

RESEARCH

Open Access



# Genome-wide identification of MBD gene family members in *Eleutherococcus senticosus* with their expression motifs under drought stress and DNA demethylation

Shuo Wang<sup>1</sup>, Jing Dong<sup>1</sup>, Xue-Lei Zhao<sup>1</sup>, Xin Song<sup>1</sup>, Yue-Hong Long<sup>1\*</sup> and Zhao-Bin Xing<sup>1\*</sup>

## Abstract

**Background** Methyl-binding domain (MBD) is a class of methyl-CpG-binding domain proteins that affects the regulation of gene expression through epigenetic modifications. *MBD* genes are not only inseparable from DNA methylation but have also been identified and validated in various plants. Although *MBD* is involved in a group of physiological processes and stress regulation in these plants, *MBD* genes in *Eleutherococcus senticosus* remain largely unknown.

**Results** Twenty *EsMBD* genes were identified in *E. senticosus*. Among the 24 chromosomes of *E. senticosus*, *EsMBD* genes were unevenly distributed on 12 chromosomes, and only one tandem repeat gene existed. Collinearity analysis showed that the fragment duplication was the main motif for *EsMBD* gene expansion. As the species of Araliaceae evolved, *MBD* genes also evolved and gradually exhibited different functional differentiation. Furthermore, cis-acting element analysis showed that there were numerous cis-acting elements in the *EsMBD* promoter region, among which light response elements and anaerobic induction elements were dominant. The expression motif analysis revealed that 60% of the *EsMBDs* were up-regulated in the 30% water content group.

**Conclusions** By comparing the transcriptome data of different saponin contents of *E. senticosus* and integrating them with the outcomes of molecular docking analysis, we hypothesized that *EsMBD2* and *EsMBD5* jointly affect the secondary metabolic processes of *E. senticosus* saponins by binding to methylated CpG under conditions of drought stress. The results of this study laid the foundation for subsequent research on the *E. senticosus* and *MBD* genes.

**Keywords** Methyl-binding domain protein, *Eleutherococcus senticosus*, Cis-acting element, Motif

\*Correspondence:

Yue-Hong Long  
longyh@ncst.edu.cn  
Zhao-Bin Xing  
xingzb@ncst.edu.cn

<sup>1</sup>College of Life Sciences, North China University of Science and Technology, Tangshan, China



© The Author(s) 2023. **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.

## Background

*Eleutherococcus senticosus* (Rupr. et Maxim) Maxim, also known as Siberian ginseng, is a valuable medicinal plant [1]. The extracts of its roots and bark have various physiological effects, such as modulation of the immune response [2]. Among the various bioactive constituents of *E. senticosus*, triterpenoid saponins, flavonoids, coumarins, and lignans play important roles in regulating these physiological activities [3]. The farnesyl pyrophosphate synthase [4] (FPS), squalene synthase (SS), squalene epoxidase [5, 6] (SE), and mevalonate pyrophosphate decarboxylase [7] (MDD) of *E. senticosus* are key enzymes that catalyze the biosynthesis of triterpenoid saponins. In previous studies, it was found that DNA methylation in the promoter regions of *FPS*, *SS*, *SE*, and *MDD* [7, 8] genes in *E. senticosus* inhibited the synthesis of the mevalonate pathway, thus reducing the accumulation of saponins.

Drought stress is usually considered a negative factor in plants and an important cause of yield loss. However, studies have shown that plants exposed to drought stress accumulate higher concentrations of secondary metabolites [9]. For example, drought stress dramatically decreases spica biomass production but increases rosmarinic acid (RA), ursolic acid (UA), and oleanolic acid (OA) contents in *Prunella vulgaris* L. [10]. In *Salvia miltiorrhiza* Bunge, water stress significantly increased the salvianolic acid B yield and decreased that of tanshinone IIA [11]. However, there are no reports on drought stress in the Araliaceae family. Drought stress is a type of adversity stress, in which plants respond to the external environment through various mechanisms. DNA methylation is one of the earliest and most studied mechanisms that regulate genome function and induce plant resistance and abiotic stress adaptation [12]. This provided a new idea for our research. Because the secondary metabolism of medicinal plants is affected by drought stress, they are inevitably subjected to epigenetic modifications. However, it is unclear how plants affect metabolite accumulation through epigenetic modifications under drought conditions.

DNA methylation is an epigenetic modification that plays an important role in regulating gene expression. Methyl-binding domain (MBD) protein is a class of methyl CpG-binding domain protein that selectively bind to fully methylated CpG dinucleotides and participate not only in the recruitment of nucleosome remodeling factors and histone deacetylases, but also in the formation of transcriptional repressor complexes that regulate gene expression through epigenetic modifications [13]. In humans, abnormalities in the *MBD* genes are strongly associated with the development of cancer and psychiatric disorders [14]. For example, overexpression of *MBD1* in the human body is usually associated with

carcinogenesis, leading to the silencing of tumor suppressor genes through the formation of the abovementioned complexes [15]. In plants, *MBD* genes also play important functions: in *Arabidopsis thaliana* (L.) Heynh. *AtMBD4* mediates phosphates uptake [16]. Similarly, *AtMBD6* is involved in RNA-mediated gene silencing [17], whereas *AtMBD9* is involved in regulating flowering time and shoot branching in *A. thaliana*, with mutant plants exhibiting an abnormal phenotype of early flowering and branching [18]. These studies on *MBD* have demonstrated their importance. Therefore, a holistic analysis of *EsMBD* in *E. senticosus* could provide a new research perspective for exploring the mechanism of DNA methylation of secondary metabolism-related genes.

## Results

### Identification and analysis of the physicochemical properties of the *EsMBD* gene family

Given the Hidden Markov Model of the *MBD* gene family (PF01429), 20 *EsMBD* genes were screened out from the whole-genome sequencing data of *E. senticosus*, and sequentially named *EsMBD1*-*EsMBD20* according to their individual positions on the chromosomes (Table 1). Based on the Hidden Markov Model of the *MBD* gene family, 10 *AeMBD* genes were screened from the whole genome sequencing data of *Aralia elata* (Miq.). Seem and consecutively named *AeMBD1*-*AeMBD10* according to their separate positions on the chromosomes. The results showed that the quantity of amino acids in the *EsMBD* gene family of proteins was in the range between 166 and 2,015 aa, and the molecular weight ranged from 18,865.60 to 226,144.33 Da, whereas the isoelectric points were below 7.00 except for *EsMBD7* (9.27), *EsMBD11* (9.39), *EsMBD14* (8.68), *EsMBD15* (7.08), and *EsMBD17* (8.27), indicating that most *EsMBD* proteins were acidic. Moreover, all *EsMBD* proteins exhibited hydrophobicity of less than 0 and an instability index greater than 40, indicating that all *EsMBD* proteins were unstable and hydrophilic.

### Chromosomal localization and collinearity analyses of *EsMBD* genes

The chromosomal localization analysis of *EsMBD* genes (Fig. 1A) showed that they were unevenly distributed in 12 of the 24 *E. senticosus* chromosomes. Chromosome No. 3 contained four *EsMBD* genes, chromosome 17 contained three *EsMBD* genes; chromosomes No.1, 8, and 12 each contained two *EsMBD* genes; and chromosome No.4, 9, 10, 11, 13, 14, and 18 each contained one *EsMBD* gene. The distribution of *EsMBD* and tandem repeats on different chromosomes of *E. senticosus* are shown in Fig. 1A, in which only one pair, *EsMBD18* and *EsMBD19*, were tandem repeat genes.

