

Supplementary Information

Figure S1. Sequence alignment of epimedium MYB1 and MYB12 proteins with several *R2R3-MYB* proteins containing the C2 repressor motif from other plant species. Identical residues are shown in black and similar residues in gray. The R2 and R3 MYB DNA-binding domains are shown in shaded text. A conserved motif for interaction with bHLH protein is indicated in line. Three conserved motifs, C1, C2 and Zn-finger are indicated in boxes for members of MYB subgroup 4, and of which the Zn-finger motif is located in the EsMYB1 protein sequence, but not in the EsMYB12 protein sequence. Abbreviation: Am, *Antirrhinum majus*; At, *Arabidopsis thaliana*; Br, *Brassica rapa*; Eg, *Eucalyptus gunnii*; Fa, *Fragaria x ananassa*; Le, *Lycopersicon esculentum*; Pt, *Pinus taeda*; Ptt, *Populus tremula x Populus tremuloides*; Vv, *Vitis vinifera*; Zm, *Zea mays*. GenBank accession numbers for the sequences: AmMYB308, P81393; AtMYB4, NP_195574; AtMYB7, NP_179263; AtMYB32, NP_195225; BrMYB4, ABQ81931; EgMYB1, CAE09058; FaMYB1, AAK84064; LeMYB27, CAA64614; PtMYB14, ABD60279; PttMYB4a, CAD98762; VvMYB4b, ABW34393; ZmMYB31, CAJ42202; ZmMYB42, CAJ42204.

