Identification of MYB Gene Family in medicinal tea tree Melaleuca alternifolia (Maiden & Betche) Cheel and Analysis of Members Regulating Terpene Biosynthesis

Xiaoning Zhang
Guangxi Foresrty Research Institute
Zhanwu Xu
Hubei University
Buming Liu
Guangxi Institute of Chinese Medicine&Pharmaceutical Science
Yufei Xiao
Guangxi Foresrty Research Institute
Ling Chai
Guangxi Institute of Chinese Medicine&Pharmaceutical Science
Lianxiang Zhong
Guangxi Foresrty Research Institute
Heqiang Huo
Mid-Florida Research and Education Center, Department of Environmental Horticulture
Li Liu
Hubei University
Hong Yang
Hubei University
Hailong Liu (hailon_liu@163.com)
Guangxi Foresrty Research Institute

Research Article

Keywords: Melaleuca alternifolia, Transcription factors (TFs), MYB gene family, Characterization, Phylogenetic Analyses, Function Predictions

Posted Date: August 24th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3262517/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Molecular Biology Reports on January 4th, 2024. See the published version at https://doi.org/10.1007/s11033-023-09019-7.

Abstract

Background

The tea tree (*Melaleuca alternifolia*) is renowned for its production of tea tree oil (TTO), an essential oil primarily composed of terpenes extracted from its shoot. MYB transcription factors, which are one of the largest TF families, play a crucial role in regulating primary and secondary metabolite synthesis. However, knowledge of the MYB gene family in *M. alternifolia* is limited.

Methods and results

Here, we conducted a comprehensive genome-wide analysis of MYB genes in *M. alternifolia*, referred to as MaMYBs, including phylogenetic relationships, structures, promoter regions, and GO annotations. Our findings classified 219 MaMYBs into four subfamilies: one 5R-MYB, four 3R-MYBs, sixty-one MYB-related, and the remaining 153 are all 2R-MYBs. Seven genes (*MYB189, MYB146, MYB44, MYB29, MYB175, MYB162*, and *MYB160*) were linked to terpenoid synthesis based on GO annotation. Phylogenetic analysis with *Arabidopsis* homologous MYB genes suggested that *MYB193* and *MYB163* may also be involved in terpenoid synthesis. Additionally, through correlation analysis of gene expression and metabolite content, we identified 42 MYB genes associated with metabolite content.

Conclusion

The results provide valuable insights into the importance of MYB transcription factors in essential oil production in *M. alternifolia.* These findings lay the groundwork for a better understanding of the MYB regulatory network and the development of novel strategies to enhance essential oil synthesis in *M. alternifolia.*

Introduction

MYB transcription factors (TFs) are a significant component of the TF families in plants. They play a pivotal role in many important biological processes, including cell morphogenesis, plant development, stress response, and regulation of metabolite synthesis. MYB proteins are differentiated by the MYB DNA-binding domain, which generally spans 50 to 53 amino acids and is highly conserved. MYB family members is classified into 1R-, 2R-, 3R-, and 4R-MYB based on their number of imperfect repeats in the MYB proteins [1]. 2R-MYB(R2R3-MYB) subfamily, containing two repeats, is the largest of the plant MYB family. 198 MYB genes have been identified in the *Arabidopsis* genome, containing 126 R2R3-MYB proteins. MYB genes have also been identified in the genomes of crop plants [2–3], orchidaceae plant[4], and woody plant species, such as Masson pine (*Pinus massoniana*) [5], sweet osmanthus (*Osmanthus fragrans*) [6], tea plant (*Camellia sinensis*)[7], poplar (*Populus trichocarpa*) [8], and eucalytus (*Eucalyptus grandis*) [9]. Some studies have focused on the regulation of secondary metabolism by MYB genes, as

demonstrated in the ornamental plant *Hedychium coronarium* (white ginger lily). In this species, it has been proposed that five MYB genes (HcMYB7, 8, 79, 145, 248) contribute to the formation of the floral scent, which is predominantly composed of terpenes and benzenoids, by regulating terpenoid and benzenoid biosynthesis. Silencing of these genes resulted in a reduction in floral volatile content [10].

In Chinese Bayberry (*Morella rubra*), the MYB transcription factor MrMYB12 has been shown to contribute to a higher content of flavanol by increasing the conversion of anthocyanins to flavonol biosynthesis. MrFLS1 is the key enzyme responsible for flavanol biosynthesis, and MrMYB12 has been demonstrated to trans-activate MrFLS1 by binding to its promoter. Overexpression of 35S : MrMYB12 in tobacco resulted in significant induction of NtFLS expression, an increase in flavanol content but a decrease in anthocyanin production in both flowers and leaves, resulting in white or light pink flowers [11].

In the aromatic plant spearmint (*Mentha spicata*), a MsMYB gene was identified as playing a role in regulating the production of essential oil in peltate glandular trichomes (PGTs). The gene is PGT-specific and negatively regulates monoterpene biosynthesis. Overexpression of MsMYB decreased monoterpene content while silencing of MsMYB increased its production [12].

Melaleuca alternifolia, native to Australia, is an economically important medicinal tea tree, renowned for its abundant tea tree oil (TTO) in its branches and leaves. Tea tree oil contains more than 100 compounds, among which, terpene-4-ol plays a key role in the antibacterial activity of tea tree oil and is of high value in the pharmaceutical and cosmetic industries. However, the function of the *Ma*MYB gene family needed to be fully explored. With the release of its genomic information [13–14], we conducted a comprehensive investigation of the MYB gene family in *Melaleuca alternifolia*. The results include the characterization of the phylogenetic relationship, gene structures, conserved motifs, and functional annotation of 219 *MaMYB* genes, as well as the evaluation of their expression patterns in different ages of *M. alternifolia*.

Material and Methods

Plant Materials

Leaves of *Melaleuca alternifolia* were collected from plantlets at four different ages: 18-day, 60-day, 1-year-old, and 3-year-old. The samples were ground in liquid nitrogen to extract RNA. The 18-day and 60-day *M. alternifolia* plantlets were grown in tissue culture (25 ℃, 16h light photoperiod, 2,500 lx illumination intensity). The 1-year-old and 3-year-old *M. alternifolia* plantlets were grown at the Guangxi Forestry Research Institute.

