

RESEARCH ARTICLE

Open Access



Transcriptomic and chemical analyses to identify candidate genes involved in color variation of sainfoin flowers

Yu Qiao¹, Qiming Cheng¹, Yutong Zhang¹, Wei Yan¹, Fengyan Yi² and Fengling Shi^{1*}

Abstract

Background: Sainfoin (*Onobrychis viciifolia* Scop) is not only a high-quality legume forage, but also a nectar-producing plant. Therefore, the flower color of sainfoin is an important agronomic trait, but the factors affecting its flower phenotype are still unclear. To gain insights into the regulatory networks associated with metabolic pathways of coloration compounds (flavonoids or anthocyanins) and identify the key genes, we conducted a comprehensive analysis of the phenotype, metabolome and transcriptome of WF and AF of sainfoin.

Results: Delphinidin, petunidin and malvidin derivatives were the main anthocyanin compounds in the AF of sainfoin. These substances were not detected in the WF of sainfoin. The transcriptomes of WF and AF in sainfoin at the S1 and S3 stages were obtained using the Illumina HiSeq4000 platform. Overall, 10,166 (4273 upregulated and 5893 downregulated) and 15,334 (8174 upregulated and 7160 downregulated) DEGs were identified in flowers at S1 and S3 stages, respectively (WF-VS-AF). KEGG pathway annotations showed that 6396 unigenes were annotated to 120 pathways and contained 866 DEGs at S1 stages, and 6396 unigenes were annotated to 131 pathways and included 1546 DEGs at the S3 stage. Nine DEGs belonging to the “flavonoid biosynthesis” and “phenylpropanoid biosynthesis” pathways involved in flower color formation were identified and verified by RT-qPCR analyses. Among these DEGs, *ACL3*, *FLS*, *ANS*, *CHS*, *DFR* and *CHI2* exhibited downregulated expression, and *F3H* exhibited upregulated expression in the WF compared to the AF, resulting in a decrease in anthocyanin synthesis and the formation of WF in sainfoin.

Conclusions: This study is the first to use transcriptome technology to study the mechanism of white flower formation in sainfoin. Our transcriptome data will be a great enrichment of the genetic information for sainfoin. In addition, the data presented herein will provide valuable molecular information for genetic breeding and provide insight into the future study of flower color polymorphisms in sainfoin.

Keywords: Sainfoin, Flavonoid, Anthocyanin, Transcriptome, DEGs, Flower color

* Correspondence: sfl0000@126.com

¹College of Grassland Resources and Environment, Key Laboratory of Forage Cultivation, Processing and High Efficient Utilization of the Ministry of Agriculture and Key Laboratory of Grassland Resources of the Ministry of Education, Inner Mongolia Agricultural University, Hohhot 010011, China
Full list of author information is available at the end of the article



© 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Sainfoin (*Onobrychis viciifolia* Scop) is a perennial herbaceous forage legume [1] that is widely distributed in temperate regions of the northern part of the world [2]. It can be used as hay, pellets, grazing and silage because of its high palatability and nutritious forage properties [2–4]. It is particularly valued for having appropriate condensed tannin content to reduce greenhouse gas emissions by preventing bloating in grazing animals [2, 5, 6]. Studies have found that sainfoin can also be used as an ornamental plant because its flowers form an erect raceme and the flowering period is 2–3 weeks [2]. Some studies have also found that sainfoin can be used as a nectar plant due to its beautiful flower petals and high sugar content [7, 8]. Therefore, studying the flower color of sainfoin is of great significance for the development of multifunctional applications.

