

ARTICLE

Open Access

Integrating full-length transcriptomics and metabolomics reveals the regulatory mechanisms underlying yellow pigmentation in tree peony (*Paeonia suffruticosa* Andr.) flowers

Xiaoning Luo¹, Daoyang Sun¹, Shu Wang¹, Sha Luo¹, Yaqi Fu¹, Lixin Niu¹, Qianqian Shi¹✉ and Yanlong Zhang¹✉

Abstract

Tree peony (*Paeonia suffruticosa* Andr.) is a popular ornamental plant in China due to its showy and colorful flowers. However, yellow-colored flowers are rare in both wild species and domesticated cultivars. The molecular mechanisms underlying yellow pigmentation remain poorly understood. Here, petal tissues of two tree peony cultivars, “High Noon” (yellow flowers) and “Roufufurong” (purple-red flowers), were sampled at five developmental stages (S1–S5) from early flower buds to full blooms. Five petal color indices (brightness, redness, yellowness, chroma, and hue angle) and the contents of ten different flavonoids were determined. Compared to “Roufufurong,” which accumulated abundant anthocyanins at S3–S5, the yellow-colored “High Noon” displayed relatively higher contents of tetrahydroxychalcone (THC), flavones, and flavonols but no anthocyanin production. The contents of THC, flavones, and flavonols in “High Noon” peaked at S3 and dropped gradually as the flower bloomed, consistent with the color index patterns. Furthermore, RNA-seq analyses at S3 showed that structural genes such as *PsC4Hs*, *PsDFRs*, and *PsUGTs* in the flavonoid biosynthesis pathway were downregulated in “High Noon,” whereas most *PsFLSs*, *PsF3Hs*, and *PsF3'Hs* were upregulated. Five transcription factor (TF) genes related to flavonoid biosynthesis were also upregulated in “High Noon.” One of these TFs, *PsMYB111*, was overexpressed in tobacco, which led to increased flavonols but decreased anthocyanins. Dual-luciferase assays further confirmed that *PsMYB111* upregulated *PsFLS*. These results improve our understanding of yellow pigmentation in tree peony and provide a guide for future molecular-assisted breeding experiments in tree peony with novel flower colors.

Introduction

Tree peony (*Paeonia suffruticosa* Andr.) belongs to the section *Moutan* in the genus *Paeonia* of the family *Paeoniaceae* and is a traditional horticultural plant in China¹. Floral color is one of the crucial ornamental characteristics of tree peony, which has more than 2000 cultivars with 9 major colors worldwide^{2,3}. The flower colors of domesticated tree peony are generally white, pink, purple, or red, but few flowers are pure yellow.

Currently, new tree peony cultivars with novel pure yellow corollas have become a breeding focus with great economic potential⁴. “High Noon,” a hybrid of *P. suffruticosa* × *Paeonia lutea*, is one of the most famous and popular tree peony cultivars. It possesses pure yellow, cup-shaped, and semidouble flowers. Due to its strong ecological adaptability, “High Noon” has been increasingly cultivated in the major production areas of tree peony around the world. However, the molecular basis of yellow coloration in tree peony remains elusive. Therefore, yellow-colored “High Noon” provided a good candidate for research on yellow coloration in tree peony flowers.

Correspondence: Qianqian Shi (shiqianqian2005@163.com) or Yanlong Zhang (zhangyanlong@nwsuaf.edu.cn)

¹College of Landscape Architecture and Art, Northwest A&F University, Yangling, China

© The Author(s) 2021



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

The chemical compounds determining flower colors mainly include flavonoids, carotenoids, chlorophylls, and their derivatives⁵. Flavonoids are the decisive pigment groups of most flower colors and are present in the vacuoles of petal epidermal cells⁶. In the flavonoid biosynthesis pathway, anthocyanins confer pink, red, purple, and blue colors to flowers and other organs, whereas chalcones and aurones are deep yellow, and flavones and flavonols are faint yellow or almost colorless. Carotenoids consist of carotenes and xanthophylls, which are present in plastids and can render plants yellow, orange, and red⁵. Chlorophylls are the predominant pigments in the green organs of plants, such as leaves and stems. Previous studies have found that tree peony flower pigments were mainly anthocyanins, including 3-*O*-glucosides and 3,5-di-*O*-glucosides of cyanidin, peonidin, and pelargonidin; flavones, including multifunctional glucosides of apigenin (Ap), luteolin (Lu), and chrysoeriol (Ch); and flavonols, including multifunctional glucosides of kaempferol (Km), quercetin (Qu), and isorhamnetin (Is)^{7–10}. In eight different color series from white to red, to yellow of tree peony, a total of 39 flavonoids (5 anthocyanins, 12 flavones, 21 flavonols, and 1 chalcone) have been identified^{11,12}. Recently, 56 flavonoids were further characterized from 15 traditional Chinese tree peony cultivars with white, pink, and red color series¹³. In yellow-colored tree peony flowers, 26 flavonoid components have been detected, among which Km, Ap, Lu glucosides, and chalcones were dominant¹¹. The main flavonoid components in yellow flowers of *P. lutea* were chalcones, flavones, and flavonols, including tetrahydrochalcone (THC), isosalipurposide (ISP), Km, Qu, Is, Ch, and Ap⁵. The production of chalconaringenin 2'-*O*-glucoside (Chalcone 2'G) was presumed to be a leading reason for yellow pigmentation in petals of *P. lutea*¹⁴.

Flavonoid synthesis can be divided into three phases. The first phase includes the transformation of phenylalanine to coumaroyl-CoA, which is shared by many secondary metabolism pathways. The second phase corresponds to the conversion of coumaroyl-CoA to dihydroflavonols, including dihydrokaempferol, dihydroquercetin, and dihydromyricetin (DHM). This process is catalyzed by a battery of enzymes covering chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonoid 3'5'-hydroxylase, flavone synthase (FNS), and flavonol synthase (FLS). The second phase is critical for yellow coloration due to its accumulation of chalcones, flavones, and flavonols. Subsequently, the third phase is the synthesis of a series of stable anthocyanins catalyzed by dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS), UDP-glucose flavonoid 3-*O*-glucosyltransferase (UGT), and anthocyanin *O*-methyltransferase^{15,16}. These structural genes are regulated by the MBW

complex, which contains MYB, bHLH transcription factors (TFs), and WD40 proteins¹⁷. In the regulation of flavonoid production, MYBs serve as key TFs and bind directly to bHLH regulators or the promoters of structural genes to activate gene expression¹⁸. As chaperone proteins, WD40 plays a role in stabilizing the MBW complex. In addition, other TFs, such as SPLs, NACs, WRKY, HY5, and ERFs, are also involved in regulating flavonoid biosynthesis^{19,20}. In particular, SPLs can disturb the formation of the MBW complex by competing for bHLHs or MYBs, thus inhibiting anthocyanin production²¹. Recently, many flavonoid synthesis-related genes have been identified in tree peony, including the structural genes *PsCHS1*, *Ps-CHI1*, *PsF3H1*, *PsDFR1*, and *PsANS1*, and the TF genes *PlWDR3*, *PlWDR18*, *PlbHLH3*, *PoMYB2*, *PsMYB12*, and *PoSPL1*^{2,14,22–28}. However, only a limited number of studies have reported candidate genes for yellow pigmentation in tree peony. For example, transgenic tobacco plants constitutively expressing *Ps-CHI1* exhibit an up to threefold increase in total flavonol and flavone levels, and a significant reduction in anthocyanin content and floral color strength compared to wild-type (WT) controls²⁴. High coexpression of *THC2'GT*, *CHI*, and *FNS II* in *P. lutea* flowers guarantees the accumulation of yellow pigments¹⁴. Notably, no TF gene regulating yellow pigmentation in tree peony has been identified.

Second-generation sequencing (SGS) technology has been widely applied to transcriptomic analysis in tree peony^{14,29–31}. Recently, the third-generation sequencing (TGS) platform single-molecule real-time sequencer PacBio RS (Pacific Biosciences of California, USA) has been developed for more advanced RNA sequencing (RNA-seq)^{32,33}. Compared to SGS, TGS has longer read lengths, higher consensus accuracy, a lower degree of bias, and simultaneous capability for epigenetic characterization³⁴. In turn, the sequence alignment errors of TGS can be algorithmically improved and corrected by SGS reads³⁵. Therefore, higher quality transcriptome assemblies can be achieved by combining SGS and TGS technologies.

In this study, we performed metabolomics and full-length transcriptomics analyses of the petal tissues of two tree peony cultivars, “High Noon” (yellow flowers) and “Roufufong” (purple–red flowers), at different flowering stages. Differentially expressed genes (DEGs) and the metabolic profiles of the two cultivars were identified and characterized. Candidate genes responsible for yellow pigmentation were further investigated by quantitative real-time PCR (qRT-PCR) analysis, subcellular localization, overexpression in tobacco, promoter sequence analysis, and dual-luciferase assays. The molecular mechanisms of yellow pigmentation in tree peony flowers were discussed.

Results

Assessment of flower color phenotypes

To characterize flower color development in tree peony, the petal tissues of the yellow-flowered cultivar “High Noon” and purple-red-flowered cultivar “Roufuring” were sampled at five developmental stages from early flower buds to full blooms (Fig. 1a, b). The petal color indices of the two cultivars were measured at S1–S5 (Fig. 1c, d). In “High Noon”, the L^* (brightness) value increased gradually from S1 to S5, indicating the elevation of petal color brightness. C^* (chroma) and b^* (yellowness) peaked at S3 and subsequently declined at S4 and S5, demonstrating that S3 was the yellowest stage. In contrast to L^* , C^* , and b^* , h (hue angle) decreased from S1 to S3, followed by a slight increase at S4 and S5. For “Roufuring”, no significant change in L^* was observed throughout the flowering process. h dropped from S1 to S3 and then increased dramatically at S4 and S5. h -Values from S3 to S5 were $\sim 0^\circ$ (360°), consistent with the purple-red color. Similar to h , C^* in “Roufuring” also decreased from S1 to S3 and increased at the later stages S4 and S5. Consistent with the flower color, b^* in the purple-red-colored

“Roufuring” was significantly lower than that in the yellow-colored “High Noon” from S1 to S5. Consistently, a^* (redness) was much higher in “Roufuring” than in “High Noon”.

Quantitative analysis of flavonoids

Significant variations in flavonoid contents were observed between the two tree peony cultivars at each developmental stage. As shown in Fig. 2a, b, the target flavonoids in “High Noon” increased dramatically from S1 to S3 and then dropped slightly at S4 and S5. At the later flowering stages S3–S5, the content of THC was significantly higher than that of other flavonoids. Ap and Km had comparable contents throughout the five flowering stages. In addition, Ch had similar production levels to Is, whereas Lu was similar to Qu. Notably, no anthocyanin was detected in “High Noon” throughout the flowering process. In contrast, three anthocyanins were detected in “Roufuring”. The dominant peonidin 3-*O*-glucoside (Pn3G) increased dramatically as the flower bloomed and reached the maximum level at S5, implying that Pn3G may contribute to purple-red flower pigmentation.