The results of the collinearity analysis of *EsMBD* genes in *E. senticosus* (Fig. 1B) showed 22,568 genes in the

**Table 1** Physicochemical properties of *EsMBD* gene family proteins

Gene Name	Gene ID	Quantity of amino acids/aa	Molecular weight/Da	Isoelectric point	Hydrophobicity	Instability index
<i>EsMBD1</i>	Ese01G001833	578	63613.66	5.25	-0.784	44.26
<i>EsMBD2</i>	Ese01G003936	339	37649.14	4.77	-1.297	54.04
<i>EsMBD3</i>	Ese03G001766	372	41867.69	5.22	-0.749	53.35
<i>EsMBD4</i>	Ese03G000746	1248	138135.81	5.10	-0.651	45.63
<i>EsMBD5</i>	Ese03G002722	338	37539.06	4.75	-1.256	43.79
<i>EsMBD6</i>	Ese03G003163	578	63631.68	5.29	-0.810	45.48
<i>EsMBD7</i>	Ese04G003192	316	35944.76	9.27	-0.708	60.01
<i>EsMBD8</i>	Ese08G003485	263	28964.99	5.15	-1.351	57.78
<i>EsMBD9</i>	Ese08G002921	368	41838.60	5.51	-0.785	59.19
<i>EsMBD10</i>	Ese09G002039	2015	226144.33	5.40	-0.459	47.53
<i>EsMBD11</i>	Ese10G001996	209	23363.52	9.39	-0.515	58.57
<i>EsMBD12</i>	Ese11G001419	228	25831.55	6.05	-1.177	52.80
<i>EsMBD13</i>	Ese12G002522	382	43366.30	5.11	-0.731	61.24
<i>EsMBD14</i>	Ese12G002001	937	103828.32	8.68	-0.810	51.91
<i>EsMBD15</i>	Ese13G002475	967	108033.86	7.08	-0.859	50.66
<i>EsMBD16</i>	Ese14G002218	332	36570.91	4.78	-1.423	52.80
<i>EsMBD17</i>	Ese17G001961	166	18865.60	8.27	-0.575	48.00
<i>EsMBD18</i>	Ese17G000257	232	26081.80	6.22	-1.219	54.92
<i>EsMBD19</i>	Ese17G000341	384	41510.17	4.48	-1.143	54.80
<i>EsMBD20</i>	Ese18G001412	345	38815.07	4.44	-0.494	40.79

genome, accounting for 62.05% of all genes, and collinearity was observed among the 17 *EsMBD* genes in the entire family. As shown in Fig. 1B, in the *E. senticosus* genome, these genes were of collinearity: *EsMBD1/EsMBD6*, *EsMBD1/EsMBD14*, *EsMBD1/EsMBD15*, *EsMBD2/EsMBD5*, *EsMBD2/EsMBD8*, *EsMBD2/EsMBD16*, *EsMBD3/EsMBD9*, *EsMBD3/EsMBD13*, *EsMBD5/EsMBD8*, *EsMBD5/EsMBD16*, *EsMBD6/EsMBD14*, *EsMBD6/EsMBD15*, *EsMBD7/EsMBD11*, *EsMBD8/EsMBD16*, *EsMBD9/EsMBD13*, *EsMBD12/EsMBD18*, and *EsMBD14/EsMBD15*, which presented many fragmental repeats among the *EsMBD* gene family.

The results of the collinearity analysis between the *EsMBD* genes of *E. senticosus* and the *AeMBD* genes of *A. elata* (Fig. 1D) showed significant collinearity between *E. senticosus* and *A. elata*, between *EsMBD* and *AeMBD* genes, and between the same *EsMBD* and different *AeMBDs*. This indicated that *MBD* genes evolved with the evolution of species in Araliaceae.

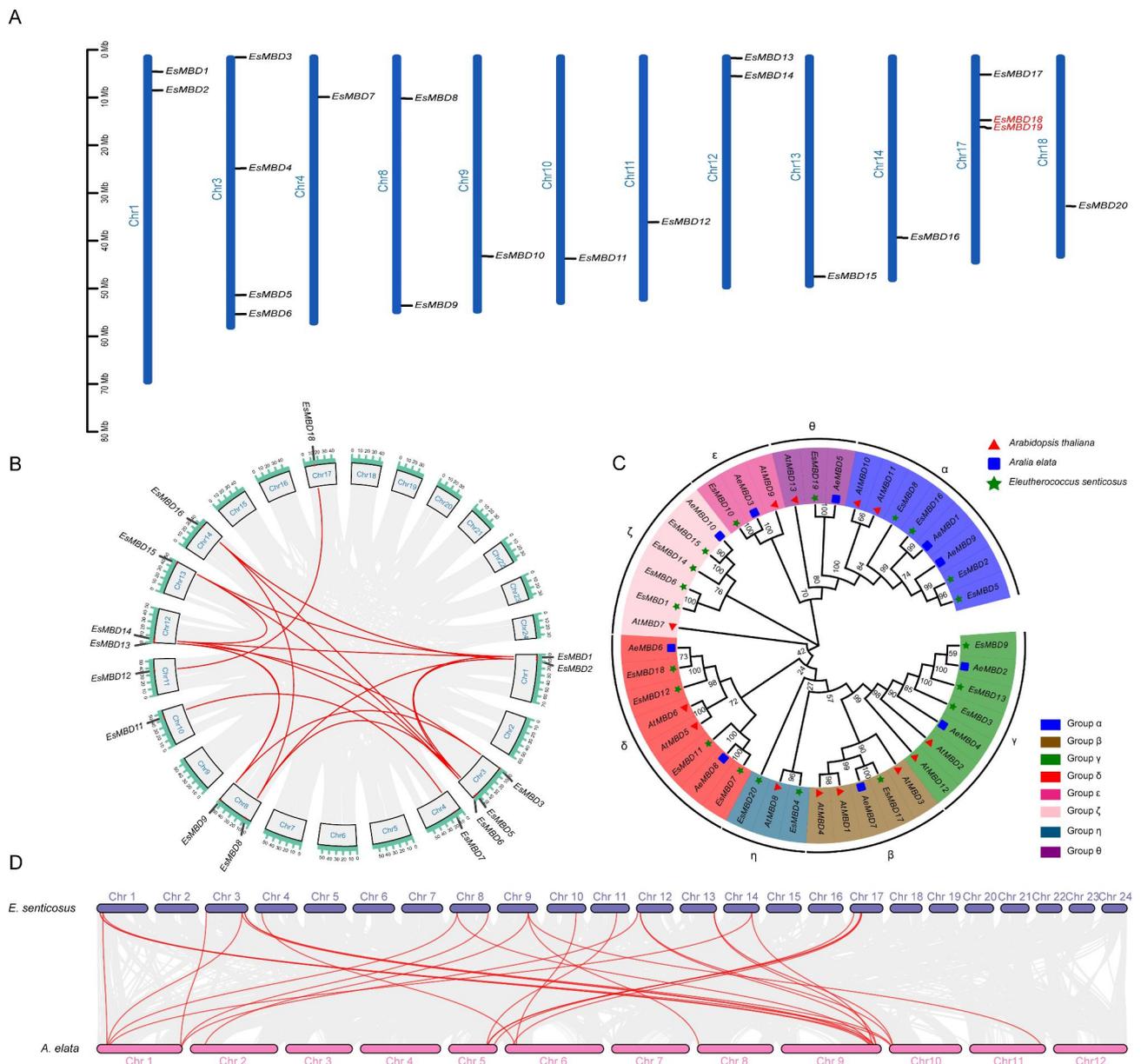
#### Phylogenetic analysis of MBD

The 20 *EsMBD* protein sequences identified from the screening and 10 *AeMBD* sequences from *A. elata* were used to construct a phylogenetic evolutionary tree, in addition to the 13 *AtMBD* protein sequences with well-defined functions in *A. thaliana*. The results (Fig. 1C) showed that the evolutionary tree was divided into eight subfamilies,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$ ,  $\eta$ , and  $\theta$ , based on the affinities between the different members of the MBD gene families of *E. senticosus*, *A. elata*, and *A. thaliana*. Except for the  $\eta$  subfamily, which contained only *EsMBD4*, *EsMBD20*

and *AtMBD8*, the *MBD* genes of *E. senticosus* and *A. elata* were both localized in the remaining subfamilies. The  $\alpha$  subfamily contained 4 genes, *EsMBD2*, *EsMBD5*, *EsMBD8*, and *EsMBD16*; the  $\beta$  subfamily contained 1 *EsMBD17* gene; the  $\delta$  subfamily contained *EsMBD7*, *EsMBD11*, *EsMBD12*, and *EsMBD18*; and the  $\gamma$  subfamily contained *EsMBD3*, *EsMBD9*, and *EsMBD13*. The  $\delta$  subfamily contained *EsMBD7*, *EsMBD11*, *EsMBD12*, and *EsMBD18*; the  $\epsilon$  subfamily contained *EsMBD10*; and the  $\zeta$  subfamily contained *EsMBD1*, *EsMBD6*, *EsMBD14*, and *EsMBD15*. The  $\eta$  subfamily contained *EsMBD4* and *EsMBD20*, whereas the  $\theta$  subfamily contained only *EsMBD19*.

