Canadian Science Publishing

ARTICLE

Analysis of *MYB* genes in four plant species and the detection of genes associated with drought resistance

Yanli Zhou, Lin Lu, Ningyawen Liu, Hua Cao, Han Li, Daping Gui, Jihua Wang, and Chengjun Zhang

Abstract: MYB transcription factor, which contains a conserved DNA binding domain, has been found in almost all eukaryotes. *MYB* genes have variable functions in plants and are involved in many pathways. We systemically analyzed the *MYB* gene family in three Ericaceae species, *Rhododendron williamsianum* Rehder & E.H. Wilson, *Rhododendron delavayi* Franch., and *Vaccinium corymbosum* L., and one outgroup, *Actinidia chinensis* Planch., with 99, 156, 480, and 185 *MYB* genes found, respectively. The *MYB* genes were classified into five types based on the number of conserved MYB motifs, and the two repeat (2R) types were dominant in all four species. The percentage of 2R type *MYB* ranged from 48.5% to 87.9% depending on the species. We further classified the conserved MYB motifs into M1, M2, and M3 types based on motif definition. We found an abundance of 3xM2 type in the 3R group, but found a species-specific type preference for 1R and 2R genes. In searching for *Arabidopsis* drought-resistant genes, we detected 34 potential candidates in four species. The expression profile of *R. delavayi* showed 11 candidate drought-resistant *RdMYB* genes, which provide a potential molecular design target for breeders. Our results describe the *MYB* gene family in these four species and could play an important role in future analyses.

Key words: Ericaceae, *Rhododendron*, blueberry, kiwi fruit, conserved motif.

Résumé : Le facteur de transcription MYB, qui comporte un domaine de liaison à l'ADN conservé, a été trouvé chez presque tous les eucaryotes. Les gènes *MYB* exercent des fonctions variables chez les plantes et sont impliqués dans de nombreuses voies. Les auteurs ont analysé de manière systématique la famille des gènes *MYB* chez trois espèces d'Ericaceae, *Rhododendron williamsianum* Rehder & E.H. Wilson, *Rhododendron delavayi* Franch. et *Vaccinium corymbosum* L., et chez un hors-groupe, *Actinidia chinensis* Planch., avec 99, 156, 480 et 185 gènes *MYB* trouvés, respectivement. Les gènes *MYB* ont été classés en cinq types en fonction du nombre de motifs MYB conservés, et les types à deux répétitions (2R) étaient dominants chez les quatre espèces. Le pourcentage de *MYB* de type 2R variait de 48,5 % à 87,9 % chez les différentes espèces. Ils ont ensuite classé les motifs MYB conservés en types M1, M2 et M3 en fonction de la définition du motif. Ils ont constaté une abondance du type 3xM2 chez le groupe 3R, mais une préférence spécifique à l'espèce quant aux types pour les gènes 1R et 2R. En recherchant les gènes de résistance à la sécheresse chez *Arabidopsis*, ils ont détecté 34 candidats potentiels chez les quatre espèces. Le profil d'expression de *R. delavayi* a montré la présence de 11 gènes *RdMYB* candidats pour la résistance à la sécheresse, qui constituent une cible moléculaire potentielle pour les sélectionneurs. Ces résultats décrivent la famille de gènes *MYB* chez ces quatre espèces qui pourraient jouer un rôle important dans les analyses futures. [Traduit par la Rédaction]

Mots-clés : Ericaceae, *Rhododendron*, myrtille, kiwi, motif conservé.

Received 16 December 2020. Accepted 20 October 2021.

Y. Zhou,* N. Liu, and D. Gui. Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Science, Kunming 650201, China.

L. Lu,* H. Cao, H. Li, and J. Wang. The Flower Research Institute, Yunnan Academy of Agricultural Sciences, Kunming 650201, China.

C. Zhang. Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Science, Kunming 650201, China; Haiyan Engineering and Technology Center, Kunming Institute of Botany, Chinese Academy of Science, Jiaxing 314300, China.

Corresponding authors: Chengjun Zhang (email: zhangchengjun@mail.kib.ac.cn) and Jihua Wang (email: wjh0505@gmail.com).

*These authors contributed equally to this work.

Introduction

Transcription factors (TFs) play a central role in developmental and metabolic programs by regulating the transcription of downstream target genes (Cao et al. 2020). v-Myb avian myeloblastosis viral oncogene homolog (*MYB*) TF genes have been found in almost all eukaryotes, which contain a conserved DNA binding domain (González-Verdejo et al. 2007). The MYB domain is composed of 1–4 imperfect repeats, and the number of repeats in their sequences represent a classification criterion, i.e., 1R, R2R3(2R), 3R, 4R, and 5R (He et al. 2016). In each repeat, the MYB domain consists of approximately 50 amino acid residues and forms three α-helices.

Systematic research into the *MYB* gene family at the genome level can be classified into three stages. The first stage was in the late 1990s when *MYB* genes were discovered using 3′-RACE and RT-PCR technology, accompanied by cloning and sequencing. For example, Gómez-Gómez et al. (2012) reported that two-repeat (R2R3) MYB subfamily is the largest group characterized in plants. Kranz et al. (1998) identified more than 100 *R2R3-MYB* genes in *Arabidopsis thaliana* (L.) Heynh., and Rabinowicz et al. (1999) found that more than 80 *R2R3-MYB* genes were expressed in maize. In this first stage, it was known that *MYB* genes were more numerous in plant genomes compared to animal genomes. The second stage involved the use of probe sets or expressed sequence tags to detect *MYB* genes before the complete genome data were available. Using this approach, Wilkins et al. (2009) and Wang et al. (2015) found 192 and 40 *MYB* genes in *Populus* and *Rehmannia glutinosa* (Gaertn.) DC., respectively. With the advent of high-throughput sequencing technology and improved genome assembly and annotation capabilities (Zhang et al. 2021), the third stage used whole genome data to uncover *MYB* genes. Using this approach, Stracke et al. (2014) employed the consensus *R2R3-MYB* DNA binding domain sequence as a query and found 70 *R2R3-MYB* genes in *Beta vulgaris* L. Du et al. (2012*b*) detected a total of 244 *R2R3-MYB* genes in soybean and further classified them into 48 subfamilies based on a phylogenetic comparative analysis with their putative orthologs, and there are more than 109 R2R3-type *MYB* genes in *Hypericum perforatum* L. (Zhou et al. 2020). Gonzalez et al. (2016) detected 186, 98, and 86 *R2R3-MYB* genes in apple, peach, and strawberry, respectively, and genes with incomplete R2 or R3 domains were removed. Du et al. (2012*a*) applied the preliminary BLASTP method to search MYB motif sequences against known maize genome. After a serial in silico calculation, they identified 157 typical non-redundant *R2R3-MYB* genes in maize; this number is much larger than in previous studies (Rabinowicz et al. 1999). Overall, using whole genome data to reveal specific species *MYB* genes is easier and more efficient.

Compared to the conserved N-terminal of MYB proteins, the highly divergent C-terminal contains a transcription activation and (or) repression domain (Katiyar et al. 2012; Mitra et al. 2021; Yang et al. 2021), which confers a regulatory role for each family member (Chen et al. 2021). It diversifies functions of these *MYB* genes in plants including primary metabolism, secondary metabolism, cell fate and identity, environmental stresses, and organ development (Wei et al. 2020; Zhu et al. 2020; Rodrigues et al. 2021; Xiao et al. 2021). Among these functions, *MYB* genes play a very important role in drought resistance. Gao et al. (2014) reported that *AtMYB20* was induced by high levels of NaCl and was suppressed by desiccation, cold stress, and abscisic acid (ABA) treatment. Su et al. (2013) found a severe filament elongation defect under drought stress in *AtMYB21* knockout plants, suggesting that appropriate stamen development under drought condition requires *AtMYB21*. Cheng et al. (2009) showed that *AtMYB21*, *AtMYB24*, and *AtMYB57* are GA-dependent, stamen-enriched genes. Genetic analysis indicated that the *myb21-t1 myb24-t1 myb57-t1* triple mutant conferred a short stamen phenotype, leading to male sterility. Palatnik et al. (2003) reported that the overexpression of a miRNA-resistant form of *MYB33* had major phenotypic consequences, which provided strong evidence for miRNA regulation of *AtMYB33*. Reyes and Chua (2007) suggested that the ABA-induced accumulation of miR159 was a homeostatic mechanism to direct *AtMYB33* and *AtMYB101* transcript degradation to desensitize hormone signaling during seedling stress responses. Park et al. (2011) found that *AtMYB52* overexpression lines were drought-tolerant, and their data suggested that *AtMYB52* is involved in the ABA response. Oh et al. (2011) and Cominelli et al. (2005) suggested that *AtMYB60* plays a crucial role in stomatal movement and enhances sensitivity to water deficit stress. Romano et al. (2012) found that *AtMYB61* was expressed in several tissues, notably the xylem, roots, and developing seeds. The loss of *AtMYB61* function decreased xylem formation, induced qualitative changes in xylem cell structure, and decreased lateral root formation. *AtMYB73* interacted with *PYL8* and promoted the growth of lateral roots, thus increasing drought tolerance (Zhao et al. 2014). Lee and Suh (2015) reported that the *AtMYB94* gene was expressed in abundance in aerial organs, and it showed a higher expression in the stem epidermis than in the stem. When seedlings were subjected to various treatments, the expression of *AtMYB94* was observed to increase approximately nine-fold under drought stress, and experiments showed that the *AtMYB94* gene activates *Arabidopsis* cuticular wax biosynthesis and might be important in the plant's response to drought stress (Lee and Suh 2015). Mengiste et al. (2003) used T-DNA insertion lines and found that *AtMYB108* (*BOS1*) interacts with the jasmonate signaling pathway, which is possibly mediated by reactive oxygen intermediates from both biotic and abiotic stress agents. Using *flp-1* and *AtMYB88* double mutants, Xie et al. (2010) showed that *AtMYB124* (*FLP*) controls water loss by regulating the stomata.