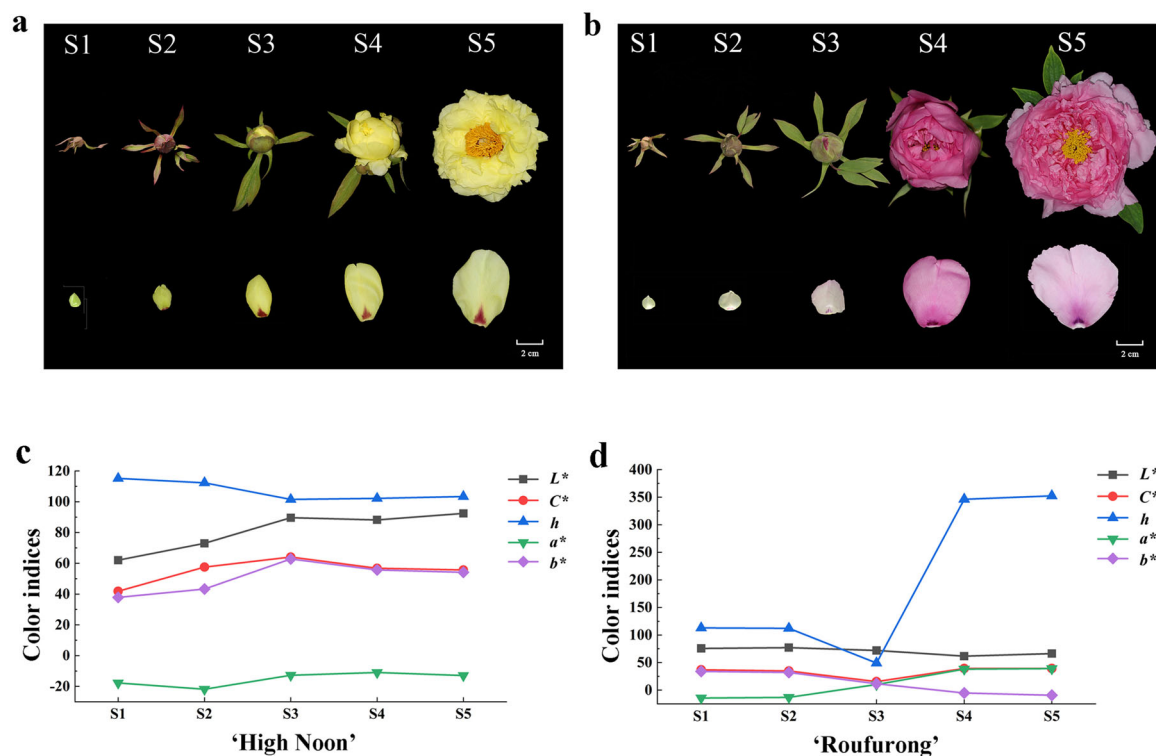


Fig. 1 Floral phenotypes and color indices of tree peony cultivars “High Noon” and “Roufuring” at five different blooming stages. **a** Floral phenotype of “High Noon”. **b** Floral phenotype of “Roufuring”. **c** Color indices of “High Noon”. **d** Color indices of “Roufuring”. S1: Stage 1, unpigmented tight bud; S2: Stage 2, slightly pigmented soft bud; S3: Stage 3, initially opened flower; S4: Stage 4, half opened flower; S5: Stage 5, fully opened and pigmented flower with exposed anthers. L^* represents the brightness; the color becomes brighter as the value increases. C^* represents chroma; the color saturation increases as the value rises. h represents the hue angle, defined as follows: 0° (360°) for purple-red, 90° for yellow, 180° for blue-green, and 270° for blue⁶⁹. a^* and b^* represent the redness and yellowness, respectively

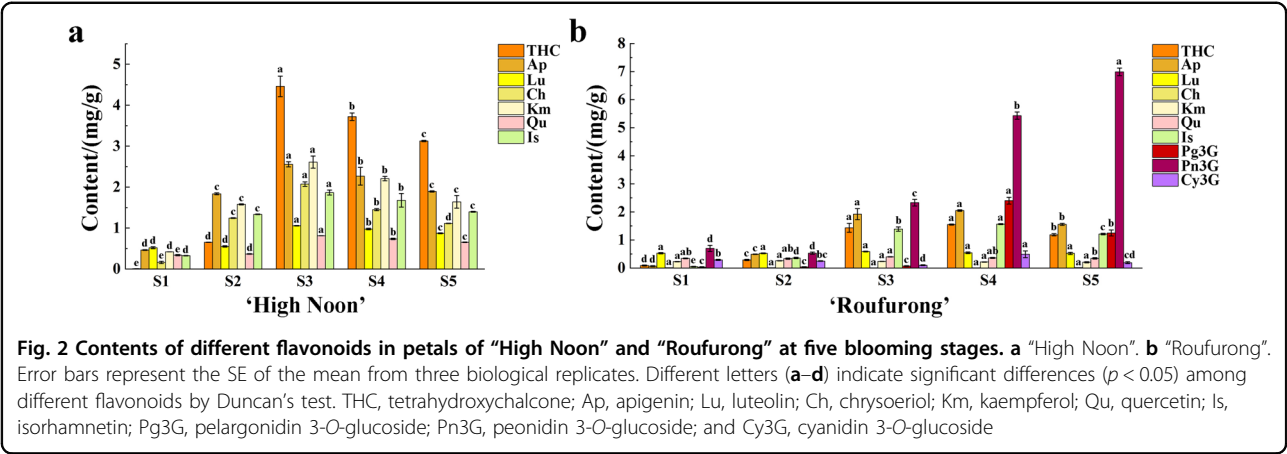


Table 1 Multiple linear regression equations of two tree peony cultivars, “High Noon” and “Roufuring”

Cultivars	Equations
“High Noon”	$L^* = 66.173 + 6.214 \text{ THC}$ ($R^2 = 0.867$, $P < 0.05$) $C^* = 41.198 + 11.574 \text{ Ch}$ ($R^2 = 0.960$, $P < 0.01$) $h = 114.597 - 3.215 \text{ THC}$ ($R^2 = 0.979$, $P < 0.01$) $a^* = -30.806 + 26.860 \text{ Lu} - 4.801 \text{ Ch}$ ($R^2 = 0.993$, $P < 0.05$) $b^* = 38.558 + 5.095 \text{ THC}$ ($R^2 = 0.980$, $P < 0.01$)
“Roufuring”	$L^* = 74.970 - 5.889 \text{ Pg3G}$ ($R^2 = 0.908$, $P < 0.05$) $C^* = 214.880 - 335.919 \text{ Lu}$ ($R^2 = 0.827$, $P < 0.05$) $h = -164.139 + 122.546 \text{ Pg3G} + 22840.617 \text{ Ch}$ ($R^2 = 0.999$, $P < 0.01$) $a^* = 7.308 + 9.695 \text{ Pn3G} - 2261.699 \text{ Ch}$ ($R^2 = 0.997$, $P < 0.05$) $b^* = 38.353 - 5.126 \text{ Pn3G} - 7.639 \text{ Ap}$ ($R^2 = 0.999$, $P < 0.01$)

a^* redness, Ap apigenin, b^* yellowness, C^* chroma, Ch chrysoeriol, $Cy3G$ cyanidin 3-O-glucoside, h hue angle, Is isorhamnetin, Km kaempferol, L^* brightness, Lu luteolin, $Pg3G$ pelargonidin 3-O-glucoside, $Pn3G$ peonidin 3-O-glucoside, Qu quercetin, THC tetrahydroxychalcone

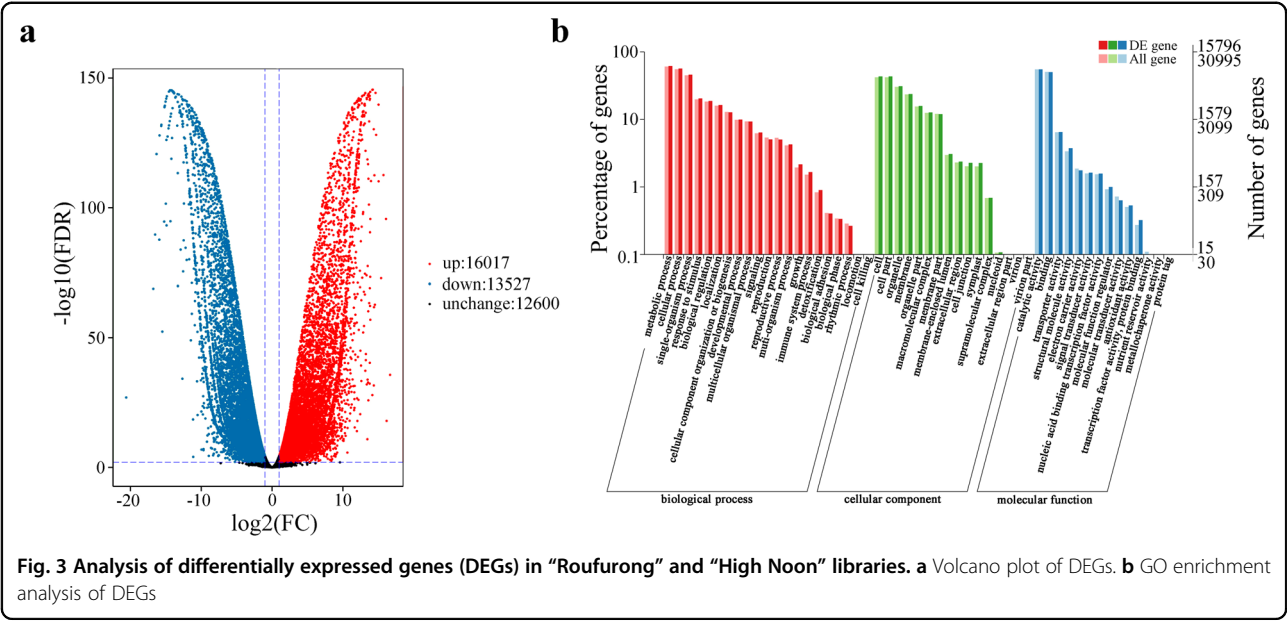
Compared to “High Noon”, the THC, Is, and Ap contents showed a quick elevation during the first four periods and a moderate decline at S5. Their variation ranges were relatively smaller than that of pelargonidin 3-O-glucoside (Pg3G). In contrast, the contents of Lu, Km, Qu, cyanidin 3-O-glucoside (Cy3G), and Ch did not display a clear developmental pattern, and thus, these components may have no significant influence on the purple–red coloration.