Identification and Classification Analysis

A combination of bioinformatics tools was used to identify MYB gene family in *M. alternifolia*. *Arabidopsis* MYB protein sequences (available at http://www.phytozome.net/phytochrome) were utilized as query sequences in the *M. alternifolia* genomic [13–14] and transcriptome data (CNCB,

PRJCA019034). 252 candidate MYB proteins were obtained and confirmed for the presence of the MYB domain using NCBI Conserved Domain Search2[15] and SMART program3[16]. The validity of the *Ma*MYBs was determined by executing multiple sequence alignment analyses of 252 *MaMYB*s with Clustal W[17] and discarding any redundant or false predicted sequences. Finally, a comprehensive analysis was performed on 219 validated *MaMYB* genes in the *M. alternifolia* genome, and their physical and chemical parameters were also evaluated using ExPASy (http://web.expasy.org).

Subcellular localization

Subcellular localization of the 219 MYB protein of *M. alternifolia* was predicted with CELLO v.2.5. (http://cello.life.nctu.edu.tw/)

Protein motif identification

We utilized the MEME tool version 3.5.7[18] to identify conserved motifs of Ma MYB proteins. The parameter settings were: a maximum numbers of motifs = 50, 6 ≤ motif width ≤ 250, and the motifs must be present in each member within the same subfamily. The motifs distribution of per sequence was set to 0 or 1. Motifs with an e-value less than 1e-20 were considered for further analysis. Finally, the MAST program was employed to search for the detected motifs in protein databases [19].

Phylogenetic Analysis

The multiple sequence alignments were used to conduct phylogenetic analysis. Protein sequences of 153 *Ma*R2R3-MYBs and 126 *At*R2R3-MYBs were utilized to build a phylogenetic tree. The sequences were first aligned using ClustalW and then manually refined. The phylogenetic tree was constructed with the Maximum Likelihood method in MEGA 7.0 [20] with 1000 bootstrap replicates. The amino acid sequence of AtMYB from the *Arabidopsis* family was obtained from the TAIR database (http://www.arabidopsis.org/).

Analysis of the relative content of terpinen-4-ol

Terpinen-4-ol was determined by qualitative analysis and manual analysis using the gas-mass spectrometry N1ST standard spectrum library. Quantitative analysis was carried out by meteorological chromatography, and its content was calculated by peak area normalization. Shimazu QP5050A was used for qualitative analysis of GC-MS. The chromatographic conditions were as follows: elastic quartz capillary column AT-35(60m *0.25mm*0.25μm) with high purity helium, injector 250℃, column temperature 250℃. Manual sampling was performed, AT = 1:50. Mass spectrum conditions: an EI ion source, ionization voltage was 70 ev, scanning standard was 33 ~ 800 amu, full scanning mode. The sample size was 0.2μL.

RNA Isolation, RNA-seq, and RT-qPCR expression analysis

RNA Isolation, sequences, and Data Analysis see [21]. The methods of total RNA extraction and analysis were performed by Luo et al [3]. And the primers were designed based on Primer Premier 5.0 software[22].

Results

Identification and Classification of Ma MYB Genes

This analysis resulted in the identification of 252 potential MYB or MYB-like repeat amino acid sequences. NCBI Conserved Domain Search as well as the SMART program were performed to confirm the existence of MYB DNA-binding domains, and Multiple sequence alignment was also conducted to ensure its sequence integrity. As a result, 33 MaMYBs containing incomplete MYB domains were excluded, leaving 219 MaMYBs, designated as MYB1 to MYB219 (Table S1).

All 219 MaMYB proteins were classified into 1R-MYBs (61), 2R-MYBs (R2R3-MYBs) (153), 3R-MYBs (4), and 5R-MYBs (1), based on the number of MYB repeats. The results indicate that R2R3-MYBs, which account for approximately 69.9% of the MYBs, are predominant in *Melaleuca alternifolia*, like most plants.

Physicochemical properties and subcellular localizations analysis

The 219 MaMYB proteins vary greatly in size and molecular mass, ranging from 74 amino acids and 8.65 kDa (MYB105) to 1,739 amino acids and 190.33 kDa (MYB143), with pI values ranging from 4.21 (MYB90) to 10.49 (MYB4). Five proteins, MYB46, MYB184, MYB57, MYB3, and MYB206 were considered stable due to their instability index ≤ 40. Subcellular localization prediction indicated that there were two hundreds and seven MaMYB proteins (over 94.5%) located in the nucleus, nine in the chloroplast, two in the cytoplasm, and only one in the endoplasmic reticulum. Additional information of the MYB proteins were shown in Table S2.

The Phylogenetic analysis

The maximum likelihood method was used to construct a phylogenetic tree, with the full-length amino acid sequences alignment of 219 MaMYBs, which were grouped into 11 subgroups designated as M1 to M11 (Figure S1 A). Subgroup M11 was the largest, containing 48 MYB members, followed by M5 and M10 (23), M8 (22), M2 and M6 (20), M1 (19), M9 (16), M3 and M7 (10), and M4 (8), respectively.

Twenty conserved motifs were verified using the MEME server (Figure S1 B). The *MaMYB* conserved motifs were located primarily in the N-terminal region. The majority of *MaMYB* proteins belong to the same subclade and share a similar composition of motifs. The motifs 3, 4, 1, 2, and 5 were the most conserved and appeared in most of the *MaMYB* proteins. Most of the *MaMYB* proteins contained more than one motifs, except for MYB102, MYB117, MYB123, and MYB158, which contained only one motif (Fig. 1A B). Detailed motif information is available in Table S3.

Gene structure analysis was performed to gain a deeper understanding of the MYB gene expansion in *M. alternifolia.* The results, including gene length, number of exons, and introns, are available in Table S4 &

Figure S1. The analysis showed that the length of MaMYB genes varied significantly among the *M. alternifolia* MYB gene family members, with MYB156 being the shortest at only 396 bp, and MYB61 being the longest at 10,712 bp (Table S4). The number of exons and introns ranged from 1- 13, and 0 - 12, respectively. Five genes (MYB116, MYB21, MYB136, MYB207, MYB213) were found to contain 10 or more introns, with MYB116 having a maximum of 12 introns (Figure 1C). The majority of the genes were disrupted by one (22%), two (54%), or three (7%) introns. However, there were also 20 intron-less MYB genes. The diversity in the number of exons or introns of the *MaMYB* gene family may contribute to their functional diversity and could be a result of gains and losses during the evolutionary process.

