

RESEARCH

Open Access

Genome-wide identification and expression analysis of DNA demethylase family in cotton



YANG Xiaomin, LU Xuke, CHEN Xiugui, WANG Delong, WANG Junjuan, WANG Shuai, GUO Lixue, CHEN Chao, WANG Xiaoge, WANG Xinlei and YE Wuwei*

Abstract

Background: DNA methylation is an important epigenetic factor that maintains and regulates gene expression. The mode and level of DNA methylation depend on the roles of DNA methyltransferase and demethylase, while DNA demethylase plays a key role in the process of DNA demethylation. The results showed that the plant's DNA demethylase all contained conserved DNA glycosidase domain. This study identified the cotton DNA demethylase gene family and analyzed it using bioinformatics methods to lay the foundation for further study of cotton demethylase gene function.

Results: This study used genomic information from diploid *Gossypium raimondii* JGI (D), *Gossypium arboreum* L. CRI (A), *Gossypium hirsutum* L. JGI (AD1) and *Gossypium barbadense* L. NAU (AD2) to *Arabidopsis thaliana*. Using DNA demethylase genes sequence of *Arabidopsis* as reference, 25 DNA demethylase genes were identified in cotton by BLAST analysis. There are 4 genes in the genome D, 5 genes in the genome A, 10 genes in the genome AD1, and 6 genes in the genome AD2. The gene structure and evolution were analyzed by bioinformatics, and the expression patterns of DNA demethylase gene family in *Gossypium hirsutum* L. were analyzed. From the phylogenetic tree analysis, the DNA demethylase gene family of cotton can be divided into four subfamilies: REPRESSOR of SILENCING 1 (ROS1), DEMETER (DME), DEMETER-LIKE 2 (DML2), and DEMETER-LIKE3 (DML3). The sequence similarity of DNA demethylase genes in the same species was higher, and the genetic relationship was also relatively close. Analysis of the gene structure revealed that the DNA demethylase gene family members of the four subfamilies varied greatly. Among them, the number of introns of ROS1 and DME subfamily was larger, and the gene structure was more complex. For the analysis of the conserved domain, it was known that the DNA demethylase family gene member has an endonuclease III (ENDO3c) domain.

Conclusion: The genes of the DNA demethylase family are distributed differently in different cotton species, and the gene structure is very different. High expression of ROS1 genes in cotton were under abiotic stress. The expression levels of ROS1 genes were higher during the formation of cotton ovule. The transcription levels of ROS1 family genes were higher during cotton fiber development.

Keywords: DNA demethylase, *Gossypium arboreum* L., *Gossypium barbadense* L., *Gossypium hirsutum* L., *Gossypium raimondii*

* Correspondence: yew158@163.com

Institute of Cotton Research of Chinese Academy of Agricultural Sciences / State Key Laboratory of Cotton Biology / Key Laboratory for Cotton Genetic Improvement, MOAA, Anyang 455000, Henan, China



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. 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.

Background

DNA methylation is an epigenetic modification widely found in bacteria, plants and animals (Chen et al. 2015; Manning et al. 2006; Zhong et al. 2013). It involves gene silencing, transposon suppression, genomic imprinting, X chromosome inactivation, cell differentiation, embryo development and other growth and development processes (Fu et al. 2014; Xie et al. 2013; Macdonald 2012; Bala et al. 2013). DNA methylation is a necessary set up for the normal growth and development of organisms. DNA methylation can affect the stability of the genome, regulate gene expression, and maintain growth and development (Wang and Xu 2014; Zhang et al. 2018; Cokus et al. 2008). It has been shown that DNA methylation is a dynamic process that can be regulated according to different development periods or environment conditions (Bartels et al. 2018). DNA methylation is regulated by different pathways to establish methylation and reverse methylation (Ja and Se 2010). Most of the DNA methylation occur on the fifth carbon atom (C5) of the cytosine in the symmetrical CG site, but also exists in the CHG and CHH (H = A, C or T) sequences in plants (Stroud et al. 2014). It has now been found that there are two methods of DNA methylation in the plant that maintain methylation and *de novo* methylation (Jullien et al. 2012). There are four types of C5-MTases in plants, including Methyltransferase (MET) family, the Chromomethylase (CMT) family, the Domains Rearranged Methyltransferase (DRM) family and Dnmt2 (Wang et al. 2016; Pavlopoulou and Kossida 2007).

There are two possibilities for the deletion of methylated cytosine, one caused by improper manipulation during DNA replication and the other by DNA demethylase activity. DNA demethylase contains a bifunctional DNA glycosidase domain (Tomkova et al. 2018). The DNA glycosidase domain not only directly cleave methylcytosine, but also cleave the DNA backbone at the abasic site. Then DNA polymerase and DNA ligase fill the base vacancy with unmodified cytosine (McCullough et al. 1989). There are four main types of DNA demethylases, ROS1, DME, DML2, and DML3. DME is unique in dicotyledons and is involved in embryo and endosperm development, and is essential for demethylation of the entire genome and transposon reactivation (Frost et al. 2018). DML2, DML3 and ROS1 are expressed in vegetative cells (Jon et al. 2007). DML2 and DML3 are capable of removing unwanted methylation at specific sites (Zhu et al. 2007). ROS1 can inhibit the methylation at gene promoters (Gong et al. 2002). ROS1b can reactivate Tos17 by remove DNA methylation (La et al. 2011). ROS1-mediated DNA demethylation can cause deconstruction of 5S rDNA chromatin, causing plants to respond to biotic and abiotic stresses,

and also prevent RNA-directed DNA methylation (Movahedi et al. 2018). DNA demethylase plays an important part in removing DNA methylation. There are related studies in *A. thaliana* and rice (Penterman et al. 2007; Choi et al. 2004; Zemach et al. 2010). As an important fiber and oil crop, cotton plays a pivotal role in China's national economy (Chen et al. 2017). It is an critical problem in cotton planting today on how to improve the quality of cotton fiber and the resistance of plants to different stresses. The results of DNA methylation research are important for studying stress resistance mechanisms and improving cotton stress resistance.