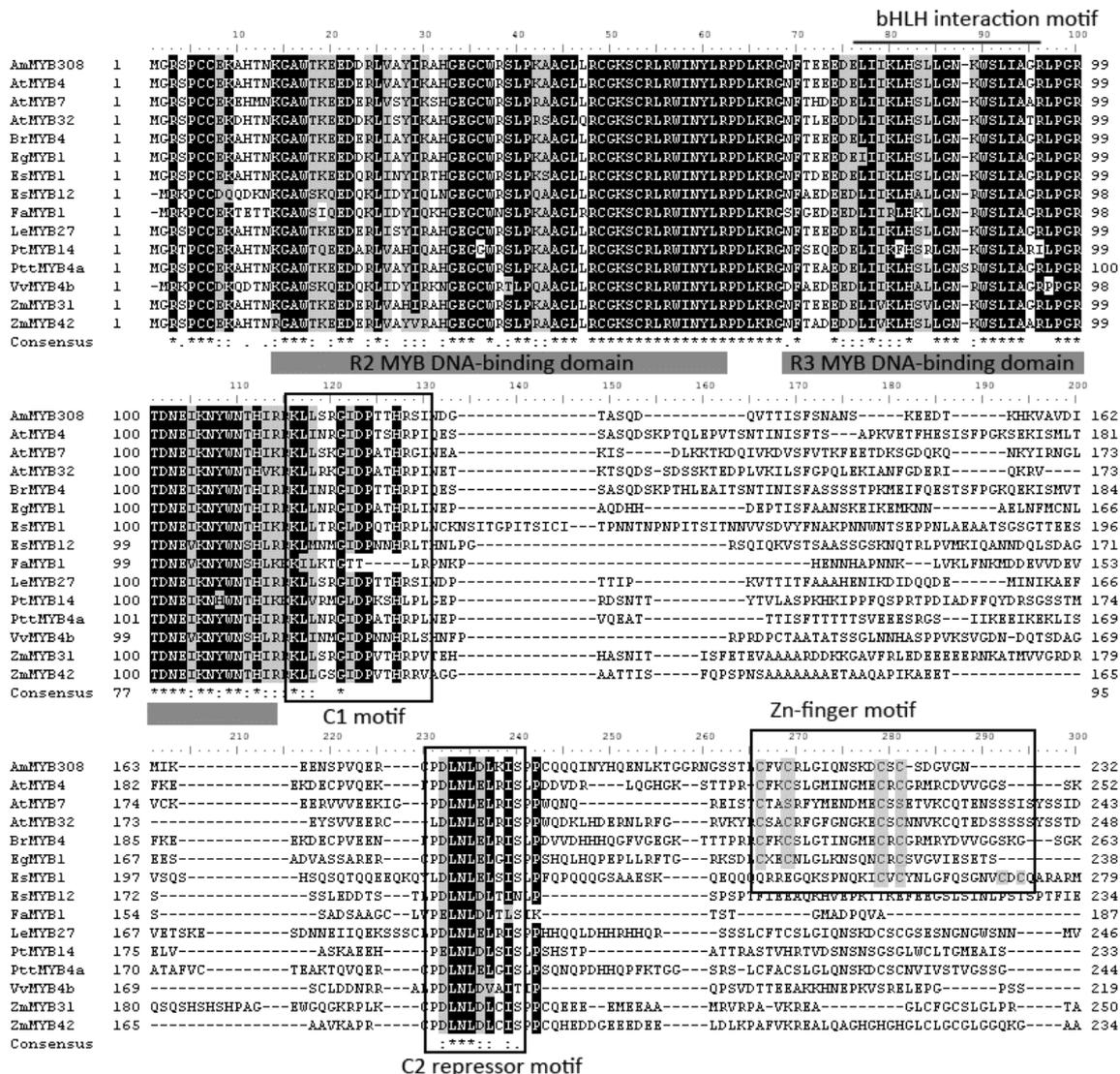


Figure S1. Cont.

```

                                     310       320       330
.....|.....|.....|.....|.....|.....|
AmMYB308  232 ----- 232
AtMYB4    253 GSDMSNGFDFLGLAKKET-TSLLCFRSLEMK 282
AtMYB7    244 ISSSNVGYDFLGLK-----TRILDFRSLEMK 269
AtMYB32   249 ISSS-ICYDFLGLMN----TRVLDFSTLEMK 274
BrMYB4    264 GSDMSNGFDFLGLAKKETNTCLFGFRSLEMK 294
EgMYB1    238 -----VGYDFLGL----KAS-VLDYRS---- 255
EsMYB1    280 VISSDNLHRFLIP-----LDA----- 295
EsMYB12   235 ETQKHVEPKTKRFEEGSSPTLLLF----- 260
FaMYB1    187 ----- 187
LeMYB27   247 SMNIMAGYDFLGL----KTNGLLDYRTLETK 273
PtMYB14   233 ----- 233
PttMYB4a  245 STSTKNGYDFLGM---KSG-VLDYRSLEMK 270
VvMYB4b   219 -----TLLLF----- 225
ZmMYB31   251 DCKCSS-SSFLCLR----TAMLDFRSLEMK 275
ZmMYB42   235 GCSCSNCHHFLCLR----TSVLDFRGLEMK 260
Consensus