#### Structure of *EsMBD* genes with analyses of conserved structural domains and motifs of the protein

The structure of *EsMBD* genes in *E. senticosus* is shown in Fig. 2A C. Of the *EsMBD* gene sequences, 70% were within 2,000 bp in length, only one was up to 6,000 bp in length, and only 20% were approximately 1,000 bp in length. Ten genes in the *EsMBD* gene family contained 5'UTR, 11 genes contained 3'UTR, and 8 genes did not have UTR sequences, namely *EsMBD1*, *EsMBD6*, *EsMBD7*, *EsMBD9*, *EsMBD10*, *EsMBD13*, *EsMBD15*, and *EsMBD20*. The *EsMBD* genes in the same branch of the evolutionary tree also presented similar structures. Ten conserved motifs of *EsMBD* were identified using the online tool MEME analysis (Fig. 2D) and named as motifs 1–10. Furthermore, *EsMBD17* contained only 2 motifs, the fewest in the *EsMBD* gene family, whereas *EsMBD10* contained the most 10 motifs. Over 50% of



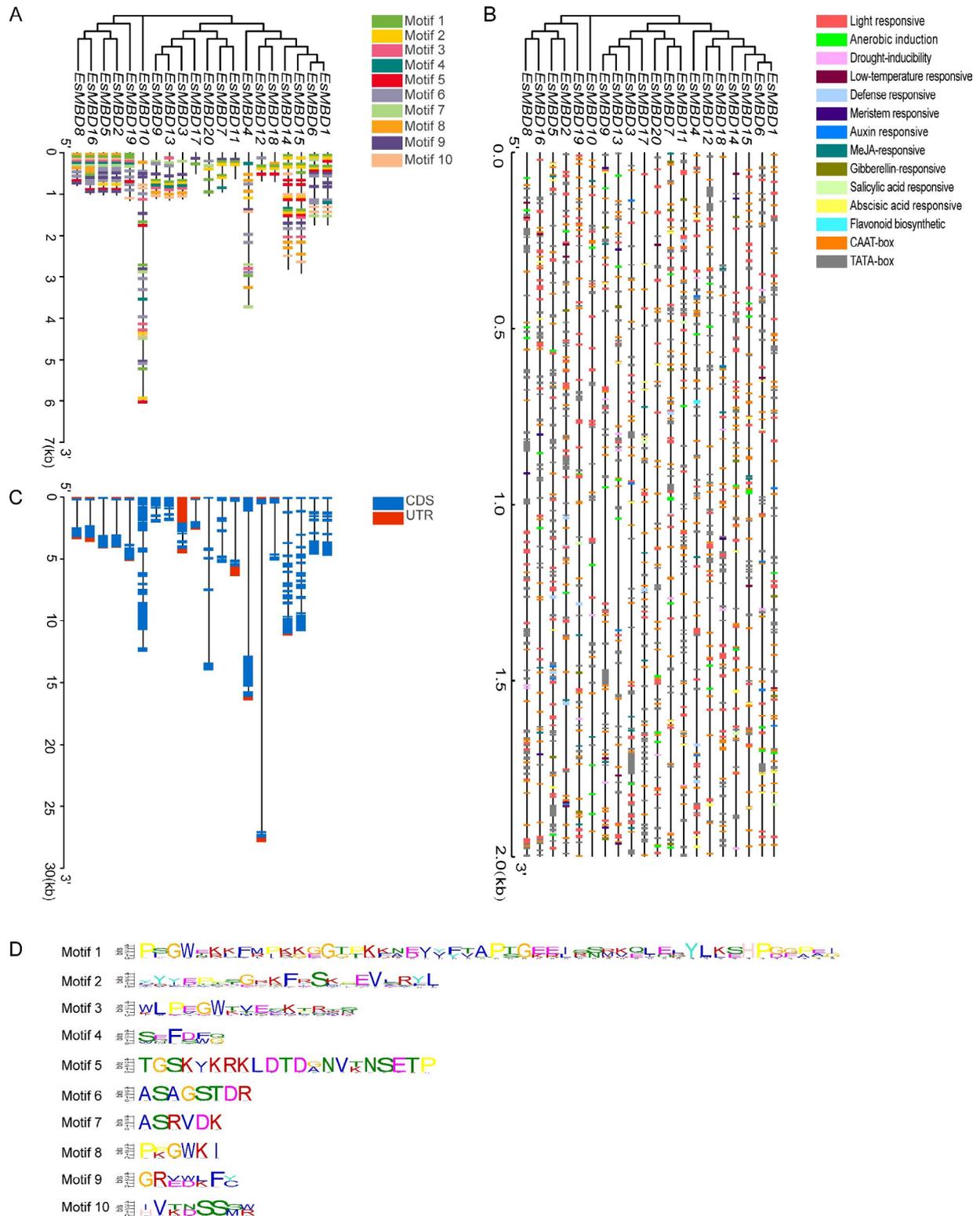
**Fig. 1** Chromosomal localization, evolutionary analysis, and collinearity analysis of *EsMBD* genes in *E. senticosus*. A: Chromosomal localization of *EsMBD* genes in *E. senticosus*. Red markers indicate tandem repeats; B: Collinearity analysis of *EsMBD* genes of *E. senticosus*; C: Phylogenetic analysis of *EsMBD* genes of *E. senticosus*; and D: Collinearity analysis of interspecies *MBD* between *E. senticosus* and *A. elata*. Red lines indicate the *MBD* genes

the *EsMBD* gene family contained motifs 1, 2, 6, 7 and 9, indicating that these five genes are relatively conserved in this family.

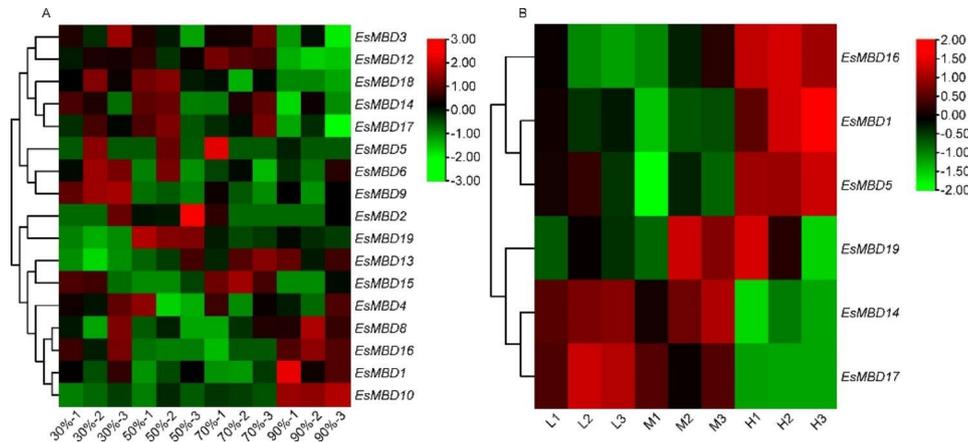
#### Analysis of cis-acting elements of the *EsMBD* gene promoter

Sequence analysis of the promoter region of the *EsMBD* gene family in *E. senticosus* (Fig. 2B) revealed that four phytohormone response elements were identified in addition to the conventionally transcriptional regulatory elements of TATA-box and CAAT-box, including the methyl jasmonate (MeJA), gibberellin, salicylic acid,

abscisic acid. Some environment-related transcriptional regulatory elements were also included, such as light response, anaerobic induction, drought inducibility, and low temperature. The number of light-response-related cis-acting elements was the highest in the promoter regions of all *EsMBD* genes (Supplementary Table S1). Additionally, cis-acting elements related to the regulation of flavonoid biosynthesis were found in the promoter region of *EsMBD4*. Therefore, we hypothesized that *EsMBD* might be involved in mediating the physiological regulation of *E. senticosus* and the synthesis of flavonoid



**Fig. 2** Gene structure, promoter analysis, and conserved motifs of *EsMBD* in *E. senticosus*. **A**: Motif analysis of *EsMBD* genes in *E. senticosus*; CDS: DNA sequence corresponding to the protein sequence one by one; UTR: non-coding segment at both ends of the mRNA molecule **B**: Promoter analysis of *EsMBD* in *E. senticosus*; **C**: Exon-intron structure analysis of *EsMBD* genes in *E. senticosus*; **D**: Conserved motifs of *EsMBD* genes in *E. senticosus*; Motif: conserved sequences with biological functions



**Fig. 3** *EsMBD* gene expression in different *E. senticosus* transcriptomes. A: *EsMBD* gene expression under drought stress. B: *EsMBD* gene expression in different saponin content groups. L, low saponin content; M, middle saponin content; H, high saponin content. Three replicates were performed for each group of samples

substances by binding to methylated cytosine under different light conditions.