Materials and methods

Identification of cotton DNA demethylase family members

Using the *Arabidopsis* DNA demethylase protein sequence (AT1G05900.2, AT2G36490, AT2G31450.1, AT3G10010.1, AT3G47830.1, AT4G34060.1, AT5G04560.2) as a motif, Blastp homologous alignments were performed in CottonFGD (<https://cottonfgd.org/>) with $P < 0.001$ and similarity $> 40\%$ with the order to identify the candidate protein and obtains the DNA demethylase family member gene locus name. Using gene locus name of candidate DNA demethylase respective cds sequence, amino acid sequence, genome sequence of identified candidate DNA demethylase were downloaded from *Gossypium arboreum* L., CRI *G. raimondii*, JGI *G. hirsutum* L., JGI *G. barbadense* L., NAU database using respective gene mapping ID. The protein sequences of the candidate genes were analyzed by using SMART (<http://smart.embl-heidelberg.de/>) to ensure that each candidate gene contained a DNA glycosidase domain. Subcellular localization prediction was performed on the Cello website (<http://cello.life.nctu.edu.tw/>). Protein analysis was performed by using ProtParam (<https://web.expasy.org/protparam/>) to obtain isoelectric points.

Cotton DNA demethylase family evolution analysis

The *Arabidopsis thaliana* amino acid sequence was used as a reference, and the $E < e^{-5}$ was used as a threshold to obtain the hairy tree in the Phytozome v12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>) database of Homologous sequence. Multi-sequence alignment (Clustal W) of DNA demethylase genomic sequence of *G. raimondii*, *G. arboreum* L., *G. hirsutum* L., and *G. barbadense* L., to *Arabidopsis* were conducted by using MEGA7.0 software, and adopted the adjacency method (Neighbor-Joining, NJ) to create a tree with Bootstrap at 1 000. The same method was also used to construct the phylogenetic tree of DNA demethylase protein family of *G. raimondii*, *G. arboreum* L., *G. hirsutum* L., *G. barbadense* L., *Populus trichocarpa* and, *A. thaliana*.

Prediction of the basic structure of DNA demethylase gene family

The basic physicochemical properties of the amino acid sequence of the family protein sequence were analyzed by the online software ProParam (<https://web.expasy.org/protparam/>) in ExPASy. The gene structure map was drawn by GSDS2.0 online (<http://gsds.cbi.pku.edu.cn/>). Motif analysis was performed by the online tool MEME (<http://meme.nbcr.net/meme/>). The physical map of the chromosome was drawn by the software Mapinspect.

Analysis of expression patterns of cotton DNA demethylase gene under stress conditions

The FPKMs (Fragments per kilobase million) of the DNA demethylase gene in *G. hirsutum* L., under cold, heat, drought and salt stress conditions, ovule development formation, and fiber development stage were obtained from the cottonFGD database (<https://cottonfgd.org/>). *G. hirsutum* L., (AD1) Genome - Texas Interim release UTX-JGI v1.1 genome assembly is made available

through a “Reserved Analyses” restriction. The FRKM value can reflect the level of gene expression and a heat-map of gene expression was obtained using the tool HemI software.

Results and analysis

Whole genome identification of cotton DNA demethylase family members

Twenty-five DNA demethylases were identified from the whole genome of cotton by multiple sequence alignment. There were four DNA demethylases genes in genome D and five DNA demethylases genes in genome A, respectively. According to their position on chromosome, they were named *GaDM1-GaDM4* and *GrDM1-GrDM5*, respectively. Ten genes of DNA demethylases were identified in genome AD1, they were named *GhDM1-GhDM10*, and six DNA demethylases genes were identified in genome AD2 named *GbDM1-GbDM6*. Most of the DNA demethylases in the 4 cotton species were located on the chromosomes, and only *GbDM5* is not on any chromosome. The DNA

Table 1 Basic characteristic of DNA MT genes in cotton genome

NAME	ID	Chromosome location	cds	AA	pl	Subcellular prediction
GaDM1	Ga01G1291	LOC:Chr01:30214974..30224853-	5 805	1 935	7.08	Outer membrane
GaDM2	Ga04G0941	LOC:Chr04:30894670..30904909+	5 679	1 892	6.10	Outer membrane
GaDM3	Ga09G1562	LOC:Chr09:72800030..72815114+	5 628	1 875	6.59	Outer membrane
GaDM4	Ga10G1602	LOC:Chr10:88365871..88371004+	1 155	384	9.47	Periplasmic
GaDM5	Ga12G0005	LOC:Chr12:52354..55562+	867	288	8.31	Cytoplasmic
GrDM1	Gorai.002G145100	LOC:Chr02:26575884..26586905+	5 805	1 935	7.56	Outer membrane
GrDM2	Gorai.008G298100	LOC:Chr08:57045889..57049432-	867	288	8.92	Cytoplasmic
GrDM3	Gorai.009G408300	LOC:Chr09:60800526..60811646-	5 682	1 893	6.22	Outer membrane
GrDM4	Gorai.011G146700	LOC:Chr11:23692355..23697691-	1 164	387	9.43	Periplasmic
GhDM1	Gohir.A01G129700	LOC:A01:40917250..40925834+	4 434	1 477	6.62	Outer membrane
GhDM2	Gohir.A04G066400	LOC:A04:50948051..50958651-	5 673	1 890	6.13	Outer membrane
GhDM3	Gohir.A09G136700	LOC:A09:70131721..70147152+	5 703	1 900	6.75	Outer membrane
GhDM4	Gohir.A10G133700	LOC:A10:69911910..69917262+	1 155	384	9.47	Outer membrane
GhDM5	Gohir.A12G276500	LOC:A12:102951986..102955383-	867	288	8.81	Cytoplasmic
GhDM6	Gohir.D01G119000	LOC:D01:27313448..27323218+	5 442	1 813	7.69	Outer membrane
GhDM7	Gohir.D04G104500	LOC:D04:36583949..36594661-	5 682	1 893	6.12	Outer membrane
GhDM8	Gohir.D09G137800	LOC:D09:42880181..42896456-	5 850	1 949	7.36	Outer membrane
GhDM9	Gohir.D10G133000	LOC:D10:25478979..25489178-	1 155	384	9.48	Periplasmic
GhDM10	Gohir.D12G277500	LOC:D12:63774832..63778258-	867	288	8.71	Cytoplasmic
GbDM1	GOBAR_AA17009	LOC:A01:39978467..39988347+	5 529	1 842	6.73	Outer membrane
GbDM2	GOBAR_AA10083	LOC:A10:69981030..69987339+	1 134	377	8.87	Periplasmic
GbDM3	GOBAR_AA31875	LOC:A12:103267706..103272791-	879	292	9.04	Cytoplasmic
GbDM4	GOBAR_DD14116	D01: 28312428–28 322 663: +	5 685	1 894	7.09	Outer membrane
GbDM5	GOBAR_DD06876	LOC:scaffold_1890..D0994272..102980+	5 691	1 896	6.76	Outer membrane
GbDM6	GOBAR_DD31695	LOC:D12:58196184..58199776-	801	266	8.31	Cytoplasmic

