Large-scale analysis of putative Euphorbiaceae R2R3-MYB transcription factors identifies a MYB involved in seed oil biosynthesis

Yunpeng Cao[1,2,3*], Tingting Fan[3], Lihu Wang[4], Lin Zhang[2,5] and Yanli Li[3*]

Abstract

Background MYB transcription factors are widely distributed in the plant kingdom and play key roles in regulatory networks governing plant metabolism and biochemical and physiological processes.

Results Here, we first determined the R2R3-MYB genes in five Euphorbiaceae genomes. The three Trp (W) residues from the first MYB domain (R2) were absolutely conserved, whereas the first W residue from the second MYB domain (R3) was preferentially mutated. The R2R3-MYBs were clustered into 48 functional subfamilies, of which 34 had both R2R3-MYBs of Euphorbiaceae species and AtMYBs, and four contained only Euphorbiaceae R2R3-MYBs. The whole-genome duplication (WGD) and/or segmental duplication (SD) played key roles in the expansion of the R2R3-MYB family. Unlike paralogous R2R3-MYB family members, orthologous R2R3-MYB members contained a higher selective pressure and were subject to a constrained evolutionary rate. *VfMYB36* was specifically expressed in fruit, and its trend was consistent with the change in oil content, indicating that it might be involved in oil biosynthesis. Overexpression experiments showed that *VfMYB36* could significantly provide linolenic acid (C18:3) content, which eventually led to a significant increase in oil content.

Conclusion Our study first provides insight into understanding the evolution and expression of R2R3-MYBs in Euphorbiaceae species, and also provides a target for the production of biomass diesel and a convenient way for breeding germplasm resources with high linolenic acid content in the future.

Keywords Euphorbiaceae, Evolutionary rate, MYB, Gene feature, Paralog, Ortholog

*Correspondence:
Yunpeng Cao
caoyunpeng@wbgcas.cn; xfcypeng@126.com
Yanli Li
liyanli17@126.com
[1]CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, 430074 Wuhan, China
[2]School of Health and Nursing, Wuchang University of Technology, Wuhan, China
[3]Forestry College, Central South University of Forestry and Technology, 410004 Changsha, Hunan, China
[4]College of Landscape and Ecological Engineering, Hebei University of Engineering, 056009 Handan, China
[5]School of Basic Medical Sciences, Hubei University of Chinese Medicine, 430065 Wuhan, China

Background

Transcription factors generally consist of at least four distinct domains, i.e. a nuclear localization signal, a DNA-binding domain, an oligomerization site, and a protein-protein binding domain [1, 2]. Many important biological processes are controlled by transcription factors that regulate gene expression [3–6]. MYB is considered to be the largest family of transcription factors in plants, which is involved in the regulation of many biological processes, such as development, differentiation, metabolism, and defense [7–13]. The common feature of MYB family proteins is that they all contain a highly conserved MYB DNA-binding domain [9, 11, 14]. This domain consists of three α-helices, of which the second and third helices form a helix-turn-helix (HTH) structure, and further studies have shown that the third helix directly interacts with the main grooves of target DNA [9, 11, 14, 15]. In addition, some studies have found that there were highly divergent activation regions at the C-terminus of the *MYB* genes, which allow for extensive regulation in different biological processes [16, 17]. The *MYB* gene family can be divided into four subgroups according to the number of adjacent MYB repeats, namely 1R-MYB (MYB-related proteins), 2R-MYB (R2R3-MYB proteins), 3R-MYB (R1R2R3-MYB proteins) and 4R-MYB (4R-like MYB proteins), which have one, two, three and four MYB DNA-binding domain repeats, respectively [14, 18]. All four types of *MYBs* have been identified, of which 1R-MYB and R2R3-MYB are the most frequent in the plant kingdom [19]. Phylogenetic studies have shown that R2R3-MYBs likely evolved from R1R2R3-MYBs by losing the R1 repeat, or from R1-MYBs by duplication R1 repeat [8, 14, 19, 20].

The first identified *MYB*, *v-MYB*, was isolated from the avian myeloblastosis virus (AMV) [21]. *C1* (*NP_001106010.1*) in *Zea mays*, the first MYB identified in plants, regulates the biosynthesis of anthocyanin [22]. Based on the analysis of molecular and genetic experiments, an increasing number of R2R3-MYBs have been characterized and studied in many plants, such as pear (*Pyrus bretschneideri*), apple (*Malus domestic*), grapevine (*Vitis vinifera*), petunia (*Petunia hybrida*), rice (*Oryza sativa*), poplar (*Populus tremuloides*), maize (*Zea may*) and *Arabidopsis thaliana* [7, 18, 20, 23]. For example, *Lycopersicon esculentum* and *A. thaliana MYB15* plays a negative regulatory role in the CBF pathway in response to cold stress [24, 25]. *PbMYB25* and *PbMYB52* from *P. bretschneideri* are considered to be regulators of lignin biosynthesis during fruit development [7]. The *O. sativa OsMYB2* gene has been shown to act as a potentially important regulator in tolerance to dehydration and salt and cold stresses [26]. *PtrMYB3* and *PtrMYB20* from *P. tremuloides* control the lignin biosynthesis by upregulating the monolignol pathway [27]. In 2015, Li et al.

reported that *Populus tomentosa PtoMYB92* as a transcriptional activator regulates the biosynthesis of lignin during secondary cell wall formation [28].

Members of R2R3-MYB family are involved in multiple plant-specific processes, raising the hypothesis that their expansion may be the cause of plant evolutionary diversity. The R2R3-MYBs have been analyzed and identified from several sequenced plants, such as pear, apple, wheat, rice, poplar, *Beta vulgaris*, rice and *A. thaliana*. Most studies were focused on MYB family identification and gene structure, phylogenetic, conserved domain, and chromosomal localization analyses. Additionally, the functions of some R2R3-MYBs have been analyzed by molecular biological experiments in model plants [4, 14]. However, the predicted functions for each subclade of the R2R3-MYB family and potential evolutionary mechanisms in the R2R3-MYB family are still uncertain.

Euphorbiaceae species including tung tree (*Vernicia fordii*), rubber tree (*Hevea brasiliensis*), physic nut (*Jatropha curcas*), cassava (*Manihot esculenta*), and castor oil plant (*Ricinus communis*) are economically important plants. These Euphorbiaceae genomes have been sequenced and published. Although the in-depth study of the R2R3-MYB family has been carried out in some model plants including poplar and *A. thalian*a, systematic analysis of the R2R3-MYB family was lacking in Euphorbiaceae. Here, we carried out a genome-wide analysis of the R2R3-MYB family to identify R2R3-MYB members and investigated family evolutionary patterns, selection pressures, and gene duplication types in five Euphorbiaceae genomes. Additionally, compared with the other four species of the family Euphorbiaceae, *V. fordii* produces tung oil containing a high proportion (~80%) of eleostearic acid and has more than 1,000 years of cultivation history in China as an ornamental plant or for the production of tung oil [29–31]. Therefore, we selected the R2R3-MYB gene from *V. fordii* for further expression analysis, aiming at mining candidate genes related to oil synthesis. Our study might help us to understand the evolution and expansion of the R2R3-MYBs in Euphorbiaceae genomes based on the comparative analysis, and also provide valuable candidate *VfMYB36* for improving oil production by marker-assisted breeding in *V. fordii*.

Results and discussion

R2R3-MYBs in Euphorbiaceae

It is well known that R2R3-MYBs can control many plant-specific processes, such as abiotic stress tolerance, plant development, disease resistance, and hormone signal transduction [8–10, 14, 17, 23, 32]. As an important family of transcription factors, R2R3-MYB members are widely present in plants, such as 129 R2R3-MYB genes in pear [7], 222 in apple [33], 192 in Poplar [34], 102 in rice [35], 55 in cucumber [36], 244 in soybean [37], 118

in grape [38], 157 in maize [23] and 124 in *A. thaliana* [18]. To obtain the R2R3-MYBs, we searched the entire genomes of *R. communis, J. curcas, M. esculenta, H. brasiliensis,* and *V. fordii* that encode proteins containing MYB domain (PF000249) by HMMER v3.2.1. At the same time, the *A. thaliana* R2R3-MYBs were also used as queries to perform BlastP against Euphorbiaceae genomes. The SMART, Pfam, and InterPro databases were used to examine the presence or completeness of MYB domain (PF000249). Finally, we identified 650 MYB genes that encode proteins containing at least two MYB domains (Fig. 1A and Table S1). Among them, 80 typical R2R3-MYBs, four 3R-MYBs, and two 4R-MYBs were scanned in *J. curcas*, 97 typical R2R3-MYBs, four