Sequence features of the MaMYB conserved domains

To investigate the sequence characteristics of homologous structural domains, multiple alignment analysis was performed of 153 homologous amino acid sequences of the MaR2R3-MYB structural domain. We obtained the frequencies of the most common amino acids within each repetitive region, as shown in Fig. 2. The longer the length of the letters at the corresponding positions, the higher the amino acid frequencies, indicating that these amino acids are more conserved.

Conserved amino acid distribution in the MYB structural domain in *Melaleuca alternifolia* is very similar to that of other plant species, with high sequence similarity observed in R2 and R3 DNA-binding structural domains. The basic region of the R2R3-MYB structural domain contains 105 basic amino acids (104–105), with each MYB repeat region containing approximately 51 or 54 amino acid residues separated by three conserved tryptophan residues.

The primary structure domains is [-W -(X 19)-W-(X 19)-W-..... -F/I-(X 20)-W-(X 18)-W-], where W, F, I and X stands for tryptophan, phenylalanine, isoleucine, and any amino acid. There are two or three evenly distributed and highly conserved Trp residues in these 2RMaMYB repeats, which serve as a label and play a key role in sequence-specific DNA binding of plant MYB proteins. In the R3 repeat, however, the first tryptophan(W) is often missing. As shown in Fig. 2 replaced by phenylalanine (F) or isoleucine (I). Like the highly conserved tryptophan residues, Glu-9, Asp-10, Leu-13, Gly-21 and Ser-40 are fully conserved in the R2 repeat, as are Glu-9, Gly-21, Ile-27, Arg-34 and Thr-35 in the R3 repeat.

Phylogenetic Analyses and Function Predictions

The R2R3-type MYB proteins comprise the largest MYB family, with the R2R3-MYB subfamily accounting for 69.86% of the MYB family. A total of 153 and 126 R2R3-MYB proteins from *M. alternifolia* and *Arabidopsis* were aligned. The results revealed that all 153 R2R3-MYB proteins from *M. alternifolia* were grouped with 126 R2R3-MYB proteins from *Arabidopsis* into 19 subclades (S1-S19) (Fig. 3) There were two species-specific clades (subclades 15 and 17) that only contained R2R3-MYB proteins from *M. alternifolia* or *Arabidopsis*. For instance, subgroup S17 contained Arabidopsis-specific genes such as *AtMYB28*, *AtMYB29*, *AtMYB34*, *AtMYB51*, *AtMYB76*, and *AtMYB122*, while subgroup S15 was species-specific in *M. alternifolia*. The R2R3-MYB transcription factor family is the most important TF family in

plants and has been well characterized in *Arabidopsis.* Gene function can be predicted by assigning genes to the same group (Table S5). For example, *AtMYB24* and *AtMYB21* in *Arabidopsis* are involved in the synthesis of sesquiterpenes [23], while the homologs *MYB193* and *MYB163* may be involved in the production of terpenoids. In addition, function prediction of the putative MaMYB proteins were carried on by GO annotation (Table S6). Among these categories, most of the genes were assigned to the 'regulation of transcription, DNA-templated' and 'binding', consistent with the characteristics of TFs. Notably, seven genes (*MYB189, MYB146, MYB44, MYB29, MYB175, MYB162, MYB160*) were annotated to 'terpenoid biosynthetic process' and 'regulation of secondary metabolite biosynthetic process', which may be associated with the Terpene Biosynthesis (Table 1).

Table 1
Putative functions of MaR2R3-MYB TFs.

M.alternifolia	*Arabidopsis*	GO annotation
MYB189	ATMYB62	1. Negative regulation of isoprenoid metabolic process
		2. Terpenoid biosynthetic process
		3. Isoprenoid biosynthetic process
		4. Isoprenoid metabolic process
		5. Terpenoid metabolic process
MYB146	ATMYBS2	6. Positive regulation of isoprenoid metabolic process
MYB44	ATMYB96	7. Regulation of isoprenoid metabolic process
MYB29	ATMYB96	
MYB175	ATMYB3	8. Regulation of secondary metabolite biosynthetic process
MYB162	ATMYB35	9. Regulation of secondary metabolic process
MYB160	ATMYB118	

2.6 Combined analysis of RNAseq expression and metabolite data

Terpinen-4-ol is the primary antibacterial active substance in tea tree oil, and its biosynthesis is tightly regulated by MYB transcription factors. HS-SPME-GC-MS method was used to analyze the content of terpene-4-ol in four different ages (L1-L4, representing 18d, 60d, 1 year, and 3 year, respectively). As shown in The Fig. 4A, the terpene-4-ol content increase first and then decrease, with a peak at the L3 period, followed by a decline. Meanwhile, our transcriptomic data showed that expression of MaMYB genes at different ages was analyzed (L1-L4, representing 18d, 60d, 1 year, and 3 year, respectively). The expression of 204 MYB genes was clustered into 6 expression patterns for trend analysis (Fig. 4B, Supplementary Table 5) and 6.8% (15 out of 219 MYB genes), including MYB8, MYB9, MYB26, MYB43, MYB53, MYB62, MYB65, MYB70, MYB71, MYB87, MYB109, MYB138, MYB164, MYB165, and MYB214

were not expressed in any different ages, indicating possible expression in other organs or under special conditions. The genes in Clusters 1, Clusters 2, and Clusters 5 showed a trend of increase followed by a decrease, while the genes in Clusters 3, Clusters 4, and Clusters 6 showed the opposite trend of decrease followed by an increase. The transcriptome and metabolite correlation analyses were performed on four different ages (Pearson's correlation test, Fig. 4C, Table S7). The expression patterns of 42 MYB genes were highly correlated with the accumulation of terpene-4-ol ($r > 0.8$), including 25 positive and 17 negative correlations.

qRT-PCR

To verify the reliability of our results, six MYB genes were randomly selected for qRT-PCR. The transcriptomic data results showed a correlation between the expression levels of six genes and the qPCR results, as shown in Fig. 5.

Table 2
The primer of six MYB genes.