demethylase protein in cotton consists of 266~1 949 amino acids, of which GhDM8 is up to 1 949 amino acids and GbDM7 contains only 266 amino acids. The isoelectric point (pI) of DNA demethylase gene in cotton ranged from 6.10 to 9.48. The isoelectric point of *GaDM2* was lowest at 6.10, and that of *GhDM9* was highest at 9.48. Subcellular localization predictions showed that most of the DNA demethylase genes in cotton were located at outer membrane. Only a few genes were located in cytoplasm and periplasm (Table 1).

Multi-sequence alignment and evolution analysis

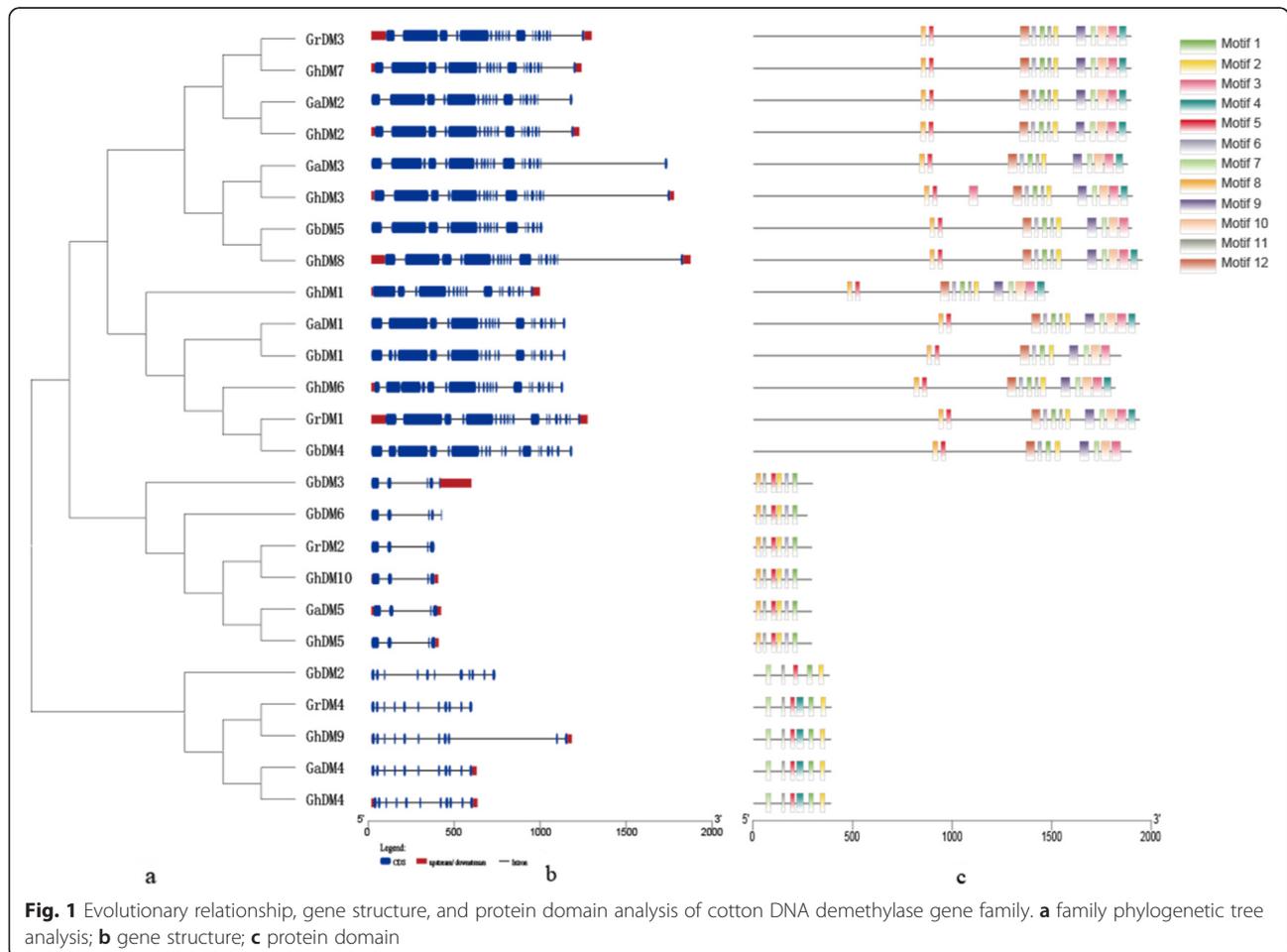
To understand the evolutionary relationship of DNA demethylases in genome A, D, AD1 and genome AD2, multiple sequence alignments were performed on 25 DNA demethylase family members and a phylogenetic tree was constructed (Fig. 1a). The DNA demethylases in cotton were divided into four subfamilies, ROS1, DME, DML2, and DML3. The ROS1 subfamily had 8 members. i.e., genome D, A, AD1, and genome AD2 with 2, 1, 4, and 1, respectively; DME had 6 members, and genome D, A, AD1, and genome AD2 had 1, 1, 2

and 2, respectively. DML2 has 6 members, genome D, A, AD1, and genome AD2 had 1, 1, 2 and 2, respectively; DML3 has 5 members, and genome D, A, AD1 and genome AD2 had 1, 1, 2 and 1, respectively.

