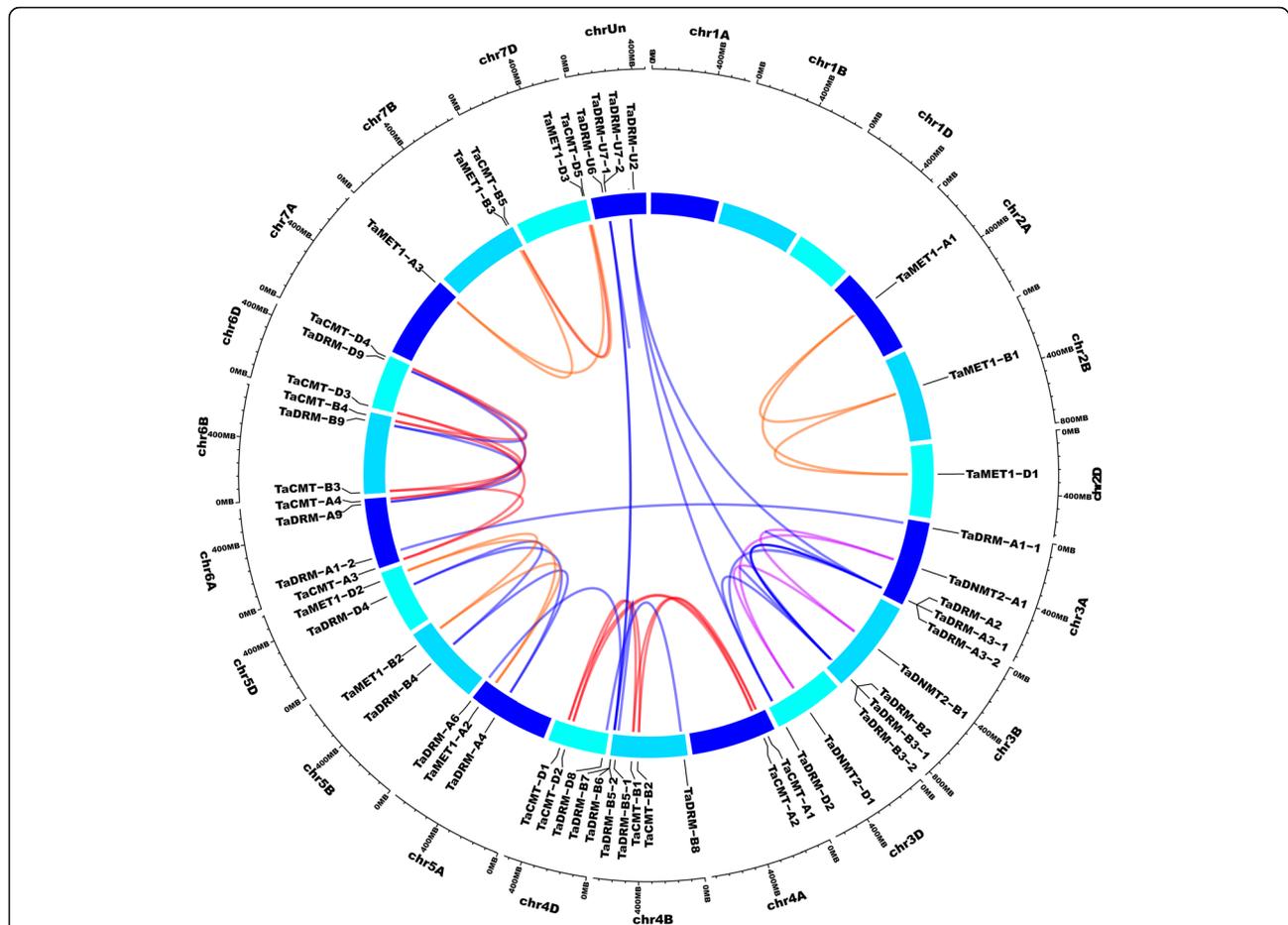
including in sequential order motifs 10, 5, 2, 6, 3, 4, 7, and 1, which were significantly different from the structure of the other three TaDMT genes. This finding was consistent with the results of the gene structures of TaDMT genes (Fig. 3c). The TaDMT genes exhibited different exon-intron organizational patterns. TaCMT genes had the largest number of exons; however, TaDRM genes had the smallest number of exons.

**DMT genes were significantly enriched in the thermosensitive genic male sterile line BS366 under sterile and fertile conditions**

A published GeneChip® data set was used to seek differently expressed genes (DEGs) under sterile and fertile conditions in the temperature-sensitive genic male sterile line BS366 anthers during meiosis [52]. The up and downregulated probe sets were used for enrichment analysis by Gene Ontology (GO). The metabolic process and DNA methylation were the major overrepresented GO terms in the biological process (Fig. 4a). DNA methylation, the most common epigenetic modification,

is an important link between phenotypes and genotypes, so we speculated that it might regulate metabolic process or other processes which eventually modulate male fertility in BS366. To further understand the function of DNA methylation, we first analyzed the function of DMT genes in wheat and the expressed transcripts of DMT genes in BS366.

To clarify the potential roles of TaDMT genes involved in male sterility, we analyzed the transcript abundance of TaDMT genes in BS366 based on the GeneChip® data. In this study, there were 13 TaDMT genes expressed, and 10 of them were differentially expressed under sterile and fertile conditions (Fig. 4b). The 9 TaDMT genes were upregulated, and only *TaMET1-D1* was downregulated under sterile conditions compared to fertile conditions. *TaMET1-B1* and *TaCMT-D2* were significantly differentially expressed in the 1.5-mm anther stage. *TaMET1-D1* was significantly differentially expressed in the 2.0-mm anther stage. *TaMET1-B1*, *TaCMT-A4*, *TaDRM-B7*, *TaMET1-B2*, *TaMET1-B3*, *TaCMT-B5*, and *TaCMT-D2* were

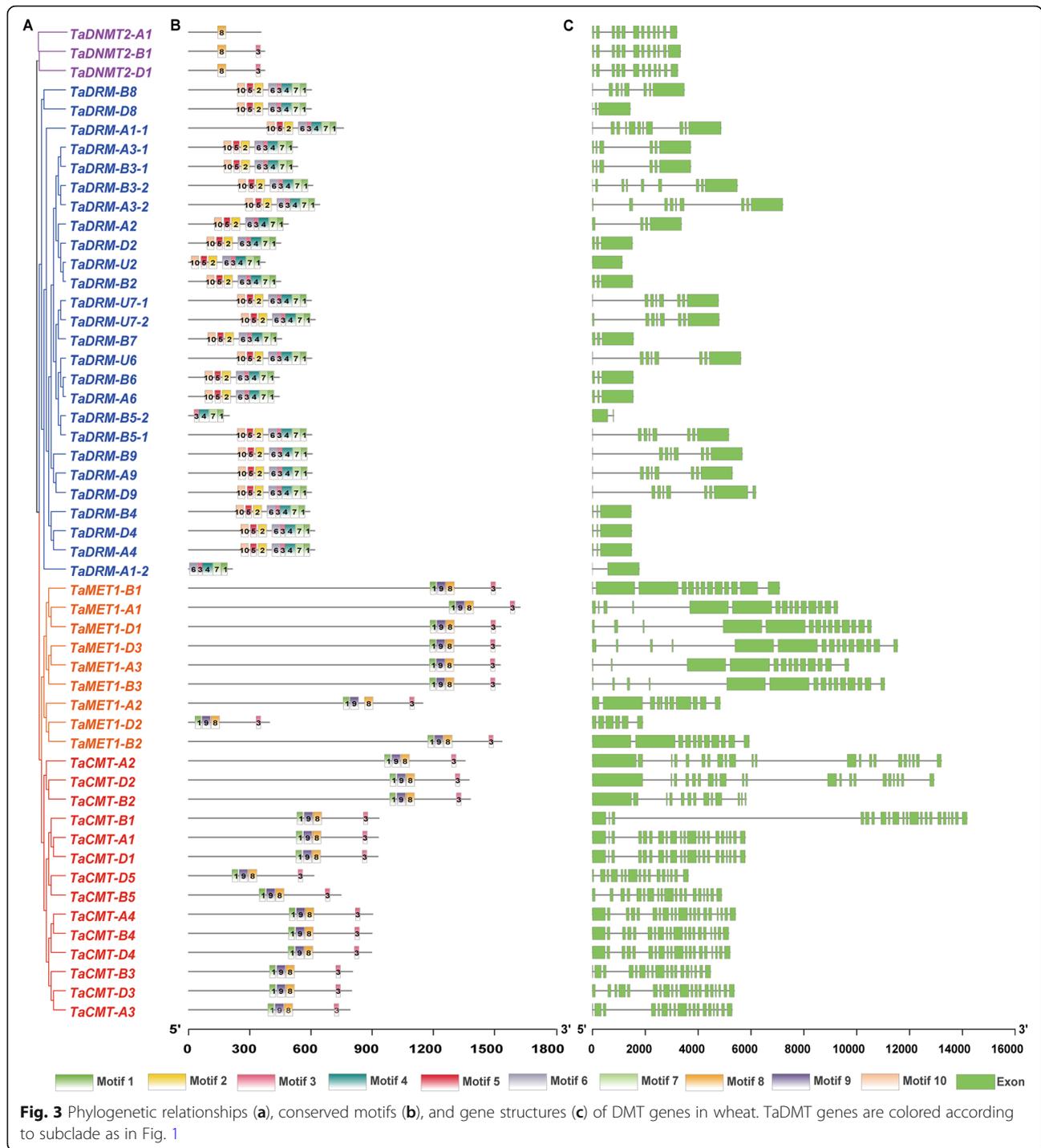


**Fig. 2** Chromosomal location and paralogs of DMT genes in the wheat genome. All TaDMT genes were mapped to their respective locus in the wheat genome in a circular diagram. Subgenomes are indicated by various shades of blue. Homeologous genes were inferred by phylogeny and linked with subfamily-specific colors as in Fig. 1 (inside the circle)