3R-MYBs, and one 4R-MYB were identified in *V. fordii*, 178 typical R2R3-MYBs, four 3R-MYBs and one 4R-MYB were identified in *M. esculenta*, 210 typical R2R3-MYBs, and eight 3R-MYBs were found in *H. brasiliensis*, and 57 typical R2R3-MYBs, three 3R-MYBs and one 4R-MYB were found in *R. communis*. We named *MYBs* from *J. curcas, R. communis, M. esculenta, H. brasiliensis* and *V. fordii* based on the order of these gene locations among the chromosomes. Here, only the primary transcript was retained when the alternative splicing events were detected for the *MYB* genes. A comparison of the number of R2R3-MYBs among these Euphorbiaceae species suggested that both *M. esculenta* and *H. brasiliensis* had
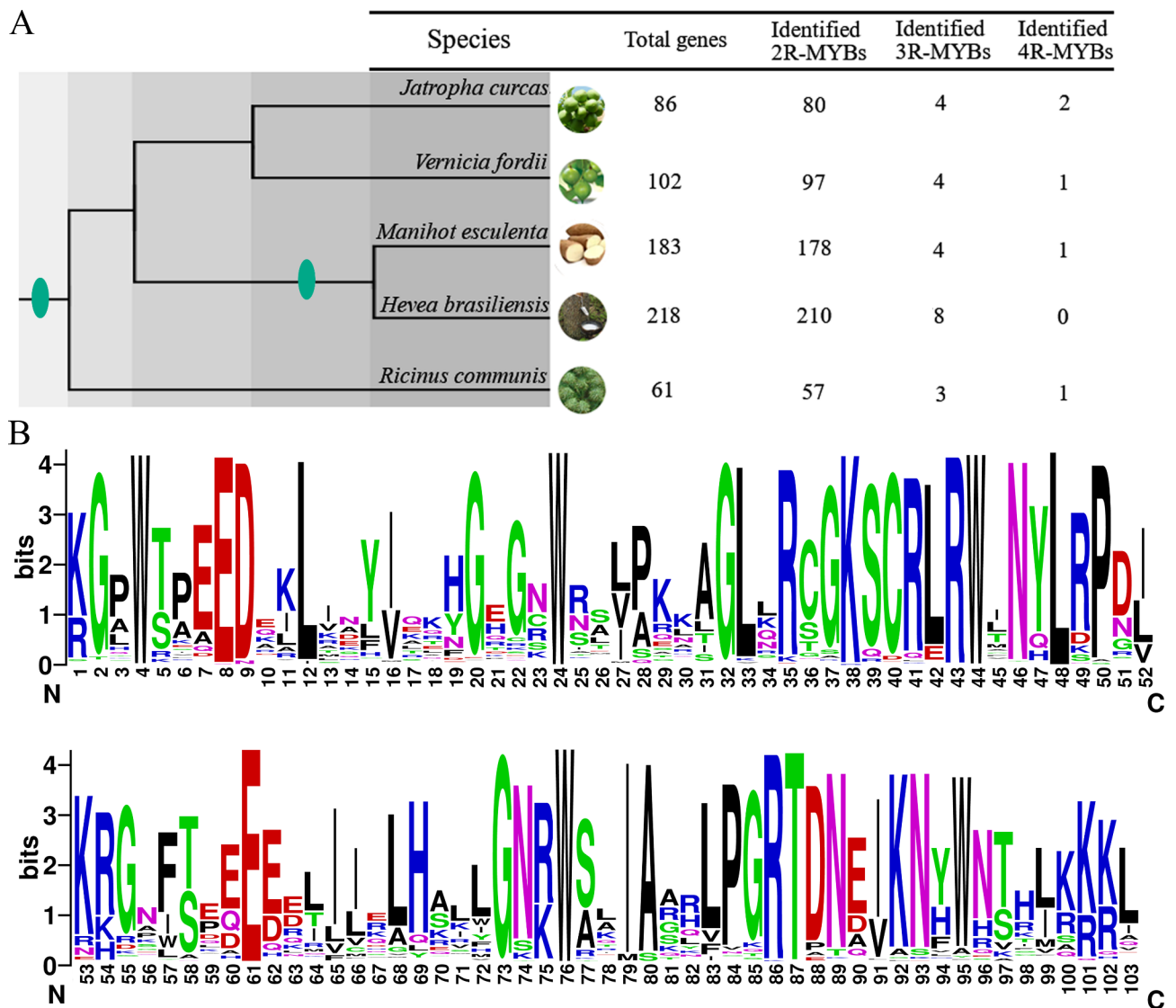

Species	Total genes	Identified 2R-MYBs	Identified 3R-MYBs	Identified 4R-MYBs
Jatropha curcas	86	80	4	2
Vernicia fordii	102	97	4	1
Manihot esculenta	183	178	4	1
Hevea brasiliensis	218	210	8	0
Ricinus communis	61	57	3	1

Fig. 1 The species tree (**A**) and domain composition (**B**) of five Euphorbiaceae genomes for identified *MYB* genes. **A**, the species tree was obtained from NCBI CommonTree (https://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi) and Timetree (http://www.timetree.org/). The whole genome duplication (WGD) was indicated by the green oval. The numbers on the right indicated the number of corresponding genes in each genome. **B**, the logos of R2 and R3 domains were generated from the multiple alignment analysis of all R2R3-MYBs

the greatest number of R2R3-MYBs, while *R. communis* possessed the fewest R2R3-MYBs (Fig. 1A).

The frequency of amino acid sequences in the R2 and R3 repeats was investigated by examining the characteristics of the conserved domain of R2R3-MYB (Fig. 1B). The results showed that the MYB domains are highly conserved and they have few insertions or deletions other than those containing basic residues, which were consistent with those of cucumber, pear, soybean, sugar beet, poplar, grape, maize, rice, *Cyclocarya paliurus*, and *A. thaliana* [7, 15, 23, 39]. The R2R3-MYB domain that contained common sequences was [-W-(X19)-W-(X19)-W-. … .-F/I-(X18)-W-(X18)-W-], which formed the primary structure of R2 and R3 repeats (Fig. 1B). Subsequently, we found that the first tryptophan (W) residue of the R3 repeat was usually replaced by hydrophobic amino acids, which was a common phenomenon in plant R2R3-MYBs [14, 18, 23]. It has been observed that the substitution of hydrophobic amino acids for tryptophan residues in animal MYB domains does not result in a significant loss of DNA-binding activity [40]. This observation suggests that the hydrophobic residues can substitute for tryptophan and maintain the function of the MYB domain, at least in terms of DNA-binding. Therefore, this change in Euphorbiaceae species might also have little or no effect on the DNA-binding activity of MYB domains. Besides the highly conserved W residues, G^{32}, R^{43}, and L^{48} in the R2 repeat, E^{61}, G^{73}, and T^{87} in the R3 repeat were also completely conserved in these MYBs from five Euphorbiaceae species. In summary, as reported in plants such as pear, apple, rice and *A. thaliana* [7, 18, 33, 35], the conserved amino acid residues mainly occurred in the second and third conserved tryptophan residues of each MYB domain (i.e., R2 or R3).

Synteny tests suggest the origin and expansion of the R2R3-MYBs in Euphorbiaceae

In plants, the common view of most studies is that R2R3-MYBs evolved from the R1R2R3-MYBs by losing a R1 domain and subsequent expansion of this gene family [41, 42]. Despite recent diversification events, several gene duplications, including tandem duplication (TD), whole-genome duplication (WGD) or segmental duplication (SD), and rearrangements, can drive the evolution of gene family members [43, 44]. To further understand which duplications were involved in the *MYBs*, we analyzed the synteny regions for this gene family among Euphorbiaceae genomes. Subsequently, all *MYB* members were assigned to five distinct duplication events: TD, WGD or SD, dispersed duplication (DD), proximal duplication (PD), and singleton. Results presented that DD and WGD or SD events were the main modes in these tested Euphorbiaceae genomes, and the TD event also contributed to the expansion of the *MYB* family.

However, we noted that the *RcMYBs* and *VfMYBs* were not involved in the PD event. These results showed that the *MYB* members were involved in different duplication models, and DD and WGD or SD had a relatively high frequency in five tested Euphorbiaceae genomes (Figure S1). Previous investigations have demonstrated that the WGD or SD event is the primary driver of *MYBs* expansion in soybean and pear [37, 45]. Combining previous reports with our study, we argue that WGD or SD events should be the most important contributors to the expansion of *MYB* members in plants.

To infer the evolutionary mechanism of the *MYB* family in five Euphorbiaceae genomes, a process similar to PGDD [46] for scanning the synteny blocks was used in this study. Totally, 25, 4, 123, 41 and 13 segmentally duplicated *MYB* pairs were identified in the *V. fordii*, *R. communis*, *M. esculenta*, *H. brasiliensis* and *J. curcas* genomes (Table S2). The homologs in the 100 kb flanking on each side of the anchored *MYBs* were also searched (Table S3). For example, in duplicated *VfMYB50-VfMYB81* pair, the number of genes in a 200 Kb window is up to 92, while the homologs in the 100 kb flanking each side of *VfMYB47-VfMYB96* reached 9 pairs in *V. fordii* genome (Fig. 2 and Table S2). The conserved synteny in R2R3-MYBs was also observed from other four Euphorbiaceae genomes, and these results suggested that the large-scale duplications played important roles in the expansion of the R2R3-MYB family in these tested genomes.