Gene structure analysis and protein domain analysis of cotton DNA demethylase family genes

Gene structure analysis is an important strategy to study genetic evolution. Analysis of the number of introns and exons in the DNA demethylase family members in D, A, AD1 and AD2 (Fig. 1b) showed that the number of DNA demethylase gene exons in cotton differed greatly. Among them, *GrDM2*, *GaDM5*, *GhDM5* and *GhDM10* have only 4 exons; *GhDMT6* has 21 exons.

The motif analysis of 25 DNA demethylases in cotton was shown in Fig. 1c. The cotton DNA demethylase gene contains 12 motif structures, of which Motif1, 2, 5, 11 together constitute a conserved ENDO3c glycosidase domain. Among different cotton genomes, the ROS1 and DME families were identical, containing Motif 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12 conserved sequences, which constitute ENDO3c, FES, Pfam: Perm-CXXC,



Pfam: RRM_DME domain. The DML2 family contains six Motif structures, namely Motif 1, 2, 5, 6, 8, and 11, which constitute the ENDO3c and HhH1 domains; the DML3 family contains six Motif structures, namely Motif 1, 2, 4, 5, 7, and 11. It constitutes the ENDO3c, HhH1, and FES domains. There were large differences in protein structure between different subfamilies, presumably due to the long-term evolution of genes.

Distribution of cotton DNA demethylase gene family members on chromosomes

The distribution of genes on chromosomes provides an important basis for studying the evolution and function of gene families. Combining the chromosomal information of the 4 genomes of cotton and the position of each DNA demethylase gene on the chromosome, the distribution map of the cotton demethylase gene on the chromosome was obtained (Fig. 2). The *G. arboreum* L. genome contains five genes, which were distributed on chromosomes 1, 4, 9, 10, and 12, respectively (Ga). The *G. raimondii* genome contains 4 GaDM genes, which were distributed on chromosomes 2, 8, 9, and 11, respectively (Gr). There were 10 GhDM genes in the *G. hirsutum* L. genome, and these 10 genes were evenly distributed on the group A and group D chromosomes. One of the genes was distributed on chromosomes 1, 4, 9, 10, and 12 in the genome A and also on same chromosomes in genome D (Gh, respectively). There were 6 GbDM genes in the *Gossypium barbadebse* L. genome and these 6 genes were unevenly distributed on the genome A and the genome D chromosomes. One of the genes was distributed on chromosomes 1, 10, and 12 in the genome A. Two genes were distributed on chromosomes 1 and 12 in the genome D, respectively, and 1 gene was mapped to scaffold_1890 (Gb).

Evolutionary relationship between cotton DNA demethylase family and other plant DNA demethylase family

Construction of phylogenetic trees revealed the homologous and evolutionary relationships of DNA demethylase genes from different species. The cotton DNA demethylase family members were aligned with the amino acid sequences of DNA demethylase members in *A. thaliana* and *P. trichocarpa*, and then a phylogenetic tree was constructed by MEGA 7.0 (Fig. 3). The results showed that the DNA demethylases in the cotton genomes had smaller evolution distance compared with other crops. The DNA demethylase in cotton is closely related to the *P. trichocarpa* on each branch, indicating that they have similar functions. *A. thaliana* and *P. trichocarpa* have a different type of gene due to evolutionary differences between species.