significantly differentially expressed in the 2.2-mm anther stage. *TaMET1-B1*, *TaMET1-B2*, *TaCMT-D2*, *TaDRM-D9*, and *TaDRM-U6* were significantly differentially expressed in the 2.5-mm anther stage. *TaMET1-B1*, *TaDRM-B7*, *TaMET1-B2*, *TaMET1-B3*, *TaCMT-B5*, *TaDRM-D9*, and *TaDRM-U6* were specifically differentially expressed in the 3.0-mm anther stage. Almost all differential expression of TaDMT genes occurred during the 2.2-, 2.5-, or 3.0-mm anther stage. From the 2.2-mm anther stage, mitosis ceased and transitioned to meiosis I according to morphological observations. The result reinforced the conclusion that cold stress contributed to aberrant cytokinesis during male meiosis I in BS366 [52]. Therefore, we speculated that the 2.2-mm, 2.5-mm, and 3.0-mm periods were the key periods affecting male sterility in BS366. Therefore, we chose *TaMET1-B1*, *TaCMT-A4*, *TaDRM-B7*, *TaMET1-B2*, *TaMET1-B3*, *TaCMT-B5*, *TaCMT-D2*, *TaDRM-D9*, and *TaDRM-U6* to further assess their expression.

#### Transcript abundance of TaDMT genes in fertile lines and hybrids of wheat under sterile and fertile conditions

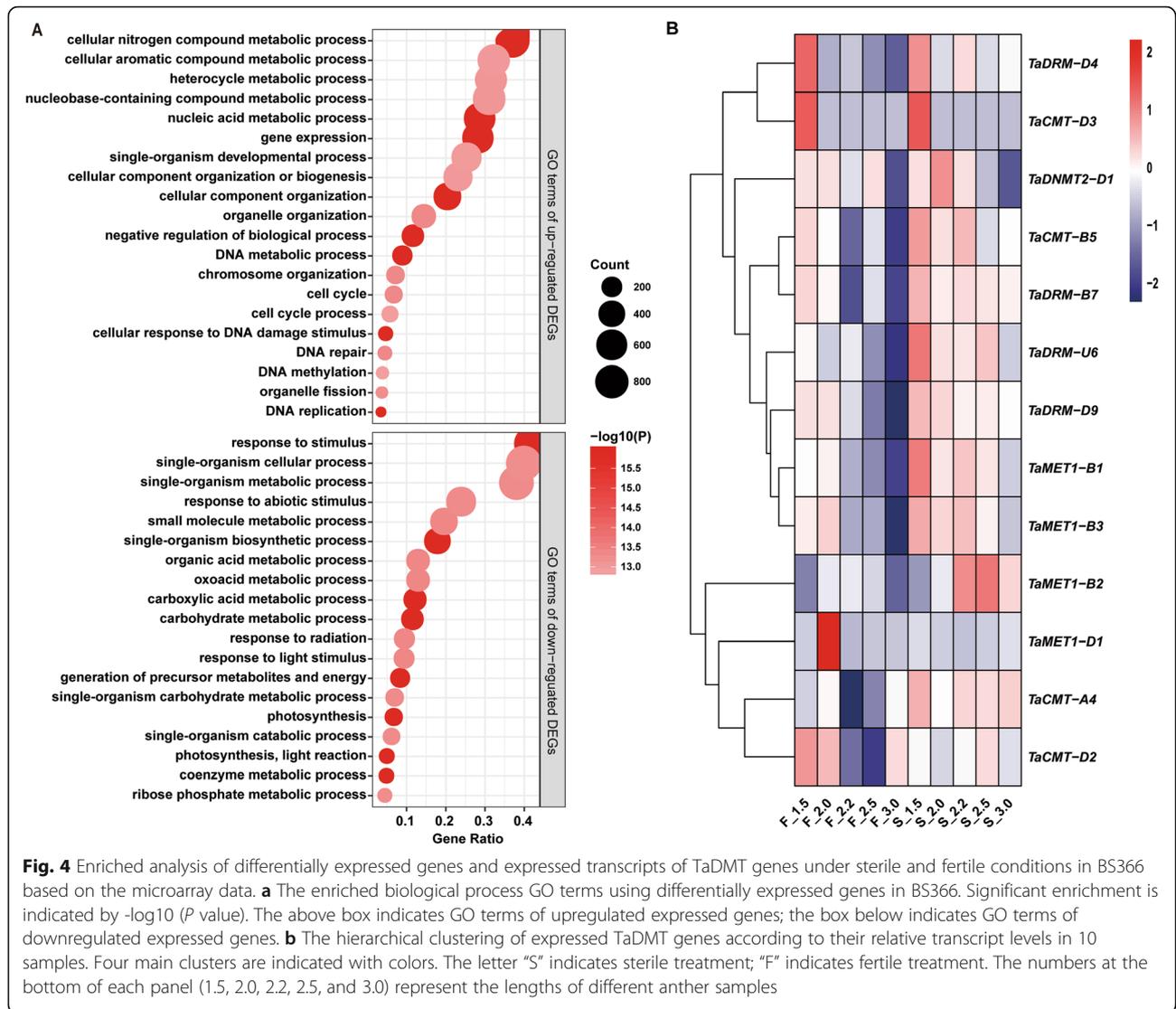
To verify the expression of the 9 TaDMT genes involved in male fertility in BS366, the anthers sampled from BS366, four fertile lines (J411, TY806, MC159, and GLDS) and three hybrids (GLDS×BS366, MC159×BS366, and TY806×BS366) at the meiosis stage were used. *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaMET1-B3*, *TaDRM-U6*, *TaDRM-B7*, and *TaDRM-D9* were variously upregulated in BS366 under sterile compared to fertile conditions, but they were downregulated or not differentially expressed in the four fertile lines (Fig. 5). *TaCMT-A4* gene was not differentially expressed in BS366 while it was differentially expressed in J411 and MC159 under sterile compared to fertile conditions (Fig. 5b). BS366 and three fertile lines (TY806, MC159, and GLDS) have been used to generate three different fertility-restored hybrids: TY806×BS366 (91.17%), MC159×BS366 (90.60%), and GLDS×BS366



(89.83%) [61]. Therefore, TY806, MC159, GLDS, and their hybrids with BS366 were considered excellent materials to assess the potential association of TaDMT genes with male sterility in BS366. Upregulated expression was detected for *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaMET1-B3*, *TaDRM-U6*, *TaDRM-B7*, and *TaDRM-D9* in GLDS×BS366, while their expression were downregulated or not different in TY806×BS366 and

MC159×BS366 under fertile and sterile conditions (Fig. 5). Bases on the above results, we speculated that *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaMET1-B3*, *TaDRM-U6*, *TaDRM-B7* and *TaDRM-D9* might be associated with male sterility in BS366.

To validate our speculation, we further analyzed the tissue-specific expression patterns of the eight TaDMT genes in roots, stems, leaves, and anthers using another