R. communis, *H. brasiliensis*, *J. curcas*, *M. esculenta* and *V. fordii* shared an ancient WGD event (i.e. γ-triplication; Ks ~1.08–1.48), while both of *M. esculent* and *H. brasiliensis* also experienced an additional recent WGD event (i.e. ~15.3 Myr ago; Ks ~0.2–0.37) [47, 48]. Thus, the synonymous substitutions per site (Ks) was further calculated to determine the evolutionary dates of WGDs or SDs associated with *MYB* genes from these five Euphorbiaceae genomes (Table S4), suggesting that *MYB* duplicated genes may be derived from the ancient WGD event and/or recent WGD event. At the same time, we also noted that the recent WGD event was attributed to the expansion of R2R3-MYB family members in both the *H. brasiliensis* and *M. esculenta* than the ancient WGD event. These results explained that there are far more members of the R2R3-MYB family in *M. esculenta* and *H. brasiliensis* than in the other three Euphorbiaceae, and also indicated that R2R3-MYB is an ancient family whose members expanded with the recent WGD.

Phylogenetic relationships among five Euphorbiaceae species

To test the phylogeny within the R2R3-MYB family, all identified R2R3-MYBs from the five Euphorbiaceae genomes were aligned with 126 *AtMYB* genes. *AtMYB* family has been determined in *A. thaliana* [18], and the
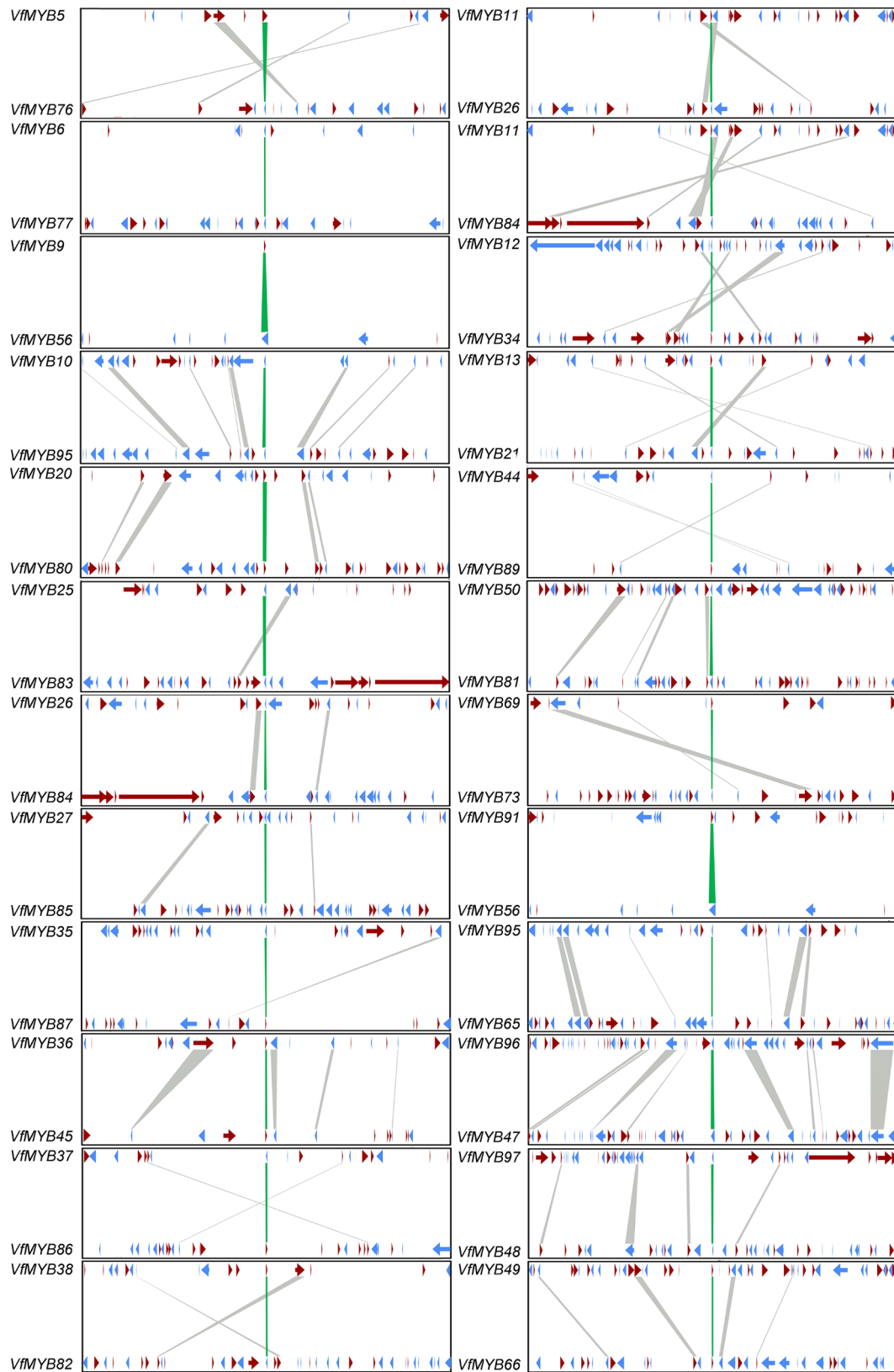
Fig. 2 Synteny blocks of *VfMYBs* in *V. fordii*. The position of the *VfMYB* was highlighted in green and the relative positions of the other flanking genes were determined by the position of the *VfMYB*. The lines indicated that the genes have synteny relationships, and the arrows represented the transcriptional orientations of these genes

family members were involved in various plant biochemical and physiological processes [8, 10, 14]. The previous results for *AtMYB* genes were also considered in our subfamily classification of the R2R3-MYBs from the five Euphorbiaceae to confirm the result of our phylogenetic reconstruction. The phylogenetic allowed us to classify the R2R3-MYBs from five Euphorbiaceae into 48 clades (C1–C48) being supported by collinearity analysis (Fig. 3 and Figure S2), indicating that these genes in each clade may possess similar functions and evolve from the same

duplicated event. According to the conservation of the amino acid motifs and DNA-binding domain, Stracke et al. (2001) have divided the *A. thaliana* R2R3-MYBs into 25 subgroups [18]. Compared with our classification of the R2R3-MYBs from Euphorbiaceae genomes, the subgroup (S) classification of *A. thaliana* was also applied to these species. Here, we clustered 48 clades, including 14 species-specific R2R3-MYB clades among the five Euphorbiaceae genomes for which there were no representatives in *A. thaliana*, indicating these genes
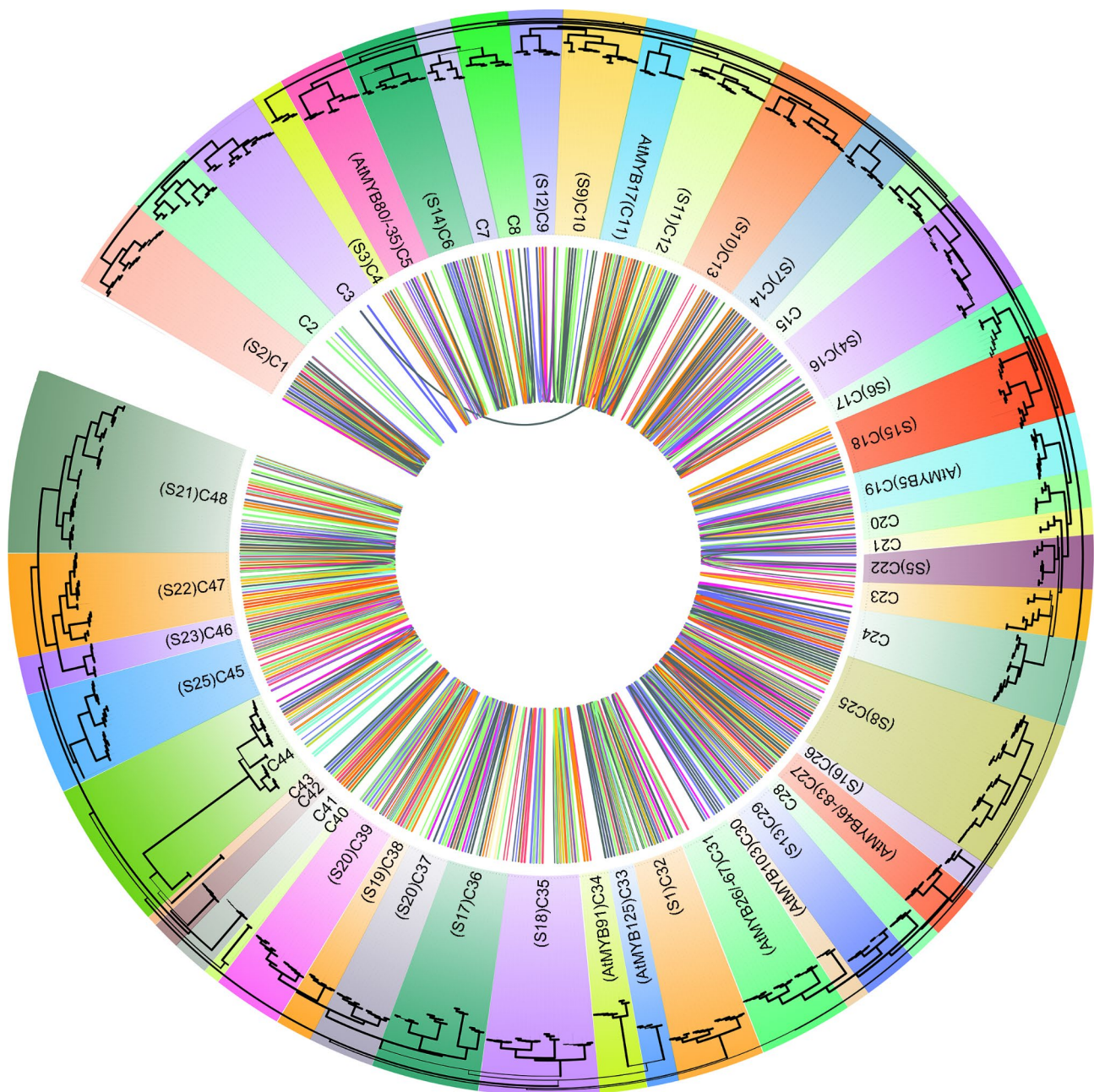