To understand the relationship between flower color and flavonoid components, multiple linear regression (MLR) analysis was performed (Table 1). In “High Noon”, L^* and b^* were positively correlated with THC levels,

indicating that the increase in the THC content led to the accumulation of yellow pigments. In contrast, the h -value was negatively correlated with THC levels. Ch had a positive effect on C^* but a negative effect on a^* . Therefore, THC and Ch may be associated with the production of yellow flowers. In “Roufuring”, Pg3G had a negative impact on L^* but a positive impact on h , suggesting that a high content of Pg3G may lead to a low brightness of the petals. In addition, C^* was negatively correlated with the Lu content, which may affect petal saturation as a copigment. Moreover, a^* was found to be positively and negatively associated with Pn3G and Ch, respectively, whereas b^* had a negative correlation with Pn3G and Ap. These results indicated that a high content of Pn3G and a low content of Ch were critical for the purple–red color of petals. Considering that THC and Ch displayed the highest accumulation levels at S3 in “High Noon”, and C^* and b^* also peaked at S3, stage S3 was further selected for subsequent transcriptome analyses to identify the candidate genes for yellow pigmentation.

Overview of transcriptome sequencing

To obtain the reference transcriptome assemblies for tree peony, ten petal samples at S1–S5 from “High Noon” and “Roufuring” were pooled for TGS sequencing. A total of 473,062 error-corrected reads of insert (ROIs) were obtained, with an average read length of 1,883 bp (Supplementary Table 1). Furthermore, 76.1% of the full-length ROIs were considered full-length nonchimeric (FLNC) reads (Supplementary Table 2) and 117,680 high-quality isoforms were obtained after filtering the low-quality consensus sequences (Supplementary Table 3). Transcriptome analyses of petal tissues of the two cultivars at S3 were carried out by the SGS approach with three biological replicates. Six libraries covered 45.88 Gb clean read data with high Q30 (>94%) (Supplementary Table 4). In total, 56,610 open reading frames (ORFs) were mapped to the assembly, of which 41,138 displayed



full-length sequences. To estimate the functions of assembled sequences, 56,974 transcripts were annotated by searching against different databases (Supplementary Table 5). The transcripts from tree peony shared the highest similarity with those from *Vitis vinifera* (27.40%) and more transcripts fell under biological process, of which metabolic process, cellular process, and single-organism process were mainly enriched with transcripts (Supplementary Fig. 1a, b). Furthermore, 27,265 transcripts were mapped to 125 metabolic pathways (Supplementary Table 6). Specifically, 115, 12, and 30 transcripts were found in the flavonoid biosynthesis pathway (ko00941), anthocyanin biosynthesis pathway (ko00942), and flavone and flavonol biosynthesis pathway (ko00944), respectively (Supplementary Table 6).

DEGs involved in flavonoid biosynthesis

To identify the candidate genes for yellow pigmentation, DEGs between “High Noon” and “Roufufurong” at S3 were identified. There were 16,017 upregulated DEGs and 13,527 downregulated DEGs in “High Noon” compared with “Roufufurong” (Fig. 3a). Gene Ontology (GO) enrichment analysis revealed that 15,796 DEGs were distributed into 51 terms comprising 21 biological processes, 16 cellular components, and 14 molecular functions. Among the largest category, biological process, the DEGs were significantly enriched in “metabolic process”, “cellular process”, and “single-organism process” (Fig. 3b). DEGs involved in flavonoid biosynthesis were screened and contained 45 structural genes (2 *PsC4Hs*, 8 *PsCHSs*, 3 *PsCHIs*, 12 *PsF3Hs*, 2 *PsF3’Hs*, 11 *PsFLSs*, 4 *PsDFRs*, and 3 *PsUFGTs*) and 5 TF genes (*PsMYB4* (F01_cb7851_c8), *PsMYB111* (F01_cb8732_c11), *PsTT8* (F01_cb3692_c7),

PsEGL3 (F01_cb8475_c28561), and *PsSPL9* (F01_cb9156_c0)). The expression profiles of these DEGs related to flavonoid biosynthesis are displayed in Fig. 4. Twelve of these genes were randomly selected for qRT-PCR to validate the transcriptome data (Supplementary Fig. 2). Their expression levels were consistent with the RNA-seq data results.

As shown in Fig. 4, structural genes at the second stage of flavonoid biosynthesis generally presented higher expression levels in “High Noon” than in “Roufufurong”. These genes included two *PsCHSs* (F01_cb9709_c117 and F01_cb9709_c59) and two *PsCHIs* (F01_cb13354_c0 and F01_cb13354_c4). Based on gene annotation, *PsCHS* functions in the synthesis of THCs, whereas *PsCHI* plays a critical role in the production of flavanones, which contribute to yellow coloration. In addition, eight putative *PsF3Hs* (F01_cb4853_c44, F01_cb4853_c20, F01_cb4853_c21, F01_cb4853_c25, F01_cb4853_c24, F01_cb4853_c6, F01_cb4853_c19, and F01_cb4853_c23), two *PsF3’Hs* (F01_cb6443_c24 and F01_cb6443_c9), and nine *PsFLSs* (F01_cb4550_c24, F01_cb15832_c6, F01_cb4550_c0, F01_cb4550_c16, F01_cb4550_c1, F01_cb15832_c2, F01_cb4550_c34, F01_cb4550_c23, and F01_cb4550_c42) also exhibited higher transcription in “High Noon”. Of these genes, *PsFLS* can divert the flavonoid pathway in favor of flavonol production. These expression variations between “High Noon” and “Roufufurong” were consistent with the observed higher levels of THC, flavones, and flavonols in “High Noon”. In contrast to the structural genes at the second stage, genes at the third stage of flavonoid biosynthesis, including four *PsDFRs* (F01_cb12671_c6, F01_cb12671_c7, F01_cb12671_c2, and F01_cb12671_c4) and three *PsUFGTs* (F01_cb7865_c4, F01_cb7865_c1, and

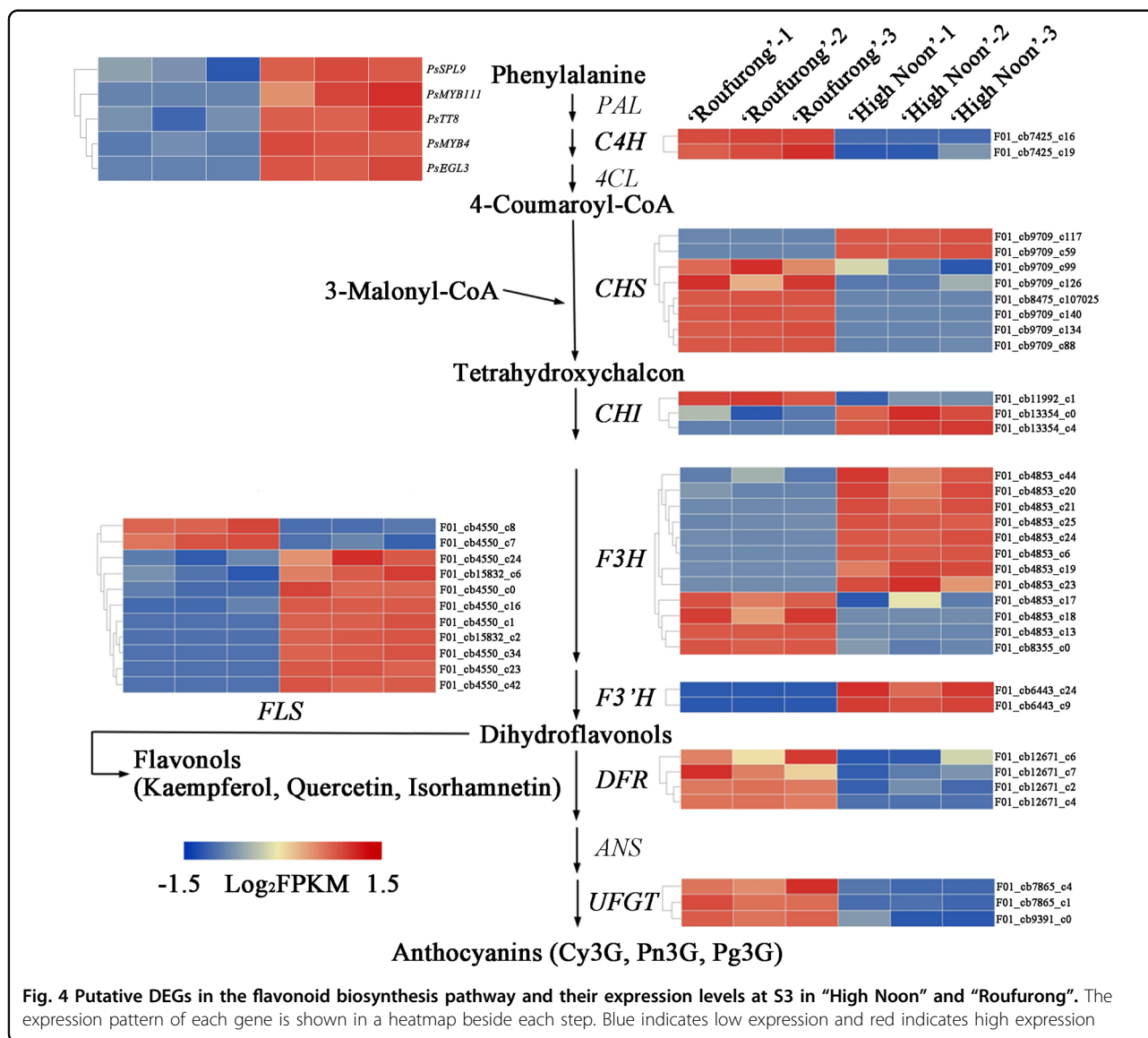


Fig. 4 Putative DEGs in the flavonoid biosynthesis pathway and their expression levels at S3 in “High Noon” and “Roufuring”. The expression pattern of each gene is shown in a heatmap beside each step. Blue indicates low expression and red indicates high expression

F01_cb9391_c0), were significantly downregulated in “High Noon”. *PsDFR* and *PsUFGT* are responsible for the biosynthesis of anthocyanins such as Cy3G, Pn3G, and Pg3G. Therefore, the downregulation of *PsDFRs* and *PsUFGTs* in “High Noon” may explain the lack of anthocyanin production in this yellow-flowered cultivar.

