Identification of the *Eutrema Salsugineum* *EsMYB90* gene important for anthocyanin biosynthesis

Yuting Qi

Shandong Normal University

Caihong Gu

Shandong Normal University

Xingjun Wang

Shandong Academy of Agricultural Sciences

Shiqing Gao

Shandong Academy of Agricultural Sciences

Changsheng Li

Shandong Academy of Agricultural Sciences

Chuanzhi Zhao

Shandong Academy of Agricultural Sciences

Chuanshun Li

Shandong Normal University

Changle Ma

Shandong Normal University

Dazhong Zhao

University of Wisconsin Milwaukee

Quan Zhang (✉ zhangquan@sdnu.edu.cn)

Shandong Normal University <https://orcid.org/0000-0002-5380-215X>

Research article

Keywords: Anthocyanin, flavonoid, *Eutrema salsugineum*, R2R3 MYB transcription factor, *EsMYB90*, transcriptional regulation, anthocyanin biosynthesis genes

Posted Date: December 4th, 2019

DOI: <https://doi.org/10.21203/rs.2.18301/v1>

License: (cc) (i) This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Plant Biology on April 28th, 2020. See the published version at <https://doi.org/10.1186/s12870-020-02391-7>.

Abstract

Background: Anthocyanins contribute to coloration and antioxidation effects in different plant tissues. MYB transcription factors have been demonstrated to be a key regulator for anthocyanin synthesis in many plants. However, little information was available about the MYB genes in the halophyte species *Eutrema salsugineum*.

Result: Here we report the identification of an important anthocyanin biosynthesis regulator *EsMYB90* from *Eutrema salsugineum*, which is a halophyte tolerant to multiple abiotic stresses. Our phylogenetic and localization analyses supported that *EsMYB90* is an R2R3 type of MYB transcription factor. Ectopic expression of *EsMYB90* in tobacco and *Arabidopsis* enhanced pigmentation and anthocyanin accumulation in various organs. The transcriptome analysis revealed that 42 genes upregulated by *EsMYB90* in *35S:EsMYB90* tobacco transgenic plants are required for anthocyanin biosynthesis. Moreover, our qRT-PCR results showed that *EsMYB90* promoted expression of early (*PAL*, *CHS*, and *CHI*) and late (*DFR*, *ANS*, and *UFGT*) anthocyanin biosynthesis genes in stems, leaves, and flowers of *35S:EsMYB90* tobacco transgenic plants.

Conclusions: Our results indicated that *EsMYB90* is a novel MYB transcription factor, which regulates anthocyanin biosynthesis genes to control anthocyanin biosynthesis. Our work provides a new tool to enhance anthocyanin production in various plants.

Background

Flavonoids which are derivatives of the phenylpropanoid/flavonoid pathway mainly contain proanthocyanidins (PAs), anthocyanins and flavonols [1–3]. As important pigments, anthocyanins are responsible for red, purple, violet and blue colors in flowers, fruits, and leaves, which determine economic traits of crops and ornamental plants [4–7]. Anthocyanins are the end products of a specific branch in the phenylpropanoid/flavonoid biosynthesis pathway. Enzymes involved in anthocyanin biosynthesis have been extensively studied in many plant species [8]. Catalyzed by phenylalanine ammonia-lyase (PAL), the initial step of the flavonoid pathway is the conversion of phenylalanine into trans-cinnamic acid [9], while chalcone synthase (CHS) catalyzes the first committed step in the flavonoid biosynthesis to form naringenin chalcone. Chalcone isomerase (CHI) cyclizes chalcone to form naringenin [8]. The naringenin is then converted into dihydrokaempferol (DHK) by flavanone 3 β -hydroxylase (F3H). DHK is further hydroxylated to dihydroquercetin (DHQ) by flavonoid 3'-hydroxylase (F3'H), or to dihydromyricetin (DHM) by flavonoid 3',5'-hydroxylase (F3'5'H). Dihydroflavonol 4-reductase (DFR) converts DHQ into leucocyanidin, which is further converted into anthocyanidins by anthocyanidin synthase (ANS). Finally, UDP-glucose: flavonoid 3-O-glucosyltransferase (UFGT) catalyzes glycosylation of anthocyanidins to form anthocyanins [8, 10–12].

MYB transcription factors play a central role in regulating expression of genes encoding major enzymes for anthocyanin biosynthesis via forming the transcriptional complex containing MYB-bHLH-WD40

(MBW) [1, 13–15]. Expression of early biosynthesis genes (EBGs) such as CHS and CHI, is regulated by MYB11, MYB12 and MYB111, whereas PAP1 to PAP4 (AtMyb75, AtMyb90, AtMyb113, and AtMyb114) control expression of late biosynthesis genes (LBGs) including DFR, ANS, and UFGT [16, 17]. In Arabidopsis, up-regulation of one of MYB75, MYB90, MYB113 and MYB114 genes is sufficient to increase anthocyanin accumulation in young leaves [15, 17]. For example, the well-known Arabidopsis AtMYB75 (PAP1) gene directs anthocyanin production in leaves, roots, flowers, and fruits [18, 19]. Overexpression of AtMyb75 in Arabidopsis and tobacco results in upregulation of PAL, CHS and DFR genes [19, 20]. Similarly, AtMYB75 induces anthocyanin production in tomato (*Solanum lycopersicum* L.) via promoting the DFR expression [18]. Furthermore, the sequence variation of AtMYB90 (PAP2) is causal for natural variation in anthocyanin accumulation [17, 21]. AtMYB90 may act together with TTG1 (a WD40 protein) and different bHLH partners including TT8, GL3 or EGL3 [15, 18, 22]. Moreover, in Arabidopsis the ternary complexes formed by R2R3-MYB, bHLH and the WD repeat protein activate the biosynthetic genes required for proanthocyanidin accumulation in the innermost cell layer of the seed coat [23]. The R2R3 MYB protein TT2 (MYB123) is also a key regulator of proanthocyanidin accumulation in developing seeds [24].