Fig. 3 Synteny and phylogenetic tree analysis among these five Euphorbiaceae species. Synteny relationships were indicated by different links among these five Euphorbiaceae species. The ML tree was constructed using IQ-tree with the best substitution model

might play special roles that were either gained in these tested five species or lost in *A. thaliana* after divergence from the last common ancestor (Fig. 3). We also noted that the remaining *A. thaliana* R2R3-MYBs have not emerged in any subgroup in *A. thaliana* (Fig. 3). These *A. thaliana* R2R3-MYB genes were divided into eight subfamilies with Euphorbiaceae R2R3-MYBs. For instance, the subfamily C23 contained *AtMYB26*, *AtMYB67* and *10 RcMYBs, 13 VfMYBs, 2 HbMYBs*, indicating that the functional roles of these gene family members further differentiated, and the R2R3-MYB family further expanded in plants.

The phylogenetic tree suggested that R2R3-MYB family members from the five Euphorbiaceae species were unevenly distributed within given subfamilies. For example, the subfamily C23 included two, three, and four R2R3-MYB genes from *R. communis, J. curcas*, and *V. fordii*, while it had eight and nine R2R3-MYB genes from *M. esculenta* and *H. brasiliensis*, respectively (Fig. 3). These phenomena were also consistent with our knowledge that the *J. curcas, R. communis*, and *V. fordii* have not undergone the recent WGD event shared by *H. brasiliensis* and *M. esculenta* during evolution. Most of these given subfamilies had R2R3-MYB genes from the five Euphorbiaceae species, indicating that these MYBs within the given subfamilies might have already existed in their common ancestral species. We noted that several subfamilies were scanned only in several particular genomes. For example, two (C3 and C7) subfamilies were present in *H. brasiliensis* and *M. esculenta*, but not in *J. curcas* and *R. communis*. These results indicated that these genes might play specialized roles that were either acquired in *M. esculenta* and *H. brasiliensis* or lost in *R. communis* and *J. curcas*. Remarkably, four *VfMYB* genes showed ambiguous placement between different phylogenetic trees or did not divide into any of the subfamilies, indicating that these R2R3-*MYB* genes might have specialized functions that were expanded or acquired in these species during evolution.

Evolutionary rates and gene features in homologs

In our study, we excluded these homologous genes, if the Ka value was nearly 0, and the Ks was more than 2 or less than 0.01, because a high Ks value suggested potential sequence saturation, and low sequence divergence could lead to unknown results [49]. Subsequently, we found that the average Ka, Ks, and Ka/Ks values of the paralogs were 0.29, 0.35, and 0.86, respectively. The average Ka, Ks, and Ka/Ks values of the orthologs were 0.33, 0.43, and, 0.65, respectively (Fig. 4A and Table S4). A comparison of the orthologs and paralogs suggested that the average Ka and Ka/Ks values of the paralogs were lower than that of the orthologs, but the Ks for paralogs were higher than that of the orthologs. Our data suggested that the

orthologs instead of paralogs had a higher selective pressure and were subject to a constrained evolutionary rate. We also performed correlation analyses to explore the relationship between Ks, Ka, and Ka/Ks values of both orthologs and paralogs (Fig. 4B). These results revealed that the Ks values were negatively correlated with the Ka/Ks in paralogs (r= -0.05, p<0.5), but this value was positively correlated with the Ka/Ks in orthologs (r=0.352, p<0.01). The positive correlation between Ka and Ks was detected in orthologs and paralogs (ortholog: r=0.607, P<0.01; paralog: r=0.464, P<0.01). At the same time, a positive correlation between Ka and Ka/Ks was found in orthologs and paralogs (ortholog: r=0.691, P<0.01; paralog: r=0.912, P<0.01). Our study suggested that the Ka/Ks values were mainly affected by Ka, but also affected to some extent by Ks in paralogs and orthologs.

Previous studies have confirmed that there is a correlation between gene features and evolutionary rate, but the correlation is different in various organisms. Here, we found that the Ks, Ka, and Ks/Ks were negatively correlated with GC1, GC2, GC3, and Fop in orthologs and paralogs, respectively (Table S5). In orthologs, the Ka, Ks, and Ka/Ks were significantly negatively correlated with GC3 and GC1, respectively, except for Ks with GC3. In paralogs, the Ka, Ks, and Ks/Ks were also negatively correlated with GC1 and GC3, but not significantly with GC3. Our data suggested that GC1 influenced both the Ks and Ka/Ks, and the GC2, GC3, and Fop also affected these three values in paralogs and orthologs to some extent.