```

Figure S2. Sequence alignment of epimedi MYB9 protein with several functionally characterized plant *R2R3-MYB* proteins. Identical residues are shown in black and similar residues in gray. The lines above the alignment indicate the R2 and R3 MYB DNA-binding domains and a conserved motif for interaction with bHLH proteins, respectively. Two conserved motifs C1 and C3 are shown in boxes. GenBank accession numbers are as follows (in parentheses): VvMYB5a (AAS68190) and VvMYB5b (AAX51291) from *Vitis vinifera*; AtMYB5 (AAC49311) from *Arabidopsis thaliana*; PH4 (AA51377) from *Petunia hybrida*; OsMYB4 (BAA23340) from *Oryza sativa*.

```

                                     10       20       30       40       50       60       70       80
VvMYB5a   1  MRNPASA-----STSK-----TPCCTKVGKLRCPWTPDEEDLLANYVKEKCEGRWTLPKRAGLLRCCGKSCRLRWNNYLR 70
VvMYB5b   1  MRNASSASAPPSSSSK----TPCCIKVGLKRCGPWTPDEEDLVANYIKKCEGRWTLPKRAGLLRCCGKSCRLRWNNYLR 75
AtMYB5    1  MMSCGGK----KPVSK---RTPCCTKMGKRCGPWTPDEEDLLWSFKKCEGRWTLPKRAGLLRCCGKSCRLRWNNYLR 73
PH4       1  MRPSSS----STTSN---KVTTPCCSKVGLKRCGPWTPDEEDLLTNYLNKCEGRWTLPKRAGLLRCCGKSCRLRWNNYLR 73
OsMYB4    1  MKRKRPA----ALRG----GEEAAAAAKKRCGPWTPDEEDVLARFVARECCDRWTLPKRAGLLRCCGKSCRLRWNNYLR 70
EsMYB9    1  MRNPSTGCSRPTASATTSTTPCCTKIGKRCGPWTPDEEDLLCDYIKKCEGRWTLPKRAGLLRCCGKSCRLRWNNYLR 80

                                     interaction domain      R2 repeat domain
                                     90       100      110      120      130      140      150      160
VvMYB5a   71  PSWKRGGIADPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLISQCIDPPTHKPLNPKPNPSP---- 145
VvMYB5b   76  PSWKRGGIADPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLISQCIDPPTHKPLNPMSSSVD---V 151
AtMYB5    74  PSWKRGGITSDPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLLRQCIDPOTHKPLDANN-----I 145
PH4       74  PSWKRGGIADPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLISHCIDPPTHKPLKSNSSSDDITNK 153