TF genes regulating flavonoid biosynthesis were also analyzed with the transcriptomes of “High Noon” and “Roufuring”. As shown in Fig. 4, all of these identified TFs were upregulated in the yellow-flowered “High Noon.” The interaction between MYB and bHLH TFs plays a key role in the regulation of structural genes in the flavonoid biosynthetic pathway¹⁷. Candidate MYB proteins related to flavonoid biosynthesis were identified by phylogenetic analyses with MYBs from *Arabidopsis*. The results showed that *PsMYB4* belonged to subgroup 4 (S4) of

Arabidopsis (Supplementary Fig. 3a). S4 MYBs in *Arabidopsis* directly interact with AtbHLHs (TT8, GL3, and EGL3) and repress the transcriptional activities of MBW complexes, thereby leading to the inhibition of anthocyanin and phenylpropanoid syntheses³⁶. *PsMYB111* was clustered to S7 (Supplementary Fig. 3a), which was shown to regulate flavonol biosynthesis in *Arabidopsis*³⁷. Sequence analysis revealed that *PsMYB4* and *PsMYB111* were identified as typical R2R3MYB proteins (Supplementary Fig. 3b, c), and *PsMYB4* contained the conserved sequence [D/E]LX₂[K/R]X₃LX₆LX₃R, which interacts with bHLH proteins. Notably, although *PsMYB111* does not interact with bHLHs, the SG7 motif [K/R][R/X][R/K]XGRT[S/X][R/G]XX[M/X]K and SG7-2 motif [W/X][L/X]LS, which are specific to flavonol biosynthesis regulators^{38,39}, were found at the C terminus (Supplementary

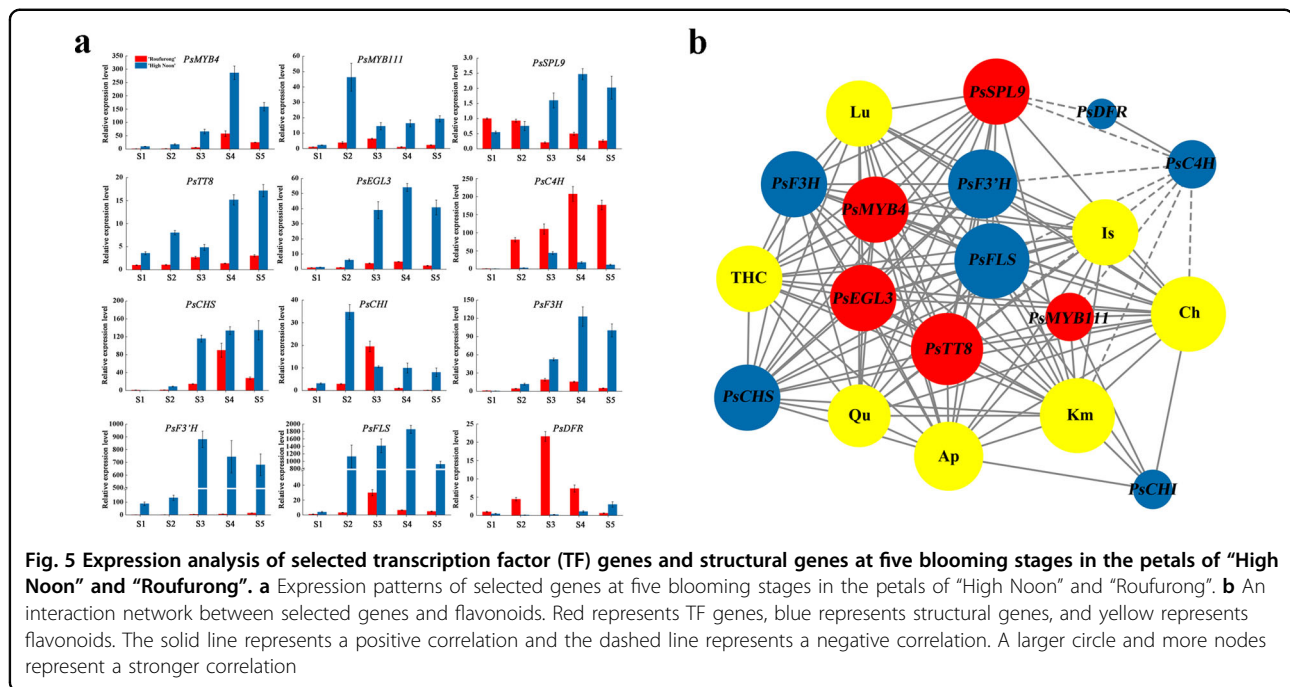


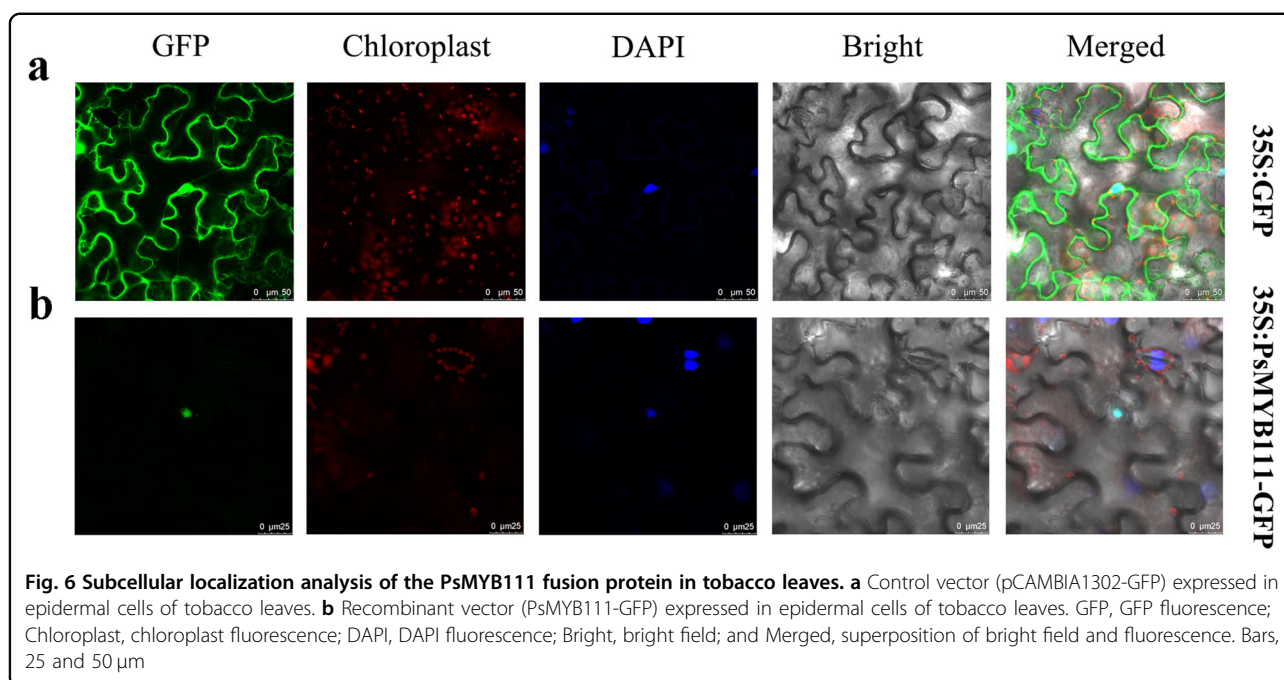
Fig. 3c). In the case of bHLHs, PsTT8 and PsEGL3 were identified in the S5 subgroup (Supplementary Fig. 4a) and were related to flavonoid biosynthesis. Both PsTT8 and PsEGL3 contained an N-terminal MYB-interacting region (MIR), a bHLH domain, and a putative ACT-like domain at the C terminus (Supplementary Fig. 4b). MIR suggests that PsTT8 and PsEGL3 may interact with MYB proteins, and the putative ACT-like domain has been shown to play a vital role in the dimerization of bHLHs⁴⁰. In addition, phylogenetic analysis suggested that PsSPL9 containing a typical squamosa promoter-binding protein domain was clustered together with AtSPL9 and AtSPL15 (Supplementary Fig. 5a, b), of which AtSPL9 has been reported to inhibit the accumulation of anthocyanin by destroying the stability of the MBW complex in *Arabidopsis*²¹.

Expression patterns of candidate DEGs related to flavonoid biosynthesis

qRT-PCR analyses were conducted to investigate the expression patterns of structural genes *PsC4H* (F01_cb7425_c19), *PsCHS* (F01_cb9709_c126), *PsCHI* (F01_cb13354_c0), *PsF3H* (F01_cb4853_c44), *PsF3'H* (F01_cb6443_c9), *PsFLS* (F01_cb4550_c24), and *PsDFR* (F01_cb12671_c6), as well as TF genes *PsMYB4*, *PsMYB111*, *PsTT8*, *PsEGL3*, and *PsSPL9* at five flowering stages, S1–S5, in “High Noon” and “Roufuring” petals (Fig. 5a). For structural genes, only *PsC4H* and *PsDFR* were expressed at higher levels in “Roufuring”, which was consistent with the accumulation of anthocyanins (Fig. 2b). In contrast, the other selected structural genes,

PsCHS, *PsCHI*, *PsF3H*, *PsF3'H*, and *PsFLS*, all displayed higher expression in “High Noon”. Furthermore, the expression patterns of these genes from S1 to S5 were generally consistent with the previously determined flavonoid accumulation pattern (Fig. 2a). *PsCHS* transcription at S3 was greatly upregulated and high expression was maintained at both S4 and S5, whereas the expression of *PsCHI* peaked at S2 and then significantly decreased at S3, S4, and S5. The high expression of *PsCHS* and the low expression of *PsCHI* led to the accumulation of THC from S3 to S5. In addition, the high expression of *PsF3H*, *PsF3'H*, and *PsFLS* at S3, S4, and S5 might have guided the metabolic flow in favor of the production of flavonols such as Km, Qu, and Is. Notably, the expression of *PsFLS* and *PsDFR* in “High Noon” was negatively related, suggesting a potential competitive mechanism between these two genes. In terms of TF genes, *PsMYB4*, *PsEGL3*, and *PsSPL9* in “High Noon” had a consistent expression pattern that first increased and then declined at later stages. This pattern was in contrast to the expression pattern of *PsDFR*. The expression patterns of *PsMYB111* and *PsTT8* were similar in “High Noon”, which peaked at S2, dropped dramatically at S3, and then increased slightly at S4 and S5. This expression pattern exactly matched that for *PsCHI* and *PsFLS*, leading to the assumption that PsMYB111 may directly regulate the transcription of these two structural genes.