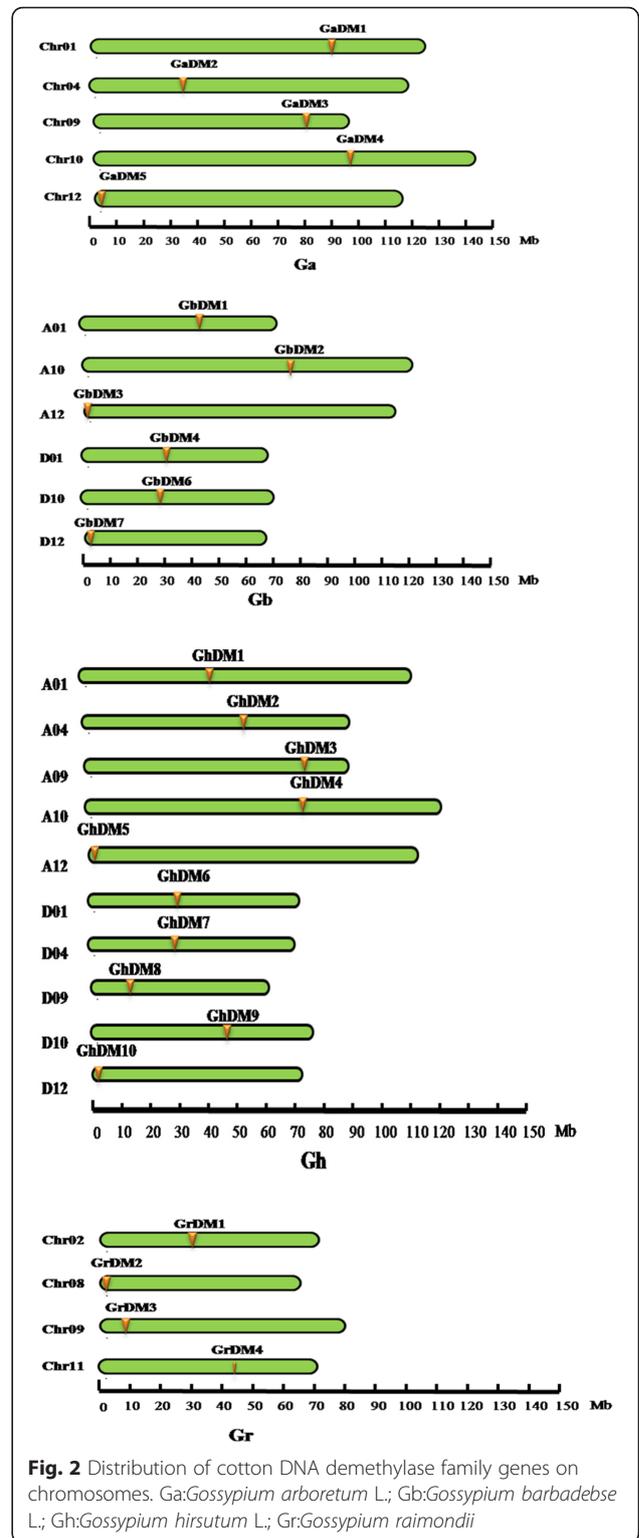
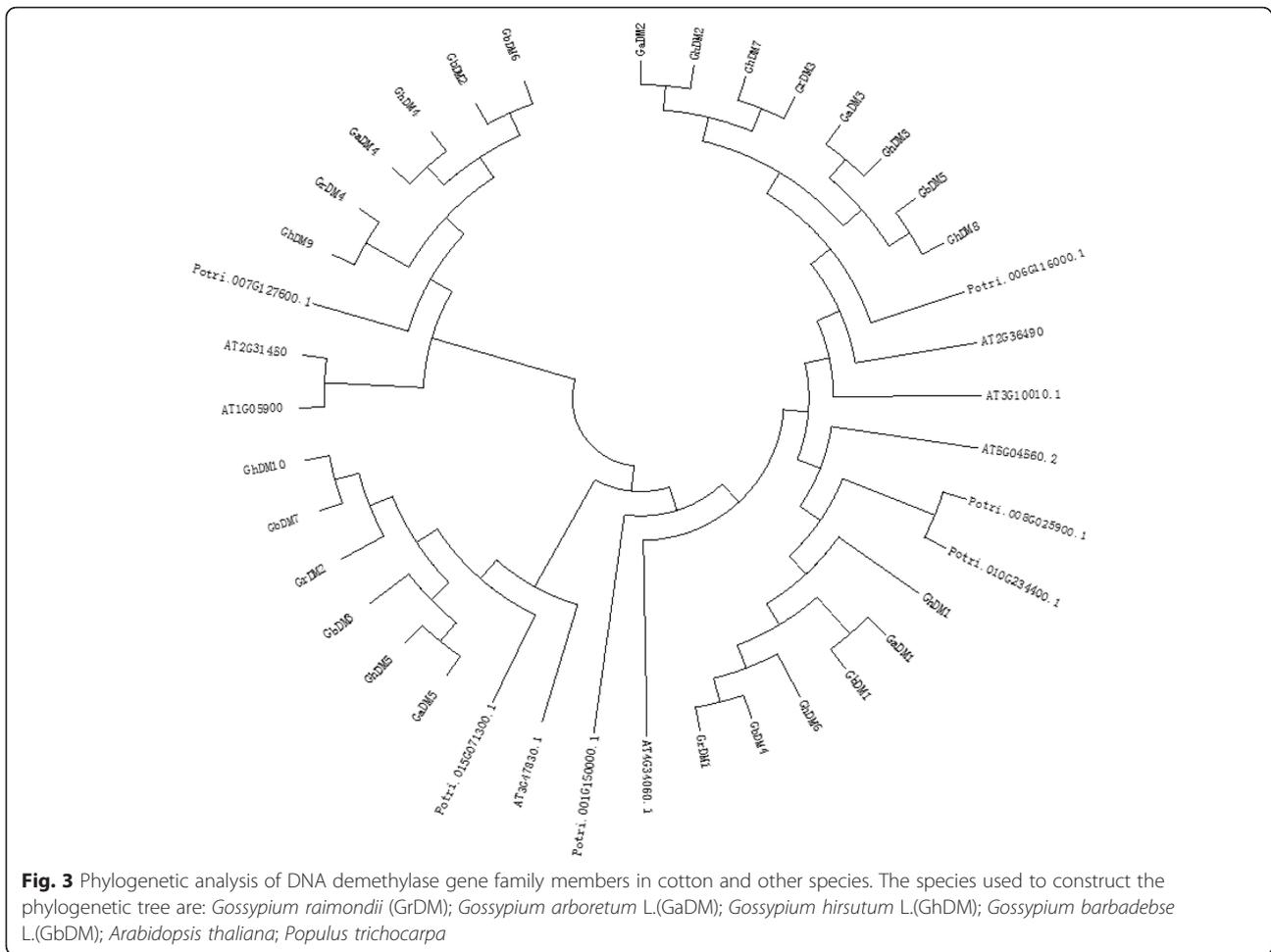


Fig. 2 Distribution of cotton DNA demethylase family genes on chromosomes. Ga:Gossypium arboreum L.; Gb:Gossypium barbadebse L.; Gh:Gossypium hirsutum L.; Gr:Gossypium raimondii

Expression of DNA methyltransferase gene in stresses and ovule formation as well as fiber formation in cotton
 The FPKM of the terpene cotton TM-1 DNA demethylase gene was downloaded from the CottonFGD



database to construct an expression map of cotton DNA demethylase gene under abiotic stresses conditions, as well as different developmental stages during ovule formation and fiber development. The results showed that three genes *GhDM2*, *GhDM4* and *GhDM7* were highly expressed under cold, heat, drought and salt stress. Under different kinds of stress, the expression levels of DNA demethylase genes were different. The expression of *GhDM2* gene was down-regulated under cold stress, but it was up-regulated under drought, heat and salt stresses. The gene expression of ROS1 and DML3 families were up-regulated, and the expression levels of DME and DML2 family genes were down-regulated when cotton was under stress. The same family gene has different expression levels under the same stress. The expression levels of *GhDM2* and *GhDM7* genes were up-regulated, but the expression levels of *GhDM3* and *GhDM8* genes were down-regulated than those of ROS1 family genes under heat stress (Fig. 4a).

