BMC Plant Biology

RESEARCH Open Access

Genome-wide identification of the *Medicago sativa* L. MYB family and its transcriptional dynamics during pollen development

Zhenfei Yan[1], Yaqi Feng[1], Qi Yan[1], Pan Xu[1], Fan Wu[1], Caibin Zhang[1] and Jiyu Zhang[1*]

Abstract

Background The myeloblastosis (MYB) gene family plays crucial roles in the development of anthers and the establishment of pollen morphology during plant growth. However, little is known about the role of MYB transcription factors in pollen development in alfalfa (*Medicago sativa* L.).

Results In this study, we identified 161 *MsMYBs* in the alfalfa genome, including 34 1R-MYBs, 123 R2R3-MYBs, 3 3R-MYBs, and 1 4R-MYBs (categorized by the number of repeats). These were classified into six subfamilies based on the phylogenetic analysis, conserved structural domains, and gene structures. All *MsMYBs* were predicted to be hydrophilic and localized in the cell nucleus. The promoter regions contained three classes of cis-regulatory elements related to pollen development, as well as a variable set of functionally diverse elements, including hormone responsiveness, growth and development, and stress responsiveness elements. A transcriptome and qRT-PCR analysis revealed 12 *MsMYBs* with anther-specific expression and exhibited distinct expression patterns. Some *MsMYBs* showed a close phylogenetic relationship with *Arabidopsis MYBs* related to pollen development, such as *MsMYB49* and *MsMYB100*, were found to be localized in the nucleus upon subcellular localization analysis. This genetic proximity suggests a potential role for these *MsMYBs* in the developmental processes of pollen.

Conclusions This study provides a comprehensive understanding of *MsMYBs* in alfalfa and elucidates their potential roles and expression patterns in pollen development.

Keywords *Medicago sativa* L., MYB gene family, Pollen development

Background

Alfalfa (*Medicago sativa* L.), a widely cultivated perennial leguminous plant and important energy crop, plays a significant role not only in animal husbandry but also in soil improvement and ecosystem services [1]. To enhance the breeding of superior alfalfa varieties and the rate of variety renewal, heterosis is a highly effective approach. However, mechanical emasculation is not only time-consuming, labor-intensive, and costly but also detrimental to plant growth, leading to reduced yields of hybrid varieties, which poses a challenge to commercial breeding and seed production [2]. To address these issues, the

*Correspondence:
Jiyu Zhang
zhangjy@lzu.edu.cn
[1]State Key Laboratory of Herbage Improvement and Grassland Agro-Ecosystems, Key Laboratory of Grassland Livestock Industry Innovation, Ministry of Agriculture and Rural Affairs, College of Pastoral Agriculture Science and Technology, Key laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, Lanzhou University, Lanzhou, Gansu 730020, China

development of efficient hybrid breeding methods is necessary. Mining genes that regulate pollen development, gene editing, and the use of male sterile lines for hybrid production is the most economical and effective method to utilize heterosis in alfalfa [3–5].

All eukaryotic organisms contain the MYB transcription factor (TF) family, which is one of the largest TF families in plants [6]. Its members contain a highly conserved MYB domain and regulate a variety of biological processes in plants, including cell differentiation, growth and development, secondary metabolism, and responses to biotic and abiotic stresses [7, 8]. MYBs can be divided into four types based on the number of repeat units: 1R-MYB (1-repeat), R2R3-MYB (2-repeats), R1R2R3-MYB (3-repeats), and 4R-MYB (4-repeats). Each repeat sequence contains 50–53 amino acids, encoding three alpha-helices, with the second and third helices forming a helix-turn-helix (HTH) structure [9]. 1-RMYBs play an important role in regulating plant transcription and maintaining chromosome structure. R2R3-MYB family genes contain two conserved R2 and R3 repeat sequences in the MYB domain, with transcriptional activation or repression abilities at the C-terminus. Its members are numerous and have a wide range of functions (e.g., cell differentiation, secondary metabolism, and environmental responses) [10]. R1R2R3-MYB subfamily genes are mainly involved in the regulation of cell differentiation and cell cycle progression [11]. The conserved structural domain of 4R-MYB subfamily genes consists of four R1/R2 repeat sequences; relatively little research has focused on this subfamily in plants.

In flowering plants, normal development of the tapetal layer, pollen wall formation, and normal dehiscence of anthers are key steps to maintain plant fertility [12]. Errors in these processes can lead to male sterility. A large number of studies have shown that MYBs are involved in the development of the tapetal layer, pollen wall formation, and anther dehiscence. After the initial formation of the tapetal layer cells, AtMYB33 and AtMYB35 are strongly expressed in tapetal layer cells, pollen mother cells, and microspore cells [13, 14]. Mutations in these genes result in the abnormal development of the tapetal layer, leading to an inability to form pollen and a male-sterile phenotype. In an *atmyb35* mutant, the number of tapetal layer cell divisions increases significantly during meiosis, resulting in a non-uniform cell layer; in addition, the tapetal layer cannot secrete callose enzymes and the tetrads are continuously squeezed by the enlarged vacuolated tapetal layer until they are completely degraded [14]. The AtMYB26 protein is present in the anthers and is directly regulated by the auxin response factor ARF8.4. MYB26 functions upstream of NST1 and NST2, directly inducing the expression of NST1 and NST2, involved in secondary thickening in the anther [15]. It is speculated that AtMYB26 promotes anther dehiscence. Studies have also found that GhMYB24 transcript levels are high in anthers and mainly in pollen. Its overexpression in *Arabidopsis* leads to anthers not dehiscing and a reduction in pollen grains, causing male sterility. Additionally, GhMYB24 interacts with intracellular GhJAZ1/2, affecting the jasmonic acid pathway, thereby affecting stamen development [16]. Therefore, MYB genes affect pollen development directly as well as indirectly through hormonal pathways.

Although the alfalfa MYB gene family members were reported in 2019 [17], a comprehensive confirmation of their chromosomal distribution and detailed identification was not available. This was due to the scarcity of genomic and annotation data for alfalfa at that time. In addition, the expression patterns of MYBs in the pollen development in alfalfa have not been determined. Therefore, to address this knowledge gap, a comprehensive whole-genome and transcriptome sequencing analyses of the MYB family in alfalfa were performed. The results of this study provide a basis for further analyzing the functions of *MsMYBs* and clarifying their roles in the pollen development process. Furthermore, this study provides reliable target genes for gene editing to generate male sterile lines.

Results

Morphological characteristics of alfalfa anthers at different developmental stages

We collected flower buds at 8 days of growth. As the buds grew, gradual morphological changes were observed (Fig. 1A). Semi-thin sections were prepared for each stage of bud development. The tapetal layer mainly appeared at stages S8 and S9, at which point the anther locule microspore cells had not developed into mature pollen grains. At stage S10, the tapetal layer cells underwent programmed cell death and were nearly undetectable; however, mature pollen grains began to appear. At stage S11, the microspore cells completely developed into mature pollen grains, and the anther locule ruptured. At stage S13, the anther outer wall ruptured, and the pollen grains were released from the anther chamber (Fig. 1B).

Identification and analysis of alfalfa *MsMYBs*

In contrast to previous studies, our genomic analysis identified 161 MYB genes in the Zhongmu No.1 genome, a discrepancy likely attributable to underlying genomic variations. These MYBs were numbered in increasing order based on the alfalfa genome sequence and are designated as *MsMYB1* to *MsMYB161*. Detailed information about these MYB is presented in