Ericaceae, or the heath family, is the largest family in Ericales, containing about 4000 species. Members of this

Table 1. Genome data and MYB gene numbers of each species.

Species name	Abbreviation	Proteins*	AtMYBs step	PF00249 step	Percentage	Reference
Rhododendron williamsianum	Rhow	23548	265	99	0.42	Soza et al. 2019
Rhododendron delavayi	Rhod	32938	473	156	0.47	Zhang et al. 2017
Vaccinium corymbosum	Vacc	128559	989	480	0.37	Colle et al. 2019
Actinidia chinensis	Kiwi	39040	419	185	0.47	Huang et al. 2013

*Links for Protein sequences: Rhow (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/009/746/105/GCA_009746105.1_ASM974610v1/GCA_009746105.1_ASM974610v1_protein.faa.gz); Rhod (ftp://parrot.genomics.cn/gigadb/pub/10.5524/100001_101000/100331/Gene/Rhododendron_delavayi.protein.fa); Vacc (https://www.vaccinium.org/vaccinium_downloads/Vaccinium_corymbosum/Vcorymbosum_Draper_genome_v1/annotation/V_corymbosum_Draper_v1.0-proteins-nameTruncated.fasta); Kiwi (http://bioinfo.bti.cornell.edu/ftp/kiwifruit/Kiwifruit_pep.fa.gz).

family occur mainly in tropical montane, Mediterranean, and temperate to arctic climates (www.britannica.com). Many important flowers and fruits belong to this family such as *Rhododendron* (azalea) and blueberry. To improve our understanding of the gene family of *R2R3-MYB*, TFs were studied in three important species of the Ericaceae family: *Rhododendron delavayi* Franch., *Rhododendron williamsianum* Rehder & E.H. Wilson, and *Vaccinium corymbosum* L. 'Draper', as well as the outgroup species *Actinidia chinensis* Planch. We performed a genome-wide analysis of *MYB* genes for these four species (including all *MYB* genes), we classified genes, we described the character of the conserved motif, and we identified potential drought-resistant genes. To the best of our knowledge, our results provide the first comprehensive analysis of the *MYB* gene family in the genome of these four species.

Materials and methods

Data resources

We obtained 125 *Arabidopsis R2R3-MYB*, two unusual *MYB*-type genes, and five 3R-MYB-type gene IDs from a previous study (Stracke et al. 2001), and then the coding domain sequences (CDS) were downloaded from TAIR (http://www.Arabidopsis.org/) (see cjb 2020 0227supplc in Supplementary Material[1]). We obtained 150 seed sequences from the PFAM database (http://pfam.xfam.org/) with the ID "PF00249" (see cjb 2020 0227suppld in Supplementary Material[1]), which contained the DNA-binding domains from MYB proteins as well as the SANT domain family. We downloaded the protein sequences of the four species from the following genome databases (Table 1): *V. corymbosum* (Colle et al. 2019) from GDB (https://www.vaccinium.org), *Actinidia chinensis* (Huang et al. 2013) from IKGC (http://bioinfo.bti.cornell.edu/ftp/kiwifruit/), *R. delavayi* (Zhang et al. 2017) from Gigadb (http://gigadb.org/), and *R. williamsianum* (Soza et al. 2019) from the NCBI assembly (https://www.ncbi.nlm.nih.gov/assembly/GCA_009746105.1/). The two *Rhododendron* datasets are also available on RPGD (http://bioinfor.kib.ac.cn/RPGD/download_genome.html).

Identification of *MYB* genes in four species

First, using the 132 *AtMYB* gene CDS sequences, we performed a BLASTX query against the local whole-genome protein database in our four species with the default parameters, set the *e* value to 0.01, and all the candidate protein sequences were collected by shell scripts. Second, using the 150 seed sequences from PFAM, we performed a BLASTP query against the local candidate protein sequences database to identify the DNA-binding domains. All candidate protein sequences with PFAM domain signals were recorded as *MYB* genes.

Detection of maximum conserved motifs from species

We employed a homemade pipeline 1 (see cjb 2020 0227supplj in Supplementary Material[1]) to obtain the maximum conserved MYB (DNA-binding) motifs from each *MYB* gene in five species (*V. corymbosum*, *Actinidia chinensis*, *R. delavayi*, *R. williamsianum*, and *Arabidopsis thaliana*). The pipeline combines several regions into the largest intervals if the 's.start to s.end' region of the different hits of one specific *AtMYB* gene overlapped. For example (see cjb 2020 0227supple in Supplementary Material[1]), the gene *AtMYB3R4* (*AT5G11510.1*) had six BLASTP hits (for simplicity, we only show six hits), the 's.start to s.end' hit regions were 32–57, 32–73, 31–75, 136–178, 84–124, and 81–123, respectively. All these hits were combined into three maximum motifs, ranging from 31–75, 81–124, and 136–178. Then, all the sequences of the maximum conserved MYB motifs were extracted according to the region information.

Classification of super-groups and the characteristics of conserved motifs

To profile the characteristics of the MYB motifs in our four species, we employed ClustalX2 (Stracke et al. 2001) using the collected maximum conserved MYB motif sequences to build a guide tree for each species together with *Arabidopsis thaliana*. Subsequently, all four guide trees were separated into three super-groups manually, based on the balance of sequence numbers and tree topology. Because of too many motif sequences in *V. corymbosum*, dozens of subgroups were presented in the tree file, and detailed information on subgroups was included in the zip file (see cjb 2020 0227supplf in Supplementary Material[1]). The maximum conserved motif sequences in each super-group for each species were aligned using MEGA7 (Kumar et al. 2016), and the major gaps were manually removed from the alignment. Finally, the multiple alignment motif

[1]Supplementary data are available with the article at https://doi.org/10.1139/cjb-2020-0227.

sequences were submitted to the WEBLOGO website (Crooks et al. 2004) (http://weblogo.berkeley.edu/logo.cgi).

Detection of drought-resistant *MYB* genes

To predict the drought-resistant *MYB* genes in our four species, we employed the reciprocal best hit method to identify orthologous genes of *MYB* genes known to possess this function in *Arabidopsis*. The *MYB* genes in *Arabidopsis* associated with drought resistance were sourced from previous studies and are listed in Table 2. To search for orthologous genes, we employed BLASTP with default parameters, except that the *e* value was set to $1e^{-5}$. In each round, the specific *AtMYB* gene sequences were used as queries and the genome protein database was used as the subject data; only the best hits were retained in subsequent analyses.

Expression analysis of drought-resistant *RdMYB* gene using transcriptome data

To investigate the changes in expression under different growing conditions, we obtained the expression profile for each potential drought-resistant *RdMYB* gene from the RPGD database (http://bioinfor.kib.ac.cn/RPGD/searchgene_expression.html) (Liu et al. 2021). The original data and expression analysis were presented in a previous study (Cai et al. 2019). The growing conditions included normal irrigation (CK), stopping irrigation for five days (D5), stopping irrigation for nine days (D9), and re-watering for six hours (REC) after 10 days of drought.

Results

Genome-wide identification of the *MYB* genes in four species

Using the two data sets from *Arabidopsis thaliana* and PFAM, we detected hundreds of candidate *MYB* genes and the maximum conserved motif features in our four species. First, using the BLASTX program with the 132 *AtMYB* CDS sequences (see cjb 2020 0227supplc in Supplementary Material[1]), we detected 265, 473, 989, and 419 *MYB* candidate genes in *R. williamsianum*, *R. delavayi*, *V. corymbosum*, and *Actinidia chinensis*, respectively (Table 1). Second, to further confirm that the MYB domains were present in the candidate genes, we used the 150 PF00249 seed sequences (see cjb 2020 0227suppld in Supplementary Material[1]) as queries, employed BLASTP with default parameters and set the *p* value to 0.01. The results indicated that there were 99, 156, 480, and 185 *MYB* genes in *R. williamsianum*, *R. delavayi*, *V. corymbosum*, and *Actinidia chinensis*, respectively (Table 1). These numbers were much smaller than those obtained from the first step. The *MYB* gene IDs (the IDs from original data as well as renamed IDs) of *R. williamsianum* are presented in Table 3, whereas the IDs of the other species are presented in the supplementary file (see cjb 2020 0227supplg in Supplementary Material[1]). The *MYB* gene sequences are presented in cjb 2020 0227supplh in Supplementary Material[1].