During the ovule formation process, the expression of *GhDM2*, *GhDM3*, *GhDM7* and *GhDM8* genes in ROS1

family were up-regulated, while the expression of *GhDM1* and *GhDM6* genes in DME family were down-regulated. The *GhDM6* gene was not expressed, while other demethylase genes were up-regulated in 3 days before-anthesis and in anthesis. *GhDM2* gene was up-regulated in 3 days before-anthesis and in 20 days post-anthesis (Fig. 4b).

The expression level of DNA demethylase gene was relatively low during fiber development, *GhDM5* and *GhDM6* were not expressed. At the 25th day of fiber formation, the *GhDM2* gene expression level was higher, but the expression of *GhDM7* gene in the ROS1 family could not be detected. At the 15th day of fiber formation, the expression of *GhDM7* gene was higher, and the expression of *GhDM2*, *GhDM3*, and *GhDM8* genes in the ROS1 family could not be detected either (Fig. 4c).

Discussion

With the expanding information of cotton genome, we used the comparative genomics research method to determine the demethylase gene in cotton, conducted

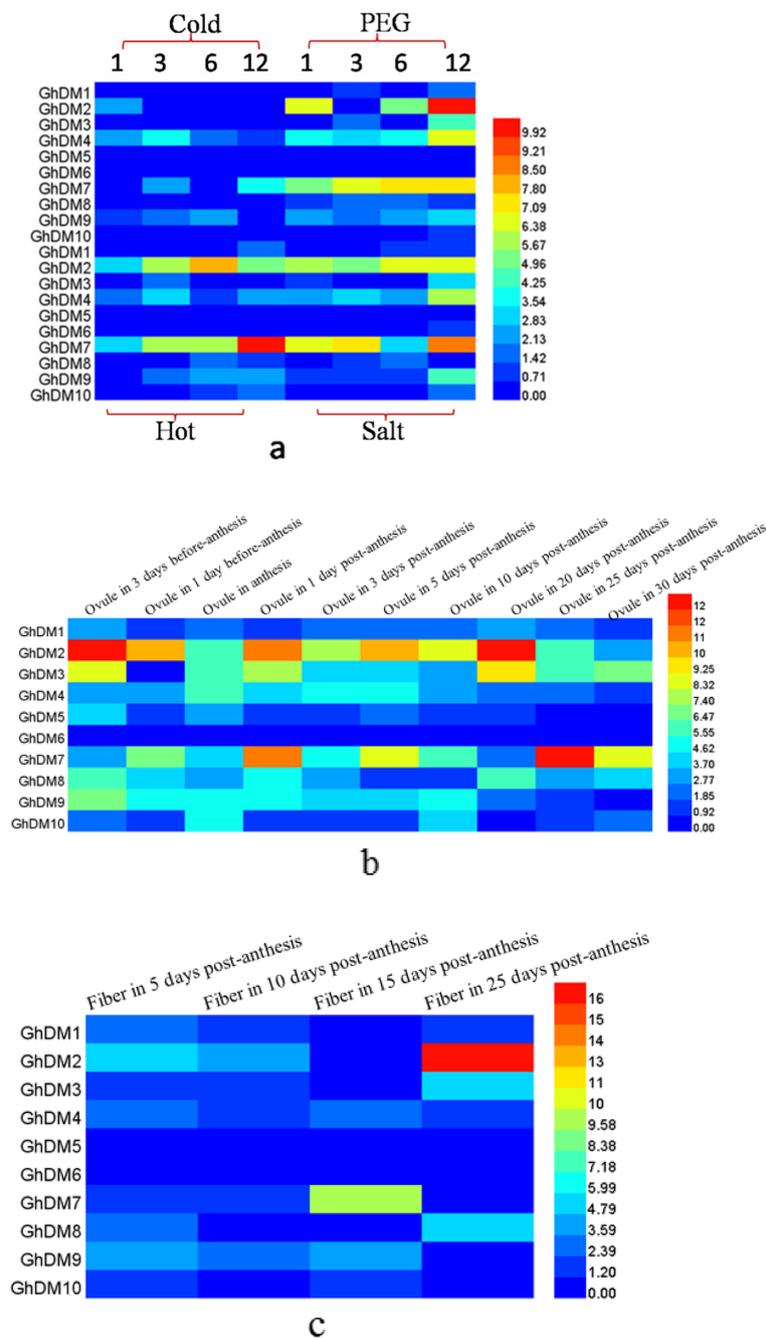


FIG. 4 Expression of cotton DNA demethylase gene. **a** abiotic stress; **b** ovule development formation process; **c** fiber development stage

sequence analysis, phylogenetic analysis, and examined expression pattern under different conditions. DNA methylation is not only involved in the regulation of gene expression but also in maintaining genome stability (Dai et al. 2014). DNA demethylase can remove methylation and regulate gene expression, which is closely related to stress resistance (Colot and Rossignol 1999). With the completion of the cotton genome sequencing work, it is convenient to study the cotton demethylase

gene from the whole genome. DNA methylation is an important epigenetic process that affects many biological processes, (Dennis 2000). DNA demethylation is a complex process which mechanism is unclear. DNA demethylase plays an important role in epigenetics. At present, it is generally believed that there are five mechanisms for DNA demethylation: base excision repair mechanism relying on DNA demethylase, base excision repair, mismatch excision repair of methyl cytidine

deamination coupling G/T, demethylation by hydrolysis and oxidative demethylation (Cao et al. 2012). DNA demethylase is essential in all mechanisms.