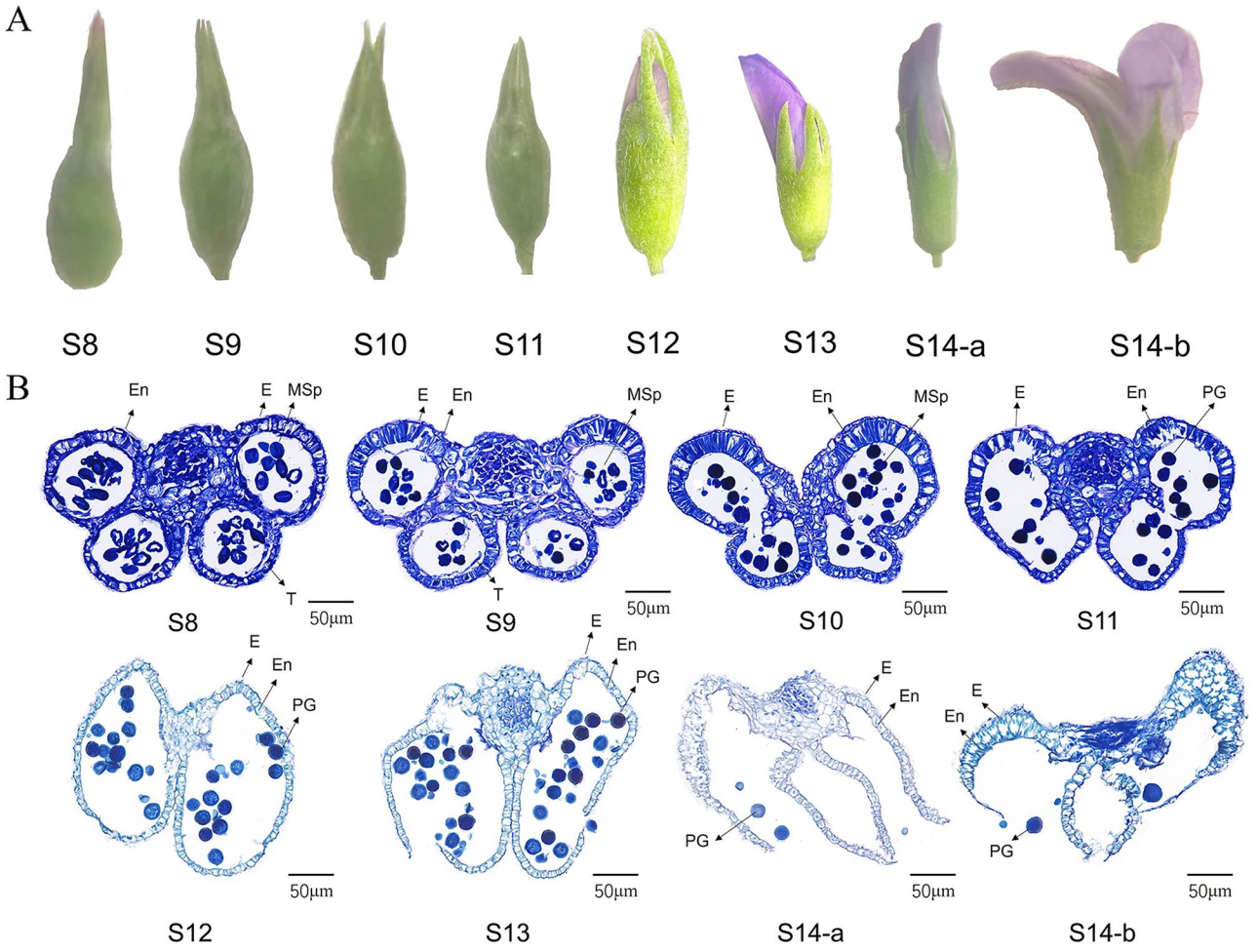

Fig. 1 Examination of alfalfa anther phenotypes. (**A**) Epigenetic morphology of flower buds at various developmental stages. (**B**) Observation of anther morphology across different developmental stages. *E*, epidermis; *En*, endothecium, *T*, tapetum, *PG*, pollen grains

Table S2. The lengths of the *MsMYB* polypeptide chains varied from 89 to 1771 amino acids, as shown in Supplementary Table 1. Additionally, the predicted molecular weights ranged from 10.2 kDa to 196.1 kDa, and the theoretical pI values ranged from 4.25 to 10.33, as indicated in Table S2. All MsMYB proteins were hydrophilic, and the proteins were predicted to be localized in the cell nucleus.

Chromosomal localization of *MsMYBs*

Through DNA sequence annotation, the 161 *MsMYBs* were distributed across eight chromosomes (Fig. 2). The analysis showed that 25 *MsMYBs* were located on chromosome 1, 13 on chromosome 2, 22 on chromosome 3, 17 on chromosome 4, 21 on chromosome 5, 14 on chromosome 6, 27 on chromosome 7, and 19 on chromosome 8. In addition, chromosomal locations of 3 *MsMYBs* could not be determined. Notably, chromosome 7 contained the most *MsMYBs* (27), while chromosome 2 had the fewest (13).

Phylogenetic, gene structure, and motif pattern analyses of *MsMYBs* in alfalfa

To explore evolutionary relationships among *MsMYB* genes, a neighbor-joining phylogenetic tree was constructed using MEGA11 based on amino acid sequences (Fig. 3A,). The *MsMYB* genes formed six clades (A–F). Clade F was the largest, with 51 *MsMYB* members. Clade A had 13 *MsMYB* genes, clade B had 18 *MsMYB* genes, clade C had 17 *MsMYB* genes, clade D had 25 *MsMYB* genes, and clade E had 37 *MsMYB* genes. To investigate the evolutionary relationships between alfalfa *MYB* genes and their Arabidopsis counterparts, we constructed a phylogenetic tree, which classified both into six distinct subgroups (Fig. S1). Notably, Clade F harbored the largest number of MYB members and included functionally characterized *AtMYB* genes. This phylogenetic context provides critical insights for inferring potential roles of *MsMYB* genes. For instance, *AtMYB35* and *AtMYB80*, located within Clade F, are well-documented to play crucial roles in anther tapetum development and pollen

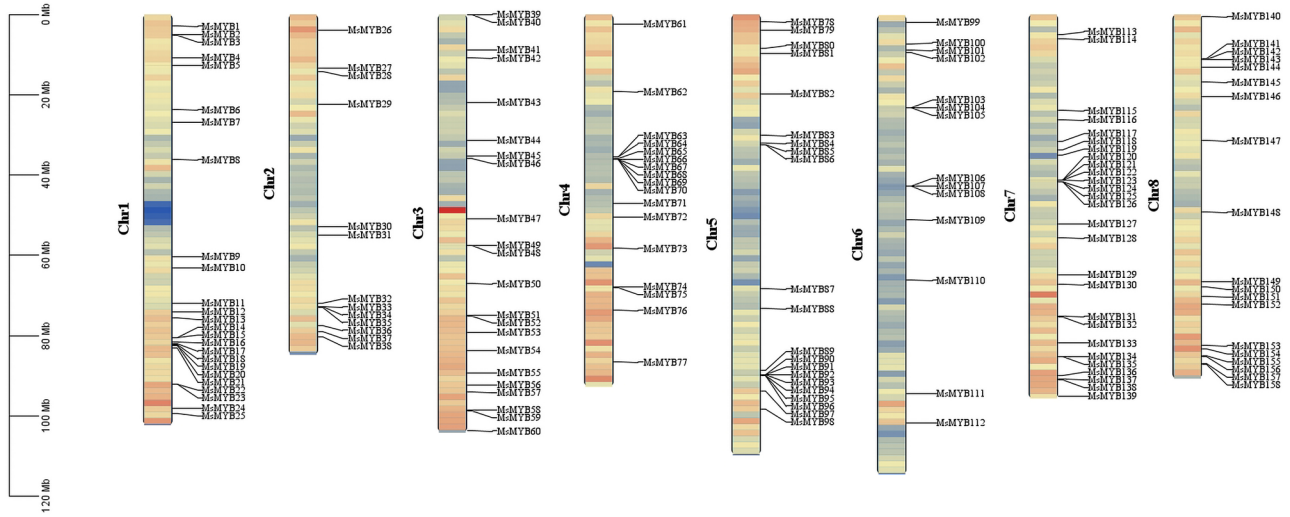

Fig. 2 Distribution of MsMYB genes across the 16 chromosomes of the coconut genome. The scale indicates the length of the coconut chromosomes

maturation. Strikingly, their closely neighboring *MsMYB* homologs (*MsMYB48*, *MsMYB49*, and *MsMYB100*) exhibit analogous genomic positioning and expression patterns, strongly suggesting conserved functional importance during anther development.

Additionally, 3D structural predictions indicated that, in most cases, members of the MYB gene family have similar 3D structures, with particularly high similarity among members within the same group (Fig. S2). For instance, the four gene family members *MsMYB22*, *MsMYB85*, *MsMYB86*, and *MsMYB133* in clade F exhibited a very high degree of similarity in three-dimensional structures.

To examine the organization of exons and introns, the genomic and cDNA sequences of all *MsMYB* members were compared. The majority of the coding sequences were interrupted by introns, with the number of introns ranging from 0 to 23 (Fig. 3A). The intron–exon structures of *MsMYBs* in clade B had the greatest diversity; *MsMYB115*, *MsMYB56*, *MsMYB81*, and *MsMYB97* did not contain introns, while *MsMYB5* had the most exons (22) and introns (23). The remaining *MsMYBs* had intron numbers ranging from 1 (*MsMYB54*) to 15 (*MsMYB159*). Clades A, C, D, E, and F had relatively similar intron–exon structures; most genes had 1 to 5 exons, while *MsMYB18* and *MsMYB52* contained 10 and 11 exons, respectively, indicating a conserved intron distribution within these *MsMYBs* lineages. The differences in the composition of MsMYBs suggest significant biases within the *M. sativa* genome. Additionally, 20 *MsMYBs* across all clades contained both 5' and 3' UTRs, while the remaining lacked either the 5' or 3' UTR, indicating that the expression of *MsMYB* genes is influenced by at least three different post-transcriptional regulatory processes.

To further understand the functions of *MsMYBs*, conserved motifs were analyzed. Ten motifs were identified. Based on the phylogenetic analysis, many closely related proteins exhibited common motifs in the same arrangement and position, suggesting that MYB members within a subgroup have similar biological functions (Fig. 3A). The most common conserved motifs in R2R3-MYB genes were motif 1, motif 2, motif 4, motif 5, and motif 6, which were 15, 19, 21, 15, and 16 amino acids (Table S3), respectively. The HTH (helix-turn-helix) structure was formed by regularly spaced tryptophan residues, consistent with previous findings in other species [18].