Flower color is one of the most important horticultural characteristics of plants in nature [9]. Flower color changes can perform important ecological functions by attracting pollinators and affecting the reproductive success of flowering plants [10] and are crucial to plant evolution [11, 12]. In addition, the color of flowers is directly or indirectly related to the agronomic traits of plants, and classic breeding methods have been widely used to develop varieties with flowers varying in both color and intensity [13, 14]. Flower color is affected by many factors, the most important of which are different kinds of plant pigments, such as flavonoids and anthocyanins [15, 16]. Anthocyanins are part of flavonoids, that are the main components to flower pigments, and they are produced by highly conserved structural and regulatory components [17, 18]. During the flowering process, somatic mutations from recessive white to pigment-reversible alleles occur, and the variegation of flowers is inevitably the result of differential gene expression regulation [19]. The anthocyanin biosynthetic pathway includes multiple metabolic processes involving seven core structural genes, such as *CHS*, *LAR*, *DFR*, and *ANS*, as well as several branching enzyme genes [20]. So far, genes associated with flower color and flavonoids have been found in many plants, such as white clover (*Trifolium repens*), alfalfa (*Medicago sativa*), white *Primula vulgaris*, and strawberry (*Fragaria × ananassa*) [14, 18, 21, 22]. However, the molecular mechanisms of the corresponding candidate genes underlying flower pigmentation in sainfoin are still unclear.

Transcriptome technology can provide unique insights into the molecular characteristics of nonmodel plants without a reference genome, especially in the study of flower color. It has been successful in many plants, for example, sheepgrass (*Leymus chinensis*), Siberian wildrye (*Elymus sibiricus*), ornamental crabapple (*Malus prunifolia*) and chrysanthemum (*Dendranthema morifolium*)

[23–26]. As far as we know, there has no report of research employing RNA-Seq to study the color of sainfoin flowers. Therefore, the mechanism of color mutation in sainfoin should be understood, and the key genes should be identified. In our study, WF of a sainfoin mutant resulting from EMS treatment and AF were used as the experimental model. Transcriptome technology, CIELAB color space and UPLC were used to assess the variation in related genes and the differences in flavonoid intermediates in the anthocyanin biosynthetic pathway that cause color transitions. The results of this study provide a theoretical basis for future sainfoin molecular breeding, provide an important molecular basis for further studies on colored-flower sainfoin and are crucial for understanding the color formation mechanism.

Results

Petal color measurements

The petal color parameters of sainfoin are shown in Table 1. The L^* value, which varies in color scale from 0 (black) to 100 (white) and represents lightness, in our study was 28.67 in WF and 21.02 in AF. The a^* value, which represents redness, was 3.58 in WF and 13.59 in AF. The b^* value, which represents blueness, was 10.75 in WF and 2.32 in AF. The C^* value, which represents color vividness, was 11.29 in WF and 13.79 in AF. The h° value, which represents basic color, was 71.58 in WF and 9.70 in AF. In summary, WF had higher ($P < 0.05$) L^* , b^* and h° values than AF. In contrast, the a^* and C^* values of WF were significantly lower than those of AF ($P < 0.05$). Therefore, these factors indicate that the petal colors of sainfoin are different.

Major classes of coloration compounds in sainfoin petals

UPLC analysis revealed that seven flavonoids and their derivatives were detected in both the WF and AF: kaempferol-3-O-rhamnosylrutinoside, rutin, kaempferol-3-O-glucoside-phenylpropanoic ester, quercetin-3-O-glucoside, kaempferol-3-O-rutinoside, isorhamnetin-3-O-rutinoside and kaempferol-3-O-glucoside (Table 2). However, three flavonoids (kaempferol-3-O-glucoside-courmaric ester, quercetin-3-O-rhamnosidyl-p-courmaric ester, kaempferol) were detected in AF but not WF. Among these compounds, only kaempferol-3-O-glucoside-phenylpropanoic ester showed a higher content in WF than AF ($P < 0.05$), and the other flavonoid and derivative contents were lower in WF than in AF ($P < 0.05$). However, five anthocyanidins and their derivatives (delphinidin-3,5-diglucoside, petunidin-3,5-diglucoside, delphinidin-3-rutinoside, petunidin-rutinoside and malvidin-rutinoside) were only detected in AF. No anthocyanidin was detected in WF. Similarly, two procyanidins were detected in our study; only