MYB transcription factors are involved in regulation of anthocyanin synthesis in many plants, such as Arabidopsis [25], snapdragon (*Antirrhinum majus*) [26], apple (*Malus × domestica*) [27, 28], chrysanthemum [10], grape hyacinth (*Muscari armeniacum*) [29], grapevine (*Vitis vinifera*) [30], Chinese bayberry (*Myrica rubra*) [6], Epimedium sagittatum [31–33], hop (*Humulus lupulus* L.) [34], tomato (*Solanum lycopersicum* L.) [18], and peach (*Prunus persica*) [14]. In addition, some MYB genes are up-regulated under various stress conditions [15, 17]. However, the transcriptional regulation of anthocyanin synthesis by MYB in stress-tolerant plants is not well studied. *Eutrema salsugineum* (salt cress), a stress-tolerant model halophyte, is highly tolerant to cold, salt, drought, oxidative stress, and nitrogen deficiency. In *Eutrema salsugineum*, many stress-tolerant related genes, such as SOS1, HKT1, and nsLTP4, have been identified [35–40]. Here, we reported our functional analysis of the MYB transcription factor EsMYB90 in anthocyanin synthesis. Our phylogenetic and localization analyses suggest that EsMYB90 is an R2R3 type of MYB transcriptional factor. Ectopic expression of EsMYB90 in tobacco and Arabidopsis led to significantly increased pigmentation and production of anthocyanins in leaves, stems, and flowers. Our further RNA-seq and qRT-PCR analyses showed that EsMYB90 promoted expression of anthocyanin early biosynthesis genes (EBGs: NtCHS, NtCHI, and NtF3H) and late biosynthesis genes (LBGs: NtDFR, NtANS, and NtUFGT) in 35S:EsMYB90 tobacco transgenic plants. Our study identified a novel MYB transcription factor, which plays an important role in plant anthocyanin biosynthesis.

Results

Database mining identifies *EsMYB90*, a candidate regulator for anthocyanin synthesis

Eutrema Salsugineum is a stress-tolerance halophyte, which produces purple flower buds after vernalization [41]. Since *MYB* genes are required for anthocyanin synthesis [13,15,29] we identified which *MYB* controls this purple phenotype in *E. salsugineum*. After comparing *MYB* genes obtained from the

transcriptome of *E. salicarpineum* with 72 MYB genes known acting as proanthocyanin (PA) and anthocyanin regulators in other plants, we found one candidate MYB gene, named as *EsMYB90*.

To determine the relationship of *EsMYB90* to characterized flavonoid and PA MYBs, we performed similarity analysis at the protein level. Our results showed that *EsMYB90* has 80.5%, 78.9%, 78.4%, 74.4%, 69.4%, 65.9%, 50% identities respectively to 7 MYB proteins, i.e. *BoMYB1*, *AtMYB90*, *BtMYB114*, *AtMYB75*, *CtMYB114*, *AtMYB113*, and *AtMYB114* (Fig. 1A). In addition, similarities between *EsMYB90* and other 10 MYB proteins range from 44.1% to 39.0% (Fig. 1A). Those MYB proteins with high similarities to *EsMYB90* belong to the class of R2R3-MYB which have a conserved DNA-binding domain (R2 and R3 repeats) in the N-terminal and a variable C-terminal region [33,42,43]. The ANDV motif (marked by red A box in Fig.1A), a characteristic identifier for anthocyanin-promoting MYBs in dicots [10], existed in *EsMYB90*, *AtMYB90*, *AtMYB75*, *AtMYB113*, *AtMYB114*, *AmROSEA1*, *StMYB113*, *EsMYBA1*(AGT39060), *VvMYBA1*, *MtMYB1*, and *FaMYB10*, while the C-terminal-conserved motif KPRPR [S/T]F for *Arabidopsis* anthocyanin-promoting MYBs [25] (marked by blue B box in Fig.1A), was only found in *EsMYB90*, *AtMYB75*, *AtMYB90*, *AtMYB113*, and *MtMYB1*. Moreover, *EsMYB90* has a conserved [D/E]Lx2[R/K]x3Lx6Lx3R motif [10] (marked by black arrows in Fig.1A), which is required for interaction with R/B-like bHLH proteins [10].

To further identify the relationship of *EsMYB90* to other MYB proteins, we generated a phylogenetic tree with 30 MYB proteins involved in anthocyanin synthesis in 18 plants. Our results demonstrated that *EsMYB90* was clustered in the clade I (Fig. 1B), which consists of *AtMYB75*, *AtMYB90*, *AtMYB113*, and *AtMYB114* that are important for anthocyanin accumulation [17,19,25,44]. *EsMYB90* has a relatively farer phylogenetic relationship to MYB proteins in clades II, III, and IV, although those MYBs promote biosynthesis of PA and anthocyanin, except for *BtMYB114* and *EsMYB5* have no research reports yet.

In summary, our results suggest that *EsMYB90* is an R2R3-MYB, which may function in proanthocyanin and anthocyanin synthesis.

***EsMYB90* is localized to the nucleus**

To test the subcellular localization of *EsMYB90*, we examined the transient expression of YFP-*EsMYB90* fusion protein in onion epidermal cells. Our results showed that YFP signals were observed in both cytoplasm and nucleus of the onion epidermal cells expressing *35S:YFP* (Fig. 2A-C), while the YFP signal was only detected in the nucleus in cells expressing *35S:YFP-EsMYB90* (Fig. 2D-F). Our result showed *EsMYB90* is localized to the nucleus, suggesting that as other MYB proteins *EsMYB90* also functions as a transcriptional regulator.

Ectopic expression of *EsMYB90* promotes anthocyanin accumulation in tobacco and *Arabidopsis*

To investigate the possible function of *EsMYB90* in anthocyanin biosynthesis, we generated the *35S:EsMYB90* construct to ectopically expression *EsMYB90* in tobacco and *Arabidopsis* (Additional file 1:

Fig.1A). Eighteen *35S:EsMYB90* transgenic tobacco and 15 *35S:EsMYB90* transgenic *Arabidopsis* plants were obtained, respectively (Additional file 1: Fig. 1B,C).

We found that in all developing stages, leaves and stems in *35S:EsMYB90* tobacco plants appeared purple-red, and the color became deepened with development (Fig. 3A-C). In addition, *35S:EsMYB90* tobacco plants produced purple-red corollas, purple-black sepals, and purple-black fruit pods, whereas wild-type corollas were pink, with green sepals and fruit pods (Fig. 3D-F). Our results from examining anthocyanin production showed that the total anthocyanin contents in three *35S:EsMYB90* tobacco lines were significantly increased in stems, young leaves (YL), mature leaves (ML), flowers, fruit pods, and mature seeds, compared with the wild type (Fig. 3G). Among L1, L2 and L4 three tested lines, the L4 transgenic line had the highest anthocyanin contents. Compared with the wild type, the total anthocyanin contents in young leaves (YL), mature leaves (ML), stems, flowers, fruits pods, and mature seeds of the L4 line were increased 95.2, 45.7, 48.8, 4.9, 17.8, and 2.6 folds, respectively (Fig.3G). These results indicate that the enhanced pigmentation in *35S:EsMYB90* tobacco plants was caused by the increased synthesis of anthocyanins.