Cis-acting regulatory elements in the MsMYB promoter regions

Genes with similar expression patterns may share common transcriptional regulatory mechanisms and binding sites for trans-acting factors. Cis-acting DNA sequences play a key role in the transcriptional control of spatiotemporal gene expression in response to developmental and environmental cues [19]. Moreover, differences in cis-acting elements among members of a gene family can provide insight into functional divergence after gene duplication. Therefore, we analyzed the promoter regions of each *MsMYB* for cis-acting regulatory elements (CAREs) by searching 2 kb upstream sequences (Fig. 4). We identified three types of CAREs: hormone responsiveness, growth and development, and stress responsiveness (Fig. S3). Hormones are crucial in the development of pollen; in the hormone response category, 11 cis-elements were detected (Fig. S4), including ABRE (34.5%) related to abscisic acid (ABA), CGTCA-motif and TGACG-motif (34.1%) related to jasmonic acid methyl ester (MeJA), four cis-elements related to auxin response (AuxRR-core, AuxRE, TGA-element, and TGA-box)

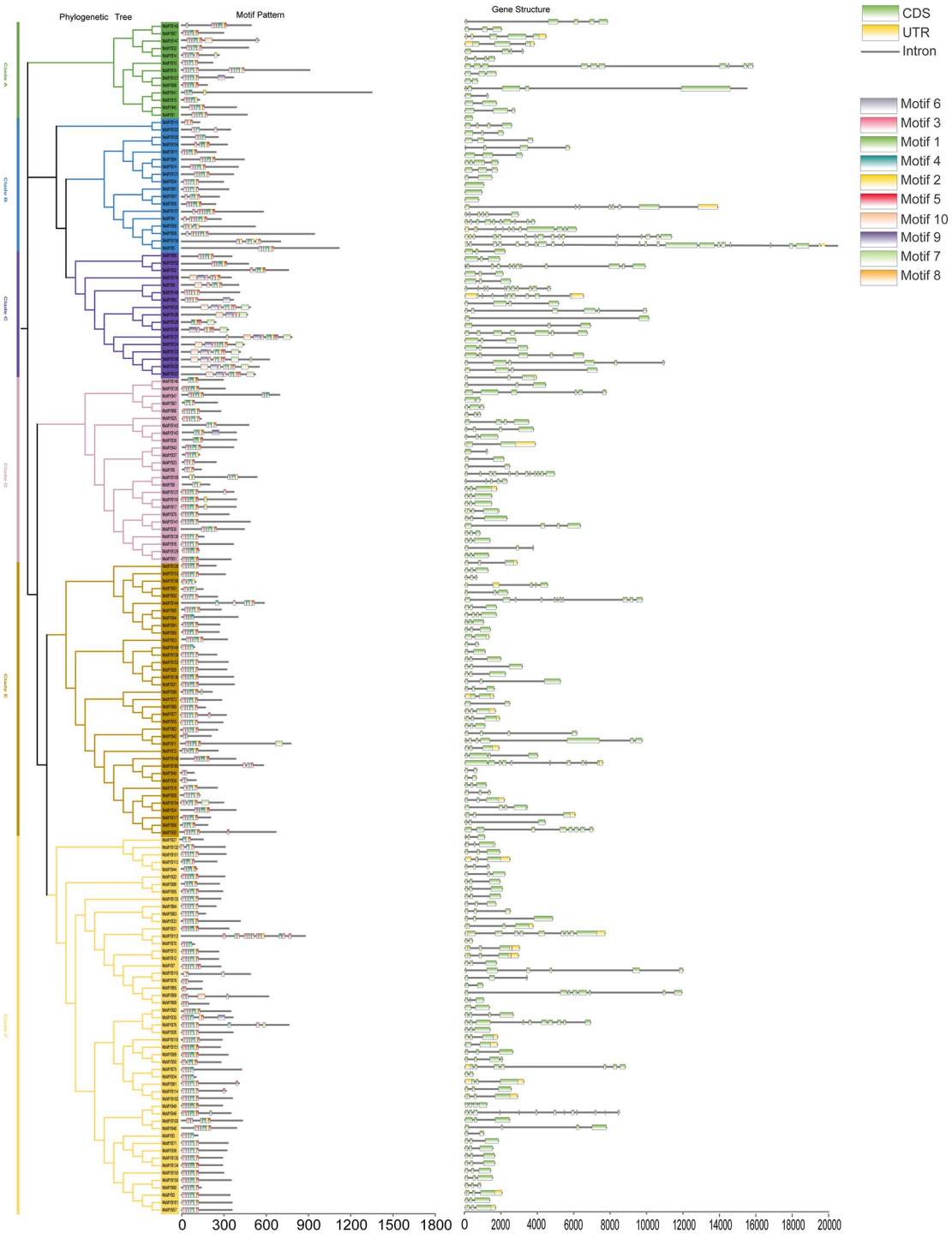

Fig. 3 Polygenic relationship, gene structure, and motif composition of the MsMYBs in alfalfa. (**A**) The phylogenetic tree was constructed using MEGA11.0 with the Neighbor-Joining (NJ) method. Ten motifs are displayed in colored rectangles. The exon-intron structures of the 161 putative MsMYB genes are shown, with exons represented by green rectangles and introns by black lines. Yellow rectangles denote untranslated regions (UTRs) of mRNA, upstream of the start codon (5′-UTR) and downstream of the stop codon (3′-UTR). (**B**) Distribution of ten conserved motifs within MsMYB genes, each represented by a numbered colored box. Sequence logos of the amino acid residues of the ten conserved motifs of MsMYB proteins are also presented

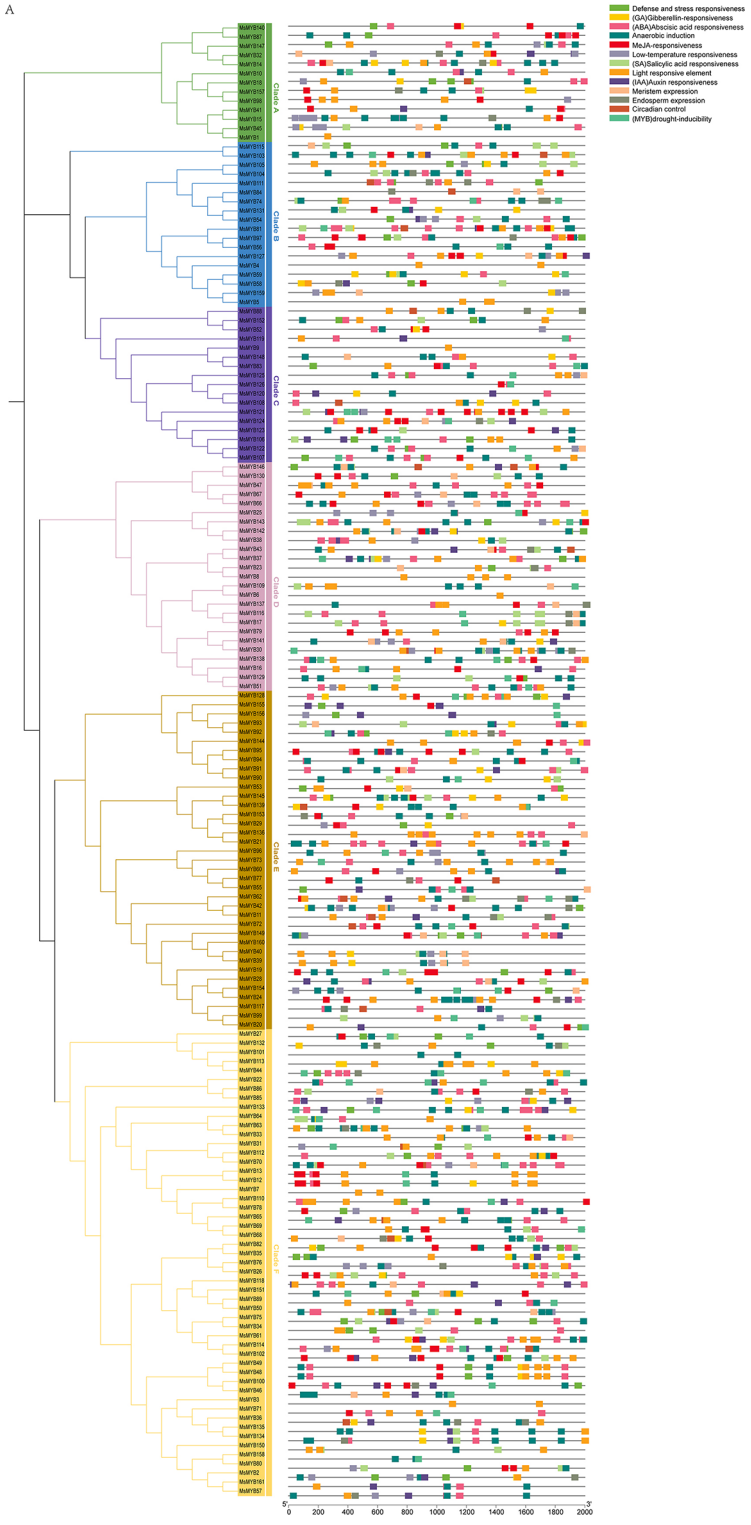

Fig. 4 Analysis of cis-regulatory element distribution in the MsMYB promoters. Cis-elements in the MsMYBs are marked with different colors

(9%), P-box and GARE-motif related to gibberellin (GA) responsiveness (5.8%), SARE and TCA-element related to salicylic acid (SA) (10.5%). Promoter regions of four *MsMYBs* contained all of these cis-regulatory elements. In the plant growth and development category, seven cis-regulatory elements were retrieved, including the Meristem expression element (CAT-box), endosperm expression element (GCN4_motif), and circadian

elements, with three *MsMYBs* containing all of these elements. In the stress responsiveness category, four cis-elements were detected, including anaerobic induction element (ARE), low-temperature responsiveness (LTR), defense and stress responsiveness (TC-rich repeats), and (MYB) drought-inducibility (MBS), with ten *MsMYBs* containing all of these elements.