Table 1 Chromatographic parameters of sainfoin petals

Materials	CIELAB coordinate				
	L*	a*	b*	C*	h°
WF	28.67 ± 0.44a	3.58 ± 0.26b	10.75 ± 0.23a	11.29 ± 0.35b	71.58 ± 1.71a
AF	21.02 ± 0.27b	13.59 ± 0.42a	2.32 ± 0.06b	13.79 ± 0.58a	9.70 ± 0.61b

L* means lightness, a* and b* means chromatic components, C* means chroma, h° means hue angle. Values with different letters in the same column are significantly different ($P < 0.05$).

proanthocyanidin was detected in WF, while catechin hydrate was detected only in AF (Table 2). In summary, the differences in the types and contents of coloration compounds in sainfoin petals were the main reasons for the color change.

De novo assembly of the sainfoin transcriptome

Twelve libraries of total RNA extracted from sainfoin petals (WF and AF) at the S1 and S3 stages were constructed for transcriptome sequencing. A total of 547,329,260 raw reads with a total of 82,099,389,000 nt were obtained. Then low-quality reads were filtered out, a total of 546,658,468 clean reads with a total of 81,319,631,904 nt were obtained from the twelve sequencing libraries for further analysis (Additional file 1). Briefly, after removing low-quality and contaminating reads, clean reads were retained for further analysis. Finally, Trinity method was used to assemble a total of 53,009 unigenes with an N50 of 1587 nt, with lengths ranging from 201 to 15,519 nt and a mean length of 903 nt (Additional file 2).

Gene annotation of the sainfoin transcriptome

In total, 31,887 unigenes (60.15% of the 53,009 total unigenes) were annotated against at least one database using BLASTx ($E\text{-value} < 1 \times 10^{-5}$) (Table 3). Among 31,858 unigenes (60.10%), 19,994 (37.72%), 16,327 (30.80%) and 12,069 (22.77%) were annotated to the Nr, Swiss-Prot, KOG and KEGG databases, respectively (Table 3). In addition, 6, 2, 15 and 8889 unigenes were annotated to only the KEGG, KOG, Swiss-Pro and Nr databases, respectively (Fig. 1a). According to the Nr database, a total of 5242 unigenes (16.45% of the total 31,858 unigenes annotated to Nr) showed homology ($1 \times E^{-20} < E\text{-value} \leq 1 \times E^{-5}$), 11,383 (35.73%) unigenes showed strong homology ($1 \times E^{-100} < E\text{-value} \leq 1E^{-20}$), and 15,233 (47.82%) unigenes showed very strong homology ($E\text{-value} \leq 1E^{-100}$) (Fig. 1b). For the distribution of the species hits obtained by BLAST, 8585 unigenes matched the homologous sequences of *Medicago truncatula*, while 8377, 3651, 2434, 1195 and 1139 unigenes matched the homologous sequences of *Cicer arietinum*, *Cajanus cajan*, *Glycine max*, *Glycine soja* and *Lupinus angustifolius*, respectively (Fig. 1c). Based on GO

Table 2 The flavonoid, anthocyanidin and procyanidin contents of petals in sainfoin

Substance		WF (µg/g)	AF (µg/g)
Flavonoids	Kaempferol-3-O-rhamnosylrutinoside	1107.23 ± 113.71b	1511.58 ± 128.24a
	Rutin	295.03 ± 22.35b	741.71 ± 111.61a
	Kaempferol-3-O-glucoside-p-coumaric ester	–	169.33 ± 31.25
	Kaempferol-3-O-glucoside-phenylpropanoic ester	379.73 ± 48.03a	232.45 ± 21.57b
	Quercetin-3-O-glucoside	29.99 ± 2.59b	476.43 ± 38.78a
	Quercetin-3-O-rhamnosidel-p-coumaric ester	–	512.69 ± 42.96
	Kaempferol-3-O-rutinoside	1816.62 ± 56.42b	3499.48 ± 392.30a
	Isorhamnetin-3-O-rutinoside	33.08 ± 2.66b	145.76 ± 24.06a
	Kaempferol-3-O-glucoside	439.76 ± 23.23b	2053.49 ± 63.21a
	Kaempferol	–	12.50 ± 2.38
Anthocyanidins	Delphinidin-3,5-diglucoside	–	584.81 ± 26.80
	Petunidin-3,5-diglucoside	–	3310.82 ± 80.96
	Delphinidin-3-rutinoside	–	8650.70 ± 409.36
	Petunidin-rutinoside	–	918.77 ± 15.99
	Malvidin-rutinoside	–	3121.98 ± 84.71
Procyanidins	(+)- Catechin hydrate	–	61.87 ± 2.98
	Proanthocyanidin	3.30 ± 0.53	–
Total		41,04.74 ± 177.65b	26,004.38 ± 1369.67a