GENE ID	FORWARD PRIMERS (5'-3')	REVERSE PRIMERS (5'-3')
MYB13	GAGAAGCAACCTGACTCATT	AGAGATGAAGTGCTTTTGGA
MYB91	TGCAGGTTCTATTGTTGATG	ATCGCTCTGTACAACGTTTC
MYB157	GTCCAGCCTCTTCGACAT	GAAAGGGGGAAGCTTAGTAG
MYB44	CATTAAGCGTGGCAACTTCA	TCTTCTTCAAGTGGGTATTCCA
MYB66	TGAAGCCAAATAGACCCTCTG	CTGTGCTCCAGTTGAGAAGC
MYB194	GCCTGATTTGAAGAGAGGTA	TTCTTCTTGATGCTGGAGTT

Discussion

TFs are abundant in plants and very important in many physiological processes, such as cell cycle, environmental response, stress response, primary metabolism, and secondary metabolism [24–27]. This study presents the first analysis of the MYB gene family in *Melaleuca alternifolia* based on the sequenced *Melaleuca alternifolia* genome.

The number of MYB genes varies greatly among species, from fewer than 100 to more than 500 [28]. We identified 219 MaMYB genes in the *M.alternifolia* genome, which falls between the number found in *Phyllostachys edulis* (85) [29] and *Gossypium hirsutum* (524) [30].

The 2R-MYB TF family is typically the most dominant in plants, while 3R-MYB is found more frequently in animals, accounting for 63.6% of MYB genes in Arabidopsis thaliana (126 out of 198) [31] 69.7% in *Brachypodium distachyon* (85 out of 122) [32] 80.1% in *Pyrus bretschneideri* (185 out of 231) [33], 86% in *Hedychium coronarium* (219 out of 255) [10], and 96.7% in both *Lotus japonicas* and *Medicago truncatula* (216 out of 223) [34]. In the *MaMYB* gene family of *M. alternifolia*, no 4R-MYB genes were

found, but instead, one 5R-MYB was identified, which is also reported in *Brassica rapa*. Additionally, typical MYB proteins have been reported in *Morella rubra* [11]and *Prunus persica* [35]. The variations in the number and type of MYB genes suggest differences in the evolution and diversification of functions among plant species.

To gain a better understanding of the evolutionary relationships among MaMYB genes, the phylogenetic relationships of the MYB gene families in *M. alternifolia* and *Arabidopsis* were investigated (Fig. 3). The phylogenetic tree indicated a high degree of conservation between the R2R3-MYB proteins of *M. alternifolia* and *Arabidopsis*, with two species-specific subgroups observed. Subgroup 17 (S17) was found only in Arabidopsis and has been reported to regulate the biosynthesis of glucosinolate, a secondary metabolite primarily found in Brassicaceae plants [31]. However, the 2R-MYB genes that cluster into S17 were not found in either *M. alternifolia* or other plant species such as *B. vulgaris* and soybean [36]. Conversely, the genes that were grouped in subgroups 6 (S6), 9 (S9), 10 (S10), 12 (S12), and 13 (S13) in *M. alternifolia* were found to be involved in the synthesis of secondary metabolites such as lignin, xylan, cellulose, anthocyanin, proanthocyanidins, and flavonol [21][31]. Therefore, it is likely that the genes in these subgroups may be related to the synthesis of secondary metabolites, and needs further in-depth research. In addition, AtMYB21 (subgroups 8) and AtMYB24 (subgroups 8) were reported to be related to the interaction with MYC2. In the flower of *Freesia hybrida* and *Arabidopsis*, they regulate the expression of terpene synthase genes [23]. AtMYB21 has also been shown to affect terpene metabolism in *Arabidopsis*. The *M. alternifolia* genes MYB193 and MYB163, which are clustered in S8, are grouped with the MYB genes of *Arabidopsis* and are proposed to participate in terpene metabolism. However, functional verification through transformation is needed to further prove the function of MYB genes in *M. alternifolia*.

The expression patterns of MYB genes can provide important insights into their function. In different organisms and tissues, MYB gene families often exhibit large abundance differences. In this study, we confirmed the transcript levels of MaMYBs during four different ages and found that they exhibited different expression patterns, with certain MaMYBs showing the highest transcript abundance at specific stages of reproductive growth. Numerous studies have shown that MYB genes are indispensable in terpene biosynthesis[23], and in this study, we conducted a correlation analysis between MYB gene expression and terpene-4-ol data and obtained 42 genes with high correlation ($r > 0.8$), 25 of which were positively and 17 negatively correlated. Interestingly, most of the positively correlated genes were 1R-MYB, while many of the negatively correlated genes were 2R-MYB. Therefore, these results provide important clues to further investigate the functional and biological roles of MaMYBs.

Conclusions

In this study, we identified and characterized the MYB gene families in *M. alternifolia* at the genome level. The analysis of the structure, phylogeny, and expression of MaMYBs not only provides a research basis for the subsequent functional study of *M. alternifolia*-related genes and the elaboration of the molecular mechanism of tea tree oil synthesis and regulation but also has important implications for the

improvement of the oil content and reduction of production cost of terpene-4-ol type *M. alternifolia* through molecular breeding. Overall, our findings provide valuable insights into a functional exploration of the *Ma*MYB gene family, which will facilitate a better understanding of MYB transcription factors for gene regulation and will support future tag-assisted breeding efforts.

Declarations

Author contributions Xiaoning Zhang, Zhanwu Xu, Hong Yang and Hailong Liu wrote the main manuscript text, Li Liu, Heqiang Huo, Buming Liu, Lianxiang Zhong, Yufei Xiao and ling Chai prepared figures and tables; All authors have read and approved to the published version of the manuscript.

Funding This study was funded by General Program of Guangxi Natural Science Foundation (2021GXNSFAA196069); Guangxi Key Laboratory of Traditional Chinese Medicine Quality Standards Open Project (No. 202002); Guangxi Characteristic Economic forest cultivation and utilization key laboratory project (JA-22-03-01).

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

Ethical Approval Not applicable.

Conflict of interest The authors declare there are no competing interest.