Five types of *MYB* genes based on the number of conserved maximum motif regions

We used the homemade pipeline 2 (see cjb 2020 0227supplj in Supplementary Material[1]) to detect 155, 275, 974, and 301 conserved maximum motif regions in *R. williamsianum*, *R. delavayi*, *V. corymbosum*, and *Actinidia chinensis*, respectively, and there were 268 conserved regions in *Arabidopsis thaliana* (Table 4). The number of motifs was almost twice the number of *MYB* genes in each species, which means that the majority of *MYB* genes had two MYB domains. Based on the number of MYB motifs inside the gene, we classified all MYB genes into 1R, 2R, 3R, 4R, and 5R types. According to the data presented in Table 3 and cjb 2020 0227supplg in the Supplementary Material[1], the 2R type *MYB* genes were dominant over all other types in our four species, with the highest percentage of 2R type being 87.9% in *V. corymbosum* (Table 4). The 1R type *MYB* genes occupied the second place, ranging from 5.6% to 47.5%. The proportion of 3R type *MYB* genes in the four species was quite stable at approximately 4%. The 4R and 5R types were rare, which is consistent with the results of previous studies. Based on the location and length information of conserved maximum motif regions in the *MYB* genes in each species (Table 3 and cjb 2020 0227supplg in Supplementary Material[1]), a motif location diagram was drawn using homemade scripts 3 (see cjb 2020 0227supplj in Supplementary Material[1]) for all *MYB* genes of *R. williamsianum*, as presented in Fig. 1, whereas those of other species are presented in Supplementary Fig. S1[1].

Characteristics of conserved maximum motifs

Based on the guide tree, all sequences of conserved maximum motifs in each species were manually separated into three super-groups (see cjb 2020 0227supplf in Supplementary Material[1]). The conserved motif sequences in each super-group for each species were aligned using the Muscle program (Edgar 2004) built in MEGA7 (Kumar et al. 2016), and then submitted to WEBLOGO (Crooks et al. 2004). The conserved motifs of *R. williamsianum* are presented in Fig. 2, whereas the others are presented in Supplementary Fig. S2[1]. We defined the conserved motifs in each species as M1, M2, and M3. Based on the information obtained from WEBLOGO, we found that (*i*) all of the M1 and M2 conserved motifs in each species contained two tryptophans (W), whereas M3 contained only one; (*ii*) the sizes of conserved motifs were stable, ranging from 45 to 51 amino acids; and (*iii*) the position of tryptophan was stable and the sequence length between two tryptophans was around 18–19 amino acids.

The profile of conserved motifs in MYB genes in each species

Based on the definitions of M1, M2, and M3, the motif location diagram (Fig. 1 and Supplementary Fig. S1[1]) clearly illustrated the profile of conserved motifs in MYB genes in each species (Table 5). The 3R genes had M1M2M3 and 3xM2 types, but most were 3xM2 type. For 1R genes, it was interesting that the four species had completely different patterns. *Rhododendron williamsianum* had only M2 and M3

Table 2. The orthologous genes of AtMYBs that had the potential drought resistance function in four species.

			Orthologous gene Name in four Ericales species				
Gene Name	Locus	References	Rhod(11)	Rhow(6)	Kiwi(7)	VccD(10)	Function
AtMYB2	*AT2G47190*	Urao et al. 1993; Abe et al. 2003					Have higher sensitivity to ABA
AtMYB13	*AT1G06180*	Kirik et al. 1998					Induced by drought and ABA treatment
AtMYB15∼ATY19	*AT3G23250*	Ding et al. 2009				*VcMYB79*	Hypersensitive to ABA, overexpression improves drought and salt tolerance
AtMYB20	*AT1G66230*	Gao et al. 2014	*RdMYB94*				Response to desiccation stress in an ABA-dependent manner
AtMYB21∼MYB3	*AT3G27810*	Su et al. 2013	*RdMYB99*		*AcMYB184*	*VcMYB303*	Required by stamen development under drought conditions
AtMYB37∼RAX1	*AT5G23000*	Yu et al. 2016					Response to ABA, drought stress
AtMYB41	*AT4G28110*	Cominelli et al. 2008					Response to drought, ABA; cell-wall and cell expansion
AtMYB52∼BW52	*AT1G17950*	Park et al. 2011	*RdMYB86*	*RwMYB94*	*AcMYB19*	*VcMYB399*	Mutant is ABA-hypersensitive, overexpression lines were drought tolerant
AtMYB60	*AT1G08810*	Oh et al. 2011	*RdMYB38*			*VcMYB293*	Involvement in stomatal regulation and root growth
AtMYB61	*AT1G09540*	Penfield et al. 2001; Romano et al. 2012	*RdMYB35*		*AcMYB39*	*VcMYB13*	Loss of *AtMYB61* function decreases xylem formation; decreases lateral root formation
AtMYB94∼MYBCP70	*AT3G47600*	Lee and Suh 2015	*RdMYB133*				Activates *WSD1* et al. genes; cuticular wax biosynthesis
AtMYB96∼MYBCOV1	*AT5G62470*	Seo et al. 2009				*VcMYB85*	Mediates ABA-auxin crosstalks in drought stress response and lateral root growth
AtMYB108∼BOS1	*AT3G06490*	Mengiste et al. 2003	*RdMYB32*	*RwMYB32*			Response from both biotic and abiotic stress
AtMYB12∼PFG1	*AT2G47460*	Nakabayashi et al. 2014		*RwMYB79*	*AcMYB52*		Overexpress help flavonoid overaccumulation, is the key to enhanced drought tolerance
AtMYB75∼PAP1, SIAA1	*AT1G56650*						
AtMYB33	*AT5G06100*	Reyes and Chua 2007	*RdMYB84*	*RwMYB65*	*AcMYB170*	*VcMYB138*	miR159 over-expression suppresses *MYB33* and *MYB101*
AtMYB101∼ATM1	*AT2G32460*		*RdMYB4*		*AcMYB182*	*VcMYB91*	transcript levels and renders plants hyposensitive to ABA
AtMYB44∼MYBR1	*AT5G67300*	Jung et al. 2008;					Sensitive to ABA, ABA-induced stomatal closure; PYL8
AtMYB73	*AT4G37260*	Zhao et al. 2014	*RdMYB154*	~~*RwMYB18*~~ *RwMYB43*	*AcMYB124*	*VcMYB383*	promotes lateral root growth by enhancing the activities of MYB77, MYB44 and MYB73
AtMYB77	*AT3G50060*						
AtMYB88	*AT2G02820*	Xie et al. 2010;		*RwMYB36*	*Achn183081 (Null)*		*flp-1 myb88* plants are significantly more susceptible to
AtMYB124∼FLP	*AT1G14350*	Lai et al. 2005	*RdMYB98*			*VcMYB358*	drought and high salt; jointly restrict divisions late in the *Arabidopsis thaliana* stomatal cell lineage

Note: The alignment data for these 15 orthologous MYB gene groups indicated that all the MYB genes had a high similarity region in the MYB domain in the N-terminal, and the similarity in the C-terminal is quite low (see cjb 2020 0227suppli in Supplementary Material[1]). This result was unexpected and implied that the function of the orthologous genes in the four species might vary.

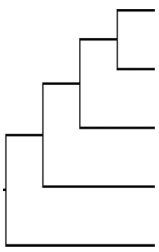
Table 3. The protein region information of maximum conserved motif.