#### Expression analysis and quantitative real-time PCR validation of *EsMBD* gene family

The expression of *EsMBD* genes under drought stress in *E. senticosus* is shown in Fig. 3A. The expression of different *EsMBD* genes varied with different water contents: *EsMBD3*, *EsMBD12*, *EsMBD14*, *EsMBD17*, and *EsMBD18* showed low expression at 90% relative water content, whereas *EsMBD6* and *EsMBD9* were highly expressed at 30% relative water content, and *EsMBD1*, *EsMBD8*, *EsMBD10*, and *EsMBD16* were highly expressed at 90% relative water content. The expression of *EsMBD13* gradually decreased with increasing water content. The expression of *EsMBD* genes in the different saponin content groups is shown in Fig. 3B. *EsMBD14* and *EsMBD17* were significantly expressed in the low and medium saponin content groups but were not significantly expressed in the high saponin content group, whereas *EsMBD1*, *EsMBD5*, and *EsMBD16* were not significantly expressed in the low and medium saponin content groups but were significantly expressed in the high saponin content group.

To verify the reliability of the transcriptome data, all differentially expressed genes were screened for qRT-PCR validation, except for four *EsMBDs*, which were not expressed in the transcript. The results showed that the trends of other genes were consistent with the transcriptome sequencing data, indicating that the transcriptome sequencing results were reliable (Fig. 4).

#### Molecular docking of *EsMBD* proteins to DNA sequences

Homology modeling of *EsMBD* proteins was performed using the SWISS-MODEL website. The GMQE ranged from 0.02 to 0.21, QMEANDisCo Global values ranged

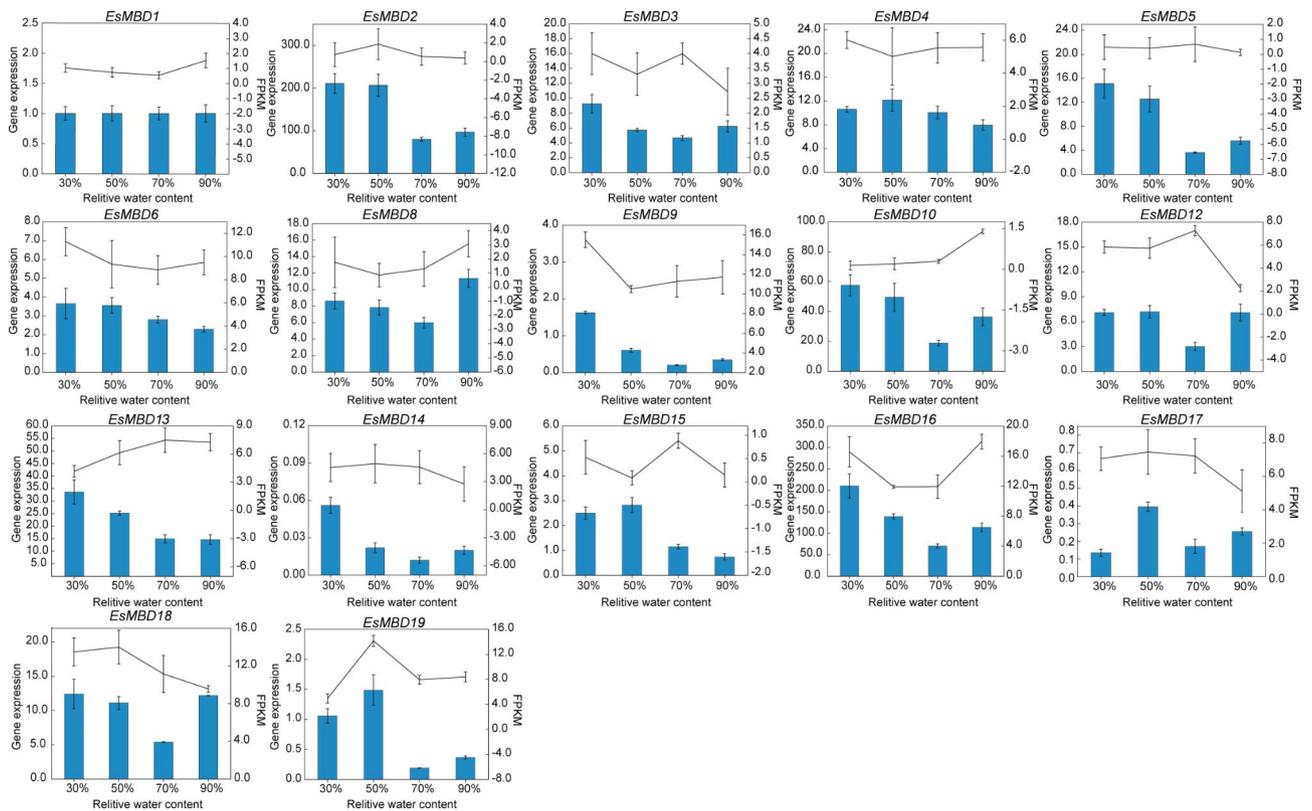
from  $0.41 \pm 0.11$  to  $0.72 \pm 0.11$ , and sequence similarity ranged between 22.22 and 73.24%. Molecular docking was performed at the HDOCK SERVER website (Fig. 5). Among the eight best predictions acquired, each *EsMBD* protein was successfully docked to the methylated DNA sequence, indicating that *EsMBD* could bind to fully methylated CpG dinucleotides and thus play a regulatory role.

#### Subcellular localization of *EsMBD* proteins

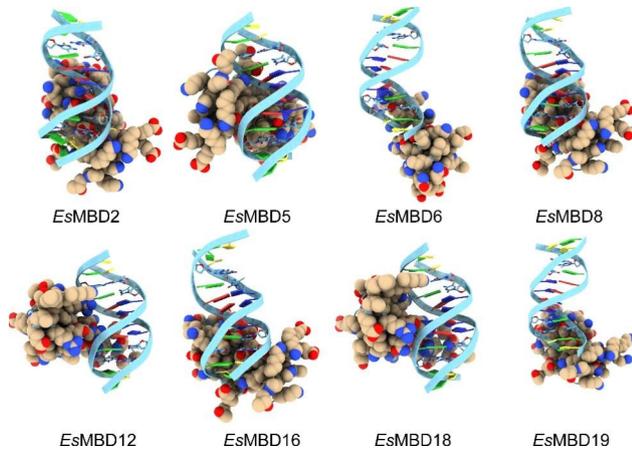
Subcellular localization can reveal the working positions of proteins at the cellular level. The WoLF PSORT website (<https://wolfpsort.hgc.jp>) was used to predict the subcellular localization of *EsMBD* proteins. The results showed that most *EsMBD* proteins were predicted to be located in the nucleus, *EsMBD12*, *EsMBD16*, *EsMBD17*, *EsMBD18*, and *EsMBD20* were predicted to be located in the cytoplasm, and *EsMBD2* and *EsMBD5* were predicted to be located in the nucleus and cytoplasm (Supplementary Table S2). To verify the prediction of subcellular localization, we selected 6 *EsMBD* proteins involved in molecular docking to fuse Green Fluorescent Protein (GFP) and expressed them instantaneously in *Allium cepa* L. Five *EsMBD* proteins were localized in the nucleus and cytoplasm, among which *EsMBD2* and *EsMBD6* were expressed in both the nucleus and cytoplasm and *EsMBD12*, *EsMBD16*, and *EsMBD18* were expressed only in the cytoplasm (Fig. 6).