OsMYB4    71  PDKRCPIADPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLIAQCIDPPTHKPLTAAADHSN-AAAA 149
EsMYB9    81  PSWKRGGITSDPEEDLILRLHRLLCNRWALACRTPCRTDNEIKMYWTHLSKRLISQCIDPPTHKPLNPMSSSS---F 156

                                     R3 repeat domain      C1 motif
                                     170      180      190      200      210      220      230      240
VvMYB5a   146  DVNAPVSKSIPNAMPNPSSSRVCHIGSNHEVKETESMENHKEPPNLDQYHSPLAADSNEMWQSADCLVTGLQSTHCTSND 225
VvMYB5b   152  KASSKAKAVMNPMPNPNPSPSEKAAANKKACNFKSDNQY-----QICAAAGNDGSANIQNSD---GSGTGLRSSNNE 220
AtMYB5    146  HKPEEIVSCKGQKYPLEPIS-----SHTDDTTVMCGDGD---SKNSINVFCEGH 191
PH4       154  LASSPPSSSKAMDLPILSPTYISSFQMEEPKINTHPG-----RITSDDQYQSMAILAE---YDLDLNIATVIE 223
OsMYB4    150  VAATSYPKAVPAKPPRTASPAAGIECSDDRARPADGCGD-----FAAMVSAADARCFEGC---FGDQPCAEADVH 217
EsMYB9    157  PLP--KPNFHTN-NPNPNSGSEREVPS-----AAGSGVIARTTDSDD---RGDHSLNWLNSN 206

                                     250      260      270      280      290      300      310      320
VvMYB5a   226  DEDDVG----FCDDTTPSSFLNLSLIM---EDVFRGMNHHQQQQQQQQQQVQVQPSNVIAPLPHPAISVQATFSSSPRTV 297
VvMYB5b   221  EDDDLN----CCDDVFPSSFLNLSLIM---EDVFRGQHHLQQQHH---GGLIAPGSDALISTSSVQSFCTGTSWEAAAMTS 290
AtMYB5    192  CYEDFG----FCDDKFPSSFLNLSLIMD--VGDPPFQNIIP-----ISQPLQMDCKDCGIVGAS 242
PH4       224  EDVEMN----CCDDVFPSSFLNLSLIM---EDMFAFCQN-----QQTNGTFQDFDFPFMSS 270
OsMYB4    218  GCFDMGSASAMVQDDDFSSFLDLSLIMDLEQLGDLRFVVEGN-----DHEHGNCEIGHGDMVMS 273
EsMYB9    207  EGDVIG----CCDDHFPSSFLNLSLIM---EDLFRVQNS-----NCMVGSHIETLLS--SARSLGCHGAVLDSALSSS 267

                                     C3 motif
                                     330      340
VvMYB5a   298  WEPAAALTSTSAPLVHDQKHSMS 320
VvMYB5b   291  TSVFSQIDHS--KRFNDQPKRF 311
AtMYB5    243  SSSLG-----HD--- 249
PH4       271  STPSS-----DQYNPS- 281
OsMYB4    274  KQ----- 275
EsMYB9    268  GAVAAAPEDEDGQRQLLDHAREKGS 290