We observed similar phenotypes in *35S:EsMYB90 Arabidopsis* transgenic plants. In comparison to the wild-type plants, the color of leaves, roots, stems, flowers, fruit pods, and seeds became light-purple to dark-purple in *35S:EsMYB90 Arabidopsis* plants (Fig. 4A-G). In particularly, seeds from *35S:EsMYB90 Arabidopsis* plants exhibited black color (Fig. 4H,I). Furthermore, the contents of anthocyanins in the roots, stems, leaves, flowers, and fruit pods at the bolting stage, and the mature seeds from three *35S:EsMYB90 Arabidopsis* transgenic lines (L1, L2, and L3) were significantly higher than that in wild-type plants (Fig. 4J).

Collectively, our results suggest that *EsMYB90* functions as a transcription factor to promote anthocyanin biosynthesis in plants.

Transcriptomic analyses show that *EsMYB90* is a key regulator in the proanthocyanidin and anthocyanin pathway

To examine the molecular mechanisms by which *EsMYB90* controls anthocyanin biosynthesis in the genome wide, we performed RNA-seq analysis using the leaves from wild-type and *35S:EsMYB90* tobacco transgenic plants. We identified 51,202 differentially expressed genes (DEGs) in the comparison of wild-type plant with *35S:EsMYB90* transgenic tobacco plants, among which 2,446 DEGs have $\log_2$ Fold Change ≥ 1 or ≤ -1 and a $P_{adj} \leq 0.05$ (Additional file 2). Furthermore, 1199 out of 2446 DEGs were up-regulated, while 1247 genes were down-regulated (Additional file 2). Moreover, 476 unique DEGs were annotated into 43 GO terms, wherein the GO terms with the top 3 of the number of DEGs encoding the binding (249 genes), catalytic activities (237 genes) and metabolic processes (236 genes) (Fig. 5A, Additional file 3).

We revealed that among 2446 DEGs, 1023 unique genes were annotated into the 128 KEGG pathways (Additional file 4). The most prominent KEGG-enriched genes are involved in secondary metabolite

biosyntheses, followed by plant hormone signaling, flavonoid biosynthesis (Additional file 1: Fig. 2). According to the q value of DEGs, the top 20 of enrichment paths included the flavonoid biosynthesis (ko00941), anthocyanin biosynthesis (ko00942) flavone and flavonol biosynthesis (ko00944), plant circadian rhythm (ko04712), and glutathione metabolism (ko00480; Fig. 5B, Additional file 5). Moreover, the anthocyanin biosynthesis pathway has the largest enrichment factor, followed by the pathway of flavonoid biosynthesis. The flavone and flavonol biosynthesis pathways also has a large enrichment factor (Fig. 5B, Additional file 5).

Mapping to the KEGG reference pathways found that a total of 57 significantly differential expression genes were assigned to five secondary metabolic pathways, i.e. phenylpropanoid biosynthesis (ko00940), flavonoid biosynthesis (ko00941), anthocyanin biosynthesis (ko00942), isoflavonoid biosynthesis (ko00943), and flavone and flavonol biosynthesis pathways (ko00944). The gene names, gene ID, and the combined functional annotations were seen in Additional file 6. Out of 57 genes, 42 genes encode PA and anthocyanin biosynthesis enzymes, such as PAL (107802063, 107761482 and 107820497), CHS (107826422, 107801774 and 107813613), CHI (107779699, 107810515 and 107825576), F3H (107770893, 107806462), DFR (107803097, 107797232), and ANS/ LDOX (107819370, 107778118, 107787193, 107787195 and 107808500). Particularly, six genes encoding UFGT (107781346, 107822886, 107781522, 107831042, 107767212 and 107819220) in the anthocyanin biosynthesis pathway (ko00942) were all strongly up-regulated (Fig. 6, Additional file 6). However, in ko00940-ko00944 pathways, only 15 genes including that encoding flavonol synthase/flavanone 3-hydroxylase (107794305, 107814657), trans-resveratrol di-O-methyltransferase-like (107785995, 107797481), and flavone 3'-O-methyltransferase 1-like (107792977), were down-regulated (Fig. 6, Additional file 6).

Taken together, our RNA-seq results demonstrated that identified DEGs are significantly enriched in the flavonoid and anthocyanin synthesis pathway (ko00941-ko00944), suggesting that *EsMYB90* play an important regulatory role in proanthocyanidin and anthocyanin synthesis.

Validation of RNA-seq results by qRT-PCR

To validate the RNA-seq results, we performed quantitative reverse transcription PCR (qRT-PCR) for 18 genes which are assigned to 5 groups related to anthocyanin biosynthesis, antioxidant production, signal transduction, transcription regulation, and ion channel in tobacco (Additional file 7). Our results showed that expression level changes of 5 anthocyanin biosynthesis genes [*NtDFR* (107803097), *NtLDOX54* (107778118), *Nt3GT12* (107781346), *Nt3GT36* (107781522), and *Nt3GT53* (107831042)] detected by qRT-PCR were in agreement with the RNA-seq data (Fig. 7A). We obtained similar RT-qPCR results from examining expression of 4 antioxidant-related genes [*NtP450* (107772738), *NtCu-ZnSOD* (107806960), *NtPOD44-1* (107827231), and *NtPOD44-2* (107797651); Fig. 7B], 4 genes encoding transcription factors [*NtbZIP* (107795590), *NtMYB3R-1* (107795213), *NtMYB4* (107802984), and *NtWRKY53* (107825953); Fig. 7C], *NtAKT2/3* (107761230) encoding a potassium channel protein (Fig. 7C), and 4 genes related to signal transduction and ion channel (*NtMAPK3* (107782983), *NtMAPK6* (107806359), *NtAX15A*

(107805986), and *NtCaM1* (107803626); Fig.7D]. Therefore, qRT-PCR results support that our transcriptome results are reliable.

***EsMYB90* promotes expression of anthocyanin biosynthetic genes in tobacco**

To further elucidate the molecular function of *EsMYB90* in proanthocyanin and anthocyanin biosynthesis, we examined expression of key anthocyanin biosynthesis genes *PAL*, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, *ANS/LDOX*, and *UFGT* in the stems (S), leaves (L) and flowers (F) from *35S:EsMYB90* tobacco transgenic lines (L2, L4) and wild-type tobacco plants at the flowering stage by qRT-PCR.

PAL is the first key enzyme in the metabolic pathway of phenylpropanoid [9]. Expression levels of *NtPAL* in stems, leaves and flowers from the L4 line increased 4.6, 6.1, and 6.3 times, respectively, than that of wild type (Fig. 8A). *CHS* catalyzes the first step of anthocyanin biosynthesis, while *CHI* catalyzes the cyclization of chalcone molecules to form naringenin [8]. Expression levels of both *NtCHS* and *NtCHI* in stems, leaves and flowers from the L4 line were significantly increased compared to that in the wild type (Fig. 8B,C). In addition, our results showed that *EsMYB90* enhanced expression of *NtF3H* and *NtF3'H* genes (Fig. 8D,E). Finally, the anthocyanin biosynthesis genes *NtDFR*, *NtANS* and *NtUFGT* which are required for anthocyanin biosynthesis at later steps were also significantly upregulated by *EsMYB90* (Fig. 8F-H).