The expression patterns of *V. fordii* R2R3-MYBs in different tissues

The *MYB* family members are thought to play key roles in many aspects, such as stress responses, and plant growth and development [7, 8, 14, 50, 51]. To get more about the possible functions of *MYB* genes in different tissues or during fruit development, we used previously publicly available RNA-seq data from our lab to investigate the transcriptional abundance of *VfMYB* genes in *V. fordii* [48]. As shown in Fig. 5, the expression patterns of *VfMYBs* were tissue-specific in *V. fordii*. Among them, most *VfMYBs* showed tissue-specific expression patterns, such as *VfMYB20, VfMYB47, VfMYB73, VfMYB96, VfMYB80, VfMYB17*, and *VfMYB23* in leaves, *VfMYB2* and *VfMYB94* in roots, indicating that these meristem-specific *VfMYBs* might play key roles in cell fate specification and organ formation. These results highlighted the functions of *VfMYBs* in meristem tissues and provided a basis for further study of functions. In our laboratory, we investigated the patterns of oil synthesis in *V. fordii* that the oil content was relatively low during the initial stages of fruit development (10 weeks after flowering: WAF), but then rose rapidly until the middle stages (20–25
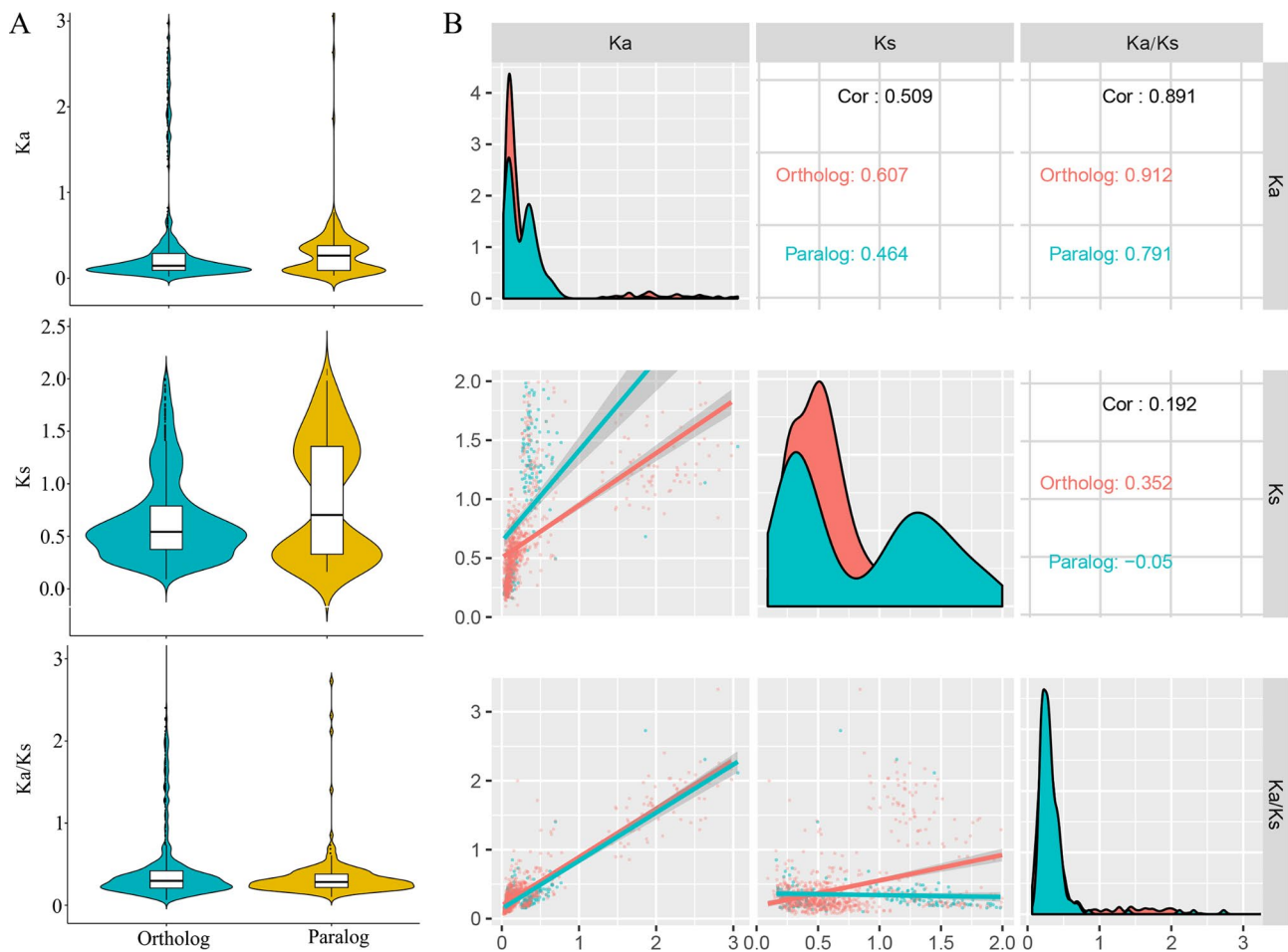
Fig. 4 Comparison of substitution rates between orthologous and paralogous *MYB* genes among these five Euphorbiaceae species. **A**, Comparative analysis of Ka, Ks and Ka/Ks of paralogous and orthologous *MYB* genes. **B**, Correlation analysis between Ka, Ks and Ka/Ks of paralogous and orthologous *MYB* genes

WAF) [48]. Finally, the oil content gradually decreased as the fruit matured [48]. This pattern was positively correlated with the expression pattern of the *VfMYB36* from subgroup C48 (Fig. 3) in developing seeds. The expression of this gene increased during the early to middle stages and then decreased as the seed matured (Fig. 6), suggesting *VfMYB36* might regulate seed-related traits. Indeed, this expressed gene *VfMYB36* (subgroup C48) was an orthologous gene of *A. thaliana AtMYB89*, which can regulate seed oil accumulation [52]. Therefore, we selected *VfMYB36* as the potential target for further study on the regulation of seed oil accumulation.

VfMYB36 involved in seed oil accumulation
The nuclear localization of transcription factors plays a key role in their performance of regulatory functions [53]. The nuclear entry process that requires one or more nuclear localization signals is affected by many factors, such as developmental stage, cell cycle, and environmental pressure. MYBs are usually located in the nucleus.

In our study, we constructed the pEarleyGate104-VfMYB36 for examining the subcellular localization of VfMYB36 (Fig. 7A). We found that the green fluorescent signal from the expressed fusion of pEarleyGate104-VfMYB36 was observed specifically in the nucleus. However, green fluorescent from the control vector was distributed throughout the whole cell (Fig. 7A).

In China, the seeds from *V. fordii* have been used to produce tung oil for more than a thousand years [54]. *MYB* genes have been confirmed to play important regulatory roles in the biosynthesis of seed oil [52, 55]. In our study, the *VfMYB36* was further selected to overexpress *A. thaliana* to study its role in oil biosynthesis. As shown in Fig. 7B, there were no significant differences in seed morphology and shape between overexpressed plants and wild-type plants, such as seed coat color, seed size, and dry weight (Fig. 7B). However, the overexpression of *VfMYB36* in *A. thaliana* resulted in a significant increase in the oil content (higher by 42.29%) of *A. thaliana* seeds (Fig. 7C). We further carried out the GC analysis
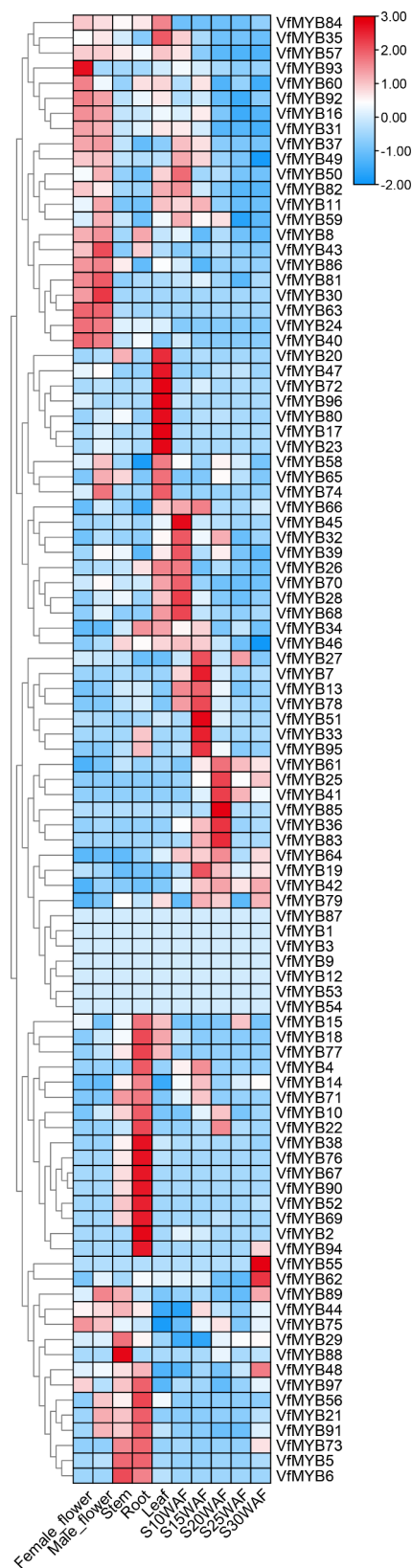
Fig. 5 Expression levels of all *VfMYBs* in different tissues. Red and blue indicated high and low expression levels, respectively

to survey whether the overexpression of *VfMYB36* in transgenic *A. thaliana* causes any qualitative change in fatty acids types. Remarkably, we found that the linolenic acid C18:3 was increased by 49.61% in transgenic lines expressing *VfMYB36* compared with WT plants. C18:3 is a basic substance that constitutes cell membranes and biological enzymes, and is also an essential nutrient for people [56]. *AtMYB89*, an orthologous gene *VfMYB36*, overexpression decreases seed oil accumulation in *A. thaliana* [52], but interestingly *VfMYB36* overexpression in *A. thaliana* leads to increased linolenic acid (C18:3). To further understand the reasons for this difference, we selected six genes (*AtWRI, AtENO1, AtBCCP1, AtKAS1, AtKCS11*, and *AtPAL2*) related to oil biosynthesis that have been reported before [52], and detected their expression patterns between transgenic and WT plants. As shown in Figure S3, we found that the overexpression of *VfMYB36* can enhance the expression of these six genes, while the overexpression of *A. thaliana AtMYB89* mainly repressed the expression of these genes [52], possibly due to functional divergence of *MYB* homologs that evolved from a common ancestor. Therefore, our research not only provided a target for the production of biomass diesel, but also provided a convenient way for breeding germplasm resources with high linolenic acid content.

Conclusion

Here, we identified the R2R3-MYBs in five Euphorbiaceae species and compared the gene numbers among these genomes. The phylogenetic analysis suggested that the majority of subfamilies had R2R3-MYBs from Euphorbiaceae species and *A. thaliana*, implying that the functions of most R2R3-MYBs are conserved and that genes within a given subfamily share recent common evolutionary origins during evolution. The paralogous and orthologous R2R3-MYB genes evolve at different rates. WGD or SD MYB genes exhibit gene expression divergence under normal growth conditions in *V. fordii*. Combining expression patterns with transgene analysis, our study provided a target for the production of biomass diesel and a convenient way for breeding germplasm resources with high linolenic acid content in the future.