	MYB gene IDs in RhoW	IDs in NCBI	Motif1 start site	Motif1 stop site	Motif1 size	Motif2 start site	Motif2 stop site	Motif2 size	Motif3 start site	Motif3 stop site	Motif3 size
1	*RwMYB29*	*KAE9452294.1*	31	78	47	83	128	45	6	25	19
2	*RwMYB70*	*KAE9462538.1*	136	183	47	188	233	45	62	130	68
3	*RwMYB22*	*KAE9450902.1*	57	103	46	109	155	46	161	206	45
4	*RwMYB62*	*KAE9461286.1*	341	387	46	395	440	45	447	495	48
5	*RwMYB23*	*KAE9450993.1*	22	67	45	1	16	15			
6	*RwMYB52*	*KAE9459385.1*	23	63	40	2	19	17			
7	*RwMYB68*	*KAE9462464.1*	27	72	45	1	21	20			
8	*RwMYB84*	*KAE9464877.1*	30	75	45	2	24	22			
9	*RwMYB90*	*KAE9465477.1*	32	75	43	4	26	22			
10	*RwMYB44*	*KAE9457292.1*	46	91	45	23	40	17			
11	*RwMYB97*	*KAE9467402.1*	55	99	44	5	49	44			
12	*RwMYB21*	*KAE9450209.1*	88	138	50	2	49	47			
13	*RwMYB94*	*KAE9467029.1*	56	101	45	4	50	46			
14	*RwMYB9*	*KAE9447984.1*	59	103	44	6	53	47			
15	*RwMYB18*	*KAE9449693.1*	64	109	45	10	58	48			
16	*RwMYB57*	*KAE9460123.1*	64	118	54	15	58	43			
17	*RwMYB87*	*KAE9465236.1*	122	172	50	23	70	47			
18	*RwMYB56*	*KAE9460081.1*	79	133	54	35	73	38			
19	*RwMYB76*	*KAE9463592.1*	135	186	51	26	74	48			
20	*RwMYB27*	*KAE9451475.1*	132	181	49	34	80	46			
21	*RwMYB1*	*KAE9444863.1*	14	84	70	45	84	39			
22	*RwMYB71*	*KAE9462861.1*	105	145	40	69	86	17			
23	*RwMYB26*	*KAE9451469.1*	97	142	45	47	91	44			
24	*RwMYB14*	*KAE9449120.1*	102	146	44	50	96	46			
25	*RwMYB99*	*KAE9468028.1*	24	97	73	61	97	36			
26	*RwMYB78*	*KAE9463847.1*	14	98	84	71	98	27			
27	*RwMYB59*	*KAE9460663.1*	14	53	39	56	100	44			
28	*RwMYB8*	*KAE9447794.1*	4	53	49	59	103	44			
29	*RwMYB86*	*KAE9465173.1*	8	54	46	60	105	45			
30	*RwMYB75*	*KAE9463514.1*	8	55	47	61	105	44			
31	*RwMYB91*	*KAE9465489.1*	8	55	47	61	106	45			
32	*RwMYB36*	*KAE9453978.1*	10	56	46	62	106	44			
33	*RwMYB95*	*KAE9467092.1*	22	106	84	68	106	38			
34	*RwMYB43*	*KAE9457199.1*	9	57	48	63	108	45			
35	*RwMYB39*	*KAE9455598.1*	116	157	41	53	110	57			
36	*RwMYB79*	*KAE9464060.1*	14	61	47	67	112	45			
37	*RwMYB10*	*KAE9448476.1*	16	62	46	68	113	45			
38	*RwMYB72*	*KAE9463058.1*	20	67	47	74	118	44			
39	*RwMYB80*	*KAE9464208.1*	135	180	45	83	129	46			
40	*RwMYB20*	*KAE9450176.1*	6	54	48	112	133	21			
41	*RwMYB17*	*KAE9449290.1*	139	184	45	87	133	46			
42	*RwMYB32*	*KAE9453288.1*	38	85	47	91	136	45			
43	*RwMYB42*	*KAE9457118.1*	40	86	46	92	137	45			
44	*RwMYB89*	*KAE9465425.1*	24	73	49	109	146	37			
45	*RwMYB7*	*KAE9447730.1*	60	107	47	111	155	44			
46	*RwMYB65*	*KAE9462071.1*	37	84	47	90	159	69			
47	*RwMYB25*	*KAE9451250.1*	31	79	48	134	181	47			
48	*RwMYB77*	*KAE9463599.1*	120	166	46	176	221	45			
49	*RwMYB88*	*KAE9465248.1*	47	93	46	202	247	45			
50	*RwMYB60*	*KAE9461019.1*	153	199	46	205	250	45			
51	*RwMYB41*	*KAE9456295.1*	565	613	48	461	500	39			
52	*RwMYB24*	*KAE9451173.1*	492	524	32	540	603	63			
53	*RwMYB37*	*KAE9454491.1*	1	21	20						
54	*RwMYB50*	*KAE9458458.1*	1	24	23						
55	*RwMYB49*	*KAE9458419.1*	4	29	25						
56	*RwMYB2*	*KAE9445881.1*	1	29	28						
57	*RwMYB54*	*KAE9459430.1*	2	30	28						

Table 3 (*concluded*).

	MYB gene IDs in RhoW	IDs in NCBI	Motif1 start site	Motif1 stop site	Motif1 size	Motif2 start site	Motif2 stop site	Motif2 size	Motif3 start site	Motif3 stop site	Motif3 size
58	*RwMYB96*	*KAE9467251.1*	8	30	22						
59	*RwMYB11*	*KAE9448737.1*	1	30	29						
60	*RwMYB58*	*KAE9460131.1*	1	31	30						
61	*RwMYB61*	*KAE9461151.1*	2	31	29						
62	*RwMYB81*	*KAE9464264.1*	1	31	30						
63	*RwMYB46*	*KAE9457769.1*	1	32	31						
64	*RwMYB45*	*KAE9457757.1*	13	36	23						
65	*RwMYB35*	*KAE9453871.1*	2	43	41						
66	*RwMYB16*	*KAE9449280.1*	10	45	35						
67	*RwMYB3*	*KAE9446530.1*	2	45	43						
68	*RwMYB33*	*KAE9453360.1*	13	51	38						
69	*RwMYB47*	*KAE9457880.1*	22	54	32						
70	*RwMYB4*	*KAE9446734.1*	14	55	41						
71	*RwMYB6*	*KAE9447092.1*	5	56	51						
72	*RwMYB66*	*KAE9462230.1*	30	58	28						
73	*RwMYB13*	*KAE9448847.1*	13	59	46						
74	*RwMYB83*	*KAE9464564.1*	14	59	45						
75	*RwMYB64*	*KAE9461750.1*	17	59	42						
76	*RwMYB55*	*KAE9459592.1*	9	62	53						
77	*RwMYB40*	*KAE9456209.1*	23	64	41						
78	*RwMYB82*	*KAE9464313.1*	11	65	54						
79	*RwMYB53*	*KAE9459393.1*	16	65	49						
80	*RwMYB93*	*KAE9466726.1*	15	66	51						
81	*RwMYB48*	*KAE9458192.1*	23	67	44						
82	*RwMYB69*	*KAE9462534.1*	11	67	56						
83	*RwMYB30*	*KAE9452688.1*	25	70	45						
84	*RwMYB34*	*KAE9453745.1*	15	71	56						
85	*RwMYB63*	*KAE9461695.1*	31	75	44						
86	*RwMYB28*	*KAE9452026.1*	14	78	64						
87	*RwMYB85*	*KAE9464878.1*	13	80	67						
88	*RwMYB92*	*KAE9465653.1*	40	83	43						
89	*RwMYB19*	*KAE9450139.1*	43	88	45						
90	*RwMYB73*	*KAE9463173.1*	42	98	56						
91	*RwMYB5*	*KAE9447041.1*	70	102	32						
92	*RwMYB12*	*KAE9448775.1*	104	151	47						
93	*RwMYB31*	*KAE9452772.1*	108	154	46						
94	*RwMYB51*	*KAE9458851.1*	200	248	48						
95	*RwMYB67*	*KAE9462340.1*	255	300	45						
96	*RwMYB15*	*KAE9449212.1*	316	348	32						
97	*RwMYB74*	*KAE9463444.1*	328	377	49						
98	*RwMYB38*	*KAE9454901.1*	868	908	40						
99	*RwMYB98*	*KAE9467983.1*	1950	1981	31						

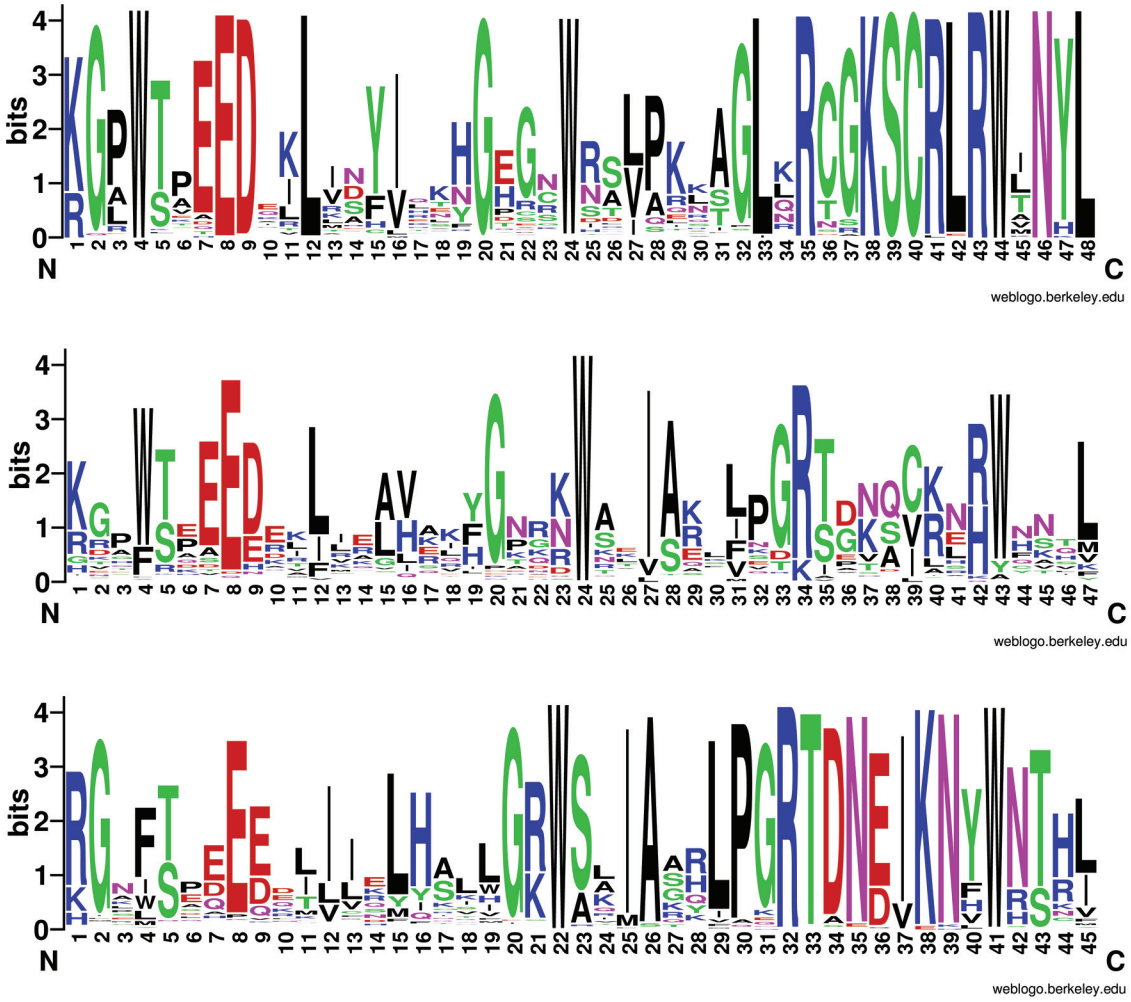
Table 4. Number of MYB genes.