#### Discussion

In the model plant *A. thaliana*, mutations are not completely random throughout the genome because of epigenetic modifications and other effects, and the mutation probability of important genes is lower than that of non-important genes [19]. *MBD* genes are inseparable from DNA methylation, and have been identified and validated



**Fig. 4** Quantitative Real-time PCR validation of *EsMBD* genes in *E. senticosus*. FPKM: Fragments Per Kilobase of exon model per Million mapped fragments



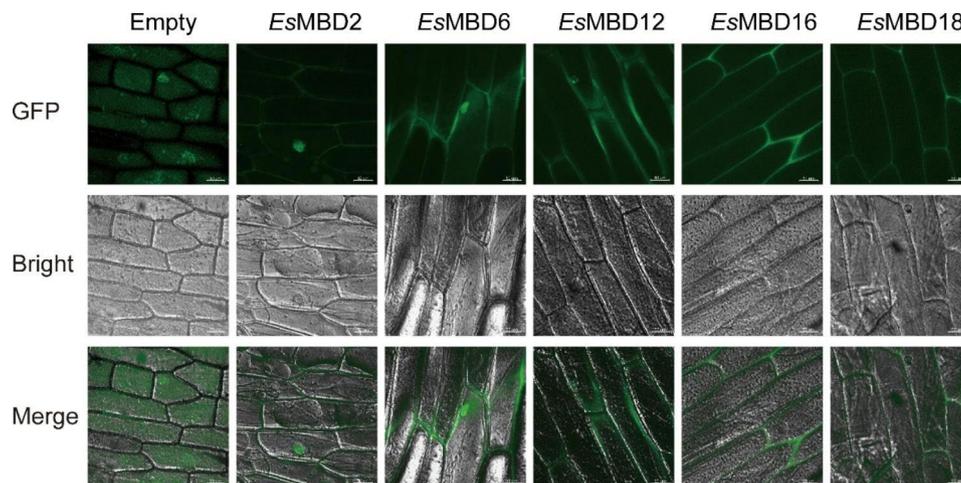
**Fig. 5** Model of *EsMBD* protein docking with methylated DNA sequences. MBDs are depicted in yellow, methylated DNA sequences are displayed in blue, and 5-methylcytosine is located between the third and ninth positions

in various plants. In these plants, MBDs are involved in a several physiological processes and stress regulation [20]. *MBD* genes are not only important for plant growth and development but also pose great relevance to humans [21]. In this study, we conducted a comprehensive bioinformatics analysis of *EsMBD* genes in *E. senticosus*, identified *EsMBD* family members by sequence similarity, and analyzed their evolutionary amplification

patterns, sequence characteristics, expression motifs, and homology to better model them, which not only provided a basis for further studies of *MBD* genes, but also laid a foundation for epigenetic modifications affecting the genome and physiological activities of the organism.

### Evolution of *MBD* genes

Gene duplication drives the evolution of genetic systems by providing a rich source of raw material. Polyploidization is a major mechanism of environmental adaptation and species formation in organisms [22], and the major mechanism of biological genome structure and gene evolution is through a single gene, chromosome, and whole-genome duplication (WGD) [23]. The sequencing of the genome of *E. senticosus* showed that in addition to the  $\gamma$  genome-wide triploidization event common to all core members of the family Pentacanthaceae, *E. senticosus* also experienced two WGD events [24]; thus, it could be hypothesized that the *E. senticosus* *MBD* gene family expanded during these two WGD events. Further analysis of the collinearity between *E. senticosus* and chromosomal localization information revealed that the genomic collinearity level of *E. senticosus* was as high as 62.05%, and there was also a significant collinearity among *EsMBD* gene families, while the *EsMBD* gene families were scattered in arrangement on chromosomes, and only one tandem repeat gene existed. Therefore, it



**Fig. 6** Subcellular localization of the five *EsMBD* proteins. Green fluorescence (GFP) signals were observed under a laser scanning confocal microscope to indicate the subcellular localization of *EsMBD*-GFP fusion proteins. The ruler in the image is 50  $\mu$ m

could be inferred that the main method of amplification of *EsMBD* gene families was gene fragment duplication. The collinearity analysis showed significant collinearity between *E. senticosus* and *A. elata*, which are both species of Araliaceae, while there was also significant collinearity between the *EsMBD* and *AeMBD* genes. Therefore, it can also be inferred that *MBD* genes evolved within the Araliaceae species before eventually turning into interspecies-specific *MBD* genes through amplification and duplication.

#### Phylogeny of the *MBD* gene family

*MBD* proteins are key determinants of DNA methylation status [25]. Twenty *EsMBD* genes and ten *AeMBD* genes were identified in *E. senticosus* and *A. elata*. Together with the 13 *AtMBD* genes in *A. thaliana*, a phylogenetic tree was constructed and branched according to a previously constructed phylogenetic tree between *Brassica napus* L. and *A. thaliana* [26]. Although there was only *AtMBD8*, *EsMBD4*, and *EsMBD20* in one branch, the *MBD* gene family members of *E. senticosus* and *A. elata* were identified in all the other branches. This might be attributed to the fact that the relatives of *E. senticosus* being closer to each other than those of *A. thaliana* and *A. elata*. The phylogenetic tree was also verified using the inter-species collinearity map of *E. senticosus* and *A. elata*. *MBD* gene family members have different abilities to bind to methylated DNA [27]. Therefore, we speculated that the functional differences in *MBD* genes among different branches might be significant, indicating that the functional differences among *MBD* proteins were gradually realized during the evolutionary process. This result is consistent with the conclusions of previous studies that *MBD* genes have a long evolutionary history, with *MBD* genes existing and beginning to differentiate

as early as the formative stages of monocotyledons and dicots [26].

#### Structure of *MBD* genes and protein-conserved motifs

UTRs are closely associated with gene expression motifs that enhance plant perception of stimuli from developmental and environmental regulatory factors [28]. In the *MBD* gene family of *E. senticosus*, UTRs were present in 12 genes, accounting for 60% of all *EsMBDs*. *EsMBDs* lacking UTR structures were in the same branch or adjacent ones, implying that the expression motifs of the *MBD* gene family shared some similarity between the same and adjacent branches, whereas the structures and differences of genes were more significant between different branches. Studies have shown that dehydration stress can induce the extension of the 3'UTR transcript, and that these stress-induced 3'UTR extensions have a new function: regulating the long non-coding RNA of its adjacent genes, which confirms the importance of UTR length in plant adaptation to stress conditions [29]. The UTR lengths of different *MBD* genes in *E. senticosus* were different, which also indicated that different genes may have functional differences and play different roles under various stresses.

*EsMBD* sequences exhibit significant conservation and inter-group specificity across adjacent branches, and conserved structural sequences are prerequisites for conserved biological functions [30]. The *EsMBD* motifs were similar in structure between adjacent branches but differed significantly between branches that were further apart from each other. It was speculated that this was due to the different functional differentiation of *MBD* genes during the process of evolution. Moreover, the 20 *EsMBD* protein sequences obtained from the screening also presented conserved motifs in most *EsMBDs*, which is consistent with the results of earlier studies comparing

conserved motifs of *MBD* in *B. napus* [26]. These conserved motifs of different *MBD* genes in *E. senticosus* also showed different similarities.

#### Promoter analysis of the *MBD* gene family

A promoter is a small region of a DNA sequence that responds to various transcription factors and triggers specific gene expression [31]. Cellular metabolism in organisms is largely regulated at the transcriptional level, and promoters are responsible for basic regulation of transcription initiation [32]. The promoters of the *MBD* genes in *E. senticosus* were found to contain a large number of cis-acting elements, and the quantity of different cis-acting elements in adjacent branches showed some similarity, suggesting that there is a certain degree of functional differentiation among the *MBD* genes in *E. senticosus*, which is consistent with a previous hypothesis that there is functional differentiation of *MBD* genes during the evolutionary process [26]. Among the various cis-acting elements of the promoter, light response elements were the most abundant, followed by anaerobic induction elements, indicating that the *EsMBD* genes of *E. senticosus* were subjected to maximum regulation by light and anaerobic conditions (Supplementary Table S1). It has been demonstrated that different light qualities like light intensities regulated the growths of roots, stems, and leaves as well as the synthesis of secondary metabolites in *E. senticosus* [33, 34]. For example, red light was found to induce the synthesis of Eleutheroside E [33] and LED lighting promoted the development of fine roots in *E. senticosus* seedlings [34]. In addition, shade treatment enhances photosynthesis in *E. senticosus* by promoting the growth and accumulation of secondary metabolites [35]. The *EsMBD* promoter regions, which contain substantial light-responsive and cis-acting elements involved in the regulation of flavonoid synthesis, were also identified. We speculated that the *EsMBD* gene of *E. senticosus* could participate in regulating the physiological activities of *E. senticosus* under certain light stress conditions, thereby affecting the primary and secondary metabolism of *E. senticosus*.