Based on the gene expression and flavonoid accumulation profiles, an interaction network between TF genes and structural genes encoding flavonoids was developed



(Fig. 5b). The results showed that *PsSPL9* was negatively correlated with *PsDFR* and *PsC4H*. A strong positive correlation was observed between *PsMYB4* and *PsEGL3*, as well as between *PsMYB111* and *PsTT8*. These four TF genes acted as hub genes in the whole interaction network. The interaction network also showed that *PsFLS* was closely related to *PsMYB111* and *PsTT8*, and directly associated with Km, Qu, and Is. Nevertheless, *PsMYB111* has been identified as devoid of bHLH interaction sites; thus, we hypothesized that *PsMYB111* might affect the production of flavonols by regulating the expression of *PsFLS* alone. To validate this hypothesis, *PsMYB111* was selected for subsequent functional analysis.

Subcellular localization analysis of *PsMYB111*

To explore the subcellular localization of *PsMYB111*, a plasmid containing *PsMYB111* fused to green fluorescent protein (GFP) was constructed and transiently transformed into *Nicotiana benthamiana* leaves. The subcellular localization of the protein was observed after 72 h. Fluorescence of the fusion protein *PsMYB111*-GFP was detected specifically in the nucleus (Fig. 6a, b), supporting the role of *PsMYB111* as a TF involved in regulating flavonol biosynthesis.

Overexpression of *PsMYB111* in tobacco

To further characterize the function of *PsMYB111* in flavonoid biosynthesis, we constructed transgenic tobacco lines overexpressing *PsMYB111* (*PsMYB111* OE). Two T₂ lines (OE-1 and OE-2) were obtained. The flower color of OE-1 and OE-2 was light pink, in contrast to the rosy red

color of the WT (Fig. 7a). OE-1 displayed a higher *PsMYB111* transcript level than OE-2 and was thus selected for further study (Fig. 7a). The color indices *L**, *a**, *C**, and *h* of tobacco petals were measured. The results showed that the *L**-value of OE-1 was significantly higher than that of WT, whereas the values of *a** and *C** decreased significantly ($P < 0.01$) (Fig. 7b). Correspondingly, the contents of THC, Ap, Lu, Km, and Qu also increased significantly in OE-1 compared to WT, whereas the contents of Pg3G and Pn3G decreased significantly ($P < 0.01$) (Fig. 7c). These results were consistent with the color phenotype of the transgenic tobacco.

To investigate whether *PsMYB111* can regulate the transcription of flavonoid biosynthetic structural genes, we analyzed the expression of selected structural genes in OE-1 petals using qRT-PCR (Fig. 7d). The results showed that the transcription of *NtCHS*, *NtCHI*, and *NtFLS* increased significantly compared to WT, whereas *NtC4H*, *NtF3H*, *NtF3'H*, and *NtDFR* were only mildly upregulated. The upregulation of *NtFLS* was the highest among the analyzed structural genes. In addition, *NtbHLH* and *NtSPL9* exhibited similar expression abundances to WT. Taken together, *PsMYB111* may regulate the expression of structural genes alone rather than in a complex and may enhance the accumulation of flavonols such as Qu and Km.

Regulation of *PsCHS* and *PsFLS* promoters by *PsMYB111*

Transgenic experiments in tobacco leaves indicated that *PsMYB111* can directly regulate *NtCHS* and *NtFLS*. To verify its regulatory role on *PsCHS* and *PsFLS*, we

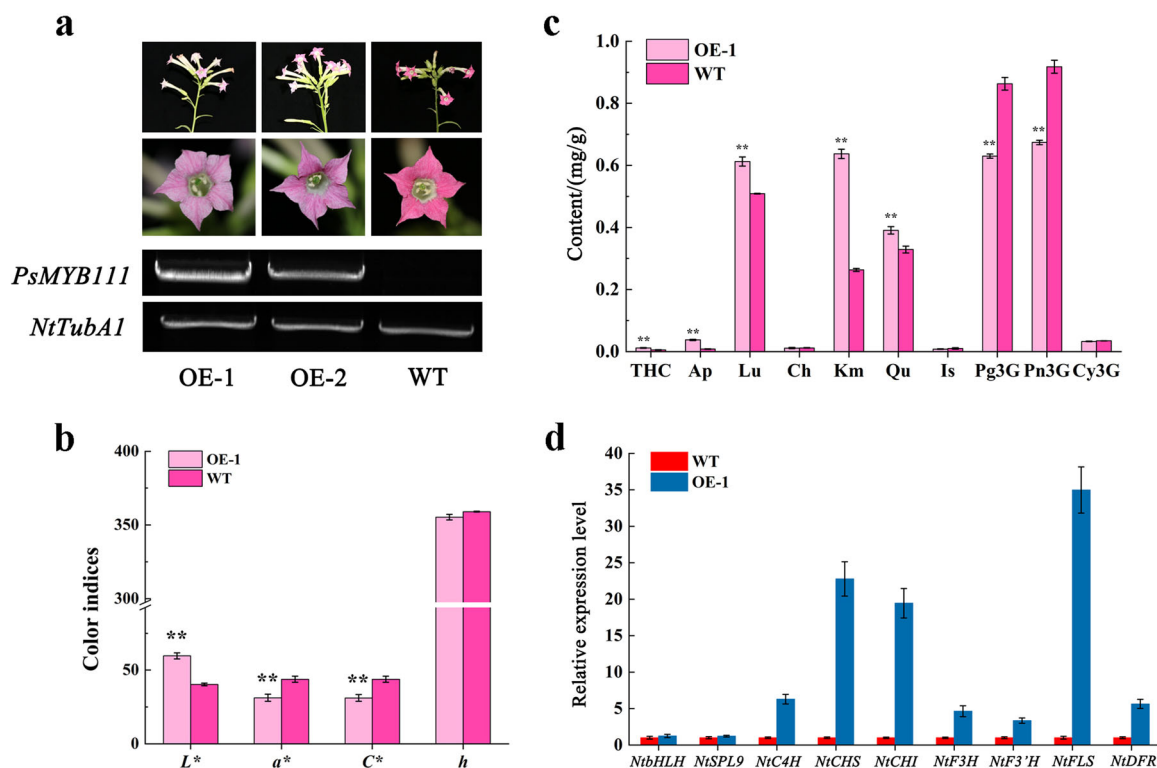


Fig. 7 Overexpression of *PsMYB111* increases the accumulation of flavonols and reduces the accumulation of anthocyanins in tobacco. a Flowers from wild type (WT) and transgenic tobacco plants, and their corresponding *PsMYB111* transcription levels. *NtTubA1* was used as an endogenous control. **b** Color indices of petals from WT and transgenic tobacco plants. **c** Flavonoid contents in petals of WT and transgenic tobacco plants. **d** Expression patterns of endogenous TF genes and flavonoid biosynthetic genes in the petals of WT and transgenic tobacco plants. Three independent experiments were performed for each sample. The data are shown as the means $\pm$ SDs. ** $p < 0.01$

sequenced the promoter sequences of *PsCHS* (750 bp) and *PsFLS* (1,010 bp), and analyzed their *cis*-acting elements (Fig. 8a). Many *cis*-acting elements related to MYB and bHLH TFs were identified, including the MYB-binding site (5'-CAACNG-3'), bHLH-binding site (5'-CATGTG-3'), and G-box (5'-CACGTG-3') (Fig. 8a), supporting their potential interaction with *PsMYB111*.

To further confirm this hypothesis, a dual-luciferase assay was performed. The pGreenII62-SK vector carrying *PsMYB111* served as an effector and the pGreenII0800 LUC vector carrying the promoters of *PsCHS* and *PsFLS* served as reporters. After coinfiltration of the TF and promoters in *N. benthamiana* leaves, the LUC/REN ratios were detected. As shown in Fig. 8b, c, significantly higher LUC/REN ratios were observed for both the *PsMYB111*-*PsCHS* and *PsMYB111*-*PsFLS* constructs compared with the corresponding controls, suggesting that *PsMYB111* can activate the promoters of both *PsCHS* and *PsFLS*. The activation intensity of *PsMYB111* on *PsFLS* (2.59-fold) was stronger than that of *PsCHS* (1.58-fold), indicating that *PsMYB111* may preferentially regulate the transcription of *PsFLS*.

Discussion

Pure yellow flowers in tree peony are extremely rare. The molecular mechanisms of yellow pigmentation in this species remain unclear. Here we selected the cultivar “High Noon” with yellow flowers as the research material, with the purple-red-flowered cultivar “Roufufurong” as a control. The floral color phenotypes and flavonoid profiles of “High Noon” petals were characterized, together with combined full-length and comparative transcriptome analyses. Candidate genes underlying yellow pigmentation were further validated by functional analyses.

Flavonoids play a key role in the coloration of different plant organs. In this study, we showed that the contents of THC, flavones (Ap, Lu, and Ch), and flavonols (Km, Qu, and Is) in the flowers of “High Noon” were markedly higher than in those of “Roufufurong”. No anthocyanin was detected in the yellow petal tissues of “High Noon”. Similar observations have been found for yellow-flowered *P. lutea* from Yunnan, which also contains abundant THC, ISP, Km, Qu, Is, Ch, and Ap⁵. Among these compounds, THC and Ch were the main pigments causing yellow color. The high amount of Chalcone 2'G has been

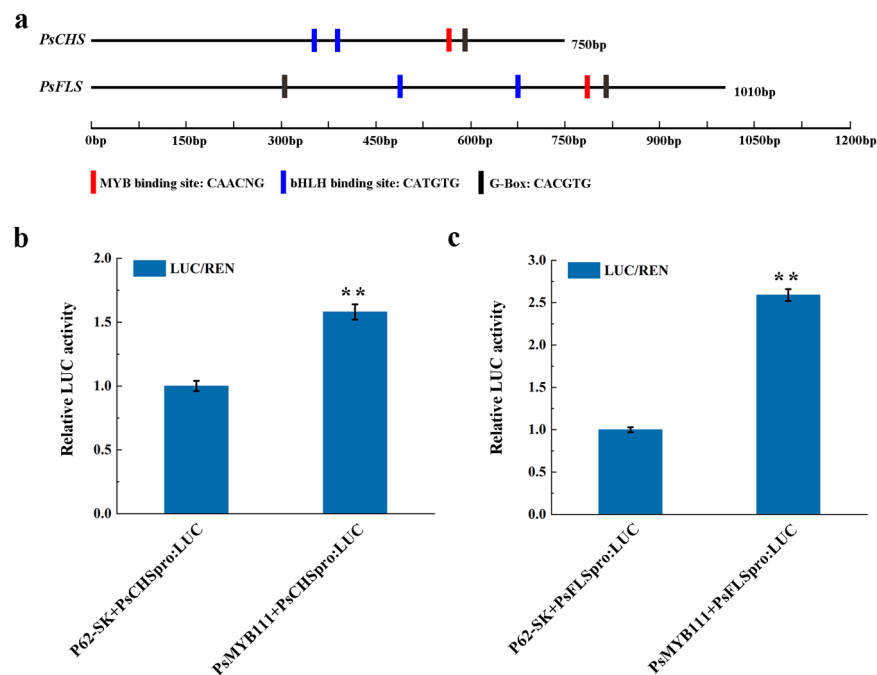


Fig. 8 Transcriptional activity analysis of PsMYB111 against the promoters of *PsCHS* and *PsFLS* of tree peony. **a** Schematic overview of *PsCHS* and *PsFLS* promoters. **b** Dual-Luciferase assays in tobacco leaves. Transformed protoplasts including only a promoter-LUC reporter construct without an effector were used as controls. The data represent the means $\pm$ SDs of three replicates from three independent experiments. $**p < 0.01$

speculated to be an important reason for the yellow-flowered phenotype in *P. lutea*¹⁴. In “Roufufurong”, the accumulation of the dominant anthocyanin Pn3G is the most likely reason for its purple–red-colored petals. It has been reported that reductions in anthocyanins increased the L^* -value and decreased the a^* -value, causing flower colors of two herbaceous peony cultivars “Coral Sunset” and “Pink Hawaiian Coral”, to change from coral to pink and then to yellow⁴¹. In addition, the sharp decrease in anthocyanins during flowering of “Jinyi Hualian” and “Xianguang” elicited colors from red to orange and then to yellow⁴². Therefore, the high contents of THC, flavones, and flavonols without anthocyanins may be essential for the yellow pigmentation of “High Noon” flowers.