Species name	1R	2R	3R	4R	5R	Numbers of MYB gene	Total MYB motifs
Rhododendron williamsianum	47 (47.5%)	48 (48.5%)	4 (4.0%)	0	0	99	155
Rhododendron delavayi	46 (29.5%)	103 (66.0%)	6 (3.8%)	0	1 (0.6%)	156	275
Vaccinium corymbosum	27 (5.6%)	422 (87.9%)	21 (4.4%)	10 (2.1%)	0	480	974
Actinidia chinensis	76 (41.1%)	102 (55.1%)	7 (3.8%)	0	0	185	301
Arabidopsis thaliana	1	124	5	1	0	131	268

Fig. 1. Motif location diagram for *Rhododendron williamsianum* with the motif size and type information. The detailed sequences (WEBLOGO) of Motif1 (M1), Motif2 (M2), and Motif3 (M3) are presented in Fig. 2. Information on other species is presented in Supplementary Fig. S1[1]. [Colour online.]

Fig. 2. The WEBLOGOs for Motif1 (M1), Motif2 (M2), and Motif3 (M3) of *Rhododendron williamsianum*. M1 and M2 had three tryptophans (W) and M1 had a more reserved motif between the second and third tryptophans (W) compared to that of M2. M3 had only two tryptophans (W), and the motif between the first and second tryptophans (W) was reserved as well. The WEBLOGO for other species were presented in Supplementary Fig. S2[1]. [Colour online.]



types, and the M3 type was the most abundant (78.7%). *Rhododendron delavayi* also had only M2 and M3 types, but the M2 type was the most abundant (91.3%). For *V. corymbosum* and *Actinidia chinensis*, M1, M2, and M3 types were all present. The ratios in *V. corymbosum* were 29.6%, 37.0%, and 33.3% for M1, M2, and M3 types, respectively, whereas those in *Actinidia chinensis* were 2.6%, 31.6%, and 65.8%, respectively. We classified the 2R genes into M2M2, M3M3, M1M2, M2M3, and M1M3 types. Interestingly, we obtained the dominant M2M2 type in *R. williamsianum*, whereas the other three species had a high ratio of M1M3 type. The M2M2 and M1M3 types were the most common, the M3M3, M1M2, and M2M3 types were very rare, and no M1M1 types were observed. The above results indicated a preference for the combination of conserved motifs among different species.

AtMYB drought-resistant gene orthologs in four species

A previous study by Baldoni et al. (2015) reported 19 AtMYB genes that were related to the drought resistance function. In addition, *AtMYB13* was reported by Kirik et al. (1998), *AtMYB37* by Yu et al. (2016), and *AtMYB73* was reported by Zhao et al. (2014) and Jung et al. (2008). Based on these drought-resistant gene sequences, we detected their orthologous genes in our four species (Table 2). Using the reciprocal best hit method, we detected 11, 6, 8, and 10 orthologous genes related to drought resistance in *R. delavayi*, *R. williamsianum*, *Actinidia chinensis*, and *V. corymbosum*, respectively. Interestingly, we found only three genes that had all the five orthologous genes in each species: *AtMYB52* (*AT1G17950*) orthologous group *RdMYB86*, *RwMYB94*, *AcMYB19*, and *VcMYB399*; *AtMYB73* (*AT4G37260*) orthologous group *RdMYB154*, *RwMYB18*, *AcMYB124*, and *VcMYB383*; and *AtMYB33* (*AT5G06100*) orthologous group *RdMYB84*, *RwMYB65*, *AcMYB170*, and *VcMYB138*. The gene tree of these three genes is presented in Supplementary Fig. S3[1]. In addition, three AtMYB genes had orthologous genes in all species except for *R. williamsianum*: *AtMYB101* (*AT2G32460*) orthologous group *RdMYB4*, *AcMYB182*, and *VcMYB91*; *AtMYB61* (*AT1G09540*) orthologous group

Table 5. Classification information of motifs in each species.

	1R			2R					3R		4R		5R
Species name	Motif1 (M1*)	Motif2 (M2)	Motif3 (M3)	M2M2	M3M3	M1M2	M2M3	M1M3	M1M2M3	3xM2	4xM2	M1M3M2M3	M1_4xM2
Rhododendron williamsianum	0 21.3%	10 78.7%	37	29 60.4%	5 10.4%	1 2.1%	0	13 27.1%	0	4	0	0	0
Rhododendron delavayi	0 91.3%	42 8.7%	4	25 24.3%	0	8 7.8%	6 5.8%	64 62.1%	1	5	0	0	1
Vaccinium corymbosum	8 29.6%	10 37.0%	9 33.3%	46 10.9%	0	0	151 35.8%	225 53.3%	4	17	8	2	0
Actinidia chinensis	2 2.6%	24 31.6%	50 65.8%	33 32.4%	0 0	7 6.9%	6 5.9%	56 54.9%	1	6	0	0	0

*M1, M2, M3 is not the same in each species.

RdMYB35, *AcMYB39*, and *VcMYB13*; and *AtMYB21* (*AT3G27810*) orthologous group *RdMYB99*, *AcMYB184*, and *VcMYB303*. All sequences for the 15 orthologous groups of drought-resistant *AtMYBs* are presented in supplementary file (see cjb 2020 0227suppli in Supplementary Material[1]). We also noticed that the *AtMYB88* (*AT2G02820*) orthologous gene *Achn183081* in *Actinidia chinensis* was not an *MYB* gene; the non-MYB part of the *Achn183081* gene had comparatively higher similarity to *AtMYB88*. We also noticed that there were seven *AtMYB* genes with no orthologous genes in our four species: *AtMYB2*, *AtMYB13*, *AtMYB37*, *AtMYB41*, *AtMYB75*, *AtMYB44*, and *AtMYB77*, indicating that these seven *AtMYB* genes might be lineage-specific genes in *Arabidopsis thaliana*. The alignment data for these 15 orthologous *MYB* gene groups indicate that all *MYB* genes have a high similarity region in the MYB domain in the N-terminal, and the similarity in the C-terminal is quite low (see cjb 2020 0227suppli in Supplementary Material[1]).

The expression profile of potential drought-resistant *RdMYB* genes

Based on the data from our previous transcriptomic analysis of drought tolerance in *R. delavayi*, we checked the expression levels of 11 *RdMYB* genes with potential drought resistance function. The final expression profile is presented in Fig. 3. Among these *RdMYB* genes, *RdMYB35*, *RdMYB38*, *RdMYB133*, and *RdMYB154* had comparatively higher expression levels, with maximum fragments per kilobase per million values ranging from tens to hundreds. The expression levels of these genes varied under the CK, D5, D9, and REC conditions. Compared to these four genes, other potential drought-resistant RdMYB genes had comparatively low expression levels, and two of them (*RdMYB4* and *RdMYB84*) had no maximum fragments per kilobase per million values.

Discussion

In this study, we performed a genome-wide analysis of four species (*V. corymbosum*, *Actinidia chinensis*, *R. delavayi*, and *R. williamsianum*) to elucidate the gene family of MYB TFs. Three species belonged to the Ericaceae family, and *Actinidia chinensis* was the outgroup species.