In summary, our results suggest that *EsMYB90* controls anthocyanin biosynthesis by promoting expression of anthocyanin biosynthesis genes, particularly LBGs.

Discussion

Anthocyanins are main contributors to the coloration of plants, and the color is an important determinant for fruit and flower quality [8, 14, 45]. Flavonoids such as anthocyanins are also capable of scavenging oxygen free radicals produced in cells, thus have antioxidant activity [46, 47]. The antioxidant activity of anthocyanins depends on the degree of B-ring hydroxylation, acylation, and glycosylation [45]. Anthocyanins, as an antioxidant, can reduce the peroxidation of lipids and delay the aging of cells. Meanwhile, anthocyanins protect plants from damage caused by biotic and abiotic stresses and allow plants to adapt to environmental changes [48]. In higher plants, PAs and anthocyanin biosynthesis are regulated by different sets of MYB-bHLH-WD40 (MBW) complexes, and the R2R3-MYBs play vital roles in transcriptional regulation of anthocyanins biosynthesis [12, 19, 20, 23, 25, 29, 32, 49]. Our results shed light on the molecular mechanism by which a novel R2R3-MYB controls PA and anthocyanin biosynthesis via promoting expression of PA and anthocyanin biosynthesis genes.

In *Arabidopsis*, MYB75, MYB90, MYB113, and MYB114 function redundantly by participating in the MBW (MYB-bHLH-WD40) complex to regulate PA and anthocyanin biosynthesis. Overexpression of one of these four MYBs is sufficient to increase anthocyanin accumulation in young leaves [15, 17]. Our study showed that *EsMYB90* is an R2R3-MYB, which contains a conserved [D/E]Lx2[R/K]x3Lx6Lx3R motif required for interaction with R/B-like bHLH proteins, and the ANDV and KPRPR [S/T]F motifs characterizing

anthocyanin-promoting MYBs (Fig. 1A). However, conserved ANDV and KPRPR [S/T]F domains were not found in other five identified MYBs (EsMYB5, EsMYB15, EsMYB106, EsMYB108, and MYB-related protein 340), and [D/E]Lx2[R/K]x3Lx6Lx3R motif were not found in other four identified MYBs (EsMYB15, EsMYB106, EsMYB108, and MYB-related protein 340) in *E. salsugineum* (Additional file 1: Fig. 3). In *Arabidopsis* and some other higher plants, the anthocyanin pathway is controlled by multiple MYB transcriptional factors [1]. For instance, at least four MBW complexes assembled with various MYBs are involved in the PA accumulation in the innermost cell layer of *Arabidopsis* seed coat [23]. Thus, our results suggest that different from other plants less MYBs regulate anthocyanin biosynthesis in *E. salsugineum* and EsMYB90 is a major player.

The regulatory mechanisms for flavonoid biosynthesis are conserved in higher plants [1]. The MYB proteins usually determine the involvement of MBW complexes in specific pathways [1, 50]. In the MBW complex, MYBs have the highest binding specificity compared with bHLH and WD40. MYB and bHLH bind to at least one of three conserved cis elements, i.e. MYB-core, AC-rich, and E/G-box, in promoters of four LBGs (late biosynthesis genes: DFR, TT19, TT12, and AHA10), which specify their expression in the seed coat of *Arabidopsis* [23, 50]. Expression of LBGs are regulated by MYB75, MYB90, MYB113 and MYB114 in *Arabidopsis* [15, 16]. Our phylogenetic analysis showed that EsMYB90 was in the same clade with *Arabidopsis* MYB75, MYB90, MYB113, and MYB114, suggesting that EsMYB90 may play a more important role in regulation of LBGs (DFR, ANS/LDOX and UFGT) expressions, which is in accordance with our research results. In addition, our results show that EsMYB90 promotes expression of early biosynthesis genes, such as PAL, CHS, CHI, and F3H which agrees with findings in other plants [1, 10, 29]. Furthermore, it is possible that EsMYB90 upregulates expressions of NtMYB and NtbHLH genes involved in anthocyanin biosynthesis, because expressions of NtMYB3 (107820930), NtMYB4 (107802984, 107769018, 107815562, 107760435), NtMYB44 (107816351, 107759750), MYB-like (107807565, 107795213), bHLH (107791671, 107791200), NtbHLH92 (107785534), NtbHLH18 (107811232), and NtbHLH35 (107774314) were significantly increased in 35S:EsMYB90 tobacco transgenic plants (Additional file 2).

Overall, our results showed that ectopic expression of the novel EsMYB90 gene can strongly induce the anthocyanin biosynthesis by promoting expression of anthocyanin biosynthesis pathway genes, particularly functioning in the LBGs. Our study also paved the way for the application of EsMYB90 to genetically engineering crops and ornamental plants.

Conclusions

EsMYB90, a R2R3 MYB transcription factor, is localized in nucleus. *35S:EsMYB90* transgenic tobacco and *Arabidopsis* showed purple-red, purple-black phenotypes, and accumulated more anthocyanin in the leaves, stems and flowers compared with wild type. The results showed that ectopic expression of *EsMYB90* in stems, leaves and flowers of transgenic tobacco could significantly enhanced the expression of anthocyanin biosynthetic genes including EBGs (*PAL*, *CHS*, and *CHI*) and LBGs (*DFR*, *ANS*, and *UFGT*),

particularly in EBGs. The study suggested that *EsMYB90* plays a key role in regulating anthocyanin biosynthesis, and it provide new clues to increase the content of anthocyanin in transgenic plants.

Methods

Plant materials and growth conditions

The seeds of *Eutrema Salsugineum* (Shandong ecotype), *Arabidopsis thaliana* (Landsberg ecotype) and tobacco (*Nicotiana tabacum*) are preserved and presented by Shandong Provincial Key Laboratory of Plant Stress Research, College of Life Science, Shandong Normal University.

Wild-type and transgenic plants of *Nicotiana tabacum* and *Arabidopsis thaliana* (Columbia-0) were grown in a mixture of vermiculite, perlite and peat moss (1:1:1) in a greenhouse with 25°C, a photoperiod of 16 h light /8 h dark. *Eutrema Salsugineum* were grown in a growth chamber with 22°C, a photoperiod of 16 h light /8 h dark and 70% relative humidity.