Synteny analysis of *MsMYBs*

Gene duplication is an important mechanism underlying functional diversity. To identify duplication events in *MsMYBs*, we performed a collinearity analysis using MCScanX software. A total of 158 *MYBs* had 37 pairs of segmental duplications, with the most duplications on chromosome 7 (Fig. 5A, Table S4). The expansion of the MsMYB family on the chromosomes indicates that whole-genome duplication (WGD) led to balanced gene drive [20]. By calculating the ratio of non-synonymous to synonymous substitutions (K_a/K_s), we estimated the rate of evolution and selective pressure of the MsMYB family. The K_a/K_s ratio was less than 1 for all genes, with values ranging from 0 to 0.91, indicating that *MsMYBs* in the *M. sativa* genome mainly evolved under purifying selection, in which deleterious mutations are removed from the population [21].

Using MCScanX, we analyzed synteny between MsMYB and MYB in three closely related species and one distant species (*Arabidopsis*, *Cicer arietinum*, *Trifolium pratense* L., and *Medicago truncatula*). The genome sizes of these species were as follows: *Arabidopsis* (115 Mb), *C. arietinum* (513 Mb), *T. pratense* (399 Mb), and *M. truncatula* (414 Mb). There were 194 pairs of homologs between *M. sativa* and *T. pratense*, followed by 192 pairs in comparisons with *M. truncatula*, 170 pairs with *C. arietinum*, and 95 pairs with *Arabidopsis* (Fig. 5B, Table S5). These results indicate that *M. sativa* is more closely related to *T. pratense* than to the other three species.

Transcriptomic profiles and differential expression of *MsMYBs* in alfalfa anthers at different developmental stages

To explore transcriptomic dynamics during the development of mature pollen grains from microspore cells, we collected anthers from stages S8 to S11 for transcriptome sequencing. Compared with levels at stage S8, 290 genes were significantly upregulated and 307 were downregulated at stage S9, 1641 genes were significantly upregulated and 1326 were downregulated at stage S10, and 3136 genes were significantly upregulated and 2145 were downregulated at stage S11. A Venn diagram showed that 193 differentially expressed genes (DEGs) were commonly upregulated and 135 DEGs were commonly downregulated at stages S9, S10, and S11 (Fig. 6A). Heatmaps and a principal component analysis (PCA) revealed that the transcriptomic profiles at S8–S10 were relatively similar and differed significantly from the transcriptome at stage S11 (Fig. 6B). Based on the state of pollen development, we speculate that the complete development of microspore cells into mature pollen grains is a key event associated with changes in anther transcriptome dynamics.

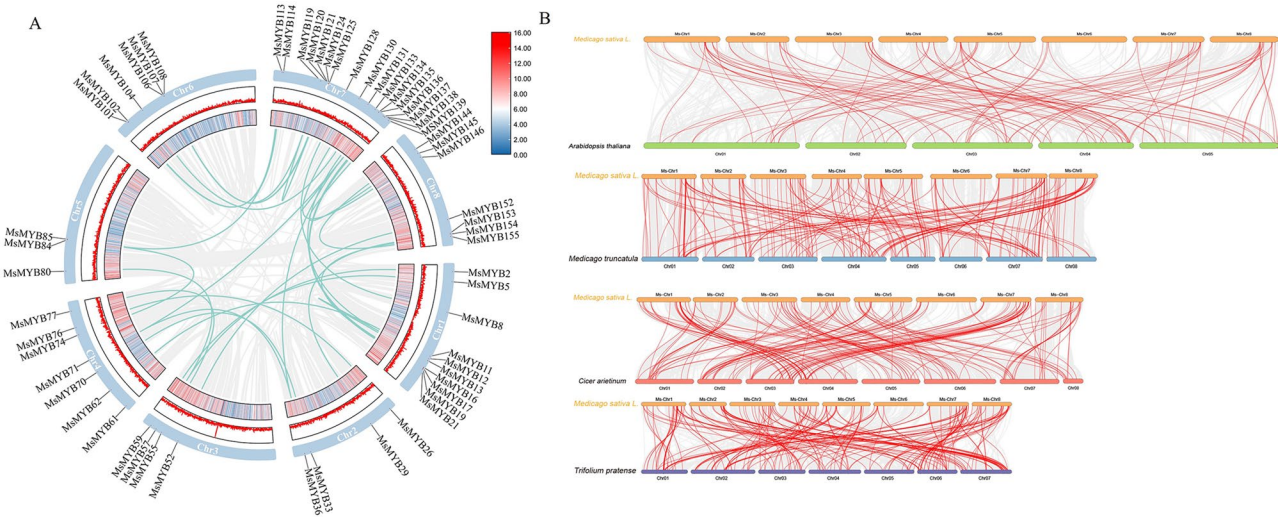

Fig. 5 Chromosome covariance analysis of MsMYB genes. (**A**) Circos plot illustrating the distribution of MsMYB genes among eight scaffold chromosomes and synteny analysis within the alfalfa genome. Annotations on the fragments represent different chromosomes, with numbers in the outer circle indicating their positions. The approximate gene positions are marked with short black lines. Paralogous MsMYBs in alfalfa are mapped to their respective locations in the circular diagram. Blue-green and grey lines indicate segmental duplication pairs between MsMYBs and within the whole genome, respectively. (**B**) Synteny analysis of MsMYB genes between alfalfa and four other species. Gray lines in the background signify collinear blocks within alfalfa and other plant genomes, while red lines highlight syntenic MsMYB gene pairs