Methods

Identification of R2R3-MYBs in Euphorbiaceae

Genome resources of *R. communis* were downloaded from the Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) database, the *J. curcas* was derived from Jatropha genome database (http://www.kazusa.or.jp/jatropha/), the *M. esculenta* was obtained from Ensembl database (http://plants.ensembl.org/Manihot_esculenta/Info/Index), the *H. brasiliensis* was downloaded from NCBI database (https://www.ncbi.nlm.nih.gov/). The
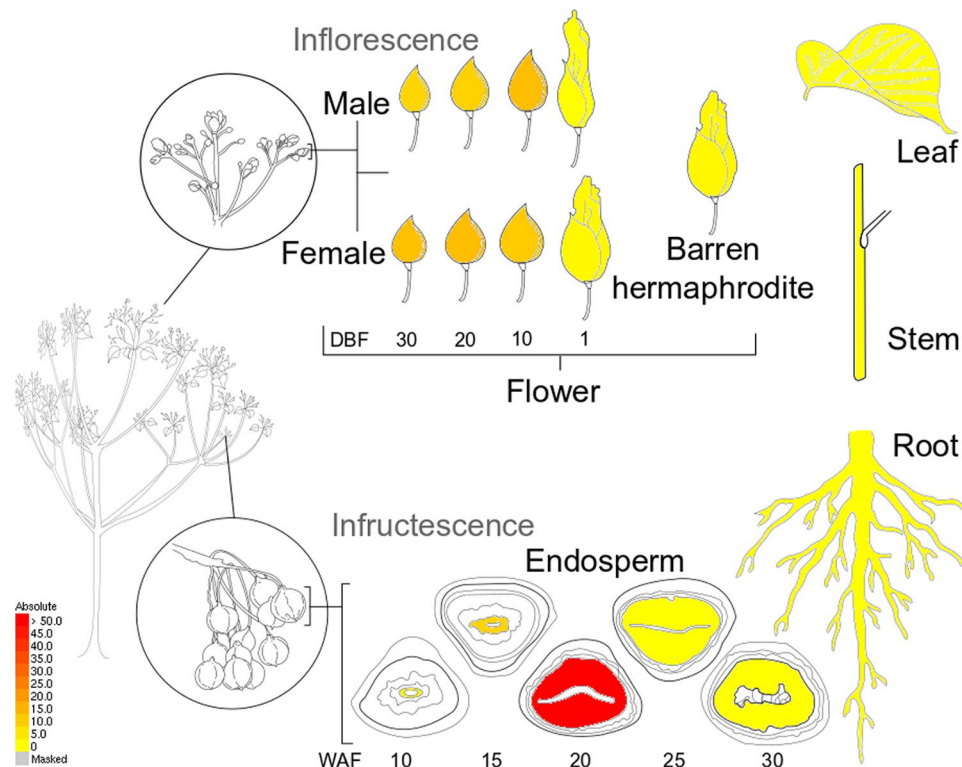
Fig. 6 Electronic fluorescent pictographic (eFP) map of *VfMYB36* in different tissues or developmental stages. Red and yellow indicated high and low expression levels, respectively

A. thaliana R2R3-MYBs were obtained from a previous study and the corresponding sequences were derived from TAIR database (www.arabidopsis.org/) [18]. Proteins with a MYB domain (PF000249) were identified by the hidden Markov model using HMMER v3.2.1 software and BlastP (E-value=1e -3) similarity searches in five Euphorbiaceae genomes: *R. communis*, *J. curcas*, *M. esculenta*, *H. brasiliensis* and *V. fordii* [57]. Subsequently, all obtained MYB sequences were again detected using Pfam database (http://pfam.xfam.org/) [58], InterPro (http://www.ebi.ac.uk/interpro/scan.html) [59], and SMART (http://smart.embl-heidelberg.de/) [60] for domain structure of MYB (PF000249). Finally, we discarded those MYB sequences that lack core or complete MYB domain (PF000249).

Phylogenetic analysis of R2R3-MYBs
The multiple sequence alignment of all full-length R2R3-MYBs was performed using MAFFT 7.0 software [61]. ModelFinder was used to estimate the best substitution model for R2R3-MYBs with maximum likelihood (ML) [62]. IQ-TREE software was used to construct the ML tree using ultrafast bootstrap approximation with 1000 and SH-aLRT test set to 1000 random addition replicates [63]. Finally, the tree was visualized with online iTol (https://itol.embl.de/) [64] and FigTree software.

Synteny and positive selection analysis of R2R3-MYBs
Paralogs are genes that are related by duplication within the same genome [65, 66]. Orthologous genes are often defined as genes that have a synteny relationship in different genomes, according to previous studies [65–67]. The intra-species and interspecies synteny analyses were performed using a similar step to that developed for the PGDD [46]. First, we searched the potential homologs using BLASTP software (E<1e-10, top five matches). Then, these results as the input object for MCScanX software were used to confirm syntenic chains with the following parameters: Max GAPs, 25; E-value, 1e-05; Overlap window, 5; GAP penalty, −1; Match score, 50; Match size, 5 [68]. Further, the "duplicate_gene_classifier" script in the MCScanX software was also used to identify the various types of duplications, including WGD/segmental, tandem, proximal, and dispersed duplication events. The KaKs_Calculator 2.0 software was used to estimate the synonymous substitution rates (Ks) and nonsynonymous substitution rates (Ka) [33]. If the Ka value was nearly 0, and the Ks value was more than 3 or less than 0.01, these duplicated genes were excluded, because a high Ks value implied the risk of saturation, and low sequence divergence could lead to unknown results. When Ka/Ks<1, >1, and =1, it suggested purifying selection, positive selection, and neutral selection, respectively.
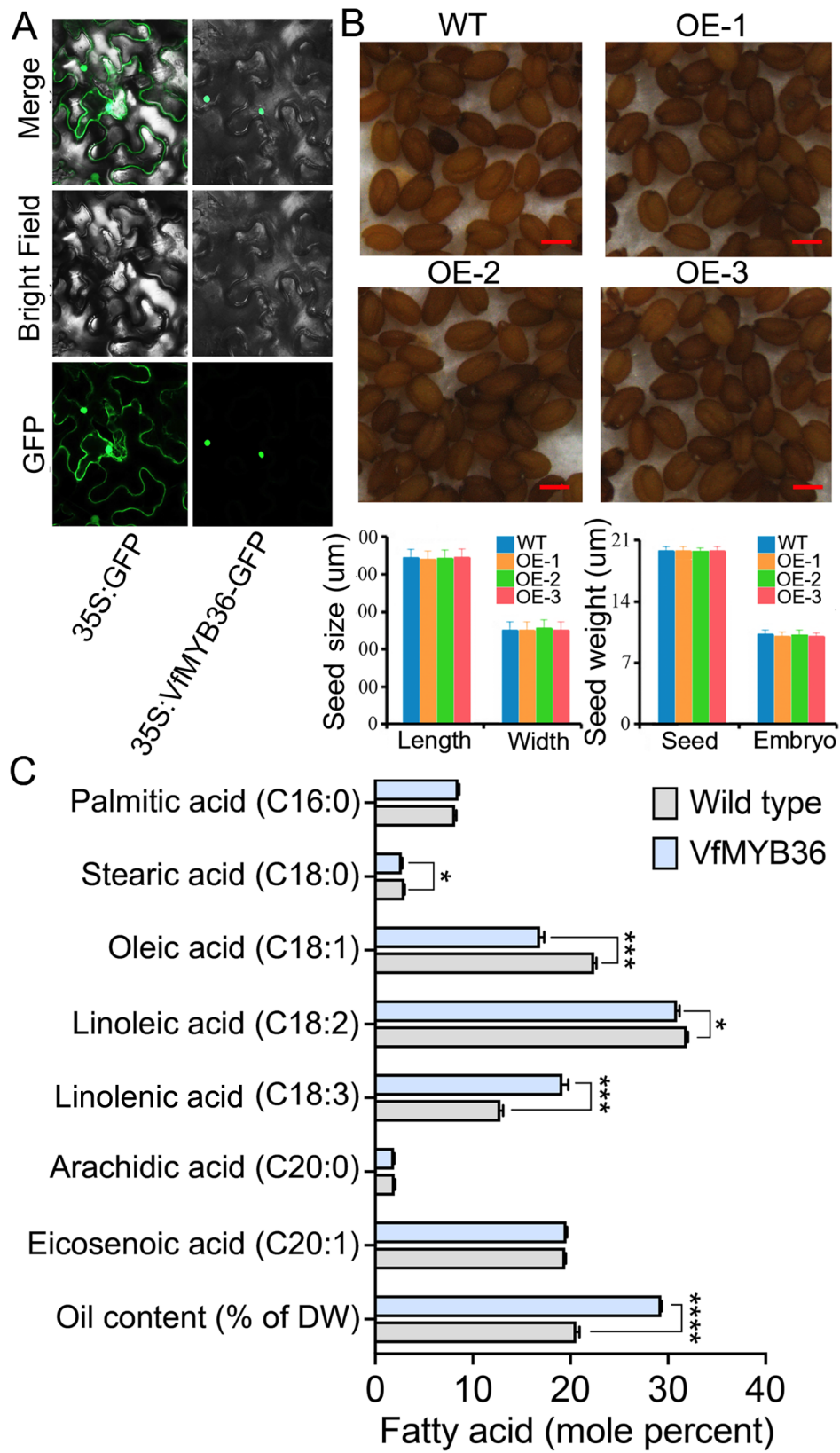
Fig. 7 Functional analysis of *VfMYB36* in *V. fordii*. **A**, Subcellular localization of the VfMYB36 fused with GFP (35 S:VfMYB36-GFP) in tobacco leaves. **B**, Microscopic observation of mature seeds randomly selected from Wild type (Col-0) and *VfMYB36* overexpression plants. **C**, The content of major fatty acid (FA) species and oil in mature seeds of Wild type (Col-0) and *VfMYB36* overexpression plants