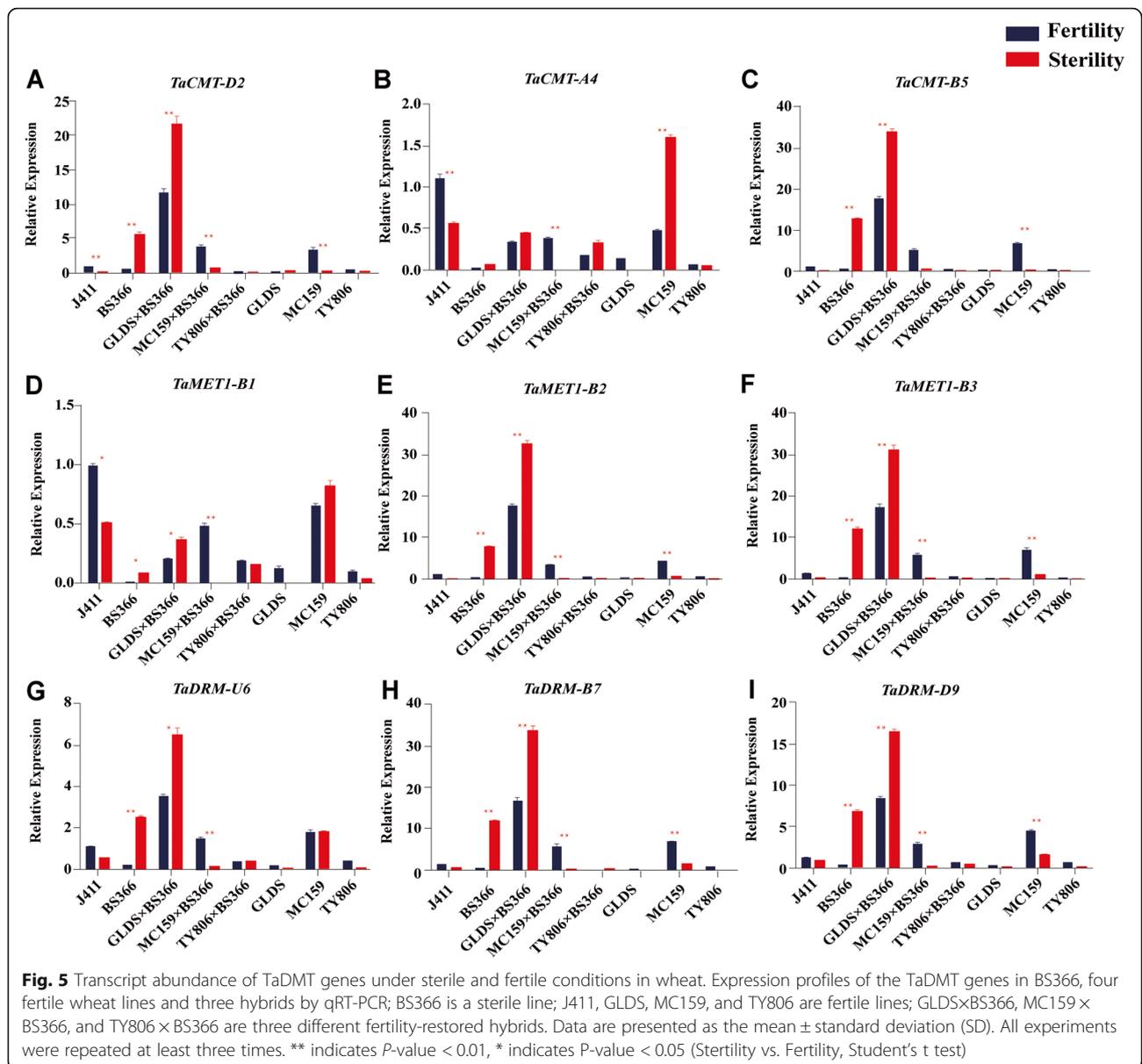
**Fig. 4** Enriched analysis of differentially expressed genes and expressed transcripts of TaDMT genes under sterile and fertile conditions in BS366 based on the microarray data. **a** The enriched biological process GO terms using differentially expressed genes in BS366. Significant enrichment is indicated by  $-\log_{10}(P)$  value. The above box indicates GO terms of upregulated expressed genes; the box below indicates GO terms of downregulated expressed genes. **b** The hierarchical clustering of expressed TaDMT genes according to their relative transcript levels in 10 samples. Four main clusters are indicated with colors. The letter “S” indicates sterile treatment; “F” indicates fertile treatment. The numbers at the bottom of each panel (1.5, 2.0, 2.2, 2.5, and 3.0) represent the lengths of different anther samples

three different fertility-restored F<sub>1</sub> hybrids 07GF83 × BS366 (97%), 07YH91–43 × BS366 (96%), and 7P395 × BS366 (89%) under sterile conditions (Fig. 6b and d). These hybrids were crossed by high-(07GF83), medium-(07YH91–43), and low-restorer lines (7P395) with BS366. Likewise, they were also prominent materials to evaluate the underlying association of TaDMT genes with male sterility in BS366. In 07GF83 × BS366, the expression levels of *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaDRM-U6*, *TaDRM-B7*, and *TaDRM-D9* in leaves were significantly higher under sterile conditions than that under fertile conditions (Fig. 7a-c). In 07YH91–43 × BS366, *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* expression were significantly higher in leaves under sterile conditions than that under fertile conditions; *TaMET1-B3* expression was significantly higher in anthers under sterile conditions than that

under fertile conditions (Fig. 7d-f). In 7P395 × BS366, *TaCMT-D2*, *TaMET1-B1*, *TaMET1-B3*, and *TaDRM-U6* expression was significantly higher in anthers under sterile conditions than that under fertile conditions (Fig. 7g-h). Considering the differing fertility of 07GF83 × BS366, 07YH91–43 × BS366, and 7P395 × BS366, *TaCMT-D2*, *TaMET1-B1* and *TaDRM-U6* displayed tissue-specific differential expression in anthers in the low fertility line 7P395 × BS366. Therefore, we concluded that *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* might take part in male sterility in BS366.

**Expressional correlation analysis among candidate TaDMT genes and the other DEGs in BS366 under fertile and sterile conditions**

To search genes that might interact with the candidate TaDMT genes, correlation analyses of expressional levels

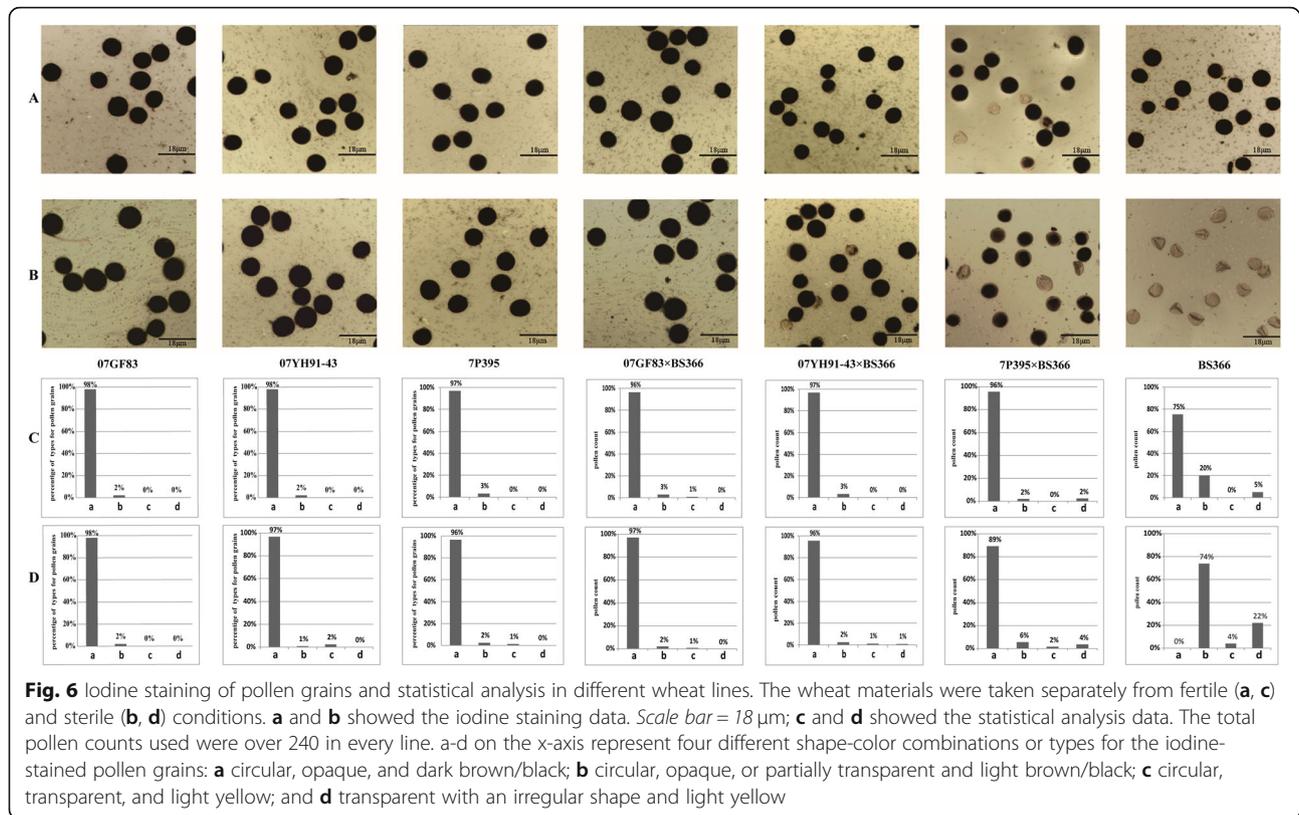


among the DEGs in BS366 under fertile and sterile conditions were conducted. The degree of correlation varied substantially among the genes, and Fig. 8 showed the highly negatively correlated genes with the candidate TaDMT genes. *TaRAFTINIA*, an important gene for pollen development [62], and *TraesCS4A02G093100* were the most highly negatively correlated with *TaMET1-B1* ( $r = -0.99$ , Fig. 8). We assumed that the interaction between *TaMET1-B1* and *TaRAFTINIA* might play an important role in the male sterility of BS366. In addition, *TaRAFTINIA* has a relatively high negative correlation with *TaDRM-U6* ( $r = -0.94$ , Fig. 8). Moreover, we analyzed the *cis*-acting elements in the promoter sequences of *TaCMT-D2*, *TaMET1-B1*, *TaDRM-U6*, and *TaRAFTINIA* (Supplemental Fig. S3).