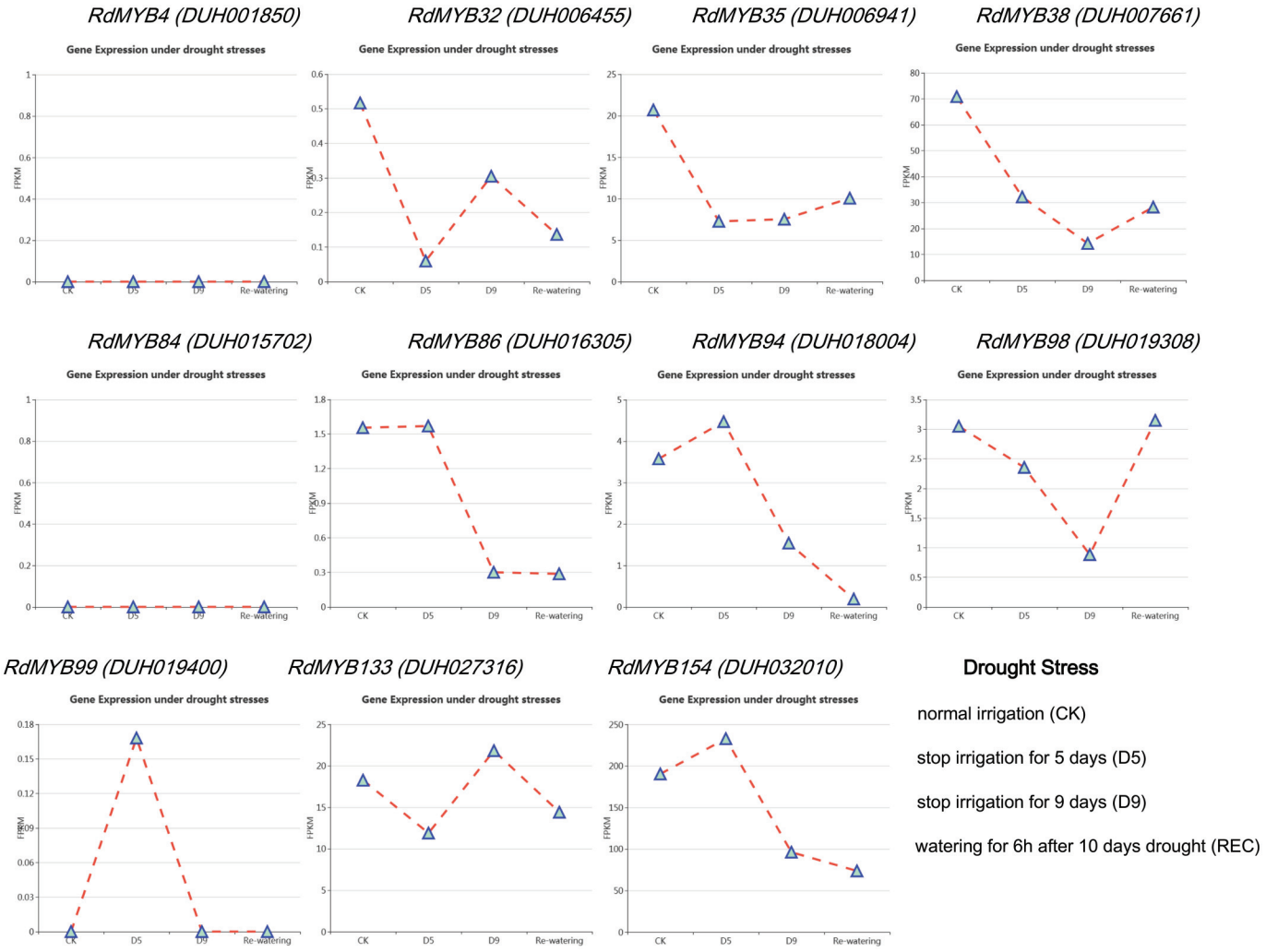
In this study, we combined BLASTP search with PFAM motif detection to obtain more stringent *MYB* gene families in our four selected species; this method was reliable. In this way, we obtained hundreds of candidate *MYB* genes first, and subsequently a small number of final *MYB* genes in each species (Table 1). Alignment results indicated that the candidate genes detected in the first step were similar to the *AtMYB* genes, but not the MYB domain, probably owing to the high similarity in the C-terminal. As a result, we were surprised to find that one of the 132 *AtMYB* genes in Stracke's article was misidentified as an *MYB* gene (Stracke et al. 2001), and that *AT4G17780* was presented as *AtMYB39*; however, it did not contain any MYB domains.

Based on the results from the PF00249 seed sequence comparison, we extracted the maximum region of *MYB* genes and recorded them as conserved MYB motifs. According to the number of conserved MYB motifs, we classified the *MYB* genes into five types: 1R, 2R, 3R, 4R, and 5R. The 2R type occupied the largest part in all four species but showed more divergent motifs (Table 5). This result was consistent with the previous idea that 2R type (R2R3) MYB proteins might predominantly be involved in plant-specific regulatory processes (Jin and Martin 1999; Stracke et al. 2001; Chen et al. 2017). However, we also noticed that 41.1% and 47.5% of genes belonged to the 1R type in *Actinidia chinensis* and *R. williamsianum*, respectively. These ratios were quite high compared to those of *R. delavayi* (29.5%), *V. corymbosum* (5.6%), and *Brachypodium distachyon* (L.) P.Beauv. (27.9%).

Significant variations in MYB family members were observed in our analyzed species (Table 4). *Rhododendron delavayi*, *Actinidia chinensis*, and *Arabidopsis thaliana* had comparable numbers of *MYB* genes. However, *V. corymbosum* had an extremely high number of *MYB* genes (480), which is possibly due to the occurrence of species-specific whole genome duplication and tandem duplication (Yang et al. 2020*a*). This result was also consistent with the highest 2R type MYB in *V. corymbosum* (Table 4). In contrast, *R. williamsianum* had much fewer *MYB* genes, corresponding to only half number of those in *R. delavayi* and *Actinidia chinensis* (Table 4), indicating that gene or ploidy loss probably happened in *R. williamsianum* during recent evolution.

For the three drought-responsive AtMYBs that had all five orthologous genes in each species, we built the gene tree using the ML method in MEGA7. The *AtMYB33* and

Fig. 3. The expression profile of 11 potential drought-resistant *RdMYB* genes under the following growing conditions: normal irrigation (CK), stopping irrigation for five days (D5), stopping irrigation for nine days (D9), and stopping irrigation for 10 days and then re-watering for six hours (REC). [Colour online.]



AtMYB52 group had gene trees identical to the species tree, suggesting their similar evolutionary patterns. *MYB33* confers drought tolerance via the ABA signaling pathway in both *Arabidopsis* and potato (Wyrzykowska et al. 2021). Park et al. (2011) also found that plants were hypersensitive to ABA and drought tolerant after overexpressing MYB52; however, the gene tree of the *AtMYB73* group was different from its species tree. We obtained the same result by using only the MYB region of the *AtMYB73* group. As the gene tree was unique, we used the *AtMYB73* homologous *RdMYB154* gene sequences to perform a BLAST search of the *R. williamsianum* genome protein sequence database and found that the best hit of *RdMYB154* was *RwMYB43*. With the new sequences, the gene tree for the *AtMYB73* group was built again, and the resulting tree was identical to the species tree. Thus, we modified the result that the orthologous gene of *AtMYB73* in *R. williamsianum* was *RwMYB43* (Table 2). We also validated the reciprocal best hit data, and the results indicated that *RwMYB18* (*KAE9449693.1*) had a longer alignment length and higher hit score than *RwMYB43* (*KAE9457199.1*), but had a lower identity, suggesting a rapid evolution of *MYB73* in diverse species. Functionally, MdMYB73 in apples was found to regulate vacuole acidification and malate accumulation (Hu et al. 2017). However, Yang et al. (2020*b*) found that AtMYB73 interacts with AtUVR8 to mediate auxin response and lateral root development. When we aligned all 15 orthologous MYB genes, higher similarities were found in the MYB domain N-terminal region, whereas those in the C-terminal were quite low (see cjb 2020 0227suppli in Supplementary Material[1]). This result was expected and implied that the functional divergence of the orthologous genes in the four species might have occurred at the C-terminus.

Conclusions

This study is the first to systematically analyze the *MYB* gene family in *V. corymbosum*, *Actinidia chinensis*, *R. delavayi*, and *R. williamsianum*. Among hundreds of *MYB* genes, we classified these genes into five types, and we classified the conserved MYB motifs into three types. Our results revealed

a clear profile of the *MYB* gene family in the four study species. In addition, we also detected potential drought-resistant genes in the four species using the reciprocal best hit method. Our results can play an important role in the future analyses of these four species.

Conflict of interest

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

Funding

This work was supported by Major Science and Technology Projects of Yunnan Province of China (2018BB014), Youth Innovation Promotion Association CAS (2021394), and Yunnan Young and Elite Talents Project.

Author contributions

Y.Z. and L.L. performed the major analyses and wrote the manuscript; N.L. performed the expression data analysis; H.C., H.L., and G.D. helped with the analyses and prepared the materials; W.J. and C.Z. designed the analyses and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank all those who provided useful discussions and technical support for our article. We would also like to thank Editage (www.editage.cn) for English language editing.

References

Abe, H., Urao, T., Ito, T., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell, **15**(1): 63–78. doi:10.1105/tpc.006130. PMID:12509522.