In this study, we investigated the structure, evolution, collinearity and expression of DNA demethylase genes in cotton. The results showed that DNA demethylase contains four conserved motif structures, which is consistent with the study in angiosperms (Liu et al. 2014). There are four types of DNA demethylase in cotton that are identical to *Arabidopsis*. The DNA demethylase gene was evenly distributed in four cotton species and evolved consistently. The ROS1 gene subfamily has been replicated several times to produce new functional and sub-functionalization of genes; this provides clues for further study of the role and mechanism of different DNA demethylase genes. Evolutionary analysis revealed that DNA demethylase genes differed greatly among different species or different families of the same species.

Plants respond rapidly to abiotic stress through DNA methylation machinery. The DNA demethylase gene plays an important role in regulating gene expression. The results showed that the DNA demethylase gene responded to cold, heat, drought and salt, abiotic stresses (Fig. 4a). The expression level of DML-like demethylase gene in *A. thaliana* was increased during stress, and the expression levels of ROS1 and DML3 demethylase family genes in cotton were higher, but DME and DML2 demethylase family were lower (Tzung-Fu et al. 2009). DNA demethylase indirectly responds to stresses by regulating DNA methylation levels (Sanchez and Paszkowski 2014). The DME gene in *Arabidopsis* is preferential expressed in the central and companion cells of the female gametophyte, which affects the development of embryo and endosperm; before the expression of *GhDM6* gene in DME family is low during cotton ovule formation (Choi et al. 2002). The function of the DNA demethylase gene changed with evolution (Agius et al. 2006).

DNA methylation is essential in regulating plant development and response to environmental stimulus, but how the DNA methylase and demethylase participate in various responses is a complex process and the mechanism is still unclear. The differential expression analysis of the demethylase gene showed that under different kinds of abiotic stress, the expression level of the demethylase gene changed greatly, and some key genes may be demethylated. The response is critical, indicating that DNA methylation is most likely involved in the effects of the environment on cotton growth and development. DNA demethylase gene expression was higher during cotton ovule formation, indicating that DNA methylation may have a regulatory role in cotton ovule formation. Therefore, this study provided some clues for the roles of DNA methylation in cotton in response to

stress, as well as its developmental role in ovule formation and fiber development, and provided a base for further exploration of epigenetic regulation mechanisms during cotton development.

Conclusions

DNA demethylase gene family plays a significant role in plant growth and development. The high expression of cotton DNA demethylase gene in abiotic stress, ovule formation and fiber development stage indicates that the demethylase family plays an important role in cotton growth and development. The results of this study laid the foundation for excavating functional genes and further studying the stress resistance mechanism of cotton.

Abbreviations

C-5: 5-methylcytosine; -CH₃: Methyl group; CMT: Chromomethylase; DNMT2: DNA methyltransferase2; DRM: Domains rearranged methyltransferase; FPKM: Fragments per kilobase million; Ga: *Gossypium arboreum* L.; Gb: *Gossypium barbadense* L.; Gh: *Gossypium hirsutum* L.; Gr: *Gossypium raimondii*; MET: Methyltransferase; pI: Isoelectric Point

Acknowledgments

Not applicable.

Authors' contributions

Yang XM and Ye WW conceived and designed the experiments; Yang XM, Wang XG and Lu XK performed the experiments and collected the data; Chen XG, Wang S, Wang DL and Wang JJ obtained funding and Guo LX, Wang XL and Chen C contributed reagents/materials/analysis tools; Yang XM and Ye WW revised the paper. All authors read and approved the final manuscript.

Authors' information

Not applicable.

Funding

This study was funded by the National Key Research and Development Program of China (2018YFD0100401).

Availability of data and materials

All data generated or analyzed in this study included in published article and additional files.

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: 18 April 2019 Accepted: 28 August 2019

Published online: 09 October 2019

References

- Agius F, Kapoor A, Zhu JK. Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci U S A*. 2006;103(31):11796–801. <https://doi.org/10.1073/pnas.0603563103>.
- Bala TN, Brahmachary M, Garg P, et al. DNA methylation profiling in X; autosome translocations supports a role for L1 repeats in the spread of X chromosome inactivation. *Hum Mol Genet*. 2013;23(5):1224–36. <https://doi.org/10.1093/hmg/ddt553>.
- Bartels A, Han Q, Nair P, et al. Dynamic DNA methylation in plant growth and development. *Int J Mol Sci*. 2018;19(7):2144. <https://doi.org/10.3390/ijms19072144>.