According to the prediction of *cis*-acting elements, *TaCMT-D2*, *TaMET1-B1*, and *TaRAFTINIA* genes contained the LTR element that was considered in response to low temperature.

**The expression and methylation levels of *TaRAFTINIA* in wheat under sterile and fertile conditions**

To verify the relationship between *TaMET1-B1* and *TaRAFTINIA*, the transcript and DNA methylation levels of *TaRAFTINIA* was further examined. The expression levels of *TaRAFTINIA* in the anthers of BS366 and its three hybrids (07GF83 × BS366, 07YH91–43 × BS366, and 7P395 × BS366) were significantly downregulated under sterile conditions than that under fertile conditions (Fig. 9a). After CpG island prediction, the



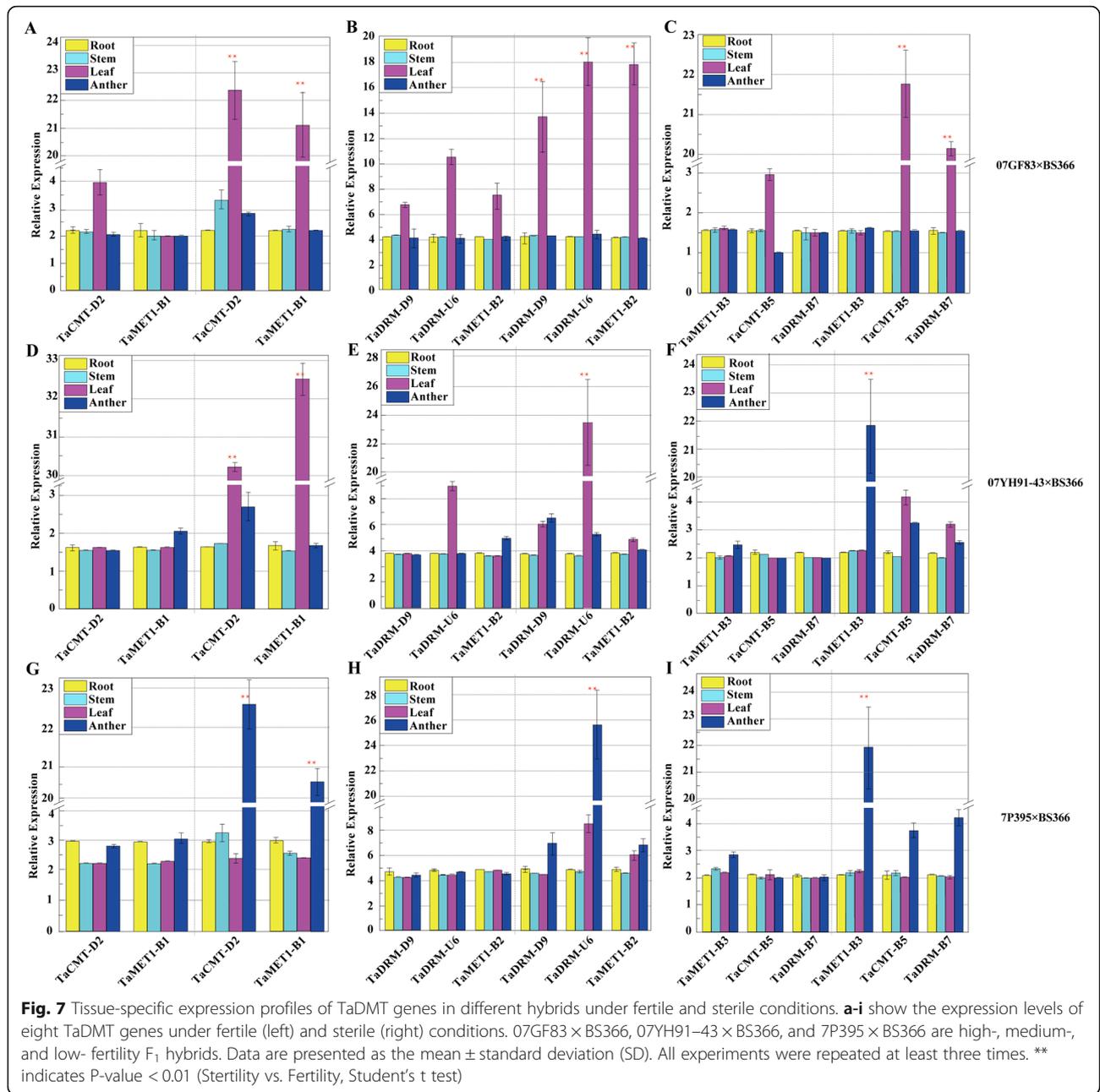
*TaRAFTINIA* sequence from -1121 to -1367 was selected to be applied with bisulfite sequencing. We identified three different DNA methylation patterns (CG, CHG, and CHH) in BS366 and 07GF83 under fertile and sterile conditions. Because very few loci of CHH type were detected, we focused on CG and CHG types in the subsequent analysis. We found that the average CG and CHG methylation levels in BS366 were with a difference by more than 20% under sterile conditions than that under fertile conditions (Fig. 9b), but not in the fertile line 07GF83. There were significant differences among four CpG sites (-1351, -1316, -1306, and -1257) with low methylation in BS366 under fertile conditions as compared with that under sterile conditions (Fig. 9c). From these results, we speculated that the expression level of *TaMET1-B1* might be negatively correlated with *TaRAFTINIA* expression and positively correlated with the methylation levels of the CpG islands in promoter regions.

## Discussion

### Structural and evolutionary features of DMT genes in wheat

DMTs play an important role in modulating DNA methylation and gene expressional levels, which are involved in various stress responses [6]. DMTs are, therefore, promising targets for crop breeding and

improvement. In this study, a total of 52 DMT genes were identified in wheat (Supplemental Table S1; Fig. 1). The number of DMT genes discovered in wheat was higher than that in *Arabidopsis* (10) and rice (7). The number of DMT genes was ~7.4 times higher in wheat than in rice. Leaving hexaploidy aside, the number of DMT genes in wheat was significantly higher than that in rice (c. 2.5-fold higher;  $\chi^2$  test,  $P = 0.02$ ). Further, we found that the strikingly high number of *TaDMT* genes was mainly a result of the significant expansion of genes in the *TaDRM* subfamily ( $\chi^2$  test,  $P = 0.01$ ). DNA methylation has long been thought to mostly control repetitive elements within genomes [63]. The insertion and deletion of repetitive elements such as transposable elements (TEs) strongly impact genome size and evolution [64]. TEs may spread rapidly, resulting in deleterious or null mutations within genomes, which present a threat to genome stability and integrity, resulting in plant structural and developmental defects [65]. To prevent this threat, DNA methylation mechanisms can be triggered by various mechanisms [66]. DNA methylation was considered one of the key components in understanding divergence in genome evolution among plants [67]. The expansion of wheat DMT genes might contribute to the large genome size of wheat. In all, 69% (12 triads) of wheat DMT genes could be assigned to 1:1:1 homoeologous groups (Fig. 2). This is significantly above the



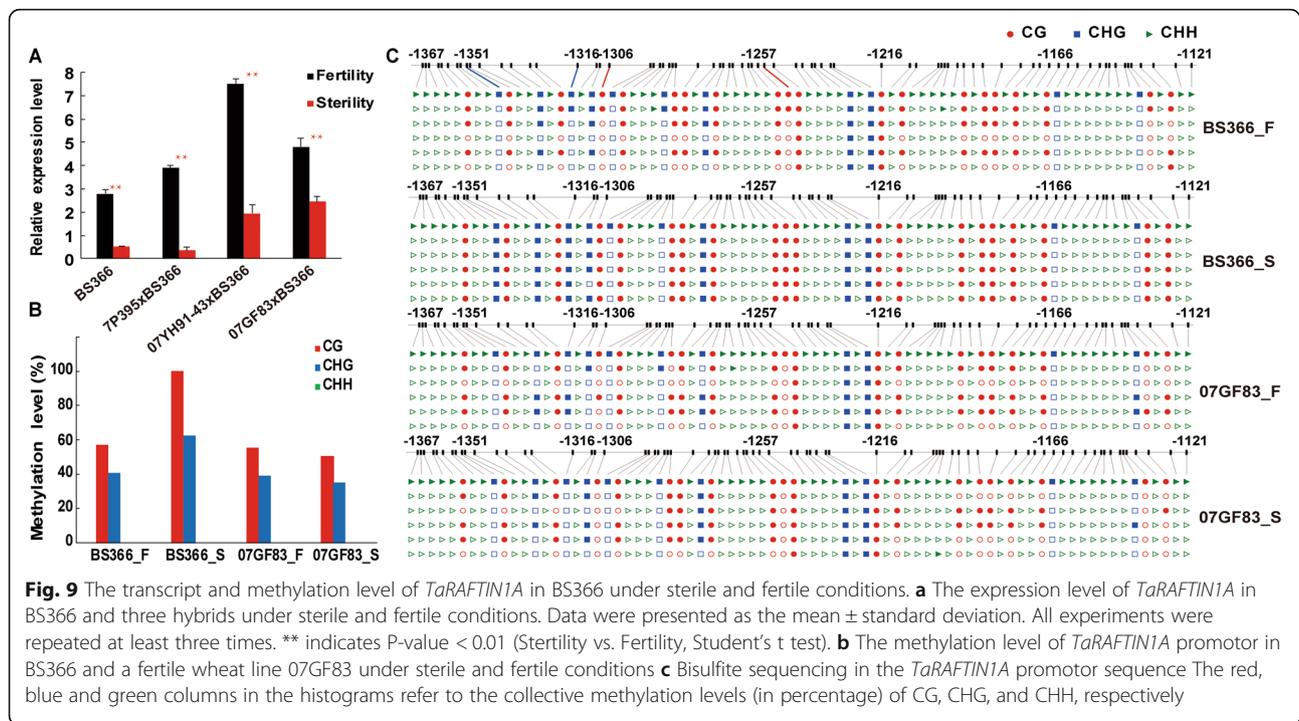
average homoeologous retention rate (35.8%) in wheat [49]. Thus, the high homoeologous retention rate can partly explain the high number of wheat DMT genes. In addition, there were 5 paralogs in the wheat genome that were all from the DRM subfamily. Hence, we deduced that tandem duplication and a high homoeologous retention rate may be the key factors in the evolution of DMT genes in wheat.