In the flavonoid biosynthetic pathway, CHS catalyzes the synthesis of a key intermediate THC, which can be further isomerized by CHI, leading to the production of flavones and flavonols¹⁴. It was previously reported that a low expression level of *PtCHI* in the *Paeonia lactiflora* cultivar “Huangjinlun” induced chalcone accumulation and produced a yellow color⁴³. In the present study, we assume that the high expression of *PsCHS* and low expression of *PsCHI* in “High Noon” during S3–S5 contributed to the large accumulation of THC, which also resulted in yellow coloring. In addition, the high expression of *PsF3H*, *PsF3'H*, and *PsFLS* from S3 to S5 in “High Noon” promoted metabolic flow to the synthesis of flavonols (Km, Qu, and Is). In white *Muscari armeniacum*

flowers, when *FLS* was upregulated, the substrates used for cyanidin synthesis were then available for the synthesis of kaempferol⁴⁴. The downregulation of *PsDFR* and *PsUFGT* is consistent with the lack of anthocyanin accumulation in “High Noon”. Similarly, observations have been made in *P. lactiflora*, in which the lower expression levels of *PtDFR*, *PLANS*, *Pl3GT*, and *Pl5GT* in the inner petals inhibited anthocyanin production, resulting in yellow pigment formation⁴⁵. Notably, the substrate competition mechanism between FLS and DFR may cause variations in anthocyanin and flavonol synthesis, as FLS strengthens dihydroflavonol flux toward flavonols and finally limits anthocyanin accumulation⁴⁶. In the present study, we also observed competition between FLS and DFR, which may have also contributed to the yellow-flowered phenotype in “High Noon”. The upregulated expression of *FLS* transcripts in “High Noon” may consume dihydroflavonols as a substrate, thereby increasing the accumulation of flavonols and reducing the accumulation of anthocyanins at the DFR branch. Similar findings have been revealed in *Paeonia ostii*, which exhibited upregulated expression of *PoFLS4* and a transition of dihydroflavonols into flavonols in nearly white flowers²⁸. The competition between FLS and DFR in *M. armeniacum* for DHM might inhibit the synthesis of delphinidin, thereby altering the ratio of flavonol to anthocyanin and furthering the elimination of blue pigmentation⁴⁴. The latest research on *Brassica oleracea* L.

italica also showed that the difference in kaempferol accumulation was likely caused by the expression level of *FLS*⁴⁷. In contrast, the increased transcription of *PIDFR* in pigmented flowers of *Pleione limprichtii* was accompanied by a decrease in *PIFLS* transcription, causing increased production of anthocyanins⁴⁸. In addition, the competition mechanism between *FLS* and *DFR* has also been reported to underlie the lack of anthocyanins in white-fruited Chinese bayberry²⁷.

MYB and bHLH are critical TFs that regulate flavonoid biosynthesis in plants. In this study, we identified several differentially expressed MYB and bHLH TFs, each of which is related to flavonoid production in tree peony. Of these, PsMYB4 was identified as a negative flavonoid regulator with bHLH interaction sites, implying that they might form a complex to negatively regulate downstream structural genes. In *Arabidopsis*, the triple mutant *Atmyb4/7/32* displayed elevated anthocyanin and phenylpropanoid accumulation compared to the WT plants, which is consistent with our results³⁶. We also observed that *PsDFR* and *PsC4H* were obviously negatively correlated with *PsSPL9*. The high expression of *PsSPL9* and low expression of *PsDFR* may inhibit anthocyanin production in “High Noon”. Similarly, the AtSPL9 TF in *Arabidopsis* was also proven to inhibit the expression of anthocyanin biosynthetic genes, particularly *DFRs*²¹. In addition, yellow coloration in the floral tissues of *P. lactiflora* is possibly under the regulation of miR156e-3p-targeted *SPL1* by suppressing *PIPAL*, *PIFLS*, *PIDFR*, *PIANS*, *PI3GT*, and *PI5GT*^{45,49}. For *PsMYB111* and *PsTT8*, although they displayed a consistent expression pattern (Fig. 5a), PsMYB111 was most likely free of bHLH interactions. Based on homology clustering, PsMYB111 belonged to the S7 MYBs of *Arabidopsis* (Supplementary Fig. 3a), which have been proven to positively regulate early biosynthetic genes in the flavonoid pathway, such as *CHSs*, *CHIs*, *F3Hs*, *F3’Hs*, and *FLSs*³⁷. AtMYB11, AtMYB12, and AtMYB111 in S7 of *Arabidopsis* function independently of bHLHs⁵⁰. In *V. vinifera*, a similar regulatory model wherein VvMYB1 was a specific activator of *VvFLS1* and resulted in flavonol accumulation has also been proposed³⁸. In this study, we observed a clear association between *PsMYB111* and *PsFLS* transcription. PsMYB111 promotes the accumulation of flavonols by individually regulating *PsFLS*. In addition, PsMYB4 and PsEGL3 may form a complex to negatively regulate some structural genes, whereas PsSPL9 may negatively regulate *PsDFR* alone and inhibit the generation of anthocyanins.

Subcellular localization analysis revealed that PsMYB111 was localized to the nucleus (Fig. 6a, b), indicating that PsMYB111 might function as a TF in the nucleus. Overexpression of *PsMYB111* in tobacco caused its floral color to change from rosy red to light pink. The contents of flavonols such as Km and Qu in transgenic

tobacco lines were significantly increased, whereas the contents of anthocyanins such as Pg3G and Pn3G were significantly decreased, thereby confirming its function in the yellow pigmentation of tree peony flowers. In *Gerbera hybrida*, *GhMYB1a* overexpression also led to a significant increase in Km-type flavonol production and a significant decrease in anthocyanin production³⁹. In our *PsMYB111*-overexpressing line, we observed an inverse correlation between flavonol and anthocyanin contents, which reflects the competition between these two metabolic fluxes. Heterologous expression of S7 MYBs can regulate the expression of flavonoid biosynthetic genes, especially by upregulating flavonol pathway genes, causing intensive flavonol synthesis and the inhibition of anthocyanin generation^{51,52}. Consistent with the increased flavonol levels, overexpression of *PsMYB111* in tobacco led to the increased expression of *NtCHS*, *NtCHI*, and *NtFLS* (Fig. 7d). It has been reported that flavonol genes (*PAL*, *CHS*, *CHI*, *F3H*, and *FLS*) could be generally upregulated in transgenic tobacco overexpressing flavonol-specific MYB TF genes^{53,54}. The overexpression of *GtMYBP3* and *GtMYBP4* identified in *Gentiana triflora* promoted the expression of flavonol biosynthesis genes in tobacco and *Arabidopsis*⁵¹. In addition, the expression of *NtCHS*, *NtF3H*, and *NtFLS* was strongly upregulated in *GhMYB1a*-overexpressing transgenic tobacco lines and *GhMYB1a* significantly activated the promoters of *NtCHS* and *NtFLS* over *GhDFR* and *GhMYB10* in gerbera³⁹. Similarly, we showed that PsMYB111 had a significant activation effect on *PsCHS* and *PsFLS* promoters, particularly *PsFLS* (Fig. 8b, c).

Taken together, our study showed that PsMYB111 may influence the accumulation of flavonols by directly regulating the expression of *PsFLS* and reducing the flux to anthocyanin synthesis, thus ultimately contributing to the formation of yellow flowers in tree peony. The present study not only provides new insights into the regulatory mechanism of flavonol biosynthesis in tree peony but also identifies a potential MYB regulator that may be applied to the molecular breeding of yellow flower tree peony cultivars. In addition, PsMYB4 may interact with PsEGL3 to reduce the synthesis of anthocyanins by negatively regulating some structural genes, whereas PsSPL9 may inhibit the accumulation of anthocyanins by negatively regulating *PsDFR* alone. The functions of these candidate regulators require further study.

Materials and methods

Plant materials

P. suffruticosa plants were grown in the Tree Peony Garden of Northwest A&F University, Shaanxi Province, China (34°26' N, 108°07' E). Two cultivars, “High Noon” with pure yellow flowers and “Roufufurong” with purple–red flowers, were used as the experimental

materials. All selected plants were grown in fields with adequate light and moisture (Fig. 1). Petal samples were collected from March to April 2018 at five flowering stages, which were characterized by Zhou et al.⁵. The color-related values (L^* , a^* , b^* , C^* , and h) of fresh petals at these five stages were measured by a tristimulus color meter (CR-400, Konica Minolta, Osaka, Japan). The materials for other tests were quickly frozen in liquid nitrogen and then stored at -80°C until further use. Tobacco plants (*N. benthamiana* and *Nicotiana tabacum*) were cultivated in an incubator at 25°C with a 16/8 h day/night photoperiod. The color-related values (L^* , a^* , C^* , and h) of *N. tabacum* petals were measured according to standard methods for tree peony.