Gene feature in homologs

The gene features, including the frequency of optimal codons (Fop), polypeptide length, and GC content at three codon positions (GC1, GC2, and GC3), were estimated to compare duplicated and singleton R2R3-MYB genes. The CodonW 1.4 (http://codonw.sourceforge.net) was used to estimate the polypeptide length and Fop. The in-house perl script was carried out to calculate the GC content.

Vernicia fordii RNA-seq data analysis

The RNA-seq data (BioProject: PRJNA503685, PRJNA445350, and PRJNA483508) were collected and retrieved from SRA database, and then these raw data were processed by trimmomatic with default parameters [69]. The HiSAT2 [70] was carried out to align the RNA-Seq clean reads to the masked *V. fordii* genome, and then the Stringtie [70, 71] was used to obtain the Fragments Per Kilobase Million (FPKM) values, as described by Cao et al. (2021) [72].

Subcellular localization

The primer VfMYB36-F, 5'-ATGGAGAGTAAGCAGT-CAAAAG-3' and VfMYB36-R 5' TCAAGAG-GCTCCTACTCCAAG-3' was used to clone the *VfMYB36* from *V. fordii* seeds. Then we used the Gateway system to produce the pEaryleyGate104-VfMYB36 for subcellular localization. The Gene Pulser Xcell (BIO-RAD, country-region USA) was used to electroporate pMDC32-VfMYB36 into Agrobacterium tumefaciens GV3101. Use the injection method to infiltrate the suspension into the tobacco leaves. The confocal laser scanning microscopy (CarlZeiss LSM710, Germany) was implemented to observe the expressed YFP signal.

Fatty acid analysis of *A. thaliana* over-expressing *VfMYB36*

The Gateway system was used to produce the over-expression vector pMDC32-VfMYB36 for the transgenic experiment. We transformed the pMDC32-VfMYB36 vector into *A. thaliana* by Agrobacterium-mediated the floral dip method [73]. Based on the protocols described by Chen et al. (2010), the gas chromatograph GC-2014 (Shimadzu, Japan) was used to measure all fatty acids [74]. Three biological repeats of all experiments were performed in our study. GraphPad Prism v9 was used for data analysis and visualization.

qRT-PCR analysis

The MiniBEST Plant RNA Extraction Kit (TaKaRa) was used to extract total RNA, and the PrimerScript RT (TaKaRa) was carried out for reverse transcription. The SYBR Green Master Mix (TaKaRa) was used to perform the qRT-PCR for three biological replicates. The primers for qRT-PCR using Primer-BLAST (https://www.ncbi. nlm.nih.gov/tools/primer-blast/index.cgi; Table S6). A $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression level as described previously [13, 75].

Abbreviations

Ks	synonymous substitutions per site
HTH	helix–turn–helix
CDS	coding sequence
DBF	day before flowering
GFP	green fluorescent protein
DIC	bright field image
WAF	week after flowering
AMV	avian myeloblastosis virus
Ka	nonsynonymous substitution rates

Declarations

Ethics approval and consent to participate
All authors confirmed that all methods were carried out in accordance with relevant guidelines and regulations under.

Consent for publication
Not applicable.

Competing interests
The authors declare that there are no conflicts of interest.