T3 generation homozygous transgenic and wild-type tobacco plants were used for analyses. After 8 weeks growth of tobacco plants at the period of 8-9 leaves, the sixth leaves of wild-type and transgenic plants were collected for transcriptome and qRT-PCR experiments. Three biological repeats for transcriptome and qRT-PCR were performed. In tobacco, stems, young leaves (YL), mature leaves (ML), flowers, and fruit pods at the flowering stage, and mature seeds were used to determine the content of anthocyanins and examine expression of anthocyanin biosynthesis genes. Similarly, *Arabidopsis* roots, stems, leaves, flowers, and fruit pods from plants growing 4 weeks, and mature seeds were used to determine the content of anthocyanins. In all cases, samples were frozen immediately in liquid nitrogen and stored at -80°C. Three repeats of all tests were conducted.

Phylogenetic analysis

A set of associated MYB protein sequences were downloaded from the NCBI and multiple sequence alignments were carried out using the DNAMAN software (Version 5.2.2). The phylogenetic tree was constructed using the MEGA6 software with the neighbor-joining method.

Subcellular localization of *EsMYB90*

The coding region of *EsMYB90* was PCR-amplified by YFPMYB-F and YFPMYB-R primers with *SacI* and *SacI* sites, respectively (Additional file 9) and then cloned in to the YFP-pCAT vector, resulting in the *35S:YFP-EsMYB90-pCAT* vector. The onion epidermal cells were transformed with the *35S:YFP-EsMYB90-pCAT* and *YFP-pCAT* (control) plasmids, respectively, using the plasmid bombardment method [51]. The transformed cells were incubated at 25 ° C for 16-24 hours, then the florescence signals were observed and recorded with a fluorescence microscope (BX51, model 7.3, Olympus, Tokyo, Japan). At least three replicates for each construct were performed.

RNA isolation and cDNA synthesis

Total RNAs were isolated from 0.5g of young leaves of *Eutrema Salsugineum* and *Nicotiana tabacum* using the TRIzol Reagent (Life Technologies) or Quick RNA isolation kit (Biotech biotechnology company, Beijing, China). The RNA concentration and purity (A260/A280) were measured with a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). First-strand cDNA was synthesized using the Rever Tra Ace qPCR RT Master Mix with gDNA Remover (Toyobo, Japan). Briefly, RNA was heat-denatured at 65 °C for 5 min and immediately placed on the ice for cooling, then 2 µL of 4×DN Master Mix, 0.5pg-0.5µg of RNA template and nuclease-free water were added to 8 µL of volume. After 5 min of incubation at 37°, 5×RT Mater Mix II was added to make 10 µL of final volume. RT reaction was carried out at 37 °C for 15 min, 50 °C for 5 min, 98 °C for 5 min to obtain 10 µl of cDNA.

Construction generation and plant transformation

The whole coding sequence of *EsMYB90* was PCR-amplified from the *Eutrema Salsugineum* cDNA using the forward primer 5'-CCGGAATTCTTTAGAATACTTATTGGTCC-3' and the reverse primer 5'-CGCGGATCCATCAGAGACAGATATTAGTTGG-3' with *EcoR* I and *Bam*H I restriction enzyme sites at the 5' and 3', respectively (Additional file 9). The resulting *EsMYB90* fragment was cloned into the pMD18-T vector (Takara, USA). After sequencing confirmation, the *EsMYB90* fragment was subcloned into the *Eco*RI - *Bam*HI sites of the pCAMBIA3301H vector, where *EsMYB90* was under the control of the *CaMV* 35S promoter. The expression vector (*35S:EsMYB90-pCAMBIA3301H*) was finally introduced into the *Agrobacterium tumefaciens* strain GV3101.

Transformation of *N. tabacum* was performed using the leaf disc method essentially as reported by Horsch *et al.* [52]. Transgenic tobacco seedlings were selected on the MS medium containing 6 mg / L of bastar and 300 mg / L of cefalexin. Transformation of *A. thaliana* Columbia-0 was performed using the floral-dipping method [53] and transformants were screened by spraying 0.1% of bastar herbicide. The presence of the transgene was further confirmed by PCR using specific primers for *EsMYB90*. The homozygous transgenic *N. tabacum* and *A. thaliana* were used for subsequent phenotypic and functional analysis.

Anthocyanin analysis

Stems, young leaves (YL), mature leaves (ML), flowers, and fruit pods of tobacco growing about 8 weeks at the flowering stage, and the mature seeds were sampled, respectively. Similarly, stems, leaves, flowers, fruit pods, and roots of *Arabidopsis* growing about 4 weeks at the bolting period, as well as the mature seeds were also collected. All materials were frozen immediately in liquid nitrogen and ground to powders. The anthocyanin content was determined using an improved method described by Neff and Chory [54]. The measurements of A_{530} and A_{657} were conducted with a spectrophotometer (UV-1800, Shimadzu). The results were calculated by the equation $(A_{530}-0.25 \times A_{657})/\text{fresh weight}$. Three replicates were performed for each sample.

RNA-seq and bioinformatic analysis

Total RNAs from the sixth leaves of wild-type and *EsMYB90* transgenic tobacco at the 7-8 leaves stage were isolated using a Quick RNA isolation kit (Bioteke Corporation, Beijing, China). The RNA library construction and sequencing were performed in the BGI Corporation (Shenzhen, China) using the BGISEQ-500 platform. Three independent biological replicates were carried out.

The low-quality reads (more than 20% of the bases qualities are lower than 10), reads with adaptors and reads with unknown bases (N bases more than 5%) were filtered to get the clean reads. The clean reads were mapped to the reference genome using HISAT [55]. Meanwhile, the clean reads were mapped to the reference transcripts using Bowtie2 [56]. The clean reads were assembled into unigenes, followed by the unigene functional annotation, *etc.*, and calculate the unigene expression levels of each sample [57]. Finally, we identified DEGs (differential expressed genes) and performed clustering analysis and functional annotations. DEGs with the GO and KEGG annotation results were classified according to the official classification, and the GO and KEGG pathway functional enrichment were performed using phyper in the R software. Transcription Factor Prediction of DEG: The ORF of each DEG were founded using getorf and aligned to TF domains (from PlntfDB) using hmmsearch [58].

Gene expression analysis by qRT-PCR

To validate the transcriptome results, the real-time qPCR was performed by the LightCycler® 96 system to examine expressions of selected genes (Additional file 9) using the total RNAs extracted from tobacco leaves used for RNA-seq. Transcriptome data was verified by comparing the result of qRT-PCR ($-\Delta\Delta CT$) with RNA-seq ($\log_2 FC$).