Baldoni, E., Genga, A., and Cominelli, E. 2015. Plant MYB transcription factors: their role in drought response mechanisms. Int. J. Mol. Sci. **16**(7): 15811–15851. doi:10.3390/ijms160715811. PMID:26184177.

Cai, Y.F., Wang, J.H., Zhang, L., Song, J., Peng, L.C., and Zhang, S.B. 2019. Physiological and transcriptomic analysis highlight key metabolic pathways in relation to drought tolerance in *Rhododendron delavayi*. Physiol. Mol. Biol. Plants, **25**(4): 991–1008. doi:10.1007/s12298-019-00685-1. PMID:31402822.

Cao, Y., Li, K., Li, Y., Zhao, X., and Wang, L. 2020. MYB transcription factors as regulators of secondary metabolism in plants. Biology, **9**(3): 61. doi:10.3390/biology9030061. PMID:21951676.

Chen, G., He, W., Guo, X., and Pan, J. 2021. Genome-wide identification, classification and expression analysis of the MYB Transcription factor family in *Petunia*. Int. J. Mol. Sci. **22**(9): 4838. doi:10.3390/ijms22094838. PMID:34063617.

Chen, S., Niu, X., Guan, Y., and Li, H. 2017. Genome-wide analysis and expression profiles of the MYB genes in *Brachypodium distachyon*. Plant Cell Physiol. **58**(10): 1777–1788. doi:10.1093/pcp/pcx115. PMID:29016897.

Cheng, H., Song, S., Xiao, L., Soo, H.M., Cheng, Z., Xie, D., and Peng, J. 2009. Gibberellin acts through jasmonate to control the expression of *MYB21*, *MYB24*, and *MYB57* to promote stamen filament growth in *Arabidopsis*. PLOS Genet. **5**(3): e1000440. doi:10.1371/journal.pgen.1000440. PMID:19325888.

Colle, M., Leisner, C.P., Wai, C.M., Ou, S., Bird, K.A., Wang, J., et al. 2019. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. Gigascience, **8**(3): giz012. doi:10.1093/gigascience/giz012. PMID:30715294.

Cominelli, E., Galbiati, M., Vavasseur, A., Conti, L., Sala, T., Vuylsteke, M., et al. 2005. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. Curr. Biol. **15** (13): 1196–1200. doi:10.1016/j.cub.2005.05.048. PMID:16005291.

Cominelli, E., Sala, T., Calvi, D., Gusmaroli, G., and Tonelli, C. 2008. Over-expression of the Arabidopsis *AtMYB41* gene alters cell expansion and leaf surface permeability. Plant J. **53**(1): 53–64. doi:10.1111/j.1365-313X.2007.03310.x. PMID:17971045.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. 2004. WebLogo: A sequence logo generator. Genome Res. **14**(6): 1188–1190. doi:10.1101/gr.849004. PMID:15173120.

Ding, Z., Li, S., An, X., Liu, X., Qin, H., and Wang, D. 2009. Transgenic expression of *MYB15* confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. J. Genet. Genomics, **36**(1): 17–29. doi:10.1016/S1673-8527(09)60003-5. PMID:19161942.

Du, H., Feng, B.-R., Yang, S.-S., Huang, Y.-B., and Tang, Y.-X. 2012*a*. The R2R3-MYB transcription factor gene family in maize. PLoS ONE, **7**(6): e37463. doi:10.1371/journal.pone.0037463. PMID:22719841.

Du, H., Yang, S.S., Liang, Z., Feng, B.R., Liu, L., Huang, Y.B., and Tang, Y.X. 2012*b*. Genome-wide analysis of the MYB transcription factor superfamily in soybean. BMC Plant Biol. **12**: 106. doi:10.1186/1471-2229-12-106. PMID:22776508.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5): 1792–1797. doi:10.1093/nar/gkh340. PMID:15034147.

Gao, S., Zhang, Y.L., Yang, L., Song, J.B., and Yang, Z.M. 2014. AtMYB20 is negatively involved in plant adaptive response to drought stress. Plant Soil, **376**(1–2): 433–443. doi:10.1007/s11104-013-1992-6.

Gómez-Gómez, L., Trapero-Mozos, A., Gómez, M.D., Rubio-Moraga, A., and Ahrazem, O. 2012. Identification and possible role of a MYB transcription factor from saffron (*Crocus sativus*). J. Plant Physiol. **169**(5): 509–515. doi:10.1016/j.jplph.2011.11.021. PMID:22297127.

Gonzalez, M., Carrasco, B., and Salazar, E. 2016. Genome-wide identification and characterization of R2R3MYB family in Rosaceae. Genomics Data, **9**: 50–57. doi:10.1016/j.gdata.2016.06.004. PMID:27408811.

González-Verdejo, C.I., Dita, M.A., Di Pietro, A., Moreno, M.T., Barandiarán, X., Rubiales, D., et al. 2007. Identification and expression analysis of a MYB family transcription factor in the parasitic plant *Orobanche ramosa*. Ann. Appl. Biol. **150**(2): 123–130. doi:10.1111/j.1744-7348.2006.00114.x.

He, Q., Jones, D.C., Li, W., Xie, F., Ma, J., Sun, R., et al. 2016. Genome-wide identification of R2R3-MYB genes and expression analyses during abiotic stress in *Gossypium raimondii*. Sci. Rep. **6**(1): 22980. doi:10.1038/srep22980. PMID:27009386.

Hu, D.-G., Li, Y.-Y., Zhang, Q.-Y., Li, M., Sun, C.-H., Yu, J.-Q., and Hao, Y.-J. 2017. The R2R3-MYB transcription factor MdMYB73 is involved in malate accumulation and vacuolar acidification in apple. Plant J. **91**(3): 443–454. doi:10.1111/tpj.13579. PMID:28423209.

Huang, S.X., Ding, J., Deng, D.J., Tang, W., Sun, H.H., Liu, D.Y., and Liu, Y.S. 2013. Draft genome of the kiwifruit *Actinidia chinensis*. Nat. Commun. **4**: 2640. doi:10.1038/ncomms3640. PMID:24136039.

Jin, H.L., and Martin, C. 1999. Multifunctionality and diversity within the plant MYB-gene family. Plant Mol. Biol. **41**(5): 577–585. doi:10.1023/A:1006319732410. PMID:10645718.

Jung, C., Seo, J.S., Han, S.W., Koo, Y.J., Kim, C.H., Song, S.I., et al. 2008. Overexpression of *AtMYB44* enhances stomatal closure to confer abiotic stress tolerance in transgenic *Arabidopsis*. Plant Physiol. **146**(2): 623–635. doi:10.1104/pp.107.110981. PMID: 18162593.

Katiyar, A., Smita, S., Lenka, S.K., Rajwanshi, R., Chinnusamy, V., and Bansal, K.C. 2012. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. BMC Genomics, **13**(1): 544. doi:10.1186/1471-2164-13-544. PMID:23050870.

Kirik, V., Kolle, K., Wohlfarth, T., Misera, S., and Baumlein, H. 1998. Ectopic expression of a novel *MYB* gene modifies the architecture of the *Arabidopsis* inflorescence. Plant J. **13**(6): 729–742. doi:10.1046/j.1365-313X.1998.00072.x. PMID:9681014.

Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R.C., et al. 1998. Towards functional characterisation of the members of the *R2 R3-MYB* gene family from *Arabidopsis thaliana*. Plant J. **16**(2): 263–276. doi:10.1046/j.1365-313x.1998.00278.x. PMID:9839469.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. **33**(7): 1870–1 874. doi:10.1093/molbev/msw054. PMID:27004904.

Lai, L.B., Nadeau, J.A., Lucas, J., Lee, E.K., Nakagawa, T., Zhao, L., et al. 2005. The *Arabidopsis* R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. Plant Cell, **17**(10): 2754–2767. doi:10.1105/tpc.105.034116. PMID:16155180.

Lee, S.B., and Suh, M.C. 2015. Cuticular wax biosynthesis is up-regulated by the MYB94 transcription factor in *Arabidopsis*. Plant Cell Physiol. **56**(1): 48–60. doi:10.1093/pcp/pcu142. PMID: 25305760.

Liu, N., Zhang, L., Zhou, Y., Tu, M., Wu, Z., Gui, D., et al. 2021. The *Rhododendron* Plant Genome Database (RPGD): a comprehensive online omics database for *Rhododendron*. BMC Genomics, **22**(1): 376. doi:10.1186/s12864-021-07704-0. PMID:34022814.

Mengiste, T., Chen, X., Salmeron, J., and Dietrich, R. 2003. The *BOTRYTIS SUSCEPTIBLE1* gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in *Arabidopsis*. Plant Cell, **15**(11): 2551–2565. doi:10.1105/tpc.014167. PMID:14555693.

Mitra, M., Agarwal, P., and Roy, S. 2021. The N-terminal MYB domains affect the stability and folding aspects of *Arabidopsis thaliana* MYB4 transcription factor under thermal stress. Protoplasma, **258**(3): 633–650. doi:10.1007/s00709-020-01590-1. PMID:33398463.