References

1. Cao Y, Li K, Li Y, Zhao X, Wang L (2020) MYB Transcription Factors as Regulators of Secondary Metabolism in Plants. Biology (Basel) 9 (3): 61
2. Chanwala J, Khadanga B, Jha DK, Sandeep IS, Dey N (2023) MYB Transcription Factor Family in Pearl Millet: Genome-Wide Identification, Evolutionary Progression and Expression Analysis under Abiotic Stress and Phytohormone Treatments. Plants (Basel) 12(2): 355
3. Luo D, Mei D, Wei W, Liu J (2023) Identification and Phylogenetic Analysis of the R2R3-MYB Subfamily in *Brassica napus*. Plants 12(4): 886
4. Zhu M, Wang Q, Tu S, Ke S, Bi Y, Ahmad S, Zhang D, Liu D, Lan S (2023) Genome-Wide Identification Analysis of the R2R3-MYB Transcription Factor Family in *Cymbidium sinense* for Insights into Drought Stress Responses. International Journal of Molecular Sciences 24 (4):3235
5. Lou X, Yao S, Chen P, Wang, D, Agassin RH, Hou Y, Zhang C, Ji K (2023) Transcriptome Identification of R2R3-MYB Gene Family Members in *Pinus massoniana* and PmMYB4 Response to Drought Stress. Forests 14 (2): 410

6. Li HY, Yue YZ, Ding WJ, Chen GW, Li L, Li YL, Shi TT, Yang XL, Wang LG (2020) Genome-Wide Identification, Classification, and Expression Profiling Reveals R2R3-MYB Transcription Factors Related to Monoterpenoid Biosynthesis in *Osmanthus fragrans*. Genes (Basel) 11 (4): 353
7. Chen X, Wang P, Gu M, Lin X, Hou B, Zheng Y, Sun Y, Jin S, Ye N(2021) R2R3-MYB transcription factor family in tea plant (*Camellia sinensis*): Genome-wide characterization, phylogeny, chromosome location, structure and expression patterns. Genomics 113 (3): 1565-1578.
8. Yang X, Li J, Guo T, Guo B, Chen Z, An X (2021) Comprehensive analysis of the R2R3-MYB transcription factor gene family in *Populus trichocarpa*. *Industrial Crops and Products* 168:113614.
9. Soler M, Camargo EL, Carocha V, Cassan-Wang H, San Clemente H, Savelli B, Hefer CA, Paiva JA, Myburg AA, Grima-Pettenati J (2015) The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. New Phytol 206 (4): 1364-77.
10. Abbas F, Ke Y, Zhou Y, Yu Y, Waseem M, Ashraf U, Wang C, Wang X, Li X, Yue Y, Yu R, Fan Y (2021) Genome-Wide Analysis Reveals the Potential Role of MYB Transcription Factors in Floral Scent Formation in *Hedychium coronarium*. Front Plant Sci 12, 623742
11. Cao Y, Jia H, Xing M, Jin R, Grierson D, Gao Z, Sun C, Chen K, Xu C, Li X (2021) Genome-Wide Analysis of MYB Gene Family in Chinese Bayberry (*Morella rubra*) and Identification of Members Regulating Flavonoid Biosynthesis. Front Plant Sci 12, 691384
12. Reddy VA, Wang Q, Dhar N, Kumar N, Venkatesh PN, Rajan C, Panicker D, Sridhar V, Mao HZ, Sarojam R (2017) Spearmint R2R3-MYB transcription factor MsMYB negatively regulates monoterpene production and suppresses the expression of geranyl diphosphate synthase large subunit (MsGPPS.LSU). Plant Biotechnol J 15 (9): 1105-1119
13. Zhang X, Chen S, Zhang Y, Xiao Y, Qin Y, Li Q, Liu L, Liu B, Chai L, Yang H, Liu H (2023) Draft genome of the medicinal tea tree *Melaleuca alternifolia*. Mol Biol Rep 50 (2): 1545-1552
14. Voelker J, Shepherd M, Mauleon R (2021) A high-quality draft genome for *Melaleuca alternifolia* (tea tree): a new platform for evolutionary genomics of myrtaceous terpene-rich species. *Gigabyte* 2021:1-15
15. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2014) CDD: NCBI's conserved domain database. Nucleic Acids Res 43: D222-6 (Database issue)
16. Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* 43:D257-60 (Database issue)
17. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res,* 31 (13): 3497-500
18. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34: W369-73 (Web Server issue)
19. Bailey TL, Gribskov M (1998) Methods and statistics for combining motif match scores. J Comput Biol 5 (2): 211-221

20. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33 (7): 1870-4
21. Pu X, Yang L, Liu L, Dong X, Chen S, Chen Z, Liu G, Jia Y, Yuan W, Liu L (2020) Genome-Wide Analysis of the MYB Transcription Factor Superfamily in *Physcomitrella patens*. Int J Mol Sci 21 (3):975
22. Lalitha S (2000) Primer premier 5. Biotech Software & Internet Report: The Computer Software. J Sci 1(6): 270-272
23. Yang Z, Li Y, Gao F, Jin W, Li S, Kimani S, Yang S, Bao T, Gao X, Wang L (2020) MYB21 interacts with MYC2 to control the expression of terpene synthase genes in flowers of *Freesia hybrida* and *Arabidopsis thaliana*. J Exp Bot 71 (14): 4140-4158
24. Kranz H, Scholz K, Weisshaar B (2000) c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. The Plant Journal 21 (2): 231-235
25. Haga N, Kato K, Murase M, Araki S, Kubo M, Demura T, Suzuki K, Müller I, Voß U, Jürgens G, Ito M (2007) R1R2R3-Myb proteins positively regulate cytokinesis through activation of KNOLLE transcription in *Arabidopsis thaliana*. Development 134 (6): 1101-1110
26. Du H, Yang S, Liang Z, Feng B, Liu L, Huang Y, Tang Y (2012) Genome-wide analysis of the MYB transcription factor superfamily in soybean. BMC Plant Biology 12, 106.
27. Cao Y, Han Y, Li D, Lin Y, Cai Y (2016) MYB Transcription Factors in Chinese Pear (*Pyrus bretschneideri* Rehd.): Genome-Wide Identification, Classification, and Expression Profiling during Fruit Development. Frontiers in Plant Science 7, 577
28. Zhou L, Yarra R, Jin L, Cao H (2020) Genome-wide identification and expression analysis of MYB gene family in oil palm (*Elaeis guineensis* Jacq.) under abiotic stress conditions. *Environmental and Experimental Botany* 180: 104245
29. Yang K, Li Y, Wang S, Xu X, Sun H, Zhao H, Li X, Gao Z (2019) Genome-wide identification and expression analysis of the MYB transcription factor in moso bamboo (*Phyllostachys edulis*). *PeerJ* 6, e6242.
30. Salih H, Gong W, He S, Sun G, Sun J, Du X (2016) Genome-wide characterization and expression analysis of MYB transcription factors in Gossypium hirsutum. BMC Genet 17:129
31. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15 (10): 573-81
32. Chen S, Niu X, Guan Y, Li H (2017) Genome-Wide Analysis and Expression Profiles of the MYB Genes in *Brachypodium distachyon*. *Plant Cell Physiol* 58 (10): 1777-1788
33. Li X, Xue C, Li J, Qiao X, Li L, Yu L, Huang Y, Wu J (2016) Genome-Wide Identification, Evolution and Functional Divergence of MYB Transcription Factors in Chinese White Pear (*Pyrus bretschneideri*). Plant Cell Physiol 57 (4): 824-47
34. Wang F, Li X (2017) Genome-wide characterization and expression analysis of MYB transcription factors in *Lotus japonicas* and *Medicago truncatula*. Genes & Genomics 39 (8): 831-842