```

Figure S3. Sequence alignment of epimedium MYB8 with other DIV-like proteins. Identical residues are shown in black and similar residues in gray. The R2 and R3 MYB DNA-binding domains are shown in lines and the “SHAQKY” conserved motif is located in the end of the R3 MYB domain. Another conserved motif is shown in box. Arrowhead indicates the insertion position of intron with the same intron phase in the DIV-like proteins. GenBank accession numbers are as follows (in parentheses): AmDIV (AAL78741) and AmDVL1 (AAL78742) from *Antirrhinum majus*; AtMYB (NP_200698) from *Arabidopsis thaliana*; BIDIV1 (ABN13121) and BIDIV2 (ABN13123) from *Bournea leiophylla*; GmMYB57 (ABH02828) from *Glycine max*; HmDIV1A (ACR09746) and HmDIV2A (ACR09741) from *Heptacodium miconioides*; MtMYB33 (ABR28332) from *Medicago truncatula*; OsMYB (BAD61069) from *Oryza sativa*; ZmDIV (ACG30643) from *Zea mays*.

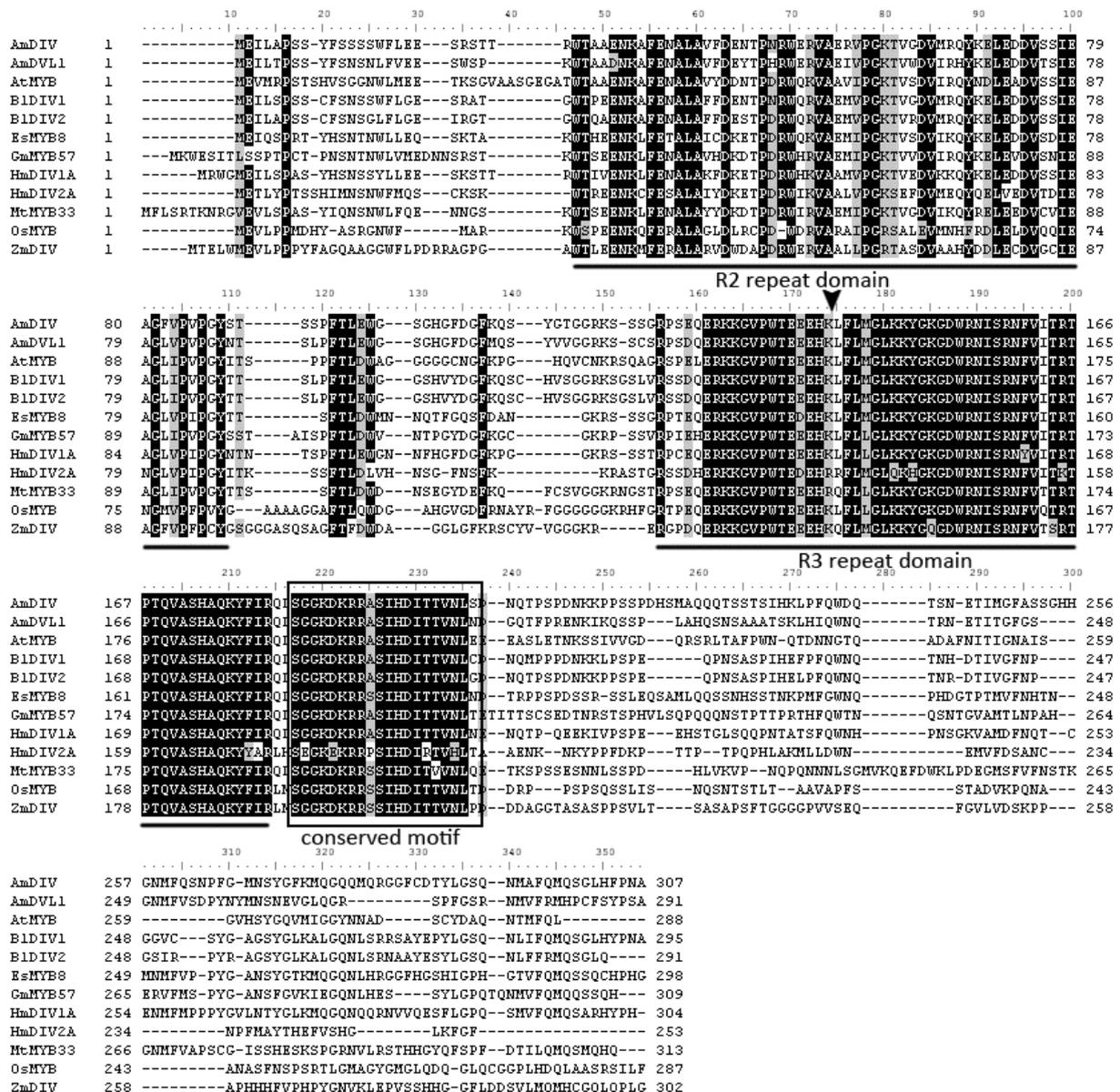


Figure S4. Phylogenetic relationships of thirteen R2R3-MYB proteins from epimedii with other characterized plant *R2R3-MYB* proteins. Topology tree is performed using the neighbor-joining method by the MEGA 4.0 version. The numbers next to the nodes are bootstrap values from 1,000 replicates. Thirteen R2R3-MYB proteins from epimedii are indicated in diamonds. The putative regulatory functions of the different R2R3-MYB proteins in the control of secondary metabolite biosynthesis or other biological processes are indicated. GenBank accession numbers of MYB proteins are shown in the figure, following the gene names.