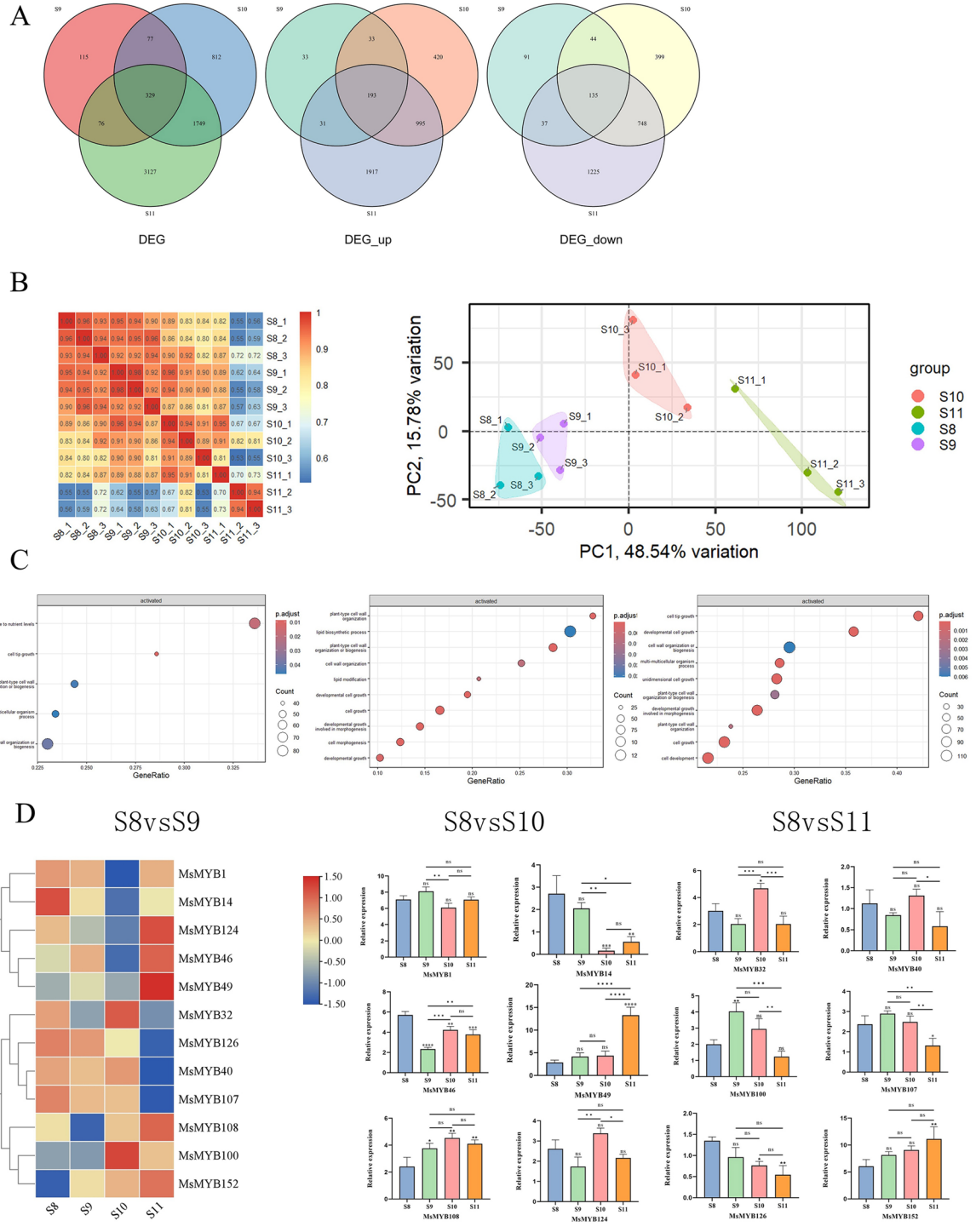

Fig. 6 Differential MsMYBs expression in anther development stages. (**A**) Venn diagram displaying the intersection of significantly up-regulated and down-regulated DEGs between stages S9, S10, and S11. (**B**) Heatmaps and principal component analysis (PCA) reveal the relationships between transcriptomes. (**C**) GO enrichment bubble chart of upregulated DEGs during stages S9, S10, and S11. (**D**) Expression patterns of MsMYBs during anther development. Color scales with Z scores beside the heatmap indicate gene expression levels, with low and high transcript abundance indicated by blue and red, respectively. Error bars represent standard deviation, and asterisks denote significant differences between control and treatments, with $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$

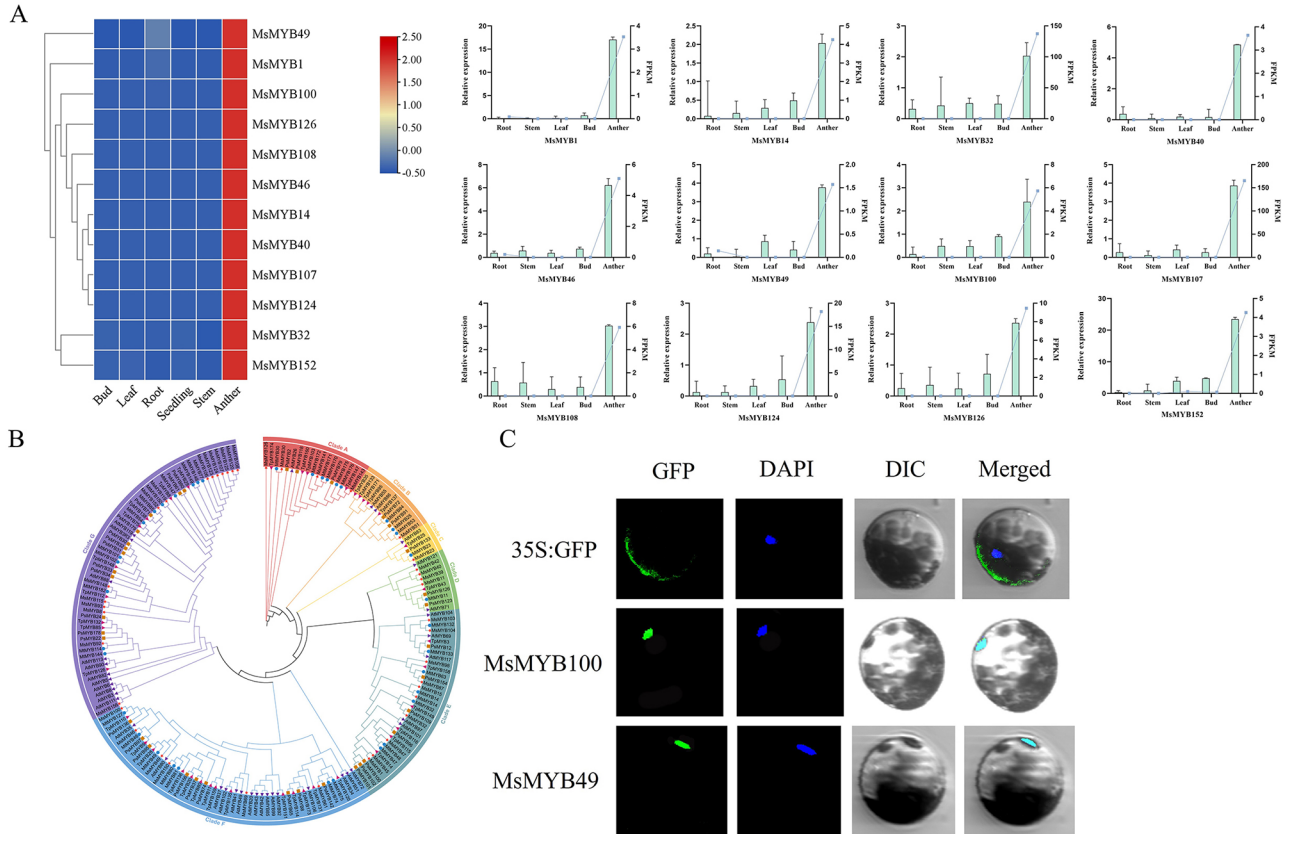

Fig. 7 Expression analysis of MsMYB genes in various tissues and phylogenetic trees of MYB proteins. (**A**) Heatmap of differentially expressed MsMYBs in leaf, root, stem, bud, and anther. Color scales with Z scores beside the heatmap indicate gene expression levels, with low and high transcript abundance indicated by blue and red, respectively. (**B**) Phylogenetic tree of MYBs in *Arabidopsis*, *Pisum sativum*, *Trifolium pratense* L., *Medicago truncatula*, and *Medicago sativa* L., classified into seven subgroups labeled from Clade A to Clade G with different colors. Different symbols represent different species. (**C**) MsMYB49 and MsMYB100 are localized in nucleus. Marker represents a known marker protein localized in nucleus

Enrichment analysis of differentially expressed genes (DEGs) revealed that, in comparison to stages S9 and S11, DEGs at stage S10 were significantly enriched in processes related to cell wall remodeling and lipid synthesis metabolism, indicating the critical nature of this stage in the maturation of pollen grains from microspore cells in alfalfa (Fig. 6C).

Heatmap and quantitative real-time PCR (qRT-PCR) analyses identified 12 alfalfa MYB (MsMYB) transcripts with significant expression changes during microspore development into mature pollen grains (Fig. 6D). Notably, *MsMYB49*, *MsMYB108*, and *MsMYB152* exhibited a progressive increase in expression throughout development, while *MsMYB14* and *MsMYB126* showed a decline. Early in anther development (S8), key processes include tapetal layer cell development and the breakdown of the tetrad callose structure. The tapetal layer synthesizes and transports nutrients to the anther chamber, supporting meiosis in pollen mother cells and microspore development. At S8, the tapetal layer commences degradation, and lipids like sporopollenin, essential for pollen wall formation, are synthesized and transported to microspore surfaces via Ubisch bodies, ensuring pollen wall stability and viability. As pollen matures, the anther chamber ruptures, releasing pollen. We hypothesize that early-expressed *MsMYB* genes participate in tapetal layer and tetrad development and degradation, late-expressed *MsMYB* genes contribute to pollen wall synthesis, and those with intermediate expression patterns may play a role in overall pollen development.

Expression patterns of *MsMYBs* in different tissues and subcellular localization

To explore the expression patterns of these 12 *MsMYB* genes across various tissues, transcriptome data for roots, leaves, stems, and buds were retrieved from the alfalfa database. The heatmap and qRT-PCR analysis revealed that these 12 *MsMYB* genes were highly expressed exclusively in anthers and minimally in other tissues (Fig. 7A, Table S6). Interestingly, analysis of the promoter cis-acting elements of these 12 *MsMYBs* revealed that they contain a significantly higher number of hormone-responsive cis-acting elements compared to other *MsMYBs*. Given the critical role of hormones in anther development and

the proper establishment of pollen wall morphology, this may represent a potential mechanism underlying the anther-specific expression of these 12 *MsMYBs*. To investigate the potential biological functions and evolutionary relationships of these *MsMYBs* during pollen development, homologs were identified in *Arabidopsis*, *Pisum sativum*, *Trifolium pratense* L., and *Medicago truncatula*, and a phylogenetic tree was constructed using MEGA11. These genes clustered with *MYBs* from other plant species, forming seven distinct clades (Fig. 7B). Previous studies have implicated *AtMYB35* and *AtMYB80* in pollen development; accordingly, we found that these genes, along with *AtMYB35* and *AtMYB80*, were grouped in clade F. *MsMYB49* and *MsMYB100* cluster within the same phylogenetic clade and exhibit high sequence homology to *AtMYB80* and *AtMYB35*, with conserved three-dimensional protein structures (Fig. S5, Fig. S6)., were subjected to subcellular localization in alfalfa protoplasts (Fig. 7C). The results mirrored those in *Arabidopsis thaliana*, with both *MsMYB49* and *MsMYB100* localized to the nucleus, suggesting potential similar functions that warrant further experimental investigation.