35. Zhang C, Ma R, Xu J, Yan J, Guo L, Song J, Feng R, Yu M (2018) Genome-wide identification and classification of MYB superfamily genes in peach. PLoS One 13 (6): e0199192
36. Stracke R, Holtgräwe D, Schneider J, Pucker B, Sörensen TR, Weisshaar B (2014)Genome-wide identification and characterisation of R2R3-MYB genes in sugar beet (*Beta vulgaris*). BMC Plant Biology (14): 249

Figures

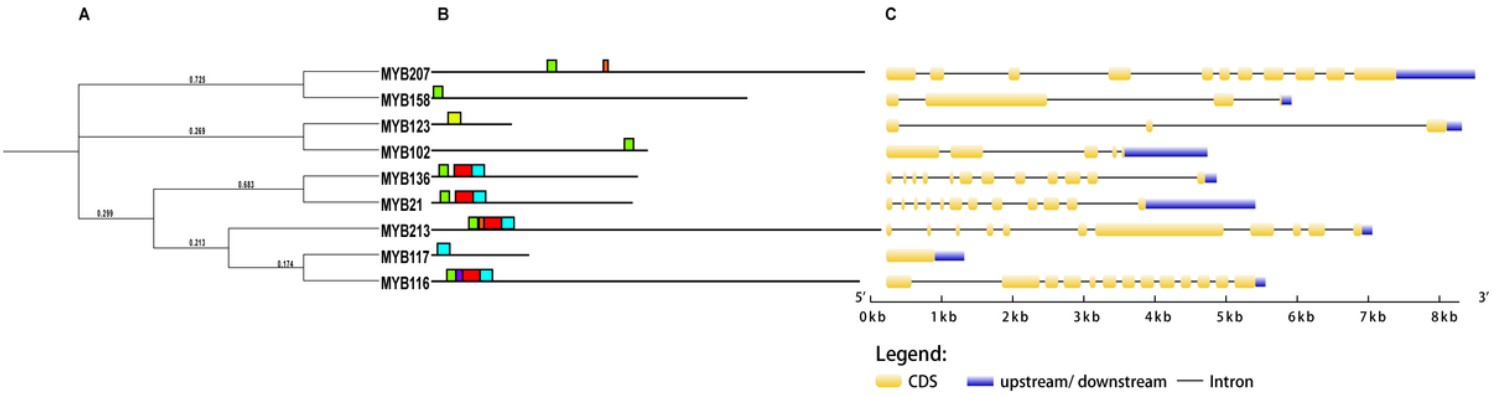

Figure 1

Analysis of phylogenetic relationships(A), conserved motifs(B), gene structure(C) of *Ma*MYBs.

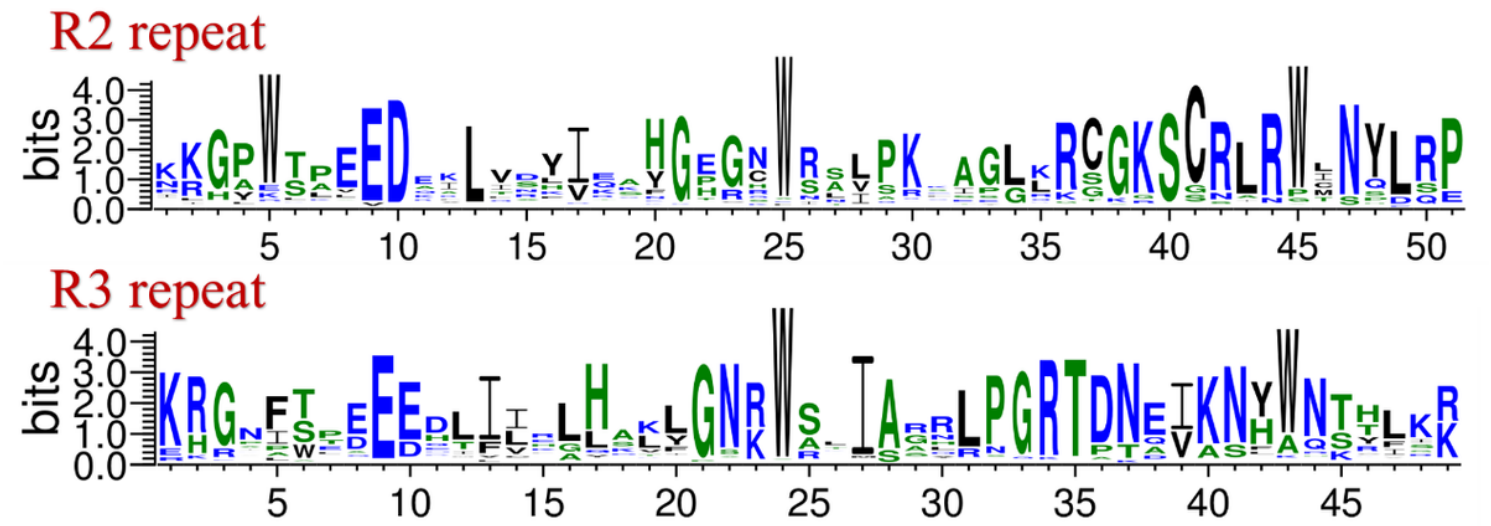

Figure 2

Conservation of R2 and R3 MYB Repeats in*M. alternifolia*. The sequence logos depict the high level of conservation across all 153 R2R3-MYB proteins. The bit score reflects the information content in every position of the sequence.