In phylogenetic classification, DMT genes were assigned to 4 conserved subfamilies, namely, MET1, CMT, DRM, and DNMT2. Except for the DRM subfamily, the numbers of genes in the remaining subfamilies

were not significantly different from the expected 3:1 ratio after factoring in the ploidy level. In all subclades, the DMT phylogeny roughly followed species phylogeny, and OsDMT genes were more closely related to the triad of wheat homologs. DMT genes diverged prior to the split of monocots and dicots, so we speculated that the functions of DMT genes may be different between dicots and monocots.

Although the number of DMT genes varied widely among different plant species, we found that they had conserved domains (Supplemental Fig. S1). All members of the DMT genes identified in this study contained the





motifs, motif 3 (Fig. 3b) was found in all the TaDMT protein except for TaDNMT2-A1. We speculated that the sequence in motif 3 might be the core sequence of the DNA-methylase domain (Fig. S2). All TaCMT and TaMET1 proteins had sequential order motifs 1, 9, 8, and 3, so these two subfamily proteins were clustered together in a small branch. Sequence analysis for conserved domains/motifs of DMT genes shed light on structural and functional conservation and/or unique features of the wheat DMT homologs.

#### DMT genes play important roles in temperature-sensitive male sterile wheat

Temperature and drought are the major abiotic stresses during the reproductive stage that threaten the growth and development of wheat. In particular, male reproductive organs are known to be more sensitive to temperature stress than female organs in higher plants [55, 71]. Temperature stress may lead to complete sterility during reproductive development [54]. The two-line system, which takes advantage of photoperiod and/or temperature-sensitive genic male sterile (P/TGMS) lines, plays a great role in hybrid seed production of wheat [51]. In the TGMS line of wheat, the pollen viability is dramatically decreased to complete sterility at a certain low temperature occurring during reproductive development [52, 53]. Recent studies have shown that DMT genes contribute to plant stress responses and development by de novo establishment and subsequent maintenance of DNA methylation [72]. However, the transcript

abundance of DMT genes in temperature-sensitive genic male sterile wheat has rarely been studied.

In this study, we investigated the transcript abundance of the TaDMT genes in the temperature-sensitive genic male sterile line BS366 based on the published GeneChip® data under sterile and fertile conditions. The results showed that there were 13 TaDMT genes that were expressed, and 10 of them were differentially expressed in BS366 under sterile and fertile conditions. Almost all the differentially expressed TaDMT genes were upregulated, and nearly all differential expression of TaDMT genes occurred during the 2.2-, 2.5- or 3.0-mm anther stage (Fig. 4b). Therefore, we speculated that the 2.2-mm, 2.5-mm and 3.0-mm periods are the key periods affecting male sterility in BS366. We further analyzed and compared the expression levels of the 9 DMT genes in three fertile wheat lines and their different fertility-restored hybrids. We found upregulated expression was detected for *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaMET1-B3*, *TaDRM-U6*, *TaDRM-B7*, and *TaDRM-D9* in the low fertility-restored hybrid GLDS $\times$ BS366, while their expression was downregulated or did not differ in TY806  $\times$  BS366 (high fertility-restored hybrid) and MC159  $\times$  BS366 (medium fertility-restored hybrid) under fertile and sterile conditions. Furthermore, the transcript levels of *TaCMT-D2*, *TaCMT-B5*, *TaMET1-B1*, *TaMET1-B2*, *TaMET1-B3*, *TaDRM-U6*, *TaDRM-B7*, and *TaDRM-D9* presented heterotic expression in GLDS $\times$ BS366, but not in other two cross combinations (Fig. 5). Two explanations were presented which

might help to decipher this result: 1) GLDS×BS366 is the least fertile combination among these three cross combinations, and the expression patterns of these genes might be correlated with fertility in the three hybrids; 2) the heterotic expression in TY806×BS366 and MC159×BS366 might be repressed due to the paternal imprinting effect. Then, the tissue-specific expression patterns of the eight TaDMT genes were conducted using another three different fertility-restored hybrids. After analysis of the transcript abundance of TaDMT genes step-by-step under sterile and fertile conditions, we concluded that *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* were the candidate TaDMT genes that might participate in male sterility in BS366.

To search target genes of the three candidate TaDMT genes, correlation analysis of the expression levels among the DEGs in BS366 under fertile and sterile conditions was conducted. We found that *TaRAFTINIA*, an important gene for pollen development, was highly negatively correlated with *TaMET1-B1* ( $r = -0.99$ ). To verify the relationship between *TaMET1-B1* and *TaRAFTINIA*, the transcript and DNA methylation levels of *TaRAFTINIA* was further examined in this study. We found that the expression levels of *TaRAFTINIA* in BS366 and its three hybrids were significantly downregulated under sterile conditions than that under fertile conditions (Fig. 9a) and the average CG and CHG methylation levels in BS366 were with a difference by more than 20% under sterile conditions than that under fertile conditions (Fig. 9b), but not in the fertile line 07GF83. From the results, we further speculated that the interaction between *TaMET1-B1* and *TaRAFTINIA* might play an important role in the male sterility of BS366. *TaMET1-A1*, *TaMET1-B1*, and *TaMET1-D1* are the three wheat homoeologs of *AtMET1* in *Arabidopsis* and *OsMET1B* in rice. Previous studies have shown that only the loss-of-function mutation of *ATMET1* resulted in almost all loss of CG methylation in *Arabidopsis thaliana* [13, 14]. Likewise, in rice, a null mutation of *OsMET1B* led to genome-wide hypomethylation [16]; however, a knockout mutant of *OsMET1A* (*OsMET1-1*) did not show alteration of DNA methylation at the assessed marker loci [17]. These results suggest the importance of *TaMET1-A1*, *TaMET1-B1*, and *TaMET1-D1* in the maintenance of CG methylation in wheat. In this study, *TaMET1-A1* was not expressed in BS366. Additionally, *TaMET1-B1* was upregulated, while *TaMET1-D1* was downregulated in BS366 under sterile conditions compared to fertile conditions. This phenomenon has also been found in the MIKC-type MADS-box genes [49]. The number of balanced transcript abundance from all three homoeologs of MIKC-type MADS-box gene triads was relatively low compared with an average assessment of all wheat genes [49]. The unbalanced

homoeolog expression pattern might predict different functions of different homoeolog alleles.

## Conclusions

In this study, DNA methyltransferase (DMT) genes in wheat (*Triticum aestivum* L.) based on genome-wide levels were identified and transcriptionally characterized to assess the regulation of DMT genes in the fertility of the temperature-sensitive genic male sterile (TGMS) wheat line BS366. We presented a detailed phylogenetic overview of wheat DMT (TaDMT) genes. Homoeolog retention for TaDMT genes was significantly above the average retention rate for whole-wheat genes. We also found extensive expansion of the TaDRM subfamily, and tandem duplications might play a crucial role in the expansion of the TaDRM subfamily. Following transcriptional analysis of TaDMT genes in the TGMS line BS366 and its hybrids with six other fertile lines under sterile and fertile conditions, we concluded that *TaCMT-D2*, *TaMET1-B1*, and *TaDRM-U6* might be involved in the male sterility of BS366. Moreover, a correlation analysis showed that *TaMET1-B1* might negatively regulate the expression of *TaRAFTINIA*, an important gene for pollen development. Our findings could provide novel insights into the effects of the DMT genes on TGMS wheat.

## Methods

### Genome-wide identification of DMT genes

Functional annotations of wheat, rice, and *Arabidopsis* genes were downloaded from the Ensembl (<http://plants.ensembl.org/info/website/ftp/index.html>). A hidden Markov model (HMM) analysis was performed in the Protein family database (Pfam) to extract HMM profiles for the DNA methylase domain (PF00145), Chromo domain (PF00385), BAH domain (PF01426), and DNMT1-RFD domain (PF12047) from all species. The software hmmer was used to search for the DMT genes ( $E < 1e-10$ ). Subsequently, Pfam and Conserved Domain Database (CDD, <https://www.ncbi.nlm.nih.gov/cdd/>) were used to confirm all the DMT genes containing the conserved domain. A total of 75 sequences were identified in the wheat genome. After excluding splice variants and keeping only the first variant, 52 genes were used for further analysis.