To examine expression of anthocyanin biosynthesis genes in different tissues, stems, leaves and flowers of tobacco at the flowering stage were sampled, respectively. The total RNAs were extracted using the total RNA rapid extraction kit (Biotech biotechnology company, China). The first-strand cDNAs were synthesized using the Rever Tra Ace qPCR RT Master Mix with gDNA Remover (Toyobo, Japan). The qPCR was performed using the LightCycler® 96 system (Roche, Switzerland; Supplementary Table S1). *Ntubi2* (*ubiquitin 2*, LOC107772211) was used as an internal reference gene. Four replicates were performed for each sample.

Abbreviations

PAs: proanthocyanidins; PAL: phenylalanine ammonia-lyase; CHS: chalcone synthase; CHI: Chalcone isomerase; DHK: dihydrokaempferol; F3H: flavanone 3 β -hydroxylase; DHQ: dihydroquercetin; F3'H: flavonoid 3'-hydroxylase; DHM: dihydromyricetin; F3'5'H: flavonoid 3',5'-hydroxylase; DFR: Dihydroflavonol 4-reductase; ANS: anthocyanidin synthase; UFGT: UDP-glucose: flavonoid 3-O-glucosyltransferase; MBW: MYB-bHLH-WD40; EBGs: early biosynthesis genes; LBGs: late biosynthesis genes; YL: young leaves; ML: mature leaves; DEGs: differentially expressed genes; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; qRT-PCR: quantitative reverse transcription PCR.

Declarations

Acknowledgments

We thank the anonymous reviewers for their valuable comments and suggestions.

Authors' contributions

QZ conceived and designed the experiments. QZ wrote the manuscript with contributions from DZ, CM, XW, SG, and CG. YQ, CG, SG, CL, CZ, and CL carried out most of the experiments. CL and CZ took care of the plants. QZ, YQ, CG and CL analyzed the data. All authors read and approved the final manuscript.

Funding

This study was supported by the National Major Science and Technology Project of China (2018ZX08009-10B), Key Technology Research and Development Program of Shandong (2019GSF107089), and Natural Science Foundation of Shandong Province (ZR2014CM041).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Additional files

- Additional file 1: Fig.1. Molecular identification of *EsMYB90* transgenic tobacco and *Arabidopsis* plants.
- Additional file 1: Fig.2. Pathways of the most enriched KEGG DEGs in RNA-Seq of *EsMYB90* transgenic tobacco.
- Additional file 1: Fig.3. Sequence alignment analysis of the *EsMYB90* and other *EsMYB* proteins in *E.salsugineum*.
- Additional file 2. Differentially expressed genes(DEGs) that meet $\log_2$ Fold Change ≥ 1 or ≤ -1 and a $P_{adj} \leq 0.05$ in RNA-Seq of *EsMYB90* transgenic tobacco.

- Additional file 3. Annotated DEGs in GO terms in RNA-Seq of *EsMYB90* transgenic tobacco.
- Additional file 4. Annotated DEGs in KEGG in RNA-Seq of *EsMYB90* transgenic tobacco.
- Additional file 5. Top 20 paths according to the RichFactor value of DEGs.
- Additional file 6. DEGs enriched in the flavonoid and anthocyanin synthesis pathways.
- Additional file 7. Detailed data for validation of RNA-Seq results using quantitative real-time PCR.
- Additional file 8. Detailed data for the expression analysis of anthocyanin biosynthesis genes in stem, leaf and flower from two *EsMYB90* transgenic tobacco lines and WT.
- Additional file 9. Sequences of the primers used in the study.

References

1. Xu WJ, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 2015;20(3):176-185.
2. Lou Q, Liu YL, Qi YY, Jiao SZ, Tian FF, Jiang L, Wang YJ. Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in *Grape hyacinth*. *J Exp Bot.* 2014;65(12):3157-3164.
3. Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: Biosynthesis, biological functions and biotechnological applications. *Front Plant Sc*2012;3(222): 222.
4. Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J Exp Bot.* 2011;62(8):2465-2483.
5. Zhang Y, Butelli E, Martin C. Engineering anthocyanin biosynthesis in plants. *Curr Opin Plant Biol.* 2014;19:81-90.
6. Niu SS, Xu CJ, Zhang WS, Zhang B, Li X, Lin-Wang K, Ferguson IB, Allan AC, Chen KS. Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (*Myrica rubra*) fruit by a R2R3 MYB transcription factor. *Planta.* 2010;231(4):887-899.
7. Petroni K, Tonelli C. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 2011;181(3):219-229.
8. Jaakola L. New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 2013;18(9):477-483.
9. Jones DH. Phenylalanine ammonia-lyase: Regulation of its induction, and its role in plant development. *Phytochem.* 1984.23(7):1349-1359.
10. Liu XF, Xiang LL, Yin XR, Grierson D, Li F, Chen KS. The identification of a MYB transcription factor controlling anthocyanin biosynthesis regulation in *Chrysanthemum* *Sci Hortic.* 2015;194:278-285.
11. Takos AM, Jaffé FW, Jacob SR, Bogs J, Robinson SP, Walker AR. Light-induced expression of a *MYB* gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* 2006;142(3):1216-1232.
12. Pérez-Díaz JR, Pérez-Díaz J, Madrid-Espinoza J, González-Villanueva E, Moreno Y, Ruiz-Lara S. New member of the R2R3-MYB transcription factors family in grapevine suppresses the anthocyanin

- accumulation in the flowers of transgenic tobacco. *Plant Mol Biol*. 2016;90(1-2): 63-76.
13. Feng SQ, Wang YL, Yang S, Xu YT, Chen X Anthocyanin biosynthesis in pears is regulated by a R2R3-MYB transcription factor PyMYB10. *Planta*. 2010;232(1):245-255.
 14. Zhou H, Peng Q, Zhao JB, Owiti A, Ren F, Liao L, Wang L, Deng XB, Jiang Q, Han Y Multiple R2R3-MYB transcription factors involved in the regulation of anthocyanin accumulation in peach flower. *Front Plant Sci*. 2016;7(28): 1557.
 15. Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* *Plant J*. 2008;53(5):814-827.
 16. Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like bHLH proteins. *Plant J*. 2004;40(1):22-34.
 17. Bac-Molenaar JA, Fradin EF, Rienstra JA, Vreugdenhil D, Keurentjes JJ. GWA mapping of anthocyanin accumulation reveals balancing selection of MYB90 in *Arabidopsis thaliana*. *PLoS ONE*. 2015;10(11):e0143212.
 18. Zuluaga DL, Gonzali S, Loreti E, Pucciariello C, Degl'Innocenti E, Guidi L, Alpi A, Perata P. *Arabidopsis thaliana* MYB75/PAP1 transcription factor induces anthocyanin production in transgenic tomato plants. *Funct Plant Biol*. 2008;35(7):606-618.
 19. Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell*. 2000;12:2383-2393.
 20. Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K. Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. *Plant J*. 2005;42(2):218-235.
 21. Ilk N, Ding J, Ihnatowicz A, Koornneef M, Reymond M. Natural variation for anthocyanin accumulation under high light and low temperature stress is attributable to the enhancer of AG-42 (*hua2*) locus in combination with production of anthocyanin pigment1 (PAP1) and PAP2. *New Phytol*. 2015;206(1):422-435.
 22. Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M. Genetics and biochemistry of seed flavonoids. *Annu Rev Plant Biol*. 2006;57(1):405-430.
 23. Xu WJ, Grain D, Bobet S, Le Gourrierc J, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in *Arabidopsis* *New Phytol*. 2014;202(1):132-144.
 24. Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L. The *Arabidopsis* *TT2* gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell*. 2001;13(9):2099-2114.
 25. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol*. 2001;4(5):447-456.

26. Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. A small family of myb-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell*. 2006;18(4):831-851.
27. Peng T, Moriguchi T. The molecular network regulating the coloration in apple. *Sci Hortic*. 2013;163:1-9.
28. Telias A, Lin-Wang K, Stevenson DE, Cooney JM, Hellens RP, Allan AC, Hoover EE, Bradeen JM. Apple skin patterning is associated with differential expression of MYB10. *BMC Plant Biol*. 2011;11(1):93.
29. Chen KL, Liu HL, Lou Q, Liu Y Ectopic expression of the grape hyacinth (*Muscari armeniacum*) R2R3-MYB transcription factor gene, MaAN2, induces anthocyanin accumulation in tobacco. *Front Plant Sci*. 2017;8:965.
30. Terrier N, Torregrosa L, Ageorges A, Vialet S, Verriès C, Cheynier V, Romieu C. Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiol*. 2009;149(2):1028-1041.
31. Huang WJ, Sun W, Lv HY, Luo M, Zeng SH, Pattanaik S, Yuan L, Wang Y. A R2R3-MYB transcription factor from *Epimedium sagittatum* regulates the flavonoid biosynthetic pathway. *PLoS ONE*. 2013;8(8):e70778.
32. Huang WJ, Khaldun ABM, Chen JJ, Zhang CJ, Lv HY, Yuan L, Wang Y. A R2R3-MYB transcription factor regulates the flavonol biosynthetic pathway in a traditional chinese medicinal plant, *Epimedium sagittatum*. *Front Plant Sci*. 2016;7:1089.
33. Huang WJ, Lv HY, Wang Y. Functional characterization of a novel R2R3-MYB transcription factor modulating the flavonoid biosynthetic pathway from *Epimedium sagittatum*. *Front Plant Sci*. 2017;8:1274.
34. Gatica-Arias A, Farag MA, Stanke M, Matoušek J, Wessjohann L, Weber G. Flavonoid production in transgenic hop (*Humulus lupulus* L.) altered by PAP1/MYB75 from *Arabidopsis thaliana* L. *Plant Cell Rep*. 2012;31(1):111-119.
35. Oh DH, Leidi E, Zhang Q, Hwang SM, Li YZ, Quintero FJ, Jiang XY, D'Urzo MP, Lee SY, Zhao YX, Bahk JD, Bressan RA, Yun DJ, Pardo JM, Bohnert HJ. Loss of halophytism by interference with SOS1 expression. *Plant Physiol*. 2009;151(11):210-222.
36. Inan G, Zhang Q, Li PH, Wang ZL, Cao ZY, Zhang H, Zhang CQ, Quist TM, Goodwin SM, Zhu JH, Shi HZ, Damsz B, Charbaji T, Gong QQ, Ma SS, Fredricksen M, Galbraith DW, Jenks MA, Rhodes D, Hasegawa PM, Bohnert HJ, Joly RJ, Bressan RA, Zhu JK. Salt cress. A halophyte and cryophyte *Arabidopsis* relative model system and its applicability to molecular genetic analyses of growth and development of extremophiles. *Plant Physiol*. 2004;135(3):1718-1737.
37. Wang JS, Zhang Q, Cui F, Hou L, Zhao SZ, Xia H, Qiu JJ, Li TT, Zhang Y, Wang XJ, Zhao C Genome-wide analysis of gene expression provides new insights into cold responses in *Thellungiella salsuginea*. *Front Plant Sci*. 2017;8:713.
38. Ali Z, Park HC, Ali A, Oh DH, Aman R, Kropornicka A, Hong H, Choi W, Chung WS, Kim WY, Bressan RA, Bohnert HJ, Lee SY, Yun DJ. TsHKT1;2, a HKT1 homolog from the extremophile *Arabidopsis* relative

Thellungiella salsuginea, shows K⁺ specificity in the presence of NaCl. Plant Physiol. 2012;158(3):1463-1474.

39. Sun W, Li Y, Zhao YX, Zhang H. The TsnsLTP4, a nonspecific lipid transfer protein involved in wax deposition and stress tolerance. Plant Mol Biol Rep. 2015;33(4):962-974.
40. Zhang Q, Zhao CZ, Li M, Sun W, Liu Y, Xia H, Sun MN, Li AQ, Li CS, Zhao SZ, Hou L, Picimbon JF, Wang XJ, Zhao Y Genome-wide identification of *Thellungiella salsuginea* microRNAs with putative roles in the salt stress response. BMC Plant Biol. 2013;13(1):180.
41. Griffith M, Timonin M, Wong AC, Gray GR, Akhter SR, Saldanha M, Rogers MA, Weretilnyk EA, Moffatt B. *Thellungiella*: an *arabidopsis*-related model plant adapted to cold temperatures. Plant Cell Environ. 2007;30(5):529-538.
42. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. Trends Plant Sci. 2010;15(10):573-581.
43. Quattrocchio F, Wing JF, van der Woude K, Mol JN, Koes R. Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J. 1998;13(4):475-488.
44. Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J. 1995;8(5):659-671.
45. Sadilova E, Stintzing FC, Carle R. Anthocyanins, colour and antioxidant properties of eggplant (*Solanum melongena* L.) and violet pepper (*Capsicum annuum* L.) peel extracts. Z Naturforsch C. 2006;61(7-8):527-535.
46. Hernández I, Alegre L, Van Breusegem F, Munné-Bosch S. How relevant are flavonoids as antioxidants in plants? Trends Plant Sci. 2009;14(3):125-132.
47. Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T, Matsuda F, Kojima M, Sakakibara H, Shinozaki K, Michael AJ, Tohge T, Yamazaki M, Saito K. Enhancement of oxidative and drought tolerance in *Arabidopsis* by overaccumulation of antioxidant flavonoids. Plant J. 2014;77(3):367-379.
48. Ahmed NU, Park JI, Jung HJ, Yang TJ, Hur Y, Nou IS. Characterization of dihydroflavonol 4-reductase (DFR) genes and their association with cold and freezing stress in *Brassica rapa*. Gene. 2014;550(1):46-55.
49. Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero JL, Schwab W, Muñoz-Blanco J, Blanco-Portales R. MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria x ananassa* J Exp Bot. 2014;65(2):401-417.
50. Heppel SC, Jaffé FW, Takos AM, Schellmann S, Rausch T, Walker AR, Bogs J. Identification of key amino acids for the evolution of promoter target specificity of anthocyanin and proanthocyanidin regulating MYB factors. Plant Mol Biol. 2013;82(4-5): 457-471.
51. Kinkema M, Fan WH, Dong X Nuclear localization of NPR1 is required for activation of *PR* gene expression. Plant Cell. 2000;12(12):2339-2350.