Discussion

MYBs belong to one of the largest families of TFs in plants. Ample research has shown that *MYBs* are involved in cell cycle regulation, anther and pollen development, plant hormone responses, and stress responses in plants [22, 23]. They are essential for plant survival and are found across a wide range of plant taxa. In this study, 161 *MsMYBs* were identified through systematic phylogenetic and homology analyses with the whole genome sequences of *Arabidopsis* and *Medicago sativa*. The *MYB* genes could be divided into six clades. Many MYB TFs have been identified in plant genomes, consistent with their key functions [24–26]. For example, 196 *MYBs* have been identified in *Arabidopsis* [27], 197 in *Populus trichocarpa* [28], 179 in *Cocos nucifera* [18], 264 in *Primulina swinglei* [29], and 293 in *Brassica rapa* [30]. Differences in gene family size among species may be due to gene family contraction or expansion in various lineages. In addition, the protein sequences of MsMYB family members exhibit different lengths, ranging from 89 (*MsMYB70*) to 1347 (*MsMYB41*) amino acids, with an average length of 362 amino acids. Their molecular weights also varied, from 10.26 kDa (*MsMYB145*) to 151.36 kDa (*MsMYB5*), with an average molecular weight of 41.10 kDa. The theoretical isoelectric point ranged from 4.25 (*MsMYB68*) to 10.33 (*MsMYB107*), with an average theoretical isoelectric point of 7.02. In a study of *Primulina eburnean* [31], the amino acid lengths ranged from 75 to 1056, with an average length of 386 amino acids, the molecular weight ranged from 18.64 kDa to 99.01 kDa, and the isoelectric point ranged from 4.92 to 10.2. This previous research and the results of our study further support the hypothesis that the MYB family shows significant evolutionary conservation across plant species.

Gene structure analyses are useful for studying the evolutionary relationships within gene families. Our analyses of gene structure and conserved protein motifs show that mosts, like those of other plant species, consisted of 2–5 exons and 1–5 introns; however, there was considerable variation in the number and length of introns and UTRs among the 161 *MsMYBs*. Similar results have been found in *Cocos nucifera* L [18]. and *Primulina eburnean* [31]. UTRs play an important role in the post-transcriptional regulation of gene expression, such as controlling the stability, localization, and translation of mRNA [32, 33]. Specific regulatory elements in introns can also affect gene expression patterns [34]. Genes with more and longer introns may have a slower rate of transcription because the splicing process takes longer to complete. This could lead to lower mRNA and protein levels as well as a longer mRNA half-life in certain tissues or developmental stages. In addition, longer introns may also contain regulatory elements that affect gene expression, such as enhancers and silencers [35]. These factors can affect the transcription rate and alter the spatiotemporal expression patterns of genes. In contrast, genes with fewer and shorter introns may have a faster transcription rate. This could lead to higher mRNA levels and potentially higher protein levels as well as a shorter half-life of mRNA in other tissues or developmental stages. However, shorter introns may also have fewer regulatory elements, which could limit responsiveness to environmental and developmental cues. Therefore, the highly differentiated exon–intron and UTR structures among *MsMYBs* may lead to functional differences. Three types of CARES were detected in the promoter regions of *MsMYBs*, i.e., those involved in hormone responsiveness, growth and development, and stress responsiveness. Previous studies have shown that hormones and stress responses have an important impact on pollen wall synthesis and anther dehiscence [36, 37], which may explain why some *MYBs* contribute to anther development.

Pollen development is a complex and intricate process. Semi-thin sections revealed that by the S10 stage, most microspore cells developed into mature pollen grains. In anthers, lipid metabolism was the most common functional group identified in a Gene Ontology (GO) enrichment analysis. Lipid metabolism, primarily lipid biosynthesis and lipid transport, plays a crucial role in reproductive development, including anther development and dehiscence, in plants [38]. Notably, lipids are mainly synthesized in the tapetal layer cells, which constitute the innermost layer of the anther wall and are transferred to the pollen or other anther wall layers to promote anther development [39]. Compared with the S9

and S11 stages, significant enrichment for lipid metabolism was observed during the S10 stage. Therefore, we inferred that S10 is a critical period for the formation of the pollen wall in alfalfa, and *MsMYB24* was related to this process in a functional enrichment analysis, suggesting that it plays a role in maintaining lipid metabolism during anther development.

We characterized the phylogenetic relationships of 12 anther-specific *MsMYBs* with homologs in four other species. The motifs of *MsMYBs* were highly conserved with those of *MYBs* from other species; however, there were interspecific differences, indicating both conservation and diversity within the MYB family. We also found that *AtMYB26*, *AtMYB35*, *AtMYB80*, and *AtMYB99*, which play important roles in pollen development, are clustered in the evolutionary tree. *AtMYB35* and *AtMYB80* affect pollen development by influencing tapetal layer development [14, 40]. *AtMYB26* promotes anther dehiscence by inducing the expression of NST1 and NST2, leading to secondary thickening within the anther [41]. Transcriptome sequencing also revealed that *AtMYB99* regulates pollen wall biosynthesis in conjunction with the NAC gene *At1g61110* [42]. Therefore, MsMYBs in adjacent clades may exhibit similar functions, and qRT-PCR subsequently confirmed their specific expression in anthers. Extensive research has shown that genes associated with tapetal layer development show decreases in expression levels after the tapetal layer undergoes programmed cell death, while genes related to pollen wall synthesis and anther dehiscence exhibit gradual increases in expression. In this study, we used qRT-PCR to evaluate the expression of 12 *MsMYBs* across four developmental stages of anthers, indicating their involvement in different stages of anther development and pollen maturation. During the process of microspore cells developing into pollen grains, *MsMYB49*, *MsMYB108* and *MsMYB152* levels increased gradually, suggesting that these MsMYBs are involved in pollen wall formation or anther dehiscence. In contrast, *MsMYB14* and *MsMYB126* levels decreased gradually, with the highest expression at the S8 stage, suggesting that these genes are involved in early anther developmental processes, such as tapetal layer cell development or tetrad callose degradation.

Through transcriptome sequencing and qRT-PCR analyses, we identified that *MsMYB49* and *MsMYB100* are specifically expressed in alfalfa anthers. Notably, *MsMYB49* exhibits a gradual increase in expression during microspore development, while *MsMYB100* shows a progressive decline in expression coinciding with tapetal programmed cell death. Phylogenetic analysis revealed that *MsMYB49* and *MsMYB100* share high sequence homology and conserved three-dimensional protein structures with *AtMYB35* and *AtMYB80* from *Arabidopsis*. Furthermore, they exhibited identical subcellular localization patterns and similar spatiotemporal expression dynamics throughout anther development [14, 40].

In *Arabidopsis*, RNA in situ hybridization demonstrated that *AtMYB35* expression in the tapetum gradually decreases during microspore maturation, whereas *AtMYB80* expression is initially low in early tapetal cells, peaks at the onset of cellular degradation, and subsequently declines. Based on these parallels in expression specificity and structural conservation, we hypothesize that *MsMYB49* and *MsMYB100* likely perform conserved functional roles in alfalfa pollen development. Future studies employing gene knockout and complementation assays in *Arabidopsis* mutants approaches will elucidate their detailed mechanistic contributions to pollen and tapetal development.

Conclusion

This study successfully identified 161 *MYBs* within the alfalfa genome through a comprehensive, genome-wide screening approach. A detailed investigation examined the genomic architecture, genetic lineages, chromosomal localization, gene duplication events, conserved motifs, and expression patterns across various tissues and anther developmental stages, pinpointing key periods for alfalfa pollen maturation. Additionally, quantitative qPCR analysis of 12 selected *MsMYBs* revealed their tissue-specific expression in anthers and distinct expression patterns at different developmental stages. These findings offer valuable insights into the fundamental functional differences among *MsMYBs* in diverse alfalfa tissues and identify them as potential candidate genes for pollen development in this significant crop. Moreover, this study provides a comprehensive understanding of the MYB gene family in alfalfa, establishing a robust foundation for further exploration of their functions and evolutionary dynamics.

Materials and methods

Plant material and semi-thin section

The "Lanmu No. 2" alfalfa variety cultivated at Lanzhou University was used. Harvesting began on the eighth day of bud development from a single alfalfa, with collections every other day. Inflorescences at different time points were fixed with formalin-acetic acid for 2–3 days and embedded in Spurr's resin. Semi-thin sections with a thickness of 1 mm were cut and stained for 5 min in a 0.01% toluidine blue/sodium borate solution at 45 °C. After rinsing off the excess solution with water, images of the sections were obtained under bright field using an Olympus BX51 fluorescence microscope. Inflorescences from day 8 to day 11 were taken, and anthers were

removed and frozen in liquid nitrogen for transcriptome sequencing.