### Maximum likelihood phylogeny, conserved motif, gene structures, and cis-acting elements of DMT genes

Related DMT sequences from wheat, rice, and *Arabidopsis* were aligned by Clustal W, and then a phylogenetic tree was inferred using MEGA X [60] software by the maximum likelihood method (bootstrap value = 1000). The motifs of DMT genes in wheat were analyzed with the MEME database (<http://meme-suite.org/tools/meme>), Parameter settings were -nmotifs 10, and the

remaining parameters were the default. Based on genetic feature format data from the wheat genome, the gene structures of DMT genes were analyzed using Ttools software [73]. To analyze the *cis*-acting elements of *TaCMT-D2*, *TaMET1-B1*, *TaDRM-U6*, and *TaRAFTINIA*, the upstream sequences (2000 bp) of the start codon were retrieved from the wheat genome database, and then the *cis*-acting elements were analyzed by using the PlantCARE tool [74].

#### Naming of DMT genes in wheat

Naming wheat DMT genes, genome location and phylogenetic relationships were considered. The name of each gene started with Ta, an abbreviation for *Triticum aestivum*, and was followed by the name of the Arabidopsis gene from this subfamily. The subsequent name was A, B or D, indicating the subgenome that the DMTs were located in. Homoeologs were identified by phylogeny. Putative homoeologs have identical gene names except for the subgenome identifier (e.g., *TaMET1-A1*, *TaMET1-B1*, and *TaMET1-D1*). Genes belonging to the same subfamily but different triads within the same subgenome were consecutively numbered (e.g., *TaCMT-A1* and *TaCMT-A2*). Paralogs were distinguished by consecutive numbers separated by a dash (e.g., *TaDRM-A1-1* and *DRM-A1-2*).

#### Plant materials, RNA isolation, and RT-PCR

Wheat (*Triticum aestivum* L.) cv. BS366, J411, GLDS, TY806, MC159, and hybrids GLDS×BS366, MC159×BS366, and TY806×BS366, 07YH91–43×BS366, 07GF83×BS366, and 7P395×BS366 were used for reverse transcription-polymerase chain reaction (RT-PCR) analysis. Wheat TGMS line BS366 (Beijing Sterility 366) was selected from a natural mutant of doubled haploid lines (offspring of Jingnong8121/E8075–7) in the experimental fields in Beijing (China, N 39°54', E 116°18'). BS366, GLDS, TY806, MC159, 07YH91–43, 07GF83, and 7P395 were bred in the Beijing Engineering and Technique Research Center for Hybrid Wheat. J411 (Jing411) was bred by Beijing Seed Company and was the offspring of Fengkang2/Changfeng 1. All the materials used in the study have the permissions for scientific research and publication. The plants were grown in the soil of plastic pots embedded in the ground according to Tang et al. [52] and they were vernalized naturally in the field. After the tip of the fourth leaf had emerged, the plants were moved to the fertility conditions, 20 °C with a 12/12-h light/dark cycle in phytotrons (Koito, Tokyo, Japan) for the entire experimental period, except during low-temperature treatment. When the tip of the flag leaf had half-emerged, part of the plants were treated for the sterility treatment, 10 °C with a 12/12-h light/dark cycle for 5 days. The anthers of J411, BS366, GLDS, TY806,

MC159 GLDS×BS366, MC159×BS366, and TY806×BS366 and the roots, stems, leaves, and anthers of 07YH91–43×BS366, 07GF83×BS366, and 7P395×BS366 were then collected and snap-frozen at –80 °C. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. First-strand cDNA was synthesized using a Takara PrimeScript™ RT Reagent Kit (Takara Biotechnology [Dalian] Co., Ltd., Dalian, China) using a random primer. An Eco Real-Time PCR System (Illumina, San Diego, CA, USA) with Takara SYBR® Premix Ex Taq™ (Tli RNase H Plus; Takara Biotechnology [Dalian] Co., Ltd.) was used for RT-PCR. *Actin* acted as an internal control for the wheat expression studies. Primers were designed using Premier 3.0 (v. 0.4.0) (Supplemental Table S2). Relative gene expression levels were calculated using the  $2^{-\Delta\Delta CT}$  method [75].

#### Expression analysis of DMT genes using GeneChip®

GeneChip® data were from Tang et al [52]. Five developmental stages categorized by anther length (1.5-mm, 2.0-mm, 2.2-mm, 2.5-mm, and 3.0-mm) refer to 10 time points from BS366, a wheat TGMS line, under fertile and sterile conditions. Probe sets were considered differentially expressed with statistical significance when *P*-values were 0.05, corresponding to the false discovery rate of 5%. Gene annotations were obtained by Batch BLASTx (E-value < e-10) using the sequences of Affymetrix probe sets [52]. Hierarchical clusters of all annotated genes and the expressed TaDMT genes were generated according to their relative transcript levels in 10 samples using the R package *heatmap*. The R package *topGO* were used to cluster the functional categories across the list of significantly enriched functional GO terms.

#### Pollen iodine staining of wheat lines

Over 240 anthers from each wheat line were used for pollen iodine staining, and pollen iodine staining was performed according to [61]: the mature anther was squeezed using small tweezers to release the pollen grains; then, 1 to 2 drops of I2-KI solution [6.7‰ (g/ml) KI and 3.3‰ (g/ml) I2] were added. Finally, the images were obtained using an Olympus BX41 laboratory microscope (Olympus, Tokyo, Japan).

#### Bisulfite sequencing

Genomic DNA for bisulfite sequencing was extracted during meiosis from the anthers of BS366 and 07GF83 plants under fertile and sterile conditions. The bisulfite modification of genomic DNA (~2 µg) from 8 samples was performed with sodium bisulfite using an EZ DNA Methylation-Gold™ Kit (Zymo Research, Orange, CA, USA). The bisulfite conversion step was extended by the

addition of the denaturation step (98 °C for 10 min) and 2.5-h incubation at 64 °C. The CpG island was predicted by EMBL-EBI database (<https://www.ebi.ac.uk/>) and primers for the target site were designed using PyroMark Assay Design 2.0 (Qiagen). The primers of *TaRAFTIN1a* were shown in Supplemental Table 2. PCR reaction was conducted with Taq Polymerase (Shanghai Sangon Biotech Corp., Shanghai, China) using an ABI Veriti™ 96-Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA). The PCR products were cloned into the pUC18-T vector (Shanghai Sangon Biotech Corp.). Ten clones per amplicon were selected for sequencing. The results were aligned by Clustal X 2.1 (<http://www.clustal.org/clustal2/>) and analyzed with CyMATE (<http://www.cymate.net/>). The methylation levels which were expressed as the percentage (%) for each of the three types of cytosines, CG, CHG, and CHH, were calculated by dividing the number of converted (methylated) cytosines by the total number of cytosines within the assay.

#### Abbreviations

DMT: DNA methyltransferase; P/TGMS: Photoperiod and/or temperature-sensitive genic male sterile; MET1: Methyltransferase 1; CMT: Chromomethylase; DRM: Domains Rearranged Methyltransferase; DNMT2: DNA methyltransferase 2; RdDM: RNA-directed DNA methylation; siRNAs: Small interfering RNAs; DEGs: Differently expressed genes; GO: Gene Ontology; TEs: Transposable elements; BAH: Bromo-adjacent homology

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07600-7>.

**Additional file 1: Table S1.** List of all DMT genes in wheat. **Table S2.** Primer information used in this study.

**Additional file 2: Fig. S1.** Conserved domain analysis of DMT genes in wheat. **Fig. S2.** Conserved motifs of DMT genes in wheat. **Fig. S3.** Analysis of the *cis*-acting elements in the promoters of *TaCMT-2*, *TaMET1-B1*, *TaDRM-U6*, and *TaRAFTIN1A*.

#### Acknowledgments

We thank Prof. Ligeng Ma (College of Life Sciences, Capital Normal University, and Beijing Key Laboratory of Plant Gene Resources and Biotechnology for Carbon Reduction and Environmental Improvement) for his valuable advice concerning our manuscript.

#### Authors' contributions

GSQ and ZCP designed the experiments; LD, FBE, LYJ, and GJ performed most of the experiments; TYM, ZLP, SRW, CZB, ZFT, and PBS completed identification of iodine dyeing and pollen fertility; WYB and CXC investigated and provided phenotype data of hybrids; LD and FBE wrote the manuscript; and GSQ, ZCP, and WAP supervised the project. All authors have read and approved the final manuscript.

#### Funding

This work was financially supported by the Beijing Academy of Agriculture and Forestry Sciences Program (grant no. KJCX20170421), Important Crops Genetically Modified New Germplasm Creation (KJCX20200205), and Beijing Postdoctoral Research Foundation (2020–22-109). The funding bodies did not participate in the design of the study, collection, analysis, and interpretation of data or in writing the manuscript.