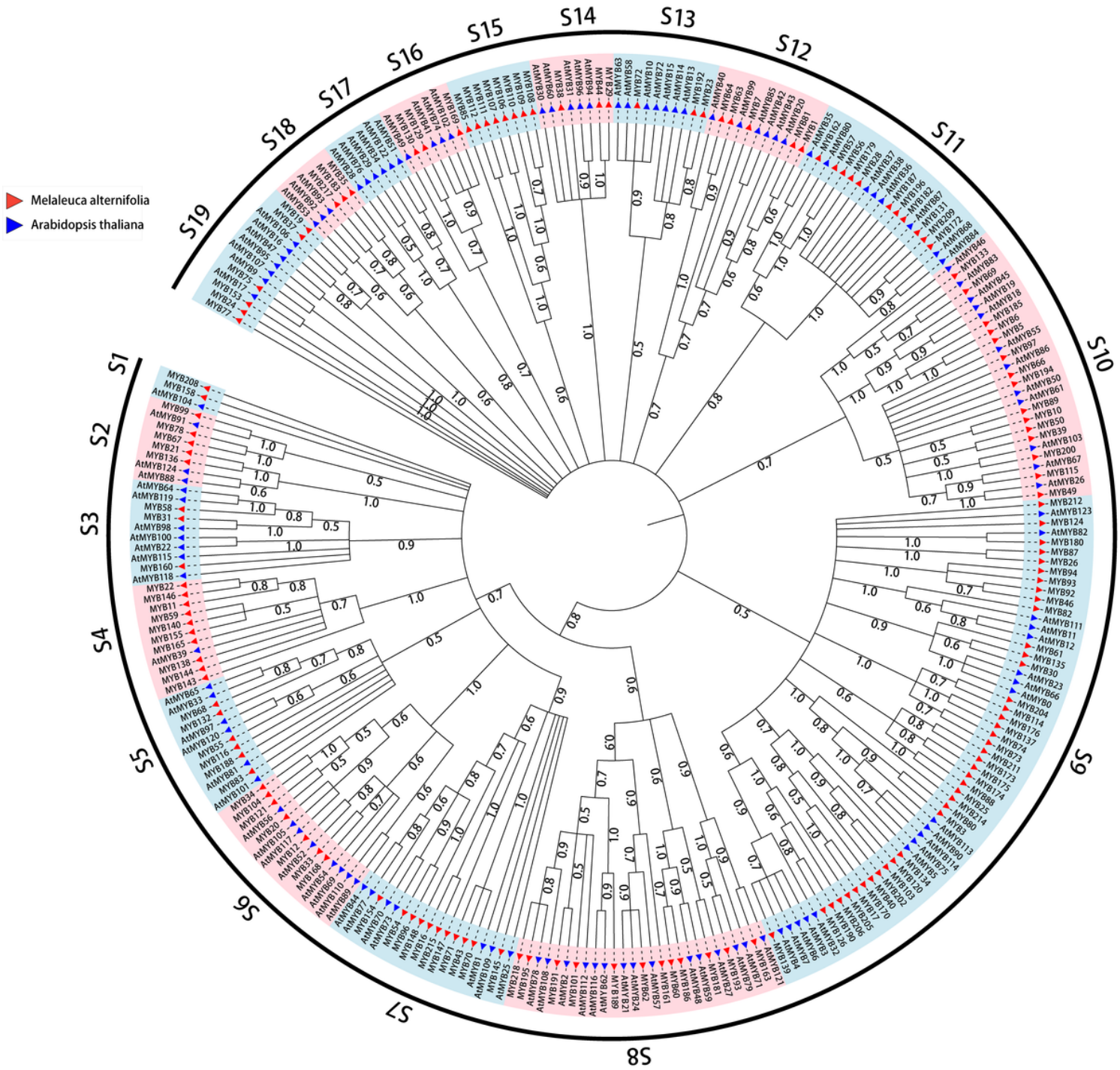

Figure 3

Phylogenetic Tree of R2R3 MYB TFs in *M. alternifolia* and *Arabidopsis*. Among them, MaMYBs are represented by a red triangle and AtMYBs are represented by a blue triangle.