Collection and identification of *MsMYB* genes in alfalfa

The reference genome and gene model annotation files were sourced from public databases [43], available at https://figshare.com/articles/dataset/Medicago_sativa_genome_and_annotation_files/12623960. To acquire all MYB sequences of alfalfa, BLASTp homology searches were conducted on the alfalfa genome dataset using the protein sequences of *Arabidopsis* MYB as the query. Candidate *MYB* genes were selected based on an E-value threshold of $< 10^{-10}$. These candidate genes' amino acid sequences were then analyzed by retrieving the Myeloblastosis oncogene family (PF00249) from the InterPro Protein Characterization database [44], accessible at https://www.ebi.ac.uk/interpro/. Subsequently, the HMMER3.0 software [45] was utilized to analyze the genes and retrieve related and aligned sequences from the alfalfa genomic dataset, available at https://www.ebi.ac.uk/Tools/hmmer/. Additional details regarding the number of amino acids, molecular weight (MW), and isoelectric point (pI) of each MYB protein were obtained through the ExPASy proteomic website [46], found at https://web.expasy.org/compute_pi/. The CELLO tool [47] was employed to predict the intracellular distribution of all MsMYB genes. The tertiary structure of each MYB protein was predicted using SWISS-MODEL [48], accessible at https://swissmodel.expasy.org/interactive.

Phylogenetic analysis

Arabidopsis AtMYB protein sequences were sourced from the TAIR database, available at https://www.arabidopsis.org/. Protein sequences for *Pisum sativum* PsMYB, *Trifolium pratense* L. TpMYB, and *Medicago truncatula* MtMYB were obtained from NCBI (https://www.ncbi.nlm.nih.gov). Multiple alignments of protein sequences from *Arabidopsis*, *Pisum sativum*, *Trifolium pratense* L., *Medicago truncatula*, and alfalfa were performed using MEGA11 software [49]. Phylogenetic trees were constructed employing the neighbor-joining (NJ), minimal evolution (ME), and maximum parsimony (MP) methods with bootstrap values based on 1000 replicates.

Exon/intron structure, conserved motifs and promoter analysis

Gene structures were visualized by aligning CDSs to Rab gene sequences using the GeneStructure Display Server [50]. Conserved motifs of MYB proteins were identified using the MEME Version 4.11.4 program (https://meme.ebi.ac.uk/meme/tools/meme) with parameters set for motif width of 6-200 and a maximum of 10 motifs. Upstream DNA sequences of 2.0 kb from the ATG were analyzed using PlantCARE [51] to predict cis-acting regulatory elements.

Chromosomal location and gene synteny analysis

Chromosomal locations for *Arabidopsis*, *Cicer arietinum*, *Trifolium pratense* L., and *Medicago truncatula MYB* genes were obtained from the NCBI database. Synteny among alfalfa MYB family members and the aforementioned plants was analyzed using MCScanX and TBtools [52, 53].

RNA isolation and transcriptome sequencing analysis

Anthers provided the source material for total RNA extraction, performed with the RC401 assay kit following the manufacturer's guidelines from Vazyme Biotech Co., Ltd. The quality of the extracted RNA was confirmed utilizing an Agilent 2100 Bioanalyzer. Construction of transcriptome libraries adhered to the Illumina RNA-seq protocol (Illumina, Inc., San Diego, CA, catalog no. RS-100-0801). Sequencing of RNA and DNA libraries was conducted using 150 bp paired-end Illumina technology, with inserts of 350 bp for the libraries. For RNA-seq, the fastp software (version 0.12.2) was employed with default parameters to filter and remove raw reads containing poly-N sequences, adapters, or exhibiting low quality [54]. High-quality filtered reads were mapped to the Zhongmu No.1 genome using Hisat2 (version 2.1.0) with the parameters (-dta, -x), followed by genetic annotation in GTF format [55]. Genes were considered expressed if they had a minimum FPKM value of greater than 1 in any sample. Differentiation of expressed genes was determined using a P-value of 0.05 and DEseq2 for differential gene analysis. FPKM values underwent a Log(FPKM + 1) transformation for comparative analysis, and heatmap generation for gene expression was completed using TBtools software [56]. Functional annotation was achieved through GSEA analysis with the GO database [57].

qRT-PCR analysis

At four developmental stages of anther growth, namely S8, S9, S10, and S11, total RNA extraction was carried out. Primers designed with Primer 5 were employed for qRT-PCR expression analysis of selected genes (Table S1), with the list provided in Supplementary Table 5. The expression of the alfalfa ACTIN gene 2 (18 S rRNA) served as an internal control. qRT-PCR assays were conducted using SYBR Premix Ex Taq II (TaKaRa), repeated a minimum of three times, and data analysis was performed using the $-2^{-\Delta\Delta Ct}$ method [58].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06542-6.

Supplementary Material 1

Supplementary Material 2

Author contributions
Conceptualization, J.Z. and Z.Y.; methodology, Z.Y.; software, Y.F and P.X.; validation, Z.Y., S.W., C.Z and Q.Y.; formal analysis, Z.Y.; investigation, Z.Y.; resources, Z.Y.; data curation, Z.Y.; writing—original draft preparation, Z.Y.; writing—review and editing, F.W.; visualization, J.Z.; supervision, J.Z.; project administration, J.Z.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding
This research was funded by the the National Key R&D Program of China, grant number 2022YFD1301002, Program of Chines Academy of Sciences, grant number XDA26030103, the Gansu Provincial Science and Technology Projects, grant number 23ZDKA013 and 24JRRA479, the Drought Meteorology Research Fund, grant number IAM201703.

Data availability
Transcriptome sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under accession number PRJNA1117211.

Declarations

Ethics approval and consent to participate
Alfalfa is not endangered or a protected species in China, and it cultivated at Lanzhou University and planted in a light incubator. The seeds are collected by Professor Jiyu Zhang in Lanzhou University. All the study procedures were carried out in accordance with relevant guidelines.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 January 2025 / Accepted: 11 April 2025
Published online: 29 April 2025