#### Availability of data and materials

The microarray data used for this study are deposited at the National Center for Biotechnology Information Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE171305.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

Received: 25 September 2020 Accepted: 11 April 2021

Published online: 29 April 2021

#### References

- Sasaki E, Kawakatsu T, Ecker JR, Nordborg M. Common alleles of CMT2 and NRPE1 are major determinants of CHH methylation variation in *Arabidopsis thaliana*. *PLoS Genet.* 2019;15(12):e1008492. <https://doi.org/10.1371/journal.pgen.1008492>.
- Xie M, Hong C, Zhang B, Lowdon RF, Xing X, Li D, et al. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nat Genet.* 2013;45(7):836–41. <https://doi.org/10.1038/ng.2649>.
- Ponger L, Li WH. Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. *Mol Biol Evol.* 2005;22(4):1119–28. <https://doi.org/10.1093/molbev/msi098>.
- Chen Zhu SZ, Zhou C, Chen L, Fu H, Li X, Lin Y, et al. Genome-wide investigation and transcriptional analysis of cytosine-5 DNA methyltransferase and DNA demethylase gene families in tea plant (*Camellia sinensis*) under abiotic stress and withering processing. *Peer J.* 2020;8:e8432. <https://doi.org/10.7717/peerj.8432>.
- Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11(3):204–20. <https://doi.org/10.1038/nrg2719>.
- Zhang HM, Lang ZB, Zhu JK. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Bio.* 2018;19(8):489–506. <https://doi.org/10.1038/s41580-018-0016-z>.
- Yang J, Lior-Hoffmann L, Wang SL, Zhang YK, Brody S. DNA cytosine methylation: structural and thermodynamic characterization of the epigenetic marking mechanism. *Biochemistry-U.S.* 2013;52(16):2828–38. <https://doi.org/10.1021/bi400163k>.
- Jullien PE, Susaki D, Yelagandula R, Higashiyama T, Berger F. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. *Curr Biol.* 2012;22(19):1825–30. <https://doi.org/10.1016/j.cub.2012.07.061>.
- Hu L, Li N, Zhang Z, Meng X, Dong Q, Xu C, et al. CG hypomethylation leads to complex changes in DNA methylation and transpositional burst of diverse transposable elements in callus cultures of rice. *Plant J.* 2020;101(1):188–203. <https://doi.org/10.1111/tpj.14531>.
- Wang PF, Gao C, Bian XT, Zhao SZ, Zhao CZ, Xia H, et al. Genome-wide identification and comparative analysis of cytosine-5 DNA methyltransferase and demethylase families in wild and cultivated peanut. *Front Plant Sci.* 2016;7:7.
- Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell.* 2013;153(1):193–205. <https://doi.org/10.1016/j.cell.2013.02.033>.
- Stroud H, Do T, Du JM, Zhong XH, Feng SH, Johnson L, et al. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol.* 2014;21(1):64–72.
- Cokus SJ, Feng SH, Zhang XY, Chen ZG, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature.* 2008;452(7184):215–9. <https://doi.org/10.1038/nature06745>.
- Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. Transgenerational stability of the *Arabidopsis* epigenome is coordinated

- by CG methylation. *Cell*. 2007;130(5):851–62. <https://doi.org/10.1016/j.cell.2007.07.007>.
15. Genger RK, Kovac KA, Dennis ES, Peacock WJ, Finnegan EJ. Multiple DNA methyltransferase genes in *Arabidopsis thaliana*. *Plant Mol Biol*. 1999;41(2):269–78. <https://doi.org/10.1023/A:1006347010369>.
  16. Hu LJ, Li N, Xu CM, Zhong SL, Lin XY, Yang JJ, et al. Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *P Natl Acad Sci USA*. 2014;111(29):10642–7. <https://doi.org/10.1073/pnas.1410761111>.
  17. Yamauchi T, Johzuka-Hisatomi Y, Fukada-Tanaka S, Terada R, Nakamura I, Iida S. Homologous recombination-mediated knock-in targeting of the *MET1a* gene for a maintenance DNA methyltransferase reproducibly reveals dosage-dependent spatiotemporal gene expression in rice. *Plant J*. 2009;60(2):386–96. <https://doi.org/10.1111/j.1365-313X.2009.03947.x>.
  18. Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, et al. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell*. 2012;151(1):167–80. <https://doi.org/10.1016/j.cell.2012.07.034>.
  19. Stroud H, Greenberg MVC, Feng SH, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell*. 2013;152(1–2):352–64. <https://doi.org/10.1016/j.cell.2012.10.054>.
  20. Haag JR, Pikaard CS. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat Rev Mol Cell Bio*. 2011;12(8):483–92. <https://doi.org/10.1038/nrm3152>.
  21. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet*. 2014;15(6):394–408. <https://doi.org/10.1038/nrg3683>.
  22. Kawashima T, Berger F. Epigenetic reprogramming in plant sexual reproduction. *Nat Rev Genet*. 2014;15(9):613–24. <https://doi.org/10.1038/nrg3685>.
  23. Martinez G, Kohler C. Role of small RNAs in epigenetic reprogramming during plant sexual reproduction. *Curr Opin Plant Biol*. 2017;36:22–8. <https://doi.org/10.1016/j.pbi.2016.12.006>.
  24. Gent JJ, Ellis NA, Guo L, Harkess AE, Yao YY, Zhang XY, et al. CHH islands: *de novo* DNA methylation in near-gene chromatin regulation in maize. *Genome Res*. 2013;23(4):628–37. <https://doi.org/10.1101/gr.146985.112>.
  25. Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, et al. Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol*. 2003;13(24):2212–7. <https://doi.org/10.1016/j.cub.2003.11.052>.
  26. Moritoh S, Eun CH, Ono A, Asao H, Okano Y, Yamaguchi K, et al. Targeted disruption of an orthologue of DOMAINS REARRANGED METHYLASE 2, *OsDRM2*, impairs the growth of rice plants by abnormal DNA methylation. *Plant J*. 2012;71(1):85–98. <https://doi.org/10.1111/j.1365-313X.2012.04974.x>.
  27. Vieira GC, Vieira GF, Sinigaglia M, Valente VLD. Linking epigenetic function to electrostatics: The DNMT2 structural model example. *PLoS One*. 2017;12(6):e0178643.
  28. Jeltsch A, Ehrenhofer-Murray A, Jurkowski TP, Lyko F, Reuter G, Anki S, et al. Mechanism and biological role of *Dnmt2* in nucleic acid methylation. *RNA Biol*. 2017;14(9):1108–23. <https://doi.org/10.1080/15476286.2016.1191737>.
  29. Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang XY, et al. Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog *Dnmt2*. *Science*. 2006;311(5759):395–8. <https://doi.org/10.1126/science.1120976>.
  30. Lin Z, Liu M, Damaris RN, Nyong'a TM, Cao D, Ou K, et al. Genome-wide DNA methylation profiling in the Lotus (*Nelumbo nucifera*) flower showing its contribution to the stamen petaloid. *Plants (Basel)*. 2019;8(5):135.
  31. Xing MQ, Zhang YJ, Zhou SR, Hu WY, Wu XT, Ye YJ, et al. Global analysis reveals the crucial roles of DNA methylation during Rice seed development. *Plant Physiol*. 2015;168(4):1417–U1554. <https://doi.org/10.1104/pp.15.00414>.
  32. Zhang B, Tieman DM, Jiao C, Xu Y, Chen K, Fei Z, et al. Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. *Proc Natl Acad Sci U S A*. 2016;113(44):12580–5. <https://doi.org/10.1073/pnas.1613910113>.
  33. Bharti P, Mahajan M, Vishwakarma AK, Bhardwaj J, Yadav SK. *AtrROS1* overexpression provides evidence for epigenetic regulation of genes encoding enzymes of flavonoid biosynthesis and antioxidant pathways during salt stress in transgenic tobacco. *J Exp Bot*. 2015;66(19):5959–69. <https://doi.org/10.1093/jxb/erv304>.
  34. Chen R, Li M, Zhang HY, Duan LJ, Sun XJ, Jiang QY, et al. Continuous salt stress-induced long non-coding RNAs and DNA methylation patterns in soybean roots. *BMC Genomics*. 2019;20(1):730.
  35. Sun LF, Miao XF, Cui J, Deng J, Wang X, Wang YF, et al. Genome-wide high-resolution mapping of DNA methylation identifies epigenetic variation across different salt stress in Maize (*Zea mays* L.). *Euphytica*. 2018;214(2):21.
  36. Van Dooren TJM, Silveira AB, Gilbault E, Jimenez-Gomez JM, Martin A, Bach L, et al. Mild drought in the vegetative stage induces phenotypic, gene expression, and DNA methylation plasticity in *Arabidopsis* but no transgenerational effects. *J Exp Bot*. 2020;71(12):3588–602. <https://doi.org/10.1093/jxb/eraa132>.
  37. Yap WS, Yeo YL, Lai KS, Wee CY. Epigenetic of DNA methylation patterns in lowland and upland Malaysian rice cultivars under induced drought and salinity. *New Biotechnol*. 2016;33:5167. <https://doi.org/10.1016/j.nbt.2016.06.1299>.
  38. Sicilia A, Scialo E, Puglisi I, Lo Piero AR. Anthocyanin biosynthesis and DNA methylation dynamics in sweet Orange fruit [*Citrus sinensis* L. (Osbeck)] under cold stress. *J Agr Food Chem*. 2020;68(26):7024–31. <https://doi.org/10.1021/acs.jafc.0c02360>.
  39. Shan XH, Wang XY, Yang G, Wu Y, Su SZ, Li SP, et al. Analysis of the DNA methylation of maize (*Zea mays* L.) in response to cold stress based on methylation-sensitive amplified polymorphisms. *J Plant Biol*. 2013;56(1):32–8. <https://doi.org/10.1007/s12374-012-0251-3>.
  40. Ebbs ML, Bender J. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell*. 2006;18(5):1166–76. <https://doi.org/10.1105/tpc.106.041400>.
  41. Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, et al. Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase mutants. *PLoS Genet*. 2012;8(8):e1002844.
  42. Teerawanichpan P, Chandrasekharan MB, Jiang YM, Narangajavana J, Hall TC. Characterization of two rice DNA methyltransferase genes and RNAi-mediated reactivation of a silenced transgene in rice callus. *Planta*. 2004;218(3):337–49. <https://doi.org/10.1007/s00425-003-1112-6>.
  43. Ahmad F, Huang X, Lan HX, Huma T, Bao YM, Huang J, et al. Comprehensive gene expression analysis of the DNA (cytosine-5) methyltransferase family in rice (*Oryza sativa* L.). *Genet Mol Res*. 2014;13(3):5159–72. <https://doi.org/10.4238/2014.July.7.9>.
  44. Qian YX, Xi YL, Cheng BJ, Zhu SW. Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize. *Plant Cell Rep*. 2014;33(10):1661–72. <https://doi.org/10.1007/s00299-014-1645-0>.
  45. Steward N, Kusano T, Sano H. Expression of *ZmMET1*, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in actively proliferating cells, but also with altered DNA methylation status in cold-stressed quiescent cells. *Nucleic Acids Res*. 2000;28(17):3250–9. <https://doi.org/10.1093/nar/28.17.3250>.
  46. Dai Y, Ni ZF, Dai J, Zhao T, Sun QX. Isolation and expression analysis of genes encoding DNA methyltransferase in wheat (*Triticum aestivum* L.). *Bba Gene Struct Expr*. 2005;1729(2):118–25. <https://doi.org/10.1016/j.bbexp.2005.04.001>.
  47. Wang Z, Li J, Chen S, Heng Y, Chen Z, Yang J, et al. Poaceae-specific MS1 encodes a phospholipid-binding protein for male fertility in bread wheat. *Proc Natl Acad Sci U S A*. 2017;114(47):12614–9. <https://doi.org/10.1073/pnas.1715570114>.
  48. Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur*. 2013;5(3):291–317. <https://doi.org/10.1007/s12571-013-0263-y>.
  49. Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R. Genome-wide analysis of MIK-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytol*. 2020;225(1):511–29. <https://doi.org/10.1111/nph.16122>.
  50. Tucker EJ, Baumann U, Koudiri A, Suchecki R, Baes M, Garcia M, et al. Molecular identification of the wheat male fertility gene *Ms1* and its prospects for hybrid breeding. *Nat Commun*. 2017;8(1):869. <https://doi.org/10.1038/s41467-017-00945-2>.
  51. Bai JF, Wang YK, Wang P, Duan WJ, Yuan SH, Sun H, et al. Uncovering male fertility transition responsive miRNA in a wheat photo-thermosensitive genic male sterile line by deep sequencing and degradome analysis. *Front Plant Sci*. 2017;8. <https://doi.org/10.3389/fpls.2017.01370>.
  52. Tang Z, Zhang L, Yang D, Zhao C, Zheng Y. Cold stress contributes to aberrant cytokinesis during male meiosis I in a wheat thermosensitive genic male sterile line. *Plant Cell Environ*. 2011;34(3):389–405. <https://doi.org/10.1111/j.1365-3040.2010.02250.x>.
  53. Li YF, Zhao CP, Zhang FT, Sun H, Sun DF. Fertility alteration in the photo-thermo-sensitive male sterile line B520 of wheat (*Triticum*