References

1. Liu L, White MJ, MacRae TH. Transcription factors and their genes in higher plants: functional domains, evolution and regulation. Eur J Biochem. 1999;262(2):247–57.
2. Yamasaki K, Kigawa T, Seki M, Shinozaki K, Yokoyama S. DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. Trends Plant Sci. 2013;18(5):267–76.
3. Chen F, Hu Y, Vannozzi A, Wu K, Cai H, Qin Y, Mullis A, Lin Z, Zhang L. The WRKY transcription factor family in model plants and crops. CRC Crit Rev Plant Sci. 2017;36(5–6):311–35.
4. Manna M, Thakur T, Chirom O, Mandlik R, Deshmukh R, Salvi P. Transcription factors as key molecular target to strengthen the drought stress tolerance in plants. Physiol Plant. 2021;172(2):847–68.
5. Chronis C, Fiziev P, Papp B, Butz S, Bonora G, Sabri S, Ernst J, Plath K. Cooperative binding of transcription factors orchestrates reprogramming. Cell. 2017;168(3):442–59.
6. Li D, Peng S, Chen S, Li Z, He Y, Ren B, Yang G. Identification and characterization of 5 walnut MYB genes in response to drought stress involved in ABA signaling. Physiol Mol Biology Plants. 2021;27(6):1323–35.
7. Cao Y, Han Y, Li D, Lin Y, Cai Y. MYB transcription factors in chinese pear (Pyrus bretschneideri Rehd.): genome-wide identification, classification, and expression profiling during fruit development. Front Plant Sci. 2016;7:577.
8. Cao Y, Li K, Li Y, Zhao X, Wang L. MYB transcription factors as regulators of secondary metabolism in plants. Biology. 2020;9(3):61.
9. Naing AH, Kim CK. Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. Plant Mol Biol. 2018;98(1):1–18.
10. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Mol Plant. 2015;8(5):689–708.
11. Wang B, Luo Q, Li Y, Yin L, Zhou N, Li X, Gan J, Dong A. Structural insights into target DNA recognition by R2R3-MYB transcription factors. Nucleic Acids Res. 2020;48(1):460–71.
12. Dong J, Hu F, Guan W, Yuan F, Lai Z, Zhong J, Liu J, Wu Z, Cheng J, Hu K. A 163-bp insertion in the Capana10g000198 encoding a MYB transcription factor causes male sterility in pepper (Capsicum annuum L). Plant J. 2023;113(3):521–35.
13. Jiang J, Lin M, Wang H, Song H, Zhang L, Huang Q, Chen R, Song C, Li G, Cao Y. Haplotype-resolved genome assembly of Bletilla striata (Thunb.) Reichb.f. to elucidate medicinal value. *The Plant Journal* 2022, n/a(n/a).
14. Martin C, Paz-Ares J. MYB transcription factors in plants. Trends Genet. 1997;13(2):67–73.
15. Ito M. Conservation and diversification of three-repeat myb transcription factors in plants. J Plant Res. 2005;118(1):61–9.
16. Rodrigues JA, Espley RV, Allan AC. Genomic analysis uncovers functional variation in the C-terminus of anthocyanin-activating MYB transcription factors. Hortic Res. 2021;8(1):1–14.
17. Liu M, Li K, Sheng S, Wang M, Hua P, Wang Y, Chen P, Wang K, Zhao M, Wang Y, et al. Transcriptome analysis of MYB transcription factors family and PgMYB genes involved in salt stress resistance in Panax ginseng. BMC Plant Biol. 2022;22(1):479.
18. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol. 2001;4(5):447–56.
19. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J. 2011;66(1):94–116.
20. Rosinski JA, Atchley WR. Molecular evolution of the myb family of transcription factors: evidence for polyphyletic origin. J Mol Evol. 1998;46(1):74–83.
21. Klempnauer K-H, Gonda TJ, Bishop JM. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. Cell. 1982;31(2):453–63.
22. Cone KC, Cocciolone SM, Burr FA, Burr B. Maize anthocyanin regulatory gene pl is a duplicate of c1 that functions in the plant. Plant Cell. 1993;5(12):1795–805.
23. Du H, Feng B-R, Yang S-S, Huang Y-B, Tang Y-X. The R2R3-MYB transcription factor gene family in maize. PLoS ONE. 2012;7(6):e37463.
24. Zhang L, Jiang X, Liu Q, Ahammed GJ, Lin R, Wang L, Shao S, Yu J, Zhou Y. The HY5 and MYB15 transcription factors positively regulate cold tolerance in tomato via the CBF pathway. Plant Cell Environ. 2020;43(11):2712–26.
25. Chinnusamy V, Zhu J, Zhu J-K. Cold stress regulation of gene expression in plants. Trends Plant Sci. 2007;12(10):444–51.
26. Yang A, Dai X, Zhang W-H. A R2R3-type MYB gene, OsMYB2, is involved in salt, cold, and dehydration tolerance in rice. J Exp Bot. 2012;63(7):2541–56.
27. McCarthy RL, Zhong R, Fowler S, Lyskowski D, Piyasena H, Carleton K, Spicer C, Ye Z-H. The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. Plant Cell Physiol. 2010;51(6):1084–90.
28. Li C, Wang X, Ran L, Tian Q, Fan D, Luo K. PtoMYB92 is a transcriptional activator of the lignin biosynthetic pathway during secondary cell wall formation in Populus tomentosa. Plant Cell Physiol. 2015;56(12):2436–46.
29. Shockey J, Rinehart T, Chen Y, Yangdong W, Zhiyong Z, Lisong H. Tung (Vernicia fordii and Vernicia montana). In: *Industrial Oil Crops* Elsevier; 2016: 243–273.
30. Kumar S, Dhillon MK, Singh M, Rathi RS, Misra AK, Rana JC. Fatty and amino acid compositions of Vernicia fordii: A source of α-eleostearic acid and methionine. 2017.
31. Li Y, Jiang L, Mo W, Wang L, Zhang L, Cao Y. AHLs' life in plants: Especially their potential roles in responding to Fusarium wilt and repressing the seed oil accumulation.International Journal of Biological Macromolecules2022.
32. Wang Y, Hou Y, Wang J, Zhao H. Analyzing lignin biosynthesis pathways in rattan using improved co-expression networks of NACs and MYBs. BMC Plant Biol. 2022;22(1):411.
33. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genom Proteom Bioinform. 2010;8(1):77–80.
34. Zhao K, Cheng Z, Guo Q, Yao W, Liu H, Zhou B, Jiang T. Characterization of the poplar R2R3-MYB gene family and over-expression of PsnMYB108 confers salt tolerance in transgenic tobacco. Front Plant Sci. 2020;11:1548.
35. Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC. Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. BMC Genomics. 2012;13(1):1–19.
36. Cheng C, Li Q, Wang X, Li Y, Qian C, Li J, Lou Q, Jahn M, Chen J. Identification and expression analysis of the CsMYB gene family in Root Knot Nematode-resistant and susceptible cucumbers. Frontiers in genetics 2020,11.
37. Du H, Yang S-S, Liang Z, Feng B-R, Liu L, Huang Y-B, Tang Y-X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. BMC Plant Biol. 2012;12(1):1–22.
38. Matus JT, Aquea F, Arce-Johnson P. Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. BMC Plant Biol. 2008;8(1):1–15.
39. Zhang Z, Zhang L, Liu Y, Shang X, Fang S. Identification and expression analysis of R2R3-MYB family genes Associated with Salt Tolerance in Cyclocarya paliurus. Int J Mol Sci. 2022;23(7):3429.
40. Zargarian L, Le Tilly V, Jamin N, Chaffotte A, Gabrielsen OS, Toma F, Alpert B. Myb – DNA recognition: role of tryptophan residues and structural changes of the minimal DNA binding domain of c-Myb. Biochemistry. 1999;38(6):1921–9.
41. Jiang C, Gu J, Chopra S, Gu X, Peterson T. Ordered origin of the typical two- and three-repeat myb genes. Gene. 2004;326:13–22.
42. Jiang C-K, Rao G-Y. Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. Plant Physiol. 2020;183(2):637–55.
43. Wu H-J, Zhang Z, Wang J-Y, Oh D-H, Dassanayake M, Liu B, Huang Q, Sun H-X, Xia R, Wu Y. Insights into salt tolerance from the genome of Thellungiella salsuginea. *Proceedings of the National Academy of Sciences* 2012, 109(30):12219–12224.
44. Manzoor MA, Li G, Abdullah M, Han W, Wenlong H, Yang Z, Xinya W, Yu Z, Xiaofeng F, Qing J. Genome-wide investigation and comparative analysis of MATE gene family in Rosaceae species and their regulatory role in abiotic stress responses in chinese pear (Pyrus bretschneideri). Physiol Plant. 2021;173(3):1163–78.
45. Li X, Xue C, Li J, Qiao X, Li L, Yu L, Huang Y, Wu J. Genome-wide identification, evolution and functional divergence of MYB transcription factors in chinese white pear (Pyrus bretschneideri). Plant Cell Physiol. 2016;57(4):824–47.
46. Lee T-H, Tang H, Wang X, Paterson AH. PGDD: a database of gene and genome duplication in plants. Nucleic Acids Res. 2012;41(D1):D1152–8.
47. Cao Y, Li Y, Wang L, Zhang L, Jiang L. Evolution and function of ubiquitin-specific proteases (UBPs): Insight into seed development roles in tung tree (Vernicia fordii). International Journal of Biological Macromolecules 2022.
48. Zhang L, Liu M, Long H, Dong W, Pasha A, Esteban E, Li W, Yang X, Li Z, Song A. Tung tree (Vernicia fordii) genome provides a resource for understanding genome evolution and improved oil production. Genom Proteom Bioinform. 2019;17(6):558–75.
49. Cao Y, Jiang L, Wang L, Cai Y. Evolutionary rate heterogeneity and functional divergence of orthologous genes in Pyrus. Biomolecules. 2019;9(9):490.

50. Wang X, Niu Y, Zheng Y. Multiple functions of MYB transcription factors in abiotic stress responses. Int J Mol Sci. 2021;22(11):6125.
51. Liu L, Chao N, Yidilisi K, Kang X, Cao X. Comprehensive analysis of the MYB transcription factor gene family in Morus alba. BMC Plant Biol. 2022;22(1):281.
52. Li D, Jin C, Duan S, Zhu Y, Qi S, Liu K, Gao C, Ma H, Zhang M, Liao Y. MYB89 transcription factor represses seed oil accumulation. Plant Physiol. 2017;173(2):1211–25.
53. Meshi T, Iwabuchi M. Plant transcription factors. Plant Cell Physiol. 1995;36(8):1405–20.
54. Zhang L, Lu S, Sun D, Peng J. Genetic variation and geographical differentiation revealed using ISSR markers in tung tree, Vernicia fordii. J Genet. 2015;94(2):5–9.
55. Singh R, Low E-TL, Ooi LC-L, Ong-Abdullah M, Nookiah R, Ting N-C, Marjuni M, Chan P-L, Ithnin M, Manaf MAA. The oil palm VIRESCENS gene controls fruit colour and encodes a R2R3-MYB. Nat Commun. 2014;5(1):1–8.
56. Li DR, Zhang YW, Zhang WX. A major breakthrough in the breeding of Brassica. Napus with high linolenic acid content, which linolenic acid content of germplasm resources exceeding 21%, and of hybrids about 15%. J Plant Sci. 2019;3(2):165–6.
57. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013;41(12):e121–1.
58. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ. Pfam: The protein families database in 2021. Nucleic Acids Research 2020.
59. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40.
60. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(D1):D493–6.
61. Katoh K, Kuma K-i, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005;33(2):511–8.
62. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9.
63. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
64. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256–9.
65. Li F, Fan K, Guo X, Liu J, Zhang K, Lu P. Genome-wide identification, molecular evolution and expression analysis of the non-specific lipid transfer protein (nsLTP) family in Setaria italica. BMC Plant Biol. 2022;22(1):547.
66. Yang J, Xiong C, Li S, Zhou C, Li L, Xue Q, Liu W, Niu Z, Ding X. Evolution patterns of NBS genes in the genus Dendrobium and NBS-LRR gene expression in D. officinale by salicylic acid treatment. BMC Plant Biol. 2022;22(1):529.
67. Ai D, Wang Y, Wei Y, Zhang J, Meng J, Zhang Y. Comprehensive identification and expression analyses of the SnRK gene family in Casuarina equisetifolia in response to salt stress. BMC Plant Biol. 2022;22(1):572.
68. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–9.
69. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
70. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc. 2016;11(9):1650–67.
71. Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.
72. Cao Y, Mo W, Li Y, Li W, Dong X, Liu M, Jiang L, Zhang L. Deciphering the roles of leucine-rich repeat receptor-like protein kinases (LRR-RLKs) in response to Fusarium wilt in the Vernicia fordii (Tung tree). Phytochemistry. 2021;185:112686.
73. Zhang X, Henriques R, Lin S-S, Niu Q-W, Chua N-H. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protoc. 2006;1(2):641–6.
74. Chen Y, Zhou G, Wang Y, Xu L. F-BOX and oleosin: additional target genes for future metabolic engineering in tung trees? Ind Crops Prod. 2010;32(3):684–6.
75. Bai Y, Liu H, Zhu K, Cheng Z-M. Evolution and functional analysis of the GRAS family genes in six Rosaceae species. BMC Plant Biol. 2022;22(1):569.

Publisher's Note