Table 3 Assembly of sainfoin Transcriptome

Annotation	Number of unigenes	Annotation percentage (%)
Annotation in Nr	31,858	60.10
Annotation in Swiss-Prot	19,994	37.72
Annotation in KOG	16,327	30.80
Annotation in KEGG	12,069	22.77
Annotation genes	31,887	60.15
without annotation genes	21,122	39.85
Total unigenes	53,009	100.00

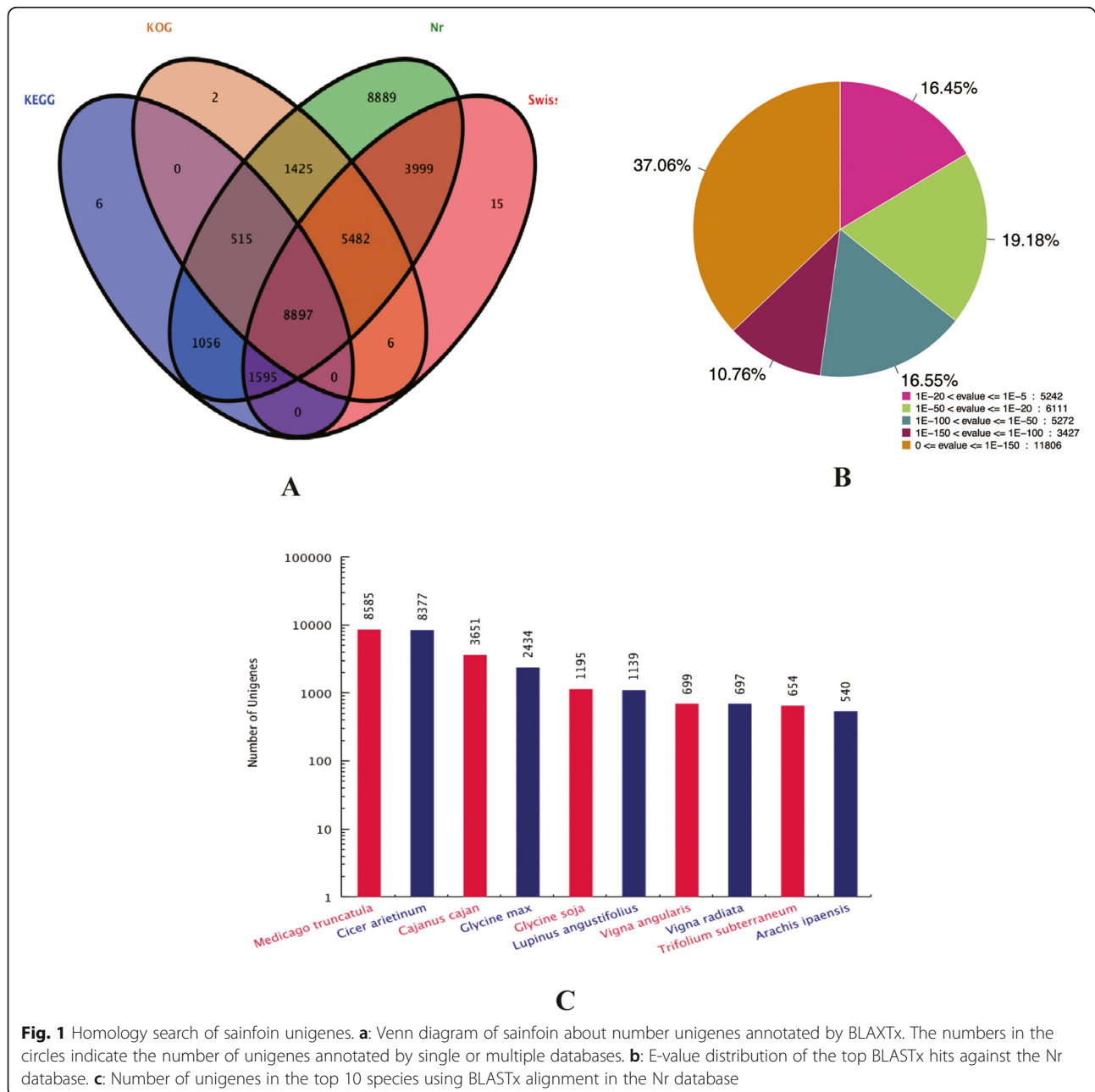
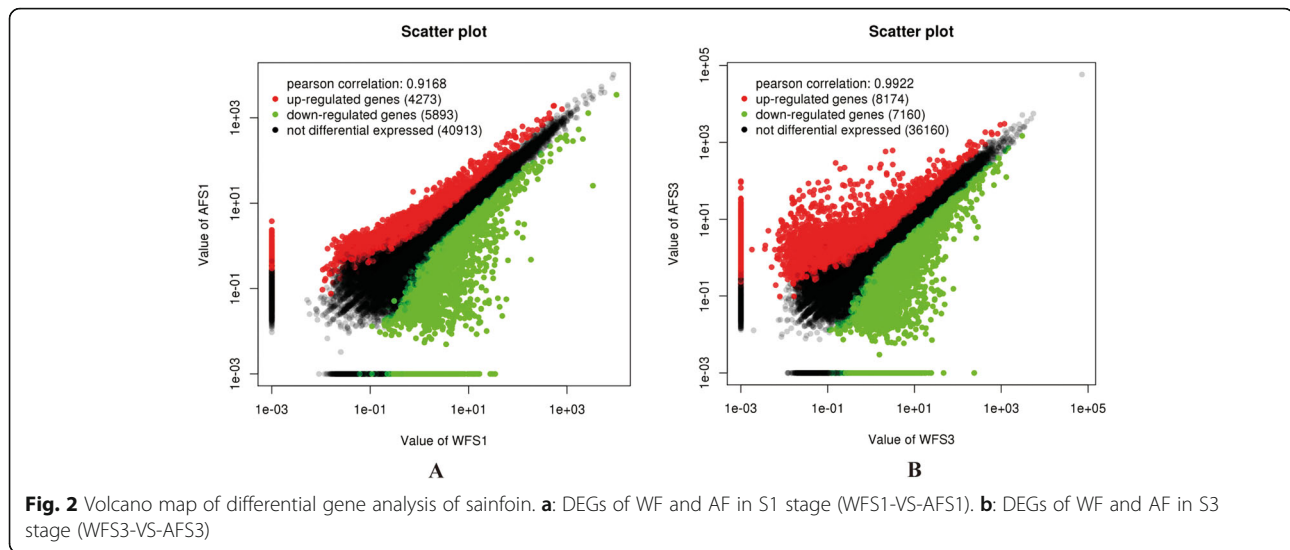


Fig. 1 Homology search of sainfoin unigenes. **a:** Venn diagram of sainfoin about number unigenes annotated by BLASTx. The numbers in the circles indicate the number of unigenes annotated by single or multiple databases. **b:** E-value distribution of the top BLASTx hits against the Nr database. **c:** Number of unigenes in the top 10 species using BLASTx alignment in the Nr database



analysis, a total of 11,558 (21.80%) unigenes were successfully annotated using GO assignments and categorized into three main categories: biological process, cellular component and molecular function (Additional file 3).