#### *EsMBD* gene expression analysis

Drought stress alters plant growth and developmental processes, causing morphological changes in plant height and leaf size. Plants also respond to drought stress by reducing water loss and enhancing water absorption capacity [36]. It has been suggested that drought induced increases in DNA methylation may be one of the main mechanisms by which plants respond to drought stress [37]. In maize, the *MBD* genes of *Zea mays* L. showed a similar expression motif under drought stress [38]. Interestingly, *EsMBD* genes in *E. senticosus* also exhibited similar expression motifs under drought stress, with some

*EsMBD* gene expression showing similar up- or down-regulation trends at different relative water contents, and most *EsMBD* genes were up-regulated under relatively dry conditions. In a previous study by Cui et al. [7], the saponin content of *E. senticosus* decreased with increasing DNA methylation ratio. Meanwhile, *EsMBD14* and *EsMBD17* were significantly expressed in the low and medium saponin content groups but were not significantly expressed in the high saponin content group. This indicates that *EsMBD14* and *EsMBD17* were not significantly expressed in the low methylation ratio group. This suggests that *EsMBD14* and *EsMBD17* are involved in regulating saponin synthesis under drought stress through DNA methylation.

#### Molecular docking model analysis and subcellular localization of *EsMBD* genes

The *MBD* protein of *E. senticosus* was revealed using homology modeling on the SWISS-MODEL website. The Global Model Quality Estimate [39] (GMQE) is a quality assessment criterion that integrates the properties of the target template alignment and template structure to improve the reliability of the quality estimate. The QmeandisCo global score [40] is the average QMEAN-DisCo score for each residue. Based on the QMEAN-DisCo score estimated for substantial models, it provides an error estimate and represents the standard deviation between the QMEANDisCo global score and underlying facts. Because the reliability of the predictions depends on the model size, the provided error estimates are calculated in line with models of sizes similar to the inputs. The global model quality estimates, together with the QMEANDisCo global score demonstrated both the accuracy and reliability of the prediction models. In homology modeling, *EsMBD* protein models were derived from the *MBD* proteins of *Homo sapiens*, and *A. thaliana* showed the highest sequence similarity of 73.24%. This confirmed our previous hypothesis that *MBD* genes evolve and are amplified within the *E. senticosus* genome. Among all *EsMBD* protein homology models, the size and morphology varied constantly. Combined with the results of motif analysis, it can be speculated that there are different degrees of functional differentiation among *EsMBDs*, whereas the study in *Glycine max* (Linn.) Merr. [41] verified that most *MBDs* were differentiated during early evolution, which was also consistent. *MBD* proteins can specifically bind to methylated cytosines, thus causing DNA methylation silencing in methylated genes [9]. All those screened from *E. senticosus* were able to dock successfully with methylated DNA after homology modeling, which not only verified the basic mode of action of *MBD* proteins but also predicted *EsMBD* binding to methylated functional genes as one of the pathways of DNA methylation regulating secondary metabolism in

*E. senticosus*. Among all models, *EsMBD2* and *EsMBD5* presented the best quality and highest sequence similarity. To further verify the function of *EsMBD* proteins, we examined the subcellular localization of 5 *EsMBDs* that responded strongly to drought stress. Previous studies have shown that *AtMBD5* protein is located in the nucleus and interacts with *AtRAN3* protein [42]. In our study, all *EsMBDs* were located in the nucleus and cytoplasm, consistent with the results of previous studies. *EsMBD* proteins are widely involved in the physiological metabolism of *E. senticosus* and play important roles in its stress response and secondary metabolic processes. This also showed that *E. senticosus* is likely to affect the synthesis of saponins and other substances through epigenetic modifications under drought stress.

## Conclusion

In this study, 20 *EsMBD* genes were identified in *E. senticosus*, which were mainly obtained by fragment duplication and found to be unevenly distributed on 12 of the 24 chromosomes. The gene structure and conserved motifs of *EsMBD* genes exhibited different degrees of similarity, with the greatest degree between adjacent branches of the evolutionary tree and substantial light-, hypoxia-, and drought-related cis-acting elements in the *EsMBD* promoter. Under drought stress, *EsMBD* genes of *E. senticosus* showed similar expression motifs. Comparing the transcriptome data and Quantitative Real time PCR results of different saponin content groups of *E. senticosus*, it was found that different *EsMBD* genes had different sensitivity to drought stress, among which *EsMBD14* and *EsMBD17* were most sensitive to drought stress. After molecular docking model analysis and subcellular localization, it can be inferred that functional differentiation occurred during the evolutionary process of *EsMBDs*. Based on the above analysis and transcriptome data, the cytosine of the functional gene promoter in the *E. senticosus* genome can be effectively methylated under various stress conditions. *EsMBD14* and *EsMBD17* combine with methylated cytosine, resulting in changes in the expression of functional factors, thus affecting the secondary metabolism process of *E. senticosus*.

## Materials and methods

### Identification and physicochemical characterization of the *EsMBD* gene family

Whole genome sequencing data for *E. senticosus* [24] were downloaded from the CNGB Sequence Archive (CNSA) database (<https://ftp.cngb.org/pub/CNSA/data3/CNP0001682/CNS0348054/CNA0019321/>). Whole genome sequencing data of *A. elatae* [43], which is the same species of Araliaceae, were downloaded from the Dryad database (<https://datadryad.org/stash/data-set/doi:https://doi.org/10.5061/dryad.69p8cz937>). The

Hidden Markov Model (PF01429) for *MBD* gene was downloaded from the PFAM database (<http://pfam.xfam.org/>). The structural domains of the MBD proteins were searched for using the HMMER program, and the E value was set to 0.001. Furthermore, the online Simple Modular Architecture Research Tool (SMART) and NCBI Batch CD-Search tool were used to confirm the conserved domains of all the candidate MBD protein sequences. We retained only sequences containing significant MBD domains in all results; therefore, the sequence of *E. senticosus* was eliminated. *MBD* gene family members were identified from the whole genomes of *E. senticosus* and *A. elata*, and named according to their different positions on the chromosome (Supplementary Table S3). The obtained *EsMBD* protein sequences were submitted to ProtParam (<https://web.expasy.org/protparam/>) and analyzed for physicochemical properties, such as amino acid content, molecular weight, isoelectric point, hydrophilicity, and instability coefficient.

### Chromosomal localization and collinearity analysis of *EsMBD* genes

For collinearity analysis, *E. senticosus* genome files and *E. senticosus* genome files with *A. elata* were used for sequence alignment using the Basic Local Alignment Search Tool (BLAST), with a cut-off E value of  $1.0 \times 10^{-10}$ . The MCScanX software was used to perform collinearity analysis based on BLASTP. TBtools [44] (<https://github.com/CJ-Chen/TBtools>) software was used to visualize collinearity.

### Phylogenetic analysis of *MBD* genes

The 13 *AtMBD* protein sequences of *A. thaliana* were downloaded from The Arabidopsis Information Resource (<https://www.arabidopsis.org/>). Twenty *EsMBD* proteins identified in *E. senticosus*, 12 *AeMBD* proteins from *A. elata*, and 13 *AtMBD* proteins from *A. thaliana* were used to construct the phylogenetic tree. We used the MBD gene family of *A. thaliana* as an outgroup, according to previous methods [26], and constructed a phylogenetic tree with *A. elata* to obtain more accurate results. ClustalW was used for multiple sequence alignment of MBD protein sequences with default parameters. The alignment results were used to construct an NJ phylogenetic tree using MEGA 11.0. Bootstrap resampling (100) was used to assess the reliability of interior branches. The bootstrap value was set at 1,000. The online tool EvolView (<https://www.evolgenius.info/evolview/>) was used to construct a phylogenetic tree.

### Gene structure and sequence analysis of the *EsMBDs*

The exon-intron structure analysis was performed using TBtools software by inputting gene annotation GFF files. The online MEME Suite was used to confirm the motifs

of *E. senticosus* EsMBD protein sequences with the following parameters: maximum number of ten motifs and optimum width of 6–50.

#### **EsMBD gene promoter's cis-acting element analysis**

The 2,000 bp nucleotide fragment upstream of the start codon of each *EsMBD* gene in the *E. senticosus* genome was screened as the promoter sequence. The cis-acting element analysis of these *EsMBD* genes was performed using the online software PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

#### **Expression analysis of EsMBD**

The transcriptome sequencing data of drought-stress-treated *E. senticosus* were downloaded from the NCBI database (Accession: SRR19962743-SRR19962757). Based on the results of a previous study by Cui et al. [7], transcriptome data of *E. senticosus* with different saponin contents (Accession: SRX13417593-SRX13417601) were downloaded from the NCBI database. Based on the different transcriptome data, differentially expressed *EsMBD* genes were screened (Supplementary Table S4-S5). The expression values (FPKM values) were used to compare the transcript levels of MBD in *E. senticosus*. The transcripts with  $\log_2FC \geq 1.5$  or  $\leq -1.5$  and  $FDR < 0.05$  were selected as transcripts that were differentially expressed. Heatmaps of *EsMBD* gene expression were drawn using the Heatmapper website (<http://www.heatmapper.ca/>), based on the expression of differentially expressed genes.