52. Horsch R, Fry JE, Hoffmann NL et al. A simple and general method for transferring genes into plants. *Science*. 1985;227:1229-1232.
53. Clough SJ, Bent AF. Floral dip : a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. 1998;16(6):735-743.
54. Neff MM, Chory J. Genetic interactions between phytochrome A, phytochrome B, and cryptochrome 1 during *Arabidopsis* Plant Physiol. 1998;118(1):27-36.
55. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Meth*. 2015;12(4):357-360.
56. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Meth*. 2012;9(4):357-359.
57. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12(1): 323.
58. 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.

Figures

Figure 1

(A) The EsMYB90 protein sequence was aligned with a set of related R2R3-MYB proteins from 10 plant species. Identical amino acids are shaded in dark blue, and the greater than or equal to the identity of 75%, 50%, and 33% are indicated in shade of pink, light blue, and yellow, respectively. R2 and R3 domains refer to two repeats of the MYB DNA binding domain. Box (A): the conserved motif ANDV in the R3 domain for dicot anthocyanin-promoting MYBs; Box (B): the C-terminal-conserved motif KPRPR [S/T]F for Arabidopsis anthocyanin-promoting MYBs; Black arrows: the specific residues of [D/E]Lx2[R/K]x3Lx6Lx3R that confer to the interaction with bHLH. (B) Phylogenetic analysis of EsMYB90

and other 29 MYB proteins. The MYB protein sequences were downloaded from the GenBank database with accession numbers showed in the diagram. EsMYB90 and EsMYB5 from *E. salsugineum* are indicated by bold blue and thin-lined blue frames. MYB proteins from the cruciferae plants in the clade I have higher identities to EsMYB90 than that in clades II to IV.

Figure 2

Transient expression of the 35S:YFP-EsMYB90 in onion epidermal cells showing EsMYB90 is localized in the nucleus. The control was 35S:YFP. (A) An onion epidermal cell expressing 35S:YFP showing YFP signals in both cytoplasm and nucleus; (B) An onion epidermal cell expressing 35S:YFP in the bright field; (C) The merged image of A and B; (D) An onion epidermal cell expressing 35S:YFP-EsMYB90 exhibiting the YFP signal only in the nucleus; (E) An onion epidermal cell expressing 35S:YFP-EsMYB90 in the bright field; (F) The merged image of D and E. Bars: 50 µm.

Figure 3

(A) Five-week-old wild-type (WT) and tobacco transgenic (T) seedlings; (B) Nine-week-old WT and tobacco T plants at the flowering stage; (C) Leaves from six-week-old WT as well as L1, L2 and L4 tobacco transgenic plants; (D) Sepal and corolla from nine-week-old WT and tobacco transgenic plants; (E) Flowers from nine-week-old WT as well as L1, L2 and L4 tobacco transgenic plants; (F) Fruit pods from nine-week-old WT and tobacco transgenic plants; (G) Anthocyanin contents in stems, young leaves (YL), mature leaves (ML), flowers, fruit pods of eight-week-old plants, and mature seeds from WT as well as L1, L2 and L4 tobacco transgenic plants. Vertical bars indicate standard errors of 3 biological replicates and the Student's t test values are indicated as c ($P < 0.001$).

Figure 4

(A) Seven-day-old wild-type (WT) and Arabidopsis transgenic (T) seedlings; (B) Three-week-old WT and Arabidopsis T plants; (C) Roots from seven-week-old WT and Arabidopsis T plants; (D) Leaves from three-week-old WT and Arabidopsis T plants; (E) Stems from seven-week-old WT and Arabidopsis T plants; (F) Flowers from seven-week-old WT and Arabidopsis T plants; (G) and (H) Fruit pods from seven-week-old WT and Arabidopsis T plants; (I) Seeds from nine-week-old WT and Arabidopsis T plants; (J) Anthocyanin contents in roots, stems, leaves, flowers, fruit pods of four-week-old plants and mature seeds, from WT as well as L1, L2 and L3 Arabidopsis transgenic plants. Vertical bars indicate standard errors of the 3 biological replicates and the Student's t test values are indicated as c (P<0.001).

Figure 5

(A) The annotated GO terms in DEGs. The X axis represents the number of DEGs. The Y axis indicates the classification of GO terms. (B) The top 20 pathways of enriched KEGG DEGs. The X axis represents the enrichment factor value, while the Y axis indicates the pathway names. The color of dots represents the q value. The size of dots represents the number of DEGs. The rich factor refers to the value of enrichment factor.

Figure 6

Purple solid circles indicate 107795677 and 107823192 genes, which also exist in the ko00944 pathway. Green diamonds indicate 107820810 and 107778746 genes, which can be also found in the ko00942 pathway. Blue triangles indicate 107781346 and 107822886 genes, which co-exist in the ko00944 pathway.

Figure 7

(A) Anthocyanin biosynthesis genes; (B) Antioxidant related genes; (C) Genes encoding transcription factor and ion channel; (D) Signaling pathway related genes. Ntubi2 (Ntubiquitin 2, LOC107772211) was used as an internal control. Each set of data were obtained from three repeats.

Figure 8

WT: Wild Type; L2 and L4: two independent 35S:EsMYB90 lines. Ntubi2 (Ntubiquitin 2, LOC107772211) was used as an internal control. (A) *NtPAL* (XM_016625506.1); (B) *NtCHS* (NM_001326166.1); (C) *NtCHI* (NM_001325287.1); (D) *NtF3H* (NM_001325083.1); (E) *NtF3'H* (NM_001325608.1); (F) *NtDFR* (NM_001325732.1); (G) *NtLDOX* (NM_001326043.1); (H) *NtUFGT* (NM_001325312.1). The $\Delta\Delta C_t$ method was used to determine the relative expression, and expression levels of tested genes in the wild type were

set to 1. Vertical bars represent standard errors of three biological replicates. b and c indicate $p < 0.01$ and $p < 0.001$, respectively.

Supplementary Files

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

- [SupplementaryFiles.zip](#)