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., et al. 2014. Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. Plant J. **77**(3): 367–379. doi:10.1111/tpj.12388. PMID:24274116.

Oh, J.E., Kwon, Y., Kim, J.H., Noh, H., Hong, S.W., and Lee, H. 2011. A dual role for MYB60 in stomatal regulation and root growth of *Arabidopsis thaliana* under drought stress. Plant Mol. Biol. **77**(1–2): 91–103. doi:10.1007/s11103-011-9796-7. PMID:21637967.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. 2003. Control of leaf morphogenesis by microRNAs. Nature, **425**(6955): 257–263. doi:10.1038/nature01958. PMID:12931144.

Park, M.Y., Kang, J-y., and Kim, S.Y. 2011. Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance. Mol. Cells, **31**(5): 447–454. doi:10.1007/s10059-011-0300-7. PMID: 21399993.

Penfield, S., Meissner, R.C., Shoue, D.A., Carpita, N.C., and Bevan, M.W. 2001. MYB61 is required for mucilage deposition and extrusion in the Arabidopsis seed coat. Plant Cell, **13**(12): 2777–2791. doi:10.1105/tpc.010265. PMID:11752387.

Rabinowicz, P.D., Braun, E.L., Wolfe, A.D., Bowen, B., and Grotewold, E. 1999. Maize R2R3 Myb genes: Sequence analysis reveals amplification in the higher plants. Genetics, **153**(1): 427–444. doi:10.1093/genetics/153.1.427. PMID:10471724.

Reyes, J.L., and Chua, N.H. 2007. ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination. Plant J. **49**(4): 592–606. doi:10.1111/j.1365-313X.2006.02980.x. PMID:17217461.

Rodrigues, J.A., Espley, R.V., and Allan, A.C. 2021. Genomic analysis uncovers functional variation in the C-terminus of anthocyanin-activating MYB transcription factors. Hortic. Res. **8**(1): 77. doi:10.1038/s41438-021-00514-1. PMID:33790254.

Romano, J.M., Dubos, C., Prouse, M.B., Wilkins, O., Hong, H., Poole, M., et al. 2012. AtMYB61, an R2R3-MYB transcription factor, functions as a pleiotropic regulator via a small gene network. New Phytol. **195**(4): 774–786. doi:10.1111/j.1469-8137.2012.04201.x. PMID:22708996.

Seo, P.J., Xiang, F., Qiao, M., Park, J.Y., Lee, Y.N., Kim, S.G., et al. 2009. The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in *Arabidopsis*. Plant Physiol. **151**(1): 275–289. doi:10.1104/pp.109.144220. PMID: 19625633.

Soza, V.L., Lindsley, D., Waalkes, A., Ramage, E., Patwardhan, R.P., Burton, J.N., et al. 2019. the rhododendron genome and chromosomal organization provide insight into shared whole-genome duplications across the heath family (Ericaceae). Genome Biol. Evol. **11**(12): 3353–3371. doi:10.1093/gbe/evz245. PMID:31702783.

Stracke, R., Werber, M., and Weisshaar, B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. Curr. Opin. Plant Biol. **4**(5): 447–456. doi:10.1016/S1369-5266(00)00199-0. PMID:11597504.

Stracke, R., Holtgrawe, D., Schneider, J., Pucker, B., Sorensen, T.R., and Weisshaar, B. 2014. Genome-wide identification and characterisation of *R2R3-MYB* genes in sugar beet (*Beta vulgaris*). BMC Plant Biol. **14**: 249. doi:10.1186/s12870-014-0249-8. PMID:25249410.

Su, Z., Ma, X., Guo, H., Sukiran, N.L., Guo, B., Assmann, S.M., and Ma, H. 2013. Flower development under drought stress: morphological and transcriptomic analyses reveal acute responses and long-term acclimation in Arabidopsis. Plant Cell, **25**(10): 3785–3807. doi:10.1105/tpc.113.115428. PMID:24179129.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S., and Shinozaki, K. 1993. An Arabidopsis *myb* homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. Plant Cell, **5**(11): 1529–1539. doi:10.2307/3869736. PMID:8312738.

Wang, F., Suo, Y., Wei, H., Li, M., Xie, C., Wang, L., et al. 2015. Identification and characterization of 40 isolated *Rehmannia glutinosa* MYB family genes and their expression profiles in response to shading and continuous cropping. Int. J. Mol. Sci. **16**(7): 15009–15030. doi:10.3390/ijms160715009. PMID:26147429.

Wei, Q., Chen, R., Wei, X., Liu, Y., Zhao, S., Yin, X., and Xie, T. 2020. Genome-wide identification of R2R3-MYB family in wheat and functional characteristics of the abiotic stress responsive gene *TaMYB344*. BMC Genomics, **21**(1): 792. doi:10.1186/s12864-020-07175-9. PMID:33183233.

Wilkins, O., Nahal, H., Foong, J., Provart, N.J., and Campbell, M.M. 2009. Expansion and diversification of the Populus R2R3-MYB family of transcription factors. Plant Physiol. **149**(2): 981–993. doi:10.1104/pp.108.132795. PMID:19091872.

Wyrzykowska, A., Bielewicz, D., Plewka, P., Sołtys-Kalina, D., Wasilewicz-Flis, I., Marczewski, W., et al. 2021. The MYB33, MYB65, and MYB101 transcription factors affect *Arabidopsis* and potato responses to drought by regulating the ABA signaling pathway. bioRxiv: 2021.2004.2007.438025. doi:10.1101/2021.04.07.438025.

Xiao, R., Zhang, C., Guo, X., Li, H., and Lu, H. 2021. MYB transcription factors and its regulation in secondary cell wall

formation and lignin biosynthesis during xylem development. Int. J. Mol. Sci. **22**(7): 3560. doi:10.3390/ijms22073560. PMID:33808132.

Xie, Z., Li, D., Wang, L., Sack, F.D., and Grotewold, E. 2010. Role of the stomatal development regulators FLP/MYB88 in abiotic stress responses. Plant J. **64**(5): 731–739. doi:10.1111/j.1365-313X.2010.04364.x. PMID:21105921.

Yang, F.-S., Nie, S., Liu, H., Shi, T.-L., Tian, X.-C., Zhou, S.-S., et al. 2020*a*. Chromosome-level genome assembly of a parent species of widely cultivated azaleas. Nat. Commun. **11**(1): 5269. doi:10.1038/s41467-020-18771-4. PMID:33077749.

Yang, J., Ruan, M., Guo, X., and Peng, M. 2021. Characterization and function analysis of Cassava MYB transcription factor *MeMYB2*. Chin. J. Trop. Crops, **42**(4): 936–944. doi:10.3969/j.issn.1000-2561.2021.04.004.

Yang, Y., Zhang, L., Chen, P., Liang, T., Li, X., and Liu, H. 2020*b*. UV-B photoreceptor UVR8 interacts with MYB73/MYB77 to regulate auxin responses and lateral root development. EMBO J. **39**(2): e101928. doi:10.15252/embj.2019101928. PMID:31777974.

Yu, Y.T., Wu, Z., Lu, K., Bi, C., Liang, S., Wang, X.F., and Zhang, D.P. 2016. Overexpression of the MYB37 transcription factor enhances abscisic acid sensitivity, and improves both drought tolerance and seed productivity in *Arabidopsis thaliana*. Plant Mol. Biol. **90**(3): 267–279. doi:10.1007/s11103-015-0411-1. PMID:26646286.

Zhang, L., Xu, P.W., Cai, Y.F., Ma, L.L., Li, S.F., Li, S.F., et al. 2017. The draft genome assembly of Rhododendron delavayi Franch. var. delavayi. Gigascience. **6**(10): 1–11. doi:10.1093/gigascience/gix076. PMID:29020749.

Zhang, W., Luo, C., Scossa, F., Zhang, Q., Usadel, B., Fernie, A.R., et al. 2021. A phased genome based on single sperm sequencing reveals crossover pattern and complex relatedness in tea plants. Plant J. **105**(1): 197–208. doi:10.1111/tpj.15051. PMID:33118252.

Zhao, Y., Xing, L., Wang, X., Hou, Y.J., Gao, J., Wang, P., et al. 2014. The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. Sci. Signal. **7**(328): ra53. doi:10.1126/scisignal.2005051. PMID:24894996.

Zhou, W., Zhang, Q., Sun, Y., Yang, L., and Wang, Z. 2020. Genome-wide identification and characterization of R2R3-MYB family in *Hypericum perforatum* under diverse abiotic stresses. Int. J. Biol. Macromol. **145**: 341–354. doi:10.1016/j.ijbiomac.2019.12.100. PMID:31857171.

Zhu, L., Guan, Y., Liu, Y., Zhang, Z., Jaffar, M.A., Song, A., et al. 2020. Regulation of flowering time in chrysanthemum by the R2R3 MYB transcription factor CmMYB2 is associated with changes in gibberellin metabolism. Hortic. Res. **7**(1): 96. doi:10.1038/s41438-020-0317-1. PMID:32637124.