To verify the accuracy of the expression levels obtained from RNA-Seq analysis, the screened *EsMBD* was validated by qRT-PCR. Primers for qRT-PCR amplification of differentially expressed *EsMBD* genes were designed using the Prime Premier software (5.0) (Supplementary Table S6). The specificity of the qRT-PCR primers was determined by analyzing the solubility curves. The qRT-PCR was performed using Talent qPCR PreMix (SYBR Green) (TIANGEN, Beijing) on an Applied Biosystems 7900HT PCR System (THERMO FISHER SCIENTIFIC, Waltham, MA, USA) with the *glyceraldehyde-3-phosphate dehydrogenase* gene (*GAPDH*) used as an internal reference gene [45]. Three biological replicates were used for each sample. The total reaction system was 10.0  $\mu\text{L}$  with 1.0  $\mu\text{L}$  enzyme, 2.9  $\mu\text{L}$  Nuclease-Free Water, 0.3  $\mu\text{L}$  each primer, 0.5  $\mu\text{L}$  cDNA, 1.0  $\mu\text{L}$  ROX Reference Dye. The reaction conditions were as follows: pre-denaturation at 95 °C for 3 min, denaturation at 95 °C for 5 s, annealing at 55 °C for 10 s, and supplemental extension at 72 °C for 15 s. The reaction was completed after 40 cycles. *EsMBD* gene expression was calculated using SDS 2.4 software through the  $2^{-\Delta\Delta C_t}$  method [46].

#### **Molecular docking of EsMBD proteins to DNA sequences**

Homology modeling of *EsMBD* proteins was performed using SWISS-MODEL (<https://swissmodel.expasy.org/>) with the Global Model Quality Estimate [39] (GMQE), QMEANDisCo [40] global score, and sequence similarity ratio. The methylated DNA sequence models (PDB ID: 5U2Z) were downloaded from the RCSB PDB website (<https://www.rcsb.org/>). *EsMBD* proteins and methylated DNA sequences were docked on the HDock SERVER website (<http://hdock.phys.hust.edu.cn/>). The molecular docking results were obtained using the UCSF ChimeraX website (<https://www.cgl.ucsf.edu/chimerax/>).

#### **Vector construction**

The cloning and restriction primers were designed using Prime premier 5.0 (Supplementary Table S7) [39]. The *EsMBD* sequences were cloned using cloning primers and enzyme digestion primers, using *E. senticosus* cDNA as a template and *Taq* DNA polymerase (TIANGEN, Beijing). The amplification conditions were as follows: pre-denaturation at 94 °C for 5 min, denaturation at 94 °C for 30s, annealing at 55 °C for 30s, extension at 72 °C for 5 min, and an additional extension at 72 °C for 5 min. A lethal fast cloning kit (TIANGEN, Beijing) was used to connect the cloned sequences with the pLB vectors and was introduced into competent *Escherichia coli* Top10 cells (TIANGEN, Beijing, China). Sequencing was performed using a SinoGenoMax (Beijing, China). The recombinant pLB plasmids were extracted using a TIANprep Mini Plasmid Kit (TIANGEN, Beijing, China). The pHG vectors and recombinant pLB plasmids were digested overnight at 37 °C with BamHI-HF, PstI-HF, and SacI-HF (NEW ENGLAND BIOLABS, Beijing, China) restriction enzymes. We used T4 DNA ligase (NEW ENGLAND BIOLABS, Beijing, China) to connect *EsMBD* genes with pHG vectors overnight at 16 °C to construct recombinant *EsMBD-GFP* expression vectors. The total reaction system was 20.0  $\mu\text{L}$  with 1.0  $\mu\text{L}$  T4 DNA Ligase, 1.0  $\mu\text{L}$  Nuclease-Free Water, 2.0  $\mu\text{L}$  T4 DNA Ligase Buffer, 11.0  $\mu\text{L}$  Insert DNA (20 ng /  $\mu\text{L}$ ), 5.0  $\mu\text{L}$  pHG vector.

#### **Subcellular localization of EsMBD proteins**

We selected *Allium cepa* L. as the infection receptor. The recombinant expression vectors were transformed into *Agrobacterium tumefaciens* (Biomed, Beijing, China) and cultured on LB solid medium containing kanamycin (100 ng/ml) and rifampicin (50 ng/ml). The positive colonies were selected and cultured overnight in LB liquid medium at 28 °C. The cells were then centrifuged at 3000 rpm for 10 min to collect the cells, which were resuspended in LB liquid medium without antibiotics. The medium contained  $\text{MgCl}_2$  (10 mmol/L) and acetosyringone (100 mmol/L), and was adjusted to an  $\text{OD}_{600}$  of 1.0.

To prepare LB solid culture medium, the outer skin of the onion was peeled off without antibiotics. Fresh and thick scales were taken and a small square with an area of 1 cm<sup>2</sup> on the inner skin was cut using scissors. The onion skin was torn off with tweezers, placed downward close to the mesophyll in LB solid culture medium without antibiotics, and sealed with a sealing film for dark culture at 28 °C for 24 h. After 24 h, the inner skin of *A. cepa* was removed and soaked in LB heavy suspension solution at room temperature for 20 min. After 20 min, tweezers were used to collect the epidermis, filter out the bacterial liquid, lay on new LB solid medium without antibiotics, and incubated for 16–24 h at 25 °C with a photoperiod of 16 h/8 h. Small pieces of onion inner skin were removed, shaken, washed with clean liquid LB medium to remove attached *A. tumefaciens*, placed on a slide, and observed under a laser scanning confocal microscope (Leica, Germany).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09191-x>.

**Additional file 1. Table S1.** *EsMBDs* promoter region cis-acting element. **Table S2.** WoLF PSORT Predicts Subcellular Localization of *EsMBD* Proteins. **Table S3.** Methyl-binding domain in the genome of *Eleutherococcus senticosus* (Rupr. et Maxim) Maxim and *Aralia elata* (Miq.) Seem. **Table S4.** Expressed MBD genes in *Eleutherococcus senticosus* (Rupr. et Maxim) Maxim of different relative water content. **Table S5.** Expressed MBD genes in *Eleutherococcus senticosus* (Rupr. et Maxim) Maxim of different saponin content. **Table S6.** Primers for qRT-PCR. **Table S7.** Primers for Vector construction and Subcellular localization.

### Acknowledgements

We would like to thank Associate Professor Min Yuan (North China University of Science and Technology) for providing pHG vectors, Lecturer Zhen-Bang Yang (North China University of Science and Technology) for his help in observing subcellular localization with confocal microscope and Chun-Liu Zuo master for providing technical assistance.

### Author contributions

These authors contributed equally (S W, J D, and X-L Z). S W, Y-H L, and Z-B X conceived and designed the experiments. X S performed the metabolome data analysis. J D and X-L Z performed the *EsMBD* genes cloning in subcellular localization. qRT-PCR was performed by X S. X-L Z performed the transcriptome data analysis. X-L Z performed the transcriptome data analysis. S W wrote the manuscript.

### Funding

This work was supported by Natural Science Foundation of Hebei Province (H2020209302), Science and Technology Project of Hebei Education Department (ZD2019075), and The National Natural Science Foundation of China (31570683).

### Data availability

The data that support the findings of this study are available from the NCBI database (Accession: SRR19962743-SRR19962757) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with the permission of Zhaobin Xing (xingzb@ncst.edu.cn).