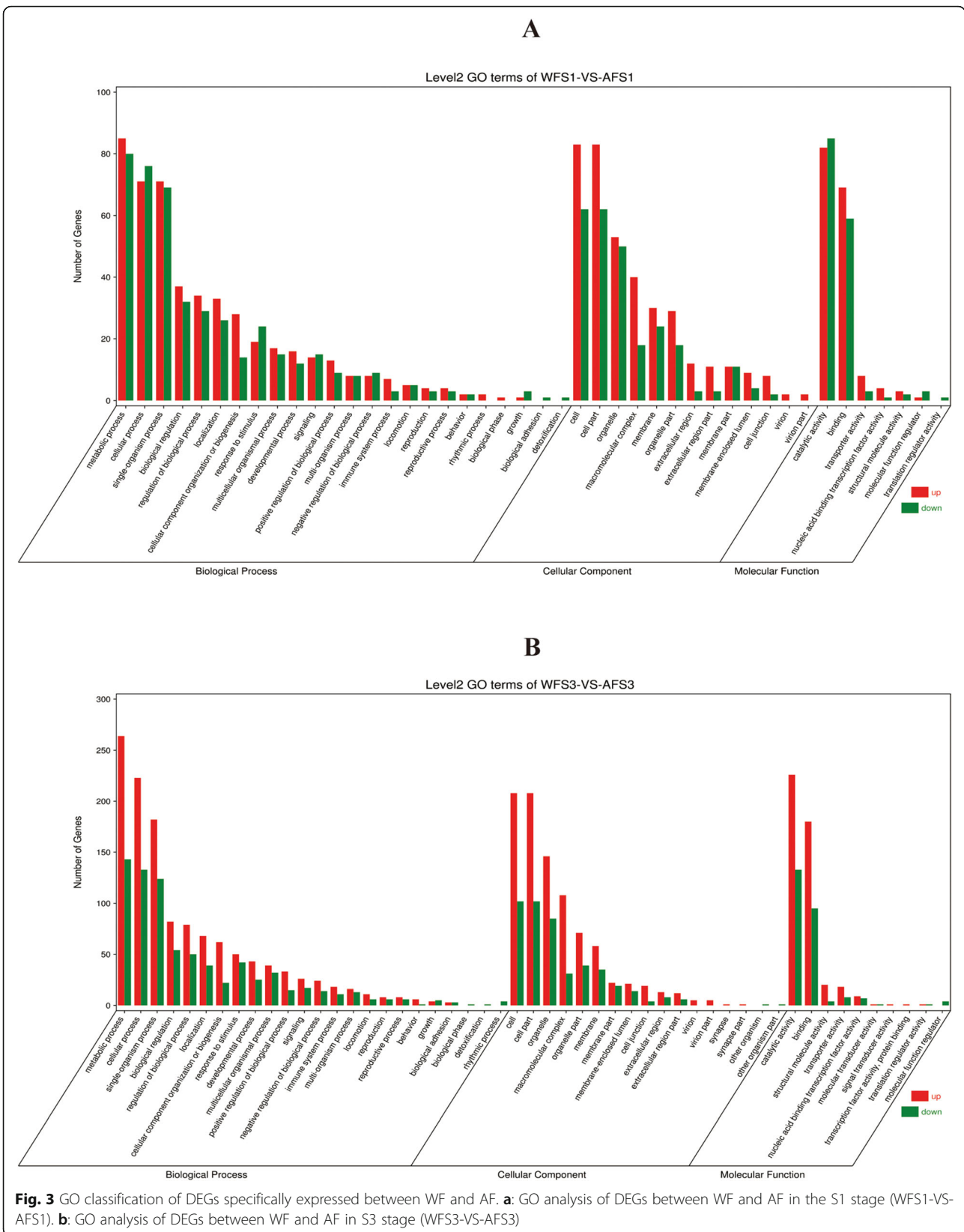
Identification and analysis of DEGs

Genes acquired by the transcriptome with a false discovery rate (FDR) < 0.05, absolute log₂ ratio ≥ 1 were selected as significant DEGs for subsequent analysis (Additional file 4, Fig. 2). To analyze the difference in flower color formation in WF and AF, we compared the DEGs of WF and AF at the same flower developmental stage. Based on these analyses, in the S1 stage, we identified 4273 upregulated DEGs and 5893 downregulated DEGs (WFS1-VS-AFS1) (Fig. 2a). Similarly, in the S3 stage, 8174 unigenes were upregulated, the other 7160 unigenes were downregulated (WFS3-VS-AFS3) (Fig. 2b). The number of DEGs in the S3 stage was greater than that in the S1 stage, indicating that with the prolongation of the growth period, the differences between the WF and AF of sainfoin increased gradually.