Measurement of flavonoids

Flavonoids were detected using a previously reported high-performance liquid chromatography (HPLC) method with some modifications¹¹. Frozen petal samples of tree peony and tobacco were promptly ground to a powder in liquid nitrogen with mortars and pestles. Approximately 300 mg of tree peony petal powder and 100 mg of tobacco petal powder were divided into two parts and dissolved in 6 ml (for tree peony)/2 ml (for tobacco) of methanol-hydrochloric acid (99:1, V/V) solution for anthocyanin detection and 6 ml (for tree peony)/2 ml (for tobacco) of methanol solution for the detection of other flavonoids. Next, the samples were leached in the dark at 4°C for 24 h and shaken and mixed once every 6 h. After 30 min of ultrasound-assisted extraction, the samples were centrifuged at 10,000 r.p.m. for 10 min to collect the liquid supernatants. Finally, each supernatant was filtered through a $0.22\ \mu\text{m}$ nylon micro-porous membrane. Ten microliters of pure supernatant was quantified by HPLC (LC-2030C 3D, Shimadzu, Kyoto, Japan) equipped with a diode array detector. A $4.6 \times 250\ \text{mm}$ C18 column (Shimadzu, Kyoto, Japan) was used. Eluent A was a 0.04% formic acid aqueous solution and eluent B was acetonitrile. The following gradient elution conditions were used: 5% B at 0 min, 40% B at 40 min, 100% B at 45 min, 100% B at 55 min, and 5% B at 60 min. The flow rate was 0.5 ml/min at a 40°C column temperature. Standards of THC, Ap, Lu, Ch, Km, Qu, Is, Cy3G, Pn3G, and Pg3G were purchased from Shanghai Yuanye Bio-Technology Co., Ltd (Shanghai, China). Mean values and SDs were calculated from three independent biological replicates.

Furthermore, MLR analysis was conducted by SPSS (version 23.0) to explore the relationship between flower colors and flavonoid components. The contents of THC, Ap, Lu, Ch, Km, Qu, Is, Cy3G, Pn3G, and Pg3G were independent variables, whereas the values of L^* , a^* , b^* , C^* , and h were dependent variables.

Library construction, sequencing, and data overview

Ten petal samples at five flowering stages of yellow- and purple-red-flowered tree peony cultivars were collected for isoform and RNA-seq. First, the total RNA of ten samples was extracted using the RNAprep Pure kit for plants (Tiangen, Beijing, China). RNA quality and integrity were detected using an RNA 6000 Nano Assay Kit in an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Subsequently, equal amounts of RNA from each sample were pooled together. A SMARTer PCR cDNA Synthesis Kit (Takara, Dalian, China) was applied to reverse-transcribe total RNA into cDNA. The sequencing process was conducted on a Pacific Bioscience RS II platform. Second, two petal samples at S3 from two tree peony cultivars with three biological replicates were sequenced on an Illumina HiSeq 2500 platform for comparative analysis.

Raw reads were filtered for ROIs through the Iso-seq pipeline and those with completed 5'- and 3'-cDNA primers, as well as poly A tails, were identified as FLNC transcripts. Furthermore, Iterative Clustering for Error Correction was conducted to screen consensus isoforms. High-quality full-length transcripts were confirmed under the criteria of postcorrection accuracy above 99%. ORFs of the transcripts were predicted using TransDecoder (<https://github.com/TransDecoder/TransDecoder/releases>). TFs were predicted by iTAK software (<https://omictools.com/itak-tool>)⁵⁵. Genes were annotated based on the following databases: Nr (NCBI nonredundant protein sequences), Pfam (protein family)⁵⁶, KOG/COG/eggNOG (Clusters of Orthologous Groups of proteins)^{57,58}, Swiss-Prot (a manually annotated and reviewed protein sequence database)⁵⁹, KEGG (Kyoto Encyclopedia of Genes and Genomes)⁶⁰, and GO⁶¹.

Differential expression analysis

Full-length transcripts sequenced by the Pacific Biosciences Sequel platform were regarded as a reference genome. RNA-seq reads of two samples were matched to the reference genome by Bowtie (v2.2.3)⁶². Quantification of gene expression was estimated by fragment per kilobase of transcripts per million mapped reads and the read counts were adjusted by the edgeR package before differential expression analysis⁶³.

DEGs between two samples based on the RNA-seq results were identified using the EBSeq R package⁶⁴. The false discovery rate (FDR) was corrected using the posterior probability values. $\text{FDR} < 0.01$ and $|\log_2(\text{fold change})| \geq 2$ were regarded as the thresholds for significant differential expression. GO enrichment analysis of DEGs was performed by the Goseq R package⁶⁵. The statistical enrichment of DEGs in KEGG pathways was tested by KOBAS software⁶⁶.

qRT-PCR analysis

DEGs putatively involved in the flavonoid biosynthesis pathway were selected for qRT-PCR analysis. Total RNA was extracted with the RNAprep Pure kit for plants (Tiangen, Beijing, China) and first-strand cDNA was synthesized using the PrimeScript™ RT Master Mix reverse transcription kit (Takara, Dalian, China). After diluting the cDNA template five times to $200 \text{ ng } \mu\text{L}^{-1}$, qRT-PCR was performed using TB Green TaKaRa Premix Ex Taq™ II (TaKaRa, Dalian, China) according to the manufacturer's instructions. The reaction took place under the following conditions: denaturation at 95°C for 15 s and 45 cycles of amplification (95°C for 5 s, 58°C for 30 s, and 72°C for 31 s). *PsUbiquitin* was used as an internal reference for the expression level normalization of DEGs. Relative expression levels were calculated by the $2^{-\Delta\Delta\text{CT}}$ method⁶⁷. The specific primers are listed in Supplementary Table 7. Three independent biological replicates were used in each qRT-PCR assay.

Identification of candidate TF genes in the flavonoid pathway

Two *MYB*, two *bHLH*, and one *SPL* genes implicated in the flavonoid biosynthesis pathway were selected. Phylogeny trees of their corresponding proteins along with 101 MYBs, 94 bHLHs, and 16 SPLs from *Arabidopsis* were constructed with MEGA 6.0 using the neighbor-joining clustering method. Sequence homology alignment was performed with DNAMAN software. Structural genes related to flavonoid synthesis were screened and clustered based on their expression profiles using TB tools.

Interaction network analysis

The interaction network was established on the basis of Pearson's correlation coefficients, which were calculated in the R environment (<https://www.r-project.org/>); correlations with a coefficient of $R \geq 0.5$ or $R \leq -0.5$ were retained. The coexpressed genes with strong interconnections were considered hub genes. The relationships between candidate genes, including TF genes and structural genes, and flavonoid components were visualized by Cytoscape (v.3.7.0).

Subcellular localization analysis of *PsMYB111*

The full-length ORF of *PsMYB111* without the termination codon was cloned into the pCambia1302-GFP vector. The primers are listed in Supplementary Table 7. Subsequently, the recombinant plasmid and control pCambia1302-GFP plasmid were transferred into *Agrobacterium* strain GV3101 by the freeze-thaw method. The *Agrobacterium* containing the target plasmid was resuspended in infiltration buffer (with 10 mM MES, 10 mM MgCl_2 , and 100 mM AS) to an OD600 of 0.4 and stationarily cultured for 2 h until infiltration. The

Agrobacterium mixture was then injected into two young leaves of each *N. benthamiana* plant from the lower epidermis via a syringe without a needle. The infiltrated plants were grown in a growth chamber under the same conditions described above for 72 h, and GFP and 4',6-diamidino-2-phenylindole fluorescence was observed under a Nikon C2-ER confocal laser scanning microscope (Nikon, Tokyo, Japan). All transient expression assays were repeated three times.

Generation of *PsMYB111*-overexpressing tobacco

After the pCambia1302-*PsMYB111* construct was transferred into the *Agrobacterium* strain GV3101, sterilized *N. tabacum* leaf disc transformation was performed following previously described methods⁶⁸. Hygromycin and kanamycin were used to screen the transgenic tobacco lines. Rooted transgenic plants were transferred to a soil mix and grown in a greenhouse until flowering. Flowers of T2 transgenic plants were used for the detection of color indices, quantification of flavonoid levels, and qRT-PCR verification as described above. Specific primers are shown in Supplementary Table 8.

Dual-luciferase assays

The full-length ORF of *PsMYB111* was cloned into a pGreenII62-SK vector (effector). The promoter sequences of *PsCHS* and *PsFLS* were isolated from the genomic DNA of "High Noon" using a genome walking kit (TaKaRa, Dalian, China) and then inserted into the pGreenII0800 LUC vector (reporter). All effector and reporter vectors were transformed into *Agrobacterium* strain GV3101 and isolated with kanamycin ($50 \text{ mg } \text{L}^{-1}$). Transient expression in *N. benthamiana* was conducted following the method described in subcellular localization assays. The enzyme activities of firefly luciferase (LUC) and *Renilla* luciferase (REN) were detected at 72 h after infiltration with a dual-luciferase assay system on GloMax® Discover (Promega, Madison, USA). Only the promoter-LUC reporter construct with no effector was regarded as a blank control. Three independent experiments were conducted for each combination and all experiments were technically repeated three times. The specific primers used for genome walking and dual-luciferase assays are listed in Supplementary Table 9.

Acknowledgements

We are particularly grateful to Dr. Yong Jia of Murdoch University for language polishing. This research was supported by the National Key R&D Program of China (Grant number 2019YFD1001505) and the National Natural Science Foundation of China (Grant number 31800599).

Author contributions

X.L., Q.S., and Y.Z. conceived the research, designed the experiments, and wrote the manuscript. S.W., X.L., S.L., and Y.F. performed the experiments, and X.L. analyzed the results. D.S. supervised the experiments. D.S. and L.N. modified the language and revised the manuscript critically. All authors contributed to the article and approved the submitted version.

Data availability

The raw sequence data reported in this study have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2017) in the National Genomics Data Center (Nucleic Acids Res 2020), Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, under accession number CRA005005, which is publicly accessible at <https://bigd.big.ac.cn/gsa>.

Conflict of interest

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

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41438-021-00666-0>.