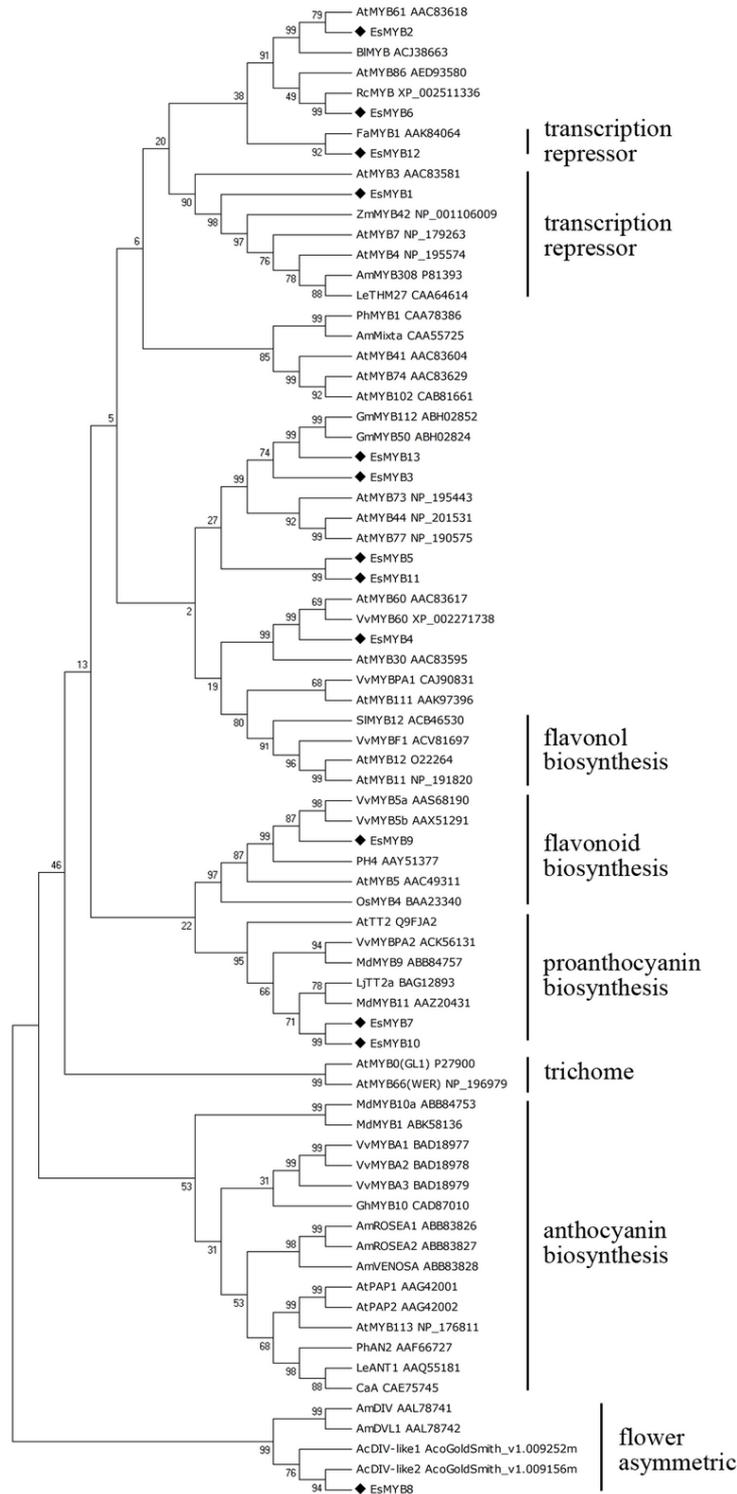


Figure S5. Band size comparison of cDNA and gDNA clones of epimedii *MYB3* and *MYB13* genes. The full-length gene specific primers were used to amplify the cDNA and corresponding genomic DNA (gDNA) clones of *EsMYB3* and *EsMYB13* genes from the cDNA template digested with DNase I and genomic DNA template. The band size is shown equal between the cDNA and gDNA amplification of both *EsMYB3* and *EsMYB13* genes.

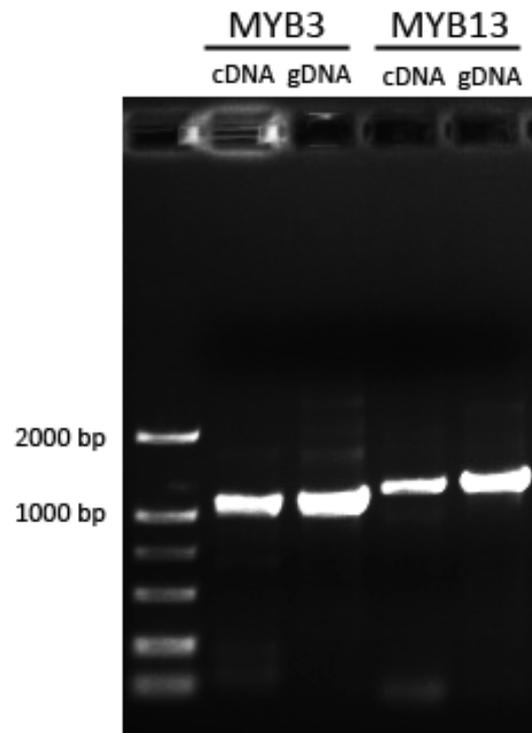


Figure S6. Phenotype observation of transgenic tobacco plants overexpressing *EsMYB9* and empty vector. Overexpression of *EsMYB9* driven by CaMV 35S promoter led to increase the anthocyanin accumulation in flowers, including petals and stamens, compared to the control plants overexpressing the empty vector.