GO analysis of DEGs

To analyze the functions of DEGs, we subjected the DEGs between WF and AF at the same flower developmental stage to enrichment analysis with GO annotation terms. In the S1 stage, a total of 6389 DEGs were divided into three ontologies: biological process, cellular component, and molecular function. For the biological process ontology, “metabolic process”, “cellular process” and “single-organism process” were the most frequent terms and were associated with 801, 684 and 557 DEGs, respectively. For “metabolic process”, there were 376 upregulated unigenes and 425 downregulated unigenes; for

“cellular process”, there were 313 upregulated unigenes and 381 downregulated unigenes; for “single-organism process”, there were 269 upregulated unigenes and 288 downregulated unigenes; and “detoxification” (1 unigene), “biological adhesion” (2 unigenes) and “growth” (5 unigenes) were infrequent. For the cellular component ontology, the DEGs were mainly enriched for “cell part” (155 upregulated unigenes, 165 down-regulated unigenes), “cell” (155 upregulated unigenes, 165 down-regulated unigenes), “membrane” (131 upregulated unigenes, 150 downregulated unigenes) and “organelle” (109 upregulated unigenes, 135 downregulated unigenes). For the molecular function ontology, the DEGs were mainly associated with the “catalytic activity” (293 upregulated unigenes, 431 downregulated unigenes) and “binding” (285 upregulated unigenes, 344 downregulated unigenes) subcategories (Additional file 5, Fig. 3a). Similarly, in the S3 stage, in total, 11,880 DEGs were divided into three ontologies. For the biological process ontology, the DEGs were also enriched for genes involved in “metabolic process” (797 upregulated unigenes, 699 downregulated unigenes), “cellular process” (704 upregulated unigenes, 600 downregulated unigenes) and “single-organism process” (539 upregulated unigenes, 450 downregulated unigenes). In the cellular component ontology, the DEGs were mainly associated with the “cell part” (384 upregulated unigenes, 261 downregulated unigenes), “cell” (384 upregulated unigenes, 261 downregulated unigenes) and “membrane” (257 upregulated unigenes, 220 downregulated unigenes) subcategories. For the molecular function ontology, the DEGs were also mainly enriched for “catalytic activity” (794 upregulated unigenes, 716 downregulated unigenes) and “binding” (607 upregulated unigenes, 514 downregulated unigenes) (Additional file 5, Fig. 3b).



KEGG pathway enrichment analysis of DEGs

To exhaustively explore the biological functions of these DEGs, we carried out an enrichment analysis based on the KEGG database. A total of 6396 unigenes with 866 DEGs were assigned to 120 KEGG pathways in the WFS1-VS-AFS1 comparison and 6396 unigenes with 1546 DEGs were mapped to 131 KEGG pathways in the WFS3-VS-AFS3 comparison (Additional file 6). In the S1 stage, the DEGs between WF and AF were significantly enriched in “flavonoid biosynthesis” (ko00941), “phenylpropanoid biosynthesis” (ko00940) and “biosynthesis of secondary metabolites” (ko01110), and DEGs between WF and AF in the S3 stage were significantly enriched in “biosynthesis of secondary metabolites” (ko01110), “linoleic acid metabolism” (ko00591) and “phenylpropanoid biosynthesis” (ko00940).