### 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: 11 August 2022 / Accepted: 15 February 2023

Published online: 22 February 2023

### References

- Li X, Chen C, Leng A, et al. Advances in the extraction, purification, structural characteristics and biological activities of *Eleutherococcus senticosus* polysaccharides: a promising medicinal and edible resource with development value. *Front Pharmacol*. 2021;12:753007.
- Jin L, Schmiech M, Gaafary EL. A comparative study on root and bark extracts of *Eleutherococcus senticosus* and their effects on human macrophages. *Phytomedicine*. 2020;68:153181.
- Huang LZ, Zhao HF, Huang BK, et al. *Acanthopanax senticosus*: review of botany, chemistry and pharmacology. *Pharmazie*. 2011;66:83–97.
- Szkopinska A, Plochocka D. Farnesyl diphosphate synthase; regulation of product specificity. *Acta Biochim Pol*. 2005;52(1):45–55.
- Abe I, Rohmer M, Prestwich GD. Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. *Chem Rev*. 1993;93(6):2189–206.
- Hwang HS, Lee H, Choi YE. Transcriptomic analysis of siberian ginseng (*Eleutherococcus senticosus*) to discover genes involved in saponin biosynthesis. *BMC Genomics*. 2015;16:180.
- Cui MH, Lin L, Zhang J, et al. Effect of DNA methylation of *Eleutherococcus senticosus* MDD Gene promoter on saponin synthesis. *Phytone-international J Experimental Bot*. 2022. <https://doi.org/10.32604/phyton.2022.021631>.
- Wang Z, Guo HY, Zhang YT, et al. DNA methylation of farnesyl pyrophosphate synthase, squalene synthase, and squalene epoxidase gene promoters and effect on the saponin content of *Eleutherococcus senticosus*. *Forests*. 2019;10(12):1053.
- Kleinwächter M, Selmar D. New insights explain that drought stress enhances the quality of spice and medicinal plants: potential applications. *Agron Sustain Dev*. 2015;35:121–31.
- Chen Y, Guo Q, Liu L, et al. Influence of fertilization and drought stress on the growth and production of secondary metabolites in *Prunella vulgaris* L. *J Med Plants Res*. 2011;5(9):1749–55.
- Liu H, Wang X, Wang D, et al. Effect of drought stress on growth and accumulation of active constituents in *Salvia miltiorrhiza* Bunge. *Ind Crops Prod*. 2011;33:84–8.
- Sun M, Yang Z, Liu L, et al. DNA methylation in plant responses and adaptation to abiotic stresses. *Int J Mol Sci*. 2022;23:6910.
- Du Q, Luu P-L, Storzaker C, et al. Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics*. 2015;7(6):1051–73.
- Gigek CO, Chen ES, Smith MAC. Methyl-CpG-Binding protein (MBD) family: epigenomic read-outs functions and roles in tumorigenesis and psychiatric diseases. *J Cell Biochem*. 2016;117(1):29–38.
- Hameed UF, Lim J, Zhang Q, et al. Transcriptional repressor domain of MBD1 is intrinsically disordered and interacts with its binding partners in a selective manner. *Sci Rep*. 2014;4:4896.
- Parida AP, Sharma A, Sharma AK. AtMBD4: a methylated DNA binding protein negatively regulates a subset of phosphate starvation genes. *J Biosci*. 2019;44(1):14.
- Parida AP, Sharma A, Sharma AK. AtMBD6, a methyl CpG binding domain protein, maintains gene silencing in Arabidopsis by interacting with RNA binding proteins. *J Biosci*. 2017;42(1):57–68.
- Peng M, Cui Y, Bi YM, et al. AtMBD9: a protein with a methyl-CpG-binding domain regulates flowering time and shoot branching in Arabidopsis. *Plant J*. 2006;46(2):282–96.
- Monroe JG, Sriktant T, Carbonell-Bejerano P, et al. Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature*. 2022;602(7895):101–5.

20. Graf G, Zemach A, Pitto L. Methyl-CpG-binding domain (MBD) proteins in plants. *Biochim Biophys Acta*. 2007;1769(5–6):287–94.
21. Zhao N, Ma D, Leong WY, et al. The methyl-CpG-binding domain (MBD) is crucial for MeCP2's dysfunction-induced defects in adult newborn neurons. *Front Cell Neurosci*. 2015;9:158.
22. Napier JD, Grabowski PP, Lovell JT, et al. A generalist-specialist trade-off between switchgrass cytotypes impacts climate adaptation and geographic range. *Proc Natl Acad Sci U S A*. 2022;119(15):e2118879119.
23. Paterson AH, Wendel JF, Gundlach H, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature*. 2012;492(7429):423–7.
24. Yang Z, Chen S, Wang S, et al. Chromosomal-scale genome assembly of *Eleutherococcus senticosus* provides insights into chromosome evolution in Araliaceae. *Mol Ecol Resour*. 2021;21(7):2204–20.
25. Jeltsch A, Broche J, Lungu C, et al. Biotechnological applications of MBD domain proteins for DNA methylation analysis. *J Mol Biol*. 2020;432(6):1816–23.
26. Xiao Y, Li M, Wang J. The impacts of allopolyploidization on Methyl-CpG-Binding domain (MBD) gene family in *Brassica napus*. *BMC Plant Biol*. 2022;22(1):103.
27. Liu K, Xu C, Lei M, et al. Structural basis for the ability of MBD domains to bind methyl-CG and TG sites in DNA. *J Biol Chem*. 2018;293(19):7344–54.
28. Srivasatava AK, Lu Y, Zinta G, et al. UTR-dependent control of gene expression in plants. *Trends Plant Sci*. 2018;23(3):248–59.
29. Sun HX, Li Y, Niu QW, et al. Dehydration stress extends mRNA 3' untranslated regions with noncoding RNA functions in *Arabidopsis*. *Genome Res*. 2017;27(8):1427–36.
30. Hertz EPT, Kruse T, Davey NE, et al. A conserved motif provides binding specificity to the PP2A-B56 phosphatase. *Mol Cell*. 2016;63(4):686–95.
31. Feng Y, Xie Z, Jiang X, et al. The applications of promoter-gene-engineered biosensors. *Sensors*. 2018;18(9):2823.
32. Xu N, Wei L, Liu J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis. *World J Microbiol Biotechnol*. 2019;35(2):33.
33. Shohael AM, Ali MB, Yu KW, et al. Effect of light on oxidative stress, secondary metabolites and induction of antioxidant enzymes in *Eleutherococcus senticosus* somatic embryos in bioreactor. *Process Biochem*. 2006;41(5):1179–85.
34. Guo S, Zhang S, Jia L, et al. Root growth of *Eleuthero* (*Eleutherococcus senticosus* [Rupr. & Maxim.] Maxim.) Seedlings cultured with chitosan oligosaccharide addition under different light spectra. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*. 2020;48(2):626–35.
35. Xu MY, Wu KX, Liu Y, et al. Effects of light intensity on the growth, photosynthetic characteristics, and secondary metabolites of *Eleutherococcus senticosus* Harms. *Photosynthetica*. 2020;58(3):881–9.
36. Anjum SA, Ashraf U, Zohaib A, et al. Growth and developmental responses of crop plants under drought stress: a review. *Zemdirbyste*. 2017;104(3):267–76.
37. Sapna H, Ashwini N, Ramesh S, et al. Assessment of DNA methylation motif under drought stress using methylation-sensitive randomly amplified polymorphism analysis in rice. *Plant Genetic Resources: Characterization and Utilization*. 2020;18(4):222–30.
38. Qian Y, Ren Q, Jiang L, et al. Genome-wide analysis of maize MBD gene family under drought stress using methylation-sensitive treatment at the seedling stage. *Plant Biotechnol Rep*. 2020;14(3):323–38.
39. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296–303.
40. Studer G, Rempfer C, Waterhouse AM, et al. QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics*. 2020;36(8):2647.
41. Coelho FS, Sangi S, Moraes JL, et al. Methyl-CpG binding proteins (MBD) family evolution and conservation in plants. *Gene*. 2022;824:146404.
42. Yano A, Kodama Y, Koike A, et al. Interaction between Methyl CpG-Binding protein and ran GTPase during cell division in tobacco cultured cells. *Ann Bot*. 2006;98(6):1179–87.
43. Wang Y, Zhang H, Ri HC, et al. Deletion and tandem duplications of biosynthetic genes drive the diversity of triterpenoids in *Aralia elata*. *Nat Commun*. 2022;13(1):2224.
44. Chen CJ, Chen H, Zhang Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant*. 2022;13(8):1194–2022.
45. Guo HY, Zhang J, Lin LM, et al. Metabolome and transcriptome analysis of eleutheroside B biosynthesis pathway in *Eleutherococcus senticosus*. *Heliyon*. 2022;8(6):e09665.
46. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods*. 2001;25(4):402–8.

## Publisher's Note

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