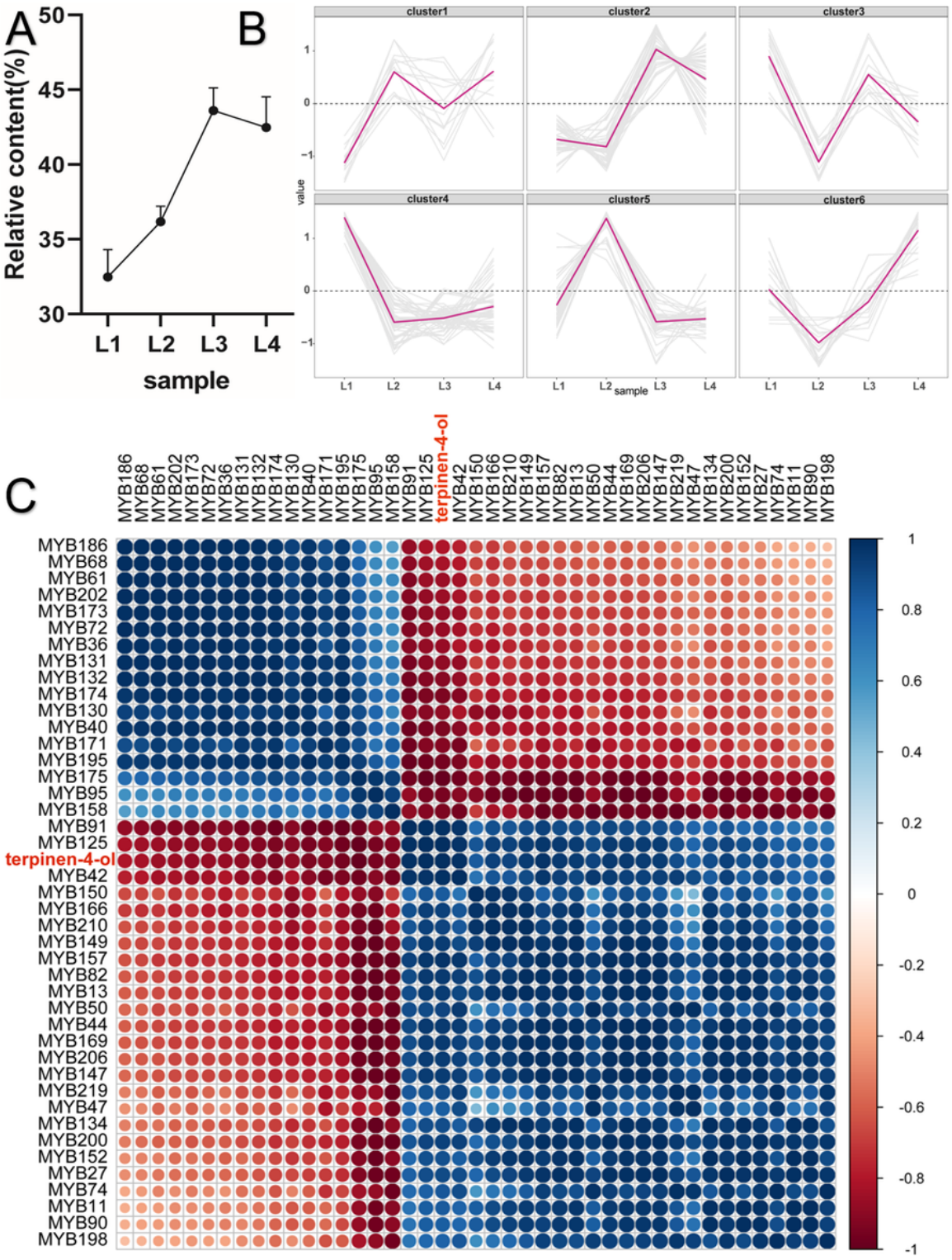

Figure 4

Identification of MYB TFs involved in regulating terpene-4-ol biosynthesis. (a) Relative content of terpene-4-ol in *M. alternifolia*. (b) Expression pattern of *MaMYB* Genes at different ages. (c) Co-expression analysis of candidate MYBs and terpene-4-ol. L1-L4, representing 18d, 60d, 1yr, and 3yr, respectively

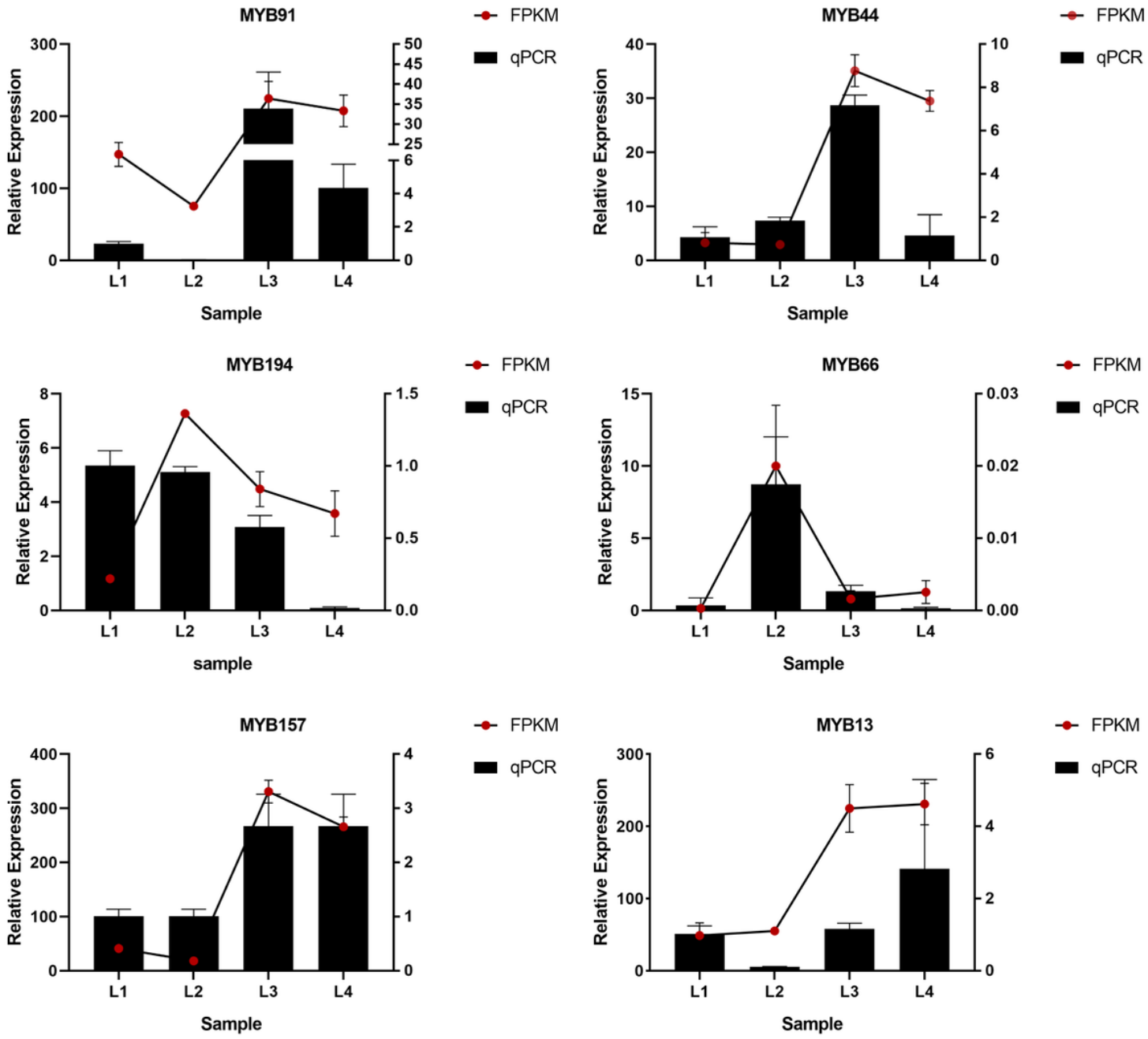

Figure 5

Expression levels of six genes and the qPCR results. L1-L4, representing 18d, 60d, 1yr, and 3yr, respectively

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryfiles.rar