Received: 15 March 2021 Revised: 8 July 2021 Accepted: 27 July 2021

Published online: 01 November 2021

References

- Li, J. J. *Chinese Tree Peony and Herbaceous Peony* (China Forestry Publishing House, 1999).
- Gu, Z. Y. et al. A novel R2R3-MYB transcription factor contributes to petal blotch formation by regulating organ-specific expression of *PsCHS* in tree peony (*Paeonia suffruticosa*). *Plant Cell Physiol.* **0**, 1–13 (2018).
- Yang, Y. et al. Germplasm resources and genetic breeding of *Paeonia*: a systematic review. *Hortic. Res.* **7**, 107 (2020).
- Han, K., Wang, E. Q. & Liang, C. A. Utilization status and prospect of wild yellow peony in peony breeding. *Northwest Hortic.* **05**, 70–71 (2018).
- Zhou, L., Wang, Y. & Peng, Z. H. Advances in study on formation mechanism and genetic engineering of yellow flowers. *Silvae Sin.* **45**, 111–119 (2009).
- Huang, J. X., Wang, L. S., Li, X. M. & Lu, Y. Q. Advances in molecular basis and evolution of floral color variation. *Chin. J. Bot.* **23**, 321–333 (2006).
- Hosoki, T., Hamada, M., Kando, T., Moriaki, R. & Inaba, K. Comparative study of anthocyanin in tree peony flowers. *J. Jpn. Soc. Hort. Sci.* **60**, 395–403 (1991).
- Wang, L. S. et al. Analysis of petal anthocyanins to investigate flower coloration of Zhongyuan (Chinese) and Daikon Island (Japanese) tree peony cultivars. *J. Plant Res.* **114**, 33–43 (2001).
- Wang, L. S. et al. Chemical taxonomy of the xibei tree peony from China by floral pigmentation. *J. Plant Res.* **117**, 47–55 (2004).
- Zhang, J. et al. Comparison of anthocyanins in non-blotches and blotches of the petals of xibei tree peony. *Sci. Hortic.* **114**, 104–111 (2007).
- Li, C. H. et al. Flavonoid composition and antioxidant activity of tree peony (*Paeonia* section *Moutan*) yellow flowers. *J. Agric. Food Chem.* **57**, 8496–8503 (2009).
- Fan, J., Zhu, W., Kang, H., Ma, H. & Tao, G. Flavonoid constituents and antioxidant capacity in flowers of different zhongyuan tree peony cultivars. *J. Funct. Foods* **4**, 147–157 (2012).
- Wang, S. L. et al. Composition of peony petal fatty acids and flavonoids and their effect on *Caenorhabditis elegans* lifespan. *Plant Physiol. Biochem.* **155**, 1–12 (2020).
- Shi, Q. Q. et al. Transcriptomic analysis of *Paeonia delavayi* wild population flowers to identify differentially expressed genes involved in purple-red and yellow petal pigmentation. *PLoS ONE* **10**, 1–20 (2015a).
- Koes, R., Verweij, W. & Quattrocchio, F. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* **10**, 236–242 (2005).
- Tanaka, Y. & Ohmiya, A. Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. *Curr. Opin. Biotechnol.* **19**, 190–197 (2008).
- Hichri, I. et al. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* **62**, 2465–2483 (2011).
- Spelt, C., Quattrocchio, F., Mol, J. & Koes, R. ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. *Plant Cell* **14**, 2121–2135 (2002).
- Zhou, H. et al. Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. *Plant J.* **82**, 105–121 (2015).
- Li, C. et al. PyWRKY26 and PybHLH3 cotargeted the PyMYB114 promoter to regulate anthocyanin biosynthesis and transport in red-skinned pears. *Hortic. Res.* **7**, 37 (2020).
- Gou, J. Y., Felippes, F. F., Liu, C. J., Weigel, D. & Wang, J. W. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell* **23**, 1512–1522 (2011).
- Zhou, L., Wang, Y. & Peng, Z. H. Molecular characterization and expression analysis of chalcone synthase gene during flower development in tree peony (*Paeonia suffruticosa*). *Afr. J. Biotechnol.* **10**, 1275–1284 (2011).
- Zhou, L., Wang, Y. & Peng, Z. H. Cloning and expression analysis of dihydroflavonol 4-reductase gene *PsDFR1* from tree peony (*Paeonia suffruticosa*). *J. Plant Physiol.* **47**, 885–892 (2011b).
- Zhou, L. et al. Overexpression of *Ps-CHL1*, a homologue of the chalcone isomerase gene from tree peony (*Paeonia suffruticosa*), reduces the intensity of flower pigmentation in transgenic tobacco. *Plant Cell Tiss. Org.* **16**, 285–295 (2014).
- Shi, Q. Q., Zhou, L., Li, K. & Wang, Y. Isolation and expression of *PlbHLH3* transcription factor genes in *Paeonia lutea*. *For. Res.* **28**, 488–496 (2015b).
- Shi, Q. Q., Zhou, L., Wang, Y., Zhai, L. J. & Li, L. Isolation and expression of *PlWDR3* and *PlWDR18* transcription factor genes in *Paeonia lutea*. *J. Northwest Forestry Univ.* **32**, 97–103 (2017).
- Shi, L. Y., Chen, X., Chen, W., Zheng, Y. H. & Yang, Z. F. Comparative transcriptomic analysis of white and red Chinese bayberry (*Myrica rubra*) fruits reveals flavonoid biosynthesis regulation. *Sci. Hortic.* **235**, 9–20 (2018).
- Gao, L. X., Yang, H. X., Liu, H. F., Yang, J. & Hu, Y. H. Extensive transcriptome changes underlying the flower color intensity variation in *Paeonia ostii*. *Front. Plant Sci.* **6**, 1205 (2016).
- Luo, J. R., Shi, Q. Q., Niu, L. X. & Zhang, Y. L. Transcriptomic analysis of leaf in tree peony reveals differentially expressed pigments genes. *Molecules* **22**, e324 (2017).
- Guo, S. et al. Complete chloroplast genome sequence and phylogenetic analysis of *Paeonia ostii*. *Molecules* **23**, e246 (2018).
- Zhang, C. et al. Transcriptomic analysis of cut tree peony with glucose supply using the RNA-Seq technique. *Plant Cell Rep.* **33**, 111–129 (2014).
- Chaisson, M. J. P. et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature* **517**, 608–611 (2014).
- Chen, X. et al. The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development. *Cell* **158**, 1187–1198 (2014).
- Huddleston, J. et al. Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res.* **24**, 688–696 (2014).
- Hackl, T., Hedrich, R., Schultz, J. & Forster, F. Proovread: large-scale high-accuracy pacbio correction through iterative short read consensus. *Bioinformatics* **30**, 3004–3011 (2014).
- Wang, X. C. et al. *Arabidopsis* MYB4 plays dual roles in flavonoid biosynthesis. *Plant J.* **101**, 3 (2020).
- Dubos, C. et al. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **15**, 573–581 (2010).
- Czemmel, S. et al. The grapevine R2R3-MYB transcription factor WMYB1 regulates flavonol synthesis in developing grape berries. *Plant Physiol.* **151**, 1513–1530 (2009).
- Zhong, C. M. et al. The R2R3-MYB transcription factor GhMYB1a regulates flavonol and anthocyanin accumulation in *Gerbera hybrida*. *Hortic. Res.* **7**, 78 (2020).
- Feller, A., Hernandez, J. M. & Grotewold, E. An ACT-like domain participates in the dimerization of several plant basic-helix-loop-helix transcription factors. *J. Biol. Chem.* **281**, 28964–28974 (2006).
- Guo, L. P. et al. Transcriptome and chemical analysis reveal putative genes involved in flower color change in *Paeonia* 'Coral Sunset'. *Plant Physiol. Biochem.* **138**, 130–139 (2019).
- Yang, Q., Yuan, H. H. & Sun, X. B. Preliminary studies on the changes of flower color during the flowering period in two tree peony cultivars. *Acta Hortic. Sin.* **42**, 930–938 (2015).
- Zhao, D. Q., Tao, J., Han, C. X. & Ge, J. T. Flower color diversity revealed by differential expression of flavonoid biosynthetic genes and flavonoid accumulation in herbaceous peony (*Paeonia lactiflora* Pall.). *Mol. Biol. Rep.* **39**, 11263–11275 (2012).
- Lou, Q. et al. Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* **65**, 3157–3164 (2014).

45. Zhao, D. Q. et al. Transcriptome sequencing of a chimaera reveals coordinated expression of anthocyanin biosynthetic genes mediating yellow formation in herbaceous peony (*Paeonia lactiflora* Pall.). *BMC Genomics* **15**, 689–706 (2014).
46. Davies, K. M. et al. Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-Reductase. *Euphytica* **131**, 259–268 (2003).
47. Duan, Y. B. et al. Genotypic variation of flavonols and antioxidant capacity in broccoli. *Food Chem.* **338**, 127997 (2021).
48. Zhang, Y. Y. et al. Comparative transcriptomics provides insight into floral color polymorphism in a *Pleione limprichtii* orchid population. *Int. J. Mol. Sci.* **21**, e247 (2020).
49. Zhao, D. Q., Wei, M. R., Shi, M., Hao, Z. J. & Tao, J. Identification and comparative profiling of miRNAs in herbaceous peony (*Paeonia lactiflora* Pall.) with red/yellow bicoloured flowers. *Sci. Rep.* **7**, 44926 (2017).
50. Stracke, R. et al. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* **50**, 660–677 (2007).
51. Nakatsuka, T. et al. Isolation and characterization of *GtMYBP3* and *GtMYBP4*, orthologues of R2R3-MYB transcription factors that regulate early flavonoid biosynthesis, in gentian flowers. *J. Exp. Bot.* **63**, 6505–6517 (2012).
52. Blanco, E. et al. Isolation and characterization of the flavonol regulator CcMYB12 from the globe artichoke [*Cynara cardunculus* var. *scolymus* (L.) Fiori]. *Front. Plant Sci.* **9**, 941 (2018).
53. Huang, W. et al. A R2R3-MYB transcription factor regulates the flavonol biosynthetic pathway in a traditional Chinese medicinal plant, *Epimedium sagittatum*. *Front. Plant Sci.* **7**, 1089 (2016).
54. Pandey, A. et al. AtMYB12 expression in tomato leads to large scale differential modulation in transcriptome and flavonoid content in leaf and fruit tissues. *Sci. Rep.* **5**, 12412 (2015).
55. Zheng, Y. et al. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* **9**, 1667–1670 (2016).
56. Finn, R. D. et al. Pfam: the protein families database. *Nucleic Acids Res.* **42**, 222–230 (2014).
57. Koonin, E. V. et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **5**, R7 (2004).
58. Huerta-Cepas, J. et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**, 286–293 (2016).
59. Bairoch, A. & Boeckmann, B. The SWISS-PROT protein sequence data bank. *Nucleic Acids Res.* **19**, 2247–2249 (1991).
60. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, 277–280 (2004).
61. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, e357 (2012).
63. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Biogeosciences* **26**, 139–140 (2010).
64. Leng, N. et al. EBSeq-HMM: a Bayesian approach for identifying gene-expression changes in ordered RNA-seq experiments. *Bioinformatics* **31**, 2614–2622 (2015).
65. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Goseq: gene ontology testing for RNA-seq datasets. *ResearchGate* **28**, e573 (2012).
66. Xie, C. et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39**, 316–322 (2011).
67. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–408 (2001).
68. Horsch, R. et al. A simple and general method for transferring genes into plants. *Science* **227**, 1229–1232 (1985).
69. Intelmann, D., Jaros, D. & Rohm, H. Identification of colour optima of commercial tomato catsup. *Eur. Food Res. Technol.* **22**, 662–666 (2005).