References

1. Li X, Hou Y, Li M, Zhang F, Yi F, Kang J, Yang Q, Long R. Overexpression of an ABA-inducible homeodomain-leucine zipper I gene MsHB7 confers salt stress sensitivity to alfalfa. IND CROP PRO.; 2022. p. 177.
2. Chang Z, Chen Z, Wang N, Xie G, Lu J, Yan W, Zhou J, Tang X, Deng XW. Construction of a male sterility system for hybrid rice breeding and seed production using a nuclear male sterility gene. Proc Natl Acad Sci. 2016;113(49):14145–50.
3. Ye Q, Meng X, Chen H, Wu J, Zheng L, Shen C, Guo D, Zhao Y, Liu J, Xue Q, et al. Construction of genic male sterility system by CRISPR/Cas9 editing from model legume to alfalfa. Plant Biotechnol J. 2022;20(4):613–5.
4. Zhou H, He M, Li J, Chen L, Huang Z, Zheng S, Zhu L, Ni E, Jiang D, Zhao B, et al. Development of commercial Thermo-sensitive genic male sterile rice accelerates hybrid rice breeding using the CRISPR/Cas9-mediated TMS5 editing system. Sci Rep. 2016;6:37395.
5. Chen X, Yang S, Zhang Y, Zhu X, Yang X, Zhang C, Li H, Feng X. Generation of male-sterile soybean lines with the CRISPR/Cas9 system. CROP J. 2021;9(6):1270–7.
6. Li J, Han G, Sun C, Sui N. Research advances of MYB transcription factors in plant stress resistance and breeding. PLANT SIGNAL BEHAV. 2019; 14(8).
7. Li J, Xu S, Mei Y, Gu Y, Sun M, Zhang W, Wang J. Genomic-wide identification and expression analysis of R2R3-MYB transcription factors related to flavonol biosynthesis in Morinda officinalis. BMC PLANT BIOL. 2023;23(1):381.
8. Ma R, Liu B, Geng X, Ding X, Yan N, Sun X, Wang W, Sun X, Zheng C. Biological function and stress response mechanism of MYB transcription factor family genes. J Plant Growth Regul. 2022;42(1):83–95.
9. Du H, Yang S-S, Liang Z, Feng B-R, Liu L, Huang Y-B, Tang Y-X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. BMC PLANT BIOL. 2012;12(1):106.
10. Tolosa LN, Zhang Z. The role of major transcription factors in solanaceous food crops under different stress conditions: current and future perspectives. PLANTS-BASEL. 2020;9(1):56.
11. Baldoni E, Genga A, Cominelli E. Plant MYB transcription factors: their role in drought response mechanisms. Int J Mol Sci. 2015;16(7):15811–51.
12. Lou Y, Zhou HS, Han Y, Zeng QY, Zhu J, Yang ZN. Positive regulation of AMS by TDF1 and the formation of a TDF1-AMS complex are required for anther development in Arabidopsis thaliana. New Phytol. 2018;217(1):378–91.
13. Millar AA, Gubler F, The Arabidopsis GAMYB-L, Genes. MYB33 and MYB65, are MicroRNA-Regulated genes that redundantly facilitate anther development. Plant Cell. 2005;17(3):705–21.
14. Zhu J, Lou Y, Xu X, Yang ZN. A genetic pathway for tapetum development and function in Arabidopsis. J Integr Plant Biol. 2011;53(11):892–900.
15. Ghelli R, Brunetti P, Napoli N, De Paolis A, Cecchetti V, Tsuge T, Serino G, Matsui M, Mele G, Rinaldi G, et al. A newly identified Flower-Specific splice variant of AUXIN RESPONSE FACTOR8 regulates stamen elongation and Endothecium lignification in Arabidopsis. Plant Cell. 2018;30(3):620–37.
16. Li Y, Jiang J, Du M-L, Li L, Wang X-L, Li X-BA. Cotton gene encoding MYB-Like transcription factor is specifically expressed in pollen and is involved in regulation of late anther/pollen development. PLANT CELL PHYSIOL. 2013;54(6):893–906.
17. Zhou Q, Jia C, Ma W, Cui Y, Jin X, Luo D, Min X, Liu Z. MYB transcription factors in alfalfa (Medicago sativa): genome-wide identification and expression analysis under abiotic stresses. PeerJ. 2019;7:e7714.
18. Li J, Guo S, Min Htwe Y, Sun X, Zhou L, Wang F, Zeng C, Chen S, Iqbal A, Yang Y. Genome-wide identification, classification and expression analysis of MYB gene family in coconut (Cocos nucifera L). Front Plant Sci. 2023;14:1263595.
19. Schmitz RJ, Grotewold E, Stam M. Cis-regulatory sequences in plants: their importance, discovery, and future challenges. Plant Cell. 2022;34(2):718–41.
20. Chen H, Zwaenepoel A, Van de Peer Y. Wgd v2: a suite of tools to uncover and date ancient polyploidy and whole-genome duplication. Bioinformatics 2024; 40(5).
21. Del Amparo R, Branco C, Arenas J, Vicens A, Arenas M. Analysis of selection in protein-coding sequences accounting for common biases. BRIEF BIOINFORM. 2021;22(5):bbaa431.
22. Song X, Yang Q, Liu Y, Li J, Chang X, Xian L, Zhang J. Genome-wide identification of Pistacia R2R3-MYB gene family and function characterization of PcMYB113 during autumn leaf coloration in Pistacia chinensis. INT J BIOL MACROMOL. 2021;192:16–27.
23. Zhang T, Cui Z, Li Y, Kang Y, Song X, Wang J, Zhou Y. Genome-Wide identification and expression analysis of MYB transcription factor superfamily in Dendrobium catenatum. Front Genet. 2021;12:714696.
24. Li Y, Lin-Wang K, Liu Z, Allan AC, Qin S, Zhang J, Liu Y. Genome-wide analysis and expression profiles of the StR2R3-MYB transcription factor superfamily in potato (Solanum tuberosum L). Int J Biol Macromol. 2020;148:817–32.
25. Liu J, Wang J, Wang M, Zhao J, Zheng Y, Zhang T, Xue L, Lei J. Genome-Wide analysis of the R2R3-MYB gene family in Fragaria X ananassa and its function identification during anthocyanins biosynthesis in Pink-Flowered strawberry. Front Plant Sci. 2021;12:702160.
26. Zhou L, Yarra R. Genome-wide identification and expression analysis of bZIP transcription factors in oil palm (Elaeis guineensis Jacq.) under abiotic stress. Protoplasma. 2022;259(2):469–83.
27. Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. BMC Genomics. 2012;13(1):544.
28. Zhang J, Zhang Y, Feng C. Genome-Wide analysis of MYB genes in Primulina Eburnea (Hance) and identification of members in response to drought stress. Int J Mol Sci 25; 2024.
29. Feng C, Ding D, Feng C, Kang M. The identification of an R2R3-MYB transcription factor involved in regulating anthocyanin biosynthesis in Primulina swinglei flowers. Gene. 2020;752:144788.
30. Yang J, Song HD, Chen Y, Chen B, Kim M, Kim P, Kawabata S, Li Y, Wang Y. A single amino acid substitution in the R2R3 conserved domain of the BrPAP1a transcription factor impairs anthocyanin production in turnip (Brassica Rapa subsp. Rapa). PLANT PHYSIOL BIOCHEM. 2021;162:124–36.
31. Zhang J, Zhang Y, Feng C. Genome-Wide analysis of MYB genes in Primulina Eburnea (Hance) and identification of members in response to drought stress. Int J Mol Sci. 2023; 25(1).
32. Mayr C. What Are 3' UTRs Doing? CSH Perspect Biol. 2019; 11(10).

33. Sangar V, Blankenberg DJ, Altman N, Lesk AM. Quantitative sequence-function relationships in proteins based on gene ontology. BMC Bioinformatics. 2007;8(1):294.
34. Jia L, Mao Y, Ji Q, Dersh D, Yewdell JW, Qian S-B. Decoding mRNA translatability and stability from the 5′ UTR. NAT STRUCT MOL BIOL. 2020;27(9):814–21.
35. Swinburne IA, Silver PA. Intron delays and transcriptional timing during development. DEV CELL. 2008;14(3):324–30.
36. Cecchetti V, Altamura MM, Brunetti P, Petrocelli V, Falasca G, Ljung K, Costantino P, Cardarelli M. Auxin controls Arabidopsis anther dehiscence by regulating endothecium lignification and jasmonic acid biosynthesis. Plant J. 2013;74(3):411–22.
37. Hong W-J, Lee SK, Kim S-H, Kim Y-J, Moon S, Kim E-J, Silva J, Jung K-H. Comparative transcriptome analysis of pollen and anther wall reveals novel insights into the regulatory mechanisms underlying anther wall development and its dehiscence in rice. PLANT CELL REP. 2022;41(5):1229–42.
38. Xiao Y, Chen Y, Charnikhova T, Mulder PPJ, Heijmans J, Hoogenboom A, Agalou A, Michel C, Morel J-B, Dreni L, et al. OsJAR1 is required for JA-regulated floret opening and anther dehiscence in rice. PLANT MOL BIOL. 2014;86(1):19–33.
39. Wan X, Wu S, Li Z, An X, Tian Y. Lipid metabolism: critical roles in male fertility and other aspects of reproductive development in plants. MOL PLANT. 2020;13(7):955–83.
40. Xu Y, Iacuone S, Li SF, Parish RW. MYB80 homologues in arabidopsis, cotton and brassica: regulation and functional conservation in tapetal and pollen development. BMC PLANT BIOL. 2014;14(1):278.
41. Yang C, Song J, Ferguson AC, Klisch D, Simpson K, Mo R, Taylor B, Mitsuda N, Wilson ZA. Transcription factor MYB26 is key to Spatial specificity in anther secondary thickening formation. Plant Physiol. 2017;175(1):333–50.
42. Cheng X-Q, Zhang X-Y, Xue F, Zhu S-H, Li Y-J, Zhu Q-H, Liu F, Sun J. Characterization and transcriptome analysis of a dominant genic male sterile cotton mutant. BMC PLANT BIOL. 2020;20(1):312.
43. Shen C, Du H, Chen Z, Lu H, Zhu F, Chen H, Meng X, Liu Q, Liu P, Zheng L, et al. The Chromosome-Level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for alfalfa research. Mol Plant. 2020;13(9):1250–61.
44. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar Gustavo A, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ et al. Pfam: The protein families database in 2021. NUCLEIC ACIDS RES. 2020; 49(D1):D412-D419.
45. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. NUCLEIC ACIDS RES. 2018;46(W1):W200–4.
46. Duvaud S, Gabella C, Lisacek F, Stockinger H, Ioannidis V, Durinx C. Expasy, the Swiss bioinformatics resource portal, as designed by its users. NUCLEIC ACIDS RES. 2021;49(W1):W216–27.
47. Chou K-C, Shen H-B. Cell-PLoc: a package of web servers for predicting subcellular localization of proteins in various organisms. NAT PROTOC. 2008;3(2):153–62.
48. Silva MA, Nascimento Junior JCD, Thomaz DV, Maia RT, Costa Amador V, Tommaso G, Coelho GD. Comparative homology of pleurotus ostreatus laccase enzyme: Swiss model or modeller?? J Biomol Struct Dyn. 2023;41(18):8927–40.
49. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7.
50. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2014;31(8):1296–7.
51. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. 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.
52. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
53. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. NUCLEIC ACIDS RES. 2012;40(7):e49–49.
54. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90.
55. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Correction: corrigendum: differential gene and transcript expression analysis of RNA-seq experiments with tophat and cufflinks. NAT PROTOC. 2014;9(10):2513–2513.
56. Au - Liu S, Au - Wang Z, Au - Zhu R, Au - Wang F, Au - Cheng Y. Au - Liu Y. Three differential expression analysis methods for RNA sequencing: Limma, EdgeR, DESeq2. JoVE 2021; (175):e62528.
57. Wang B, van der Kloet F, Kes M, Luirink J, Hamoen LW. Improving gene set enrichment analysis (GSEA) by using regulation directionality. Microbiol Spectr. 2024;12(3):e0345623.
58. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods. 2001;25(4):402–8.

Publisher's note