- Cao L, Zhang Y, Lan X, et al. Progress in study on mechanism of DNA demethylation. *Adv Mod Biomed Sci.* 2012;12(1):160–2. <https://doi.org/10.13241/j.cnki.pmb.2012.01.002>.
- Chen W, Kong J, Qin C, et al. Requirement of chromomethylase3 for somatic inheritance of the spontaneous tomato epimutation colourless non-ripening. *Sci Rep.* 2015;5(9192):9192. <https://doi.org/10.1038/srep09192>.
- Chen X, Lu X, Shu N, et al. Targeted mutagenesis in cotton (*Gossypium hirsutum* L.) using the CRISPR/Cas9 system. *Sci Rep.* 2017;7:44304. <https://doi.org/10.1038/srep44304>.
- Choi Y, Gehring M, Johnson L, et al. Demeter, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell.* 2002;110(1):33–42. [https://doi.org/10.1016/s0092-8674\(02\)00807-3](https://doi.org/10.1016/s0092-8674(02)00807-3).
- Choi Y, Harada JJ, Goldberg RB, et al. An invariant aspartic acid in the DNA glycosylase domain of demeter is necessary for transcriptional activation of the imprinted meeda gene. *Proc Natl Acad Sci U S A.* 2004;101(19):7481–6. <https://doi.org/10.1073/pnas.0402328101>.
- Cokus SJ, Feng S, Zhang X, 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>.
- Colot V, Rossignol JL. Eukaryotic DNA methylation as an evolutionary device. *Bioessays.* 1999;21(5):402. [https://doi.org/10.1002/\(sici\)1521-1878\(199905\)21:5<402::aid-bies7>3.0.co;2-b](https://doi.org/10.1002/(sici)1521-1878(199905)21:5<402::aid-bies7>3.0.co;2-b).
- Dai HF, Eu H, A LGL, et al. Analysis of salt tolerance of cotton of different genotypes at seedling stage and screening of its identification indexes. *Sci Agric Sin.* 2014;47(7):1290–300. <https://doi.org/10.3864/j.issn.0578-1752.2014.07.005>.
- Dennis ES. DNA methylation, a key regulator of plant development and other processes. *Curr Opin Genet Dev.* 2000;10(2):217. [https://doi.org/10.1016/S0959-437X\(00\)00061-7](https://doi.org/10.1016/S0959-437X(00)00061-7).
- Frost JM, Kim MY, Park GT, et al. FACT complex is required for DNA demethylation at heterochromatin during reproduction in *Arabidopsis*. *Proc Natl Acad Sci U S A.* 2018;115(20):E4720–9. <https://doi.org/10.1073/pnas.1713333115>.
- Fu Y, Kawabe A, Etcheverry M, et al. Mobilization of a plant transposon by expression of the transposon-encoded anti-silencing factor. *EMBO J.* 2014; 32(17):2407–17. <https://doi.org/10.1038/emboj.2013.169>.
- Gong Z, Morales-Ruiz T, Ariza RR, et al. ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell.* 2002;111(6): 803–14. [https://doi.org/10.1016/s0092-8674\(02\)01133-9](https://doi.org/10.1016/s0092-8674(02)01133-9).
- JA L, SE J. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010;11(3):204–20. [https://doi.org/10.1016/s0092-8674\(02\)01133-9](https://doi.org/10.1016/s0092-8674(02)01133-9).
- Jon P, Daniel Z, Hoe HJ, et al. DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci.* 2007;104(16):6752–7. <https://doi.org/10.1073/pnas.0701861104>.
- Jullien PE, Susaki D, Yelagandula R, et al. 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>.
- La H, Ding B, Mishra GP, et al. A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon Tos17 and promotes its transposition in rice. *Proc Natl Acad Sci.* 2011;108(37):15498–503. <https://doi.org/10.1073/pnas.1112704108>.
- Liu Q, Xue Q, Xu J. Phylogenetic analysis of DNA demethylase genes in angiosperm. *Hereditas.* 2014;36(3):276–85. <https://doi.org/10.3724/SP.J.1005.2014.0276>.
- Macdonald WA. Epigenetic mechanisms of genomic imprinting: common themes in the regulation of imprinted regions in mammals, plants, and insects. *Genet Res Int.* 2012;2012(1):585024. <https://doi.org/10.1155/2012/585024>.
- Manning K, Tör M, Poole M, et al. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat Genet.* 2006;38(8):948–52. <https://doi.org/10.1038/ng1841>.
- Mccullough AK, And MLD, Lloyd RS. Initiation of base excision repair: glycosylase mechanisms and structures. *Annu Rev Biochem.* 1989;68(1):255–85. <https://doi.org/10.1146/annurev.biochem.68.1.255>.
- Movahedi A, Ming S, Zhang J, et al. Functional analyses of PtROS1 -RNAi in poplars and evaluation of its effect on DNA methylation. *J Plant Biol.* 2018; 61(61):227–40. <https://doi.org/10.1007/s12374-017-0410-7>.
- Pavlopoulou A, Kossida S. Plant cytosine-5 DNA methyltransferases: structure, function, and molecular evolution. *Genomics.* 2007;90(4):530–41. <https://doi.org/10.1016/j.ygeno.2007.06.011>.
- Penterman J, Zilberman D, Jin HH, et al. DNA demethylation in the *Arabidopsis* genome. *Proc Natl Acad Sci.* 2007;104(16):6752–7. <https://doi.org/10.1073/pnas.0701861104>.
- 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. <https://doi.org/10.1371/journal.pgen.1004806>.
- Stroud H, Do T, Du J, et al. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat Struct Mol Biol.* 2014;21(1):64–72. <https://doi.org/10.1038/nsmb.2735>.
- Tomkova M, McClellan M, Kriaucionis S, et al. DNA replication and associated repair pathways are involved in the mutagenesis of methylated cytosine. *DNA Repair.* 2018;62:1–7. <https://doi.org/10.1016/j.dnarep.2017.11.005>.
- Tzung-Fu H, Ibarra CA, Pedro S, et al. Genome-wide demethylation of *Arabidopsis* endosperm. *Science.* 2009;324(5933):1451–4. <https://doi.org/10.1126/science.1172417>.
- Wang P, Gao C, Bian X, 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. <https://doi.org/10.3389/fpls.2016.00007>.
- Wang R, Xu J. Genomic DNA methylation and histone methylation. *Hereditas.* 2014;36(3):191–9. <https://doi.org/10.3724/SP.J.1005.2014.0191>.
- Xie M, Hong C, Zhang B, et al. DNA hypomethylation within specific transposable element families associates with tissue-specific enhancer landscape. *Nature Genet.* 2013;45(7):836. <https://doi.org/10.1038/ng.2649>.
- Zemang A, Kim MY, Silva P, et al. Local DNA hypomethylation activates genes in rice endosperm. *Proc Natl Acad Sci.* 2010;107(43):18729–34. <https://doi.org/10.2307/25748557>.
- Zhang H, Lang Z, Zhu JK. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol.* 2018;19(Suppl. 1). <https://doi.org/10.1038/cr.2011.23>.
- Zhong S, Fei Z, Chen YR, et al. Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat Biotechnol.* 2013;31(2):154–9. <https://doi.org/10.1038/nbt.2462>.
- Zhu J, Kapoor A, Sridhar VV, et al. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in *Arabidopsis*. *Curr Biol.* 2007;17(1):54–9. <https://doi.org/10.1016/j.cub.2006.10.059>.

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