- aestivum* L.). *Euphytica*. 2006;151(2):207–13. <https://doi.org/10.1007/s10681-006-9141-4>.
54. Zinn KE, Tunc-Ozdemir M, Harper JF. Temperature stress and plant sexual reproduction: uncovering the weakest links. *J Exp Bot*. 2010;61(7):1959–68. <https://doi.org/10.1093/jxb/erq053>.
  55. Min L, Li Y, Hu Q, Zhu L, Gao W, Wu Y, et al. Sugar and auxin signaling pathways respond to high-temperature stress during anther development as revealed by transcript profiling analysis in cotton. *Plant Physiol*. 2014; 164(3):1293–308. <https://doi.org/10.1104/pp.113.232314>.
  56. Ali A, Li Y, Chen H, Xu PZ, Zhang HY, Chen XQ, et al. Effect of male sterile and fertile cytoplasm on nuclear DNA methylation in hybrid rice. *Acta Physiol Plant*. 2019;41(6):1–7.
  57. Chen B, Zhang Y, Lu YL, Wang J, Zhang SZ, Lan H, et al. DNA methylation analysis of sterile and fertile CMS-C hybrids and their parents in maize. *J Plant Biochem Biot*. 2016;25(1):3–11. <https://doi.org/10.1007/s13562-015-0298-6>.
  58. Ba QS, Zhang GS, Niu N, Ma SC, Wang JW. Cytoplasmic effects on DNA methylation between male sterile lines and the maintainer in wheat (*Triticum aestivum* L.). *Gene*. 2014;549(1):192–7. <https://doi.org/10.1016/j.gene.2014.01.001>.
  59. Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. 2018;361(6403):661.
  60. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–9. <https://doi.org/10.1093/molbev/msy096>.
  61. Li X, Gao S, Tang Y, Li L, Zhang F, Feng B, et al. Genome-wide identification and evolutionary analyses of bZIP transcription factors in wheat and its relatives and expression profiles of anther development related *TabZIP* genes. *BMC Genomics*. 2015;16(1):976. <https://doi.org/10.1186/s12864-015-2196-7>.
  62. Wang A, Xia Q, Xie W, Datla R, Selvaraj G. The classical Ubisch bodies carry a sporophytically produced structural protein (RAFTIN) that is essential for pollen development. *Proc Natl Acad Sci U S A*. 2003;100(24):14487–92. <https://doi.org/10.1073/pnas.2231254100>.
  63. Brautigam K, Cronk Q. DNA methylation and the evolution of developmental complexity in plants. *Front Plant Sci*. 2018;9. <https://doi.org/10.3389/fpls.2018.01447>.
  64. Lee SJ, Kim NS. Transposable elements and genome size variations in plants. *Genomics Inform*. 2014;12(3):87–97. <https://doi.org/10.5808/GI.2014.12.3.87>.
  65. Deniz O, Frost JM, Branco MR. Regulation of transposable elements by DNA modifications. *Nat Rev Genet*. 2019;20(7):432. <https://doi.org/10.1038/s41576-019-0117-3>.
  66. Freitag M, Selker EU. Controlling DNA methylation: many roads to one modification. *Curr Opin Genet Dev*. 2005;15(2):191–9. <https://doi.org/10.1016/j.gde.2005.02.003>.
  67. Mbichi RW, Wang QF, Wan T. RNA directed DNA methylation and seed plant genome evolution. *Plant Cell Rep*. 2020;39(8):983–96. <https://doi.org/10.1007/s00299-020-02558-4>.
  68. Wang RB, Lou J, Li J. A mobile restriction modification system consisting of methylases on the IncA/C plasmid. *Mobile DNA-Uk*. 2019;10:26.
  69. Bewick AJ, Niederhuth CE, Ji L, Rohr NA, Griffin PT, Leebens-Mack J, et al. The evolution of CHROMOMETHYLASES and gene body DNA methylation in plants. *Genome Biol*. 2017;18(1):65. <https://doi.org/10.1186/s13059-017-1195-1>.
  70. Yarychivskaya O, Shahabuddin Z, Comfort N, Boulard M, Bestor TH. BAH domains and a histone-like motif in DNA methyltransferase 1 (*DNMT1*) regulate de novo and maintenance methylation in vivo. *J Biol Chem*. 2018; 293(50):19466–75. <https://doi.org/10.1074/jbc.RA118.004612>.
  71. Sanchez DH, Paszkowski J. Heat-induced release of epigenetic silencing reveals the concealed role of an imprinted plant gene. *PLoS Genet*. 2014;10(11):e1004806.
  72. Kumar R, Chauhan PK, Khurana A. Identification and expression profiling of DNA methyltransferases during development and stress conditions in Solanaceae. *Funct Integr Genomics*. 2016;16(5):513–28. <https://doi.org/10.1007/s10142-016-0502-3>.
  73. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*. 2020;13(8):1194–202. <https://doi.org/10.1016/j.molp.2020.06.009>.
  74. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. 2002;30(1):325–7. <https://doi.org/10.1093/nar/30.1.325>.
  75. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>−(Delta Delta C(T))</sup> method. *Methods*. 2001;25(4):402–8. <https://doi.org/10.1006/meth.2001.1262>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