In these pathways, the DEGs related to direct or indirect effects on flower color were predicted. Flavonoids, anthocyanins and their derivatives are the main flower color pigments, so we identified a total of three metabolic pathways with eleven genes that control the biosynthesis of flavonoids and anthocyanins (Table 4, Additional file 7). In the “phenylpropanoid biosynthesis” (ko00940) pathway, one DEG (*4CL3*) was annotated. In the “flavonoid biosynthesis” (ko00941) pathway, eight DEGs (*LAR*, *ANR*, *FLS*, *ANS*, *CHS*, *DFR*, *CHI2* and *F3H*) were annotated. In the “flavonoid and flavonol biosynthesis” pathway (ko00944), two DEGs (*FG3* and *PMAT1*) were annotated. All of these genes were used to analyze the expression pattern of the flower color change in sainfoin.

Quantitative real-time PCR analysis of DEGs related to defoliation

To test the reliability and reproducibility of the RNA-Seq data, gene-specific primers were designed for eleven candidate DEGs. The endogenous reference (Additional file 8) was JZ818469 gene. RNA samples

extracted from petals of WF and AF of sainfoin were used as templates, and selected genes related to flower coloration at the S1 and S3 stages were validated based on RT-qPCR. Among the candidate DEGs, only *F3H* had higher expression levels in WF than AF, the other 8 genes of expression levels were all lower in WF (Fig. 4). The significant difference in RT-qPCR data between WF and AF of sainfoin at the S1 and S3 stages was analyzed by t-test, and the results of RT-qPCR exhibited expression patterns almost identical to the RNA-Seq data patterns, which proved the reliability of the RNA-Seq data. In addition, the expression levels of two other genes (*FG3* and *PMAT1*) detected via RT-qPCR were inconsistent with the RNA-Seq data, so we did not show them in Fig. 4.

Discussion

Variations in anthocyanin components and color levels between AF and WF

Color mutants are widely used in horticulture and other crops. It has been found that there are many factors leading to flower color mutations in plants, such as ion beam mutations, gamma irradiation and EMS mutagenesis [27, 28]. Among them, flower color mutations induced by EMS have been widely used in cucumber [29], rice [30] and black cumin [31]. However, there is no report on the application of EMS mutagenesis to cause the flower color changes in sainfoin. In this study, we used EMS mutagenesis to obtain WF materials, with AF as control, using HPLC and CLELAB methods to study the chemical substances and phenotypes of AF and WF. At present, the identification and quantitative research on flavonoids and anthocyanins in sainfoin mainly focus on leaves, and less on petals [32, 33]. Regos [34] reported that sainfoin flower buds contained isorhamnetin derivatives, quercetin derivatives, rutin and catechin. In our study, in addition to the above substances, we also detected kaempferol and its derivatives, delphinidin

Table 4 Partial KEGG pathways associated with flower color of sainfoin

Function	Gene	Enzymes	ID	Enzymes no.
Phenylpropanoid biosynthesis	<i>4CL3</i>	4-coumarate: CoA ligase	Unigene0034144	[EC: 6.2.1.12]
Flavonoid biosynthesis	<i>LAR</i>	Anthocyanidin reductase	Unigene0025084	[EC: 1.17.1.3]
	<i>ANR</i>	Leucoanthocyanidin reductase	Unigene0029546	[EC: 1.3.1.77]
	<i>FLS</i>	Flavonol synthase	Unigene0029125	[EC: 1.14.20.6]
	<i>ANS</i>	Anthocyanidin synthase	Unigene0014621	[EC: 1.14.20.4]
	<i>CHS</i>	Chalcone synthase	Unigene0028334	[EC: 2.3.1.74]
	<i>DFR</i>	Dihydroflavonol 4-reductase	Unigene0023972	[EC: 1.1.1.219]
	<i>CHI2</i>	Chalcone isomerase	Unigene0035272	[EC: 5.5.1.6]
	<i>F3H</i>	Flavanone-3-hydroxylase	Unigene0006821	[EC: 1.14.11.9]
Flavone and flavonol biosynthesis	<i>FG3</i>	Flavonol-3-O-glucoside	Unigene0039000	[EC: 2.4.1.239]
	<i>PMAT1</i>	Isoflavone 7-O-glucoside-6"-O-malonyltransferase	Unigene0022146	[EC: 2.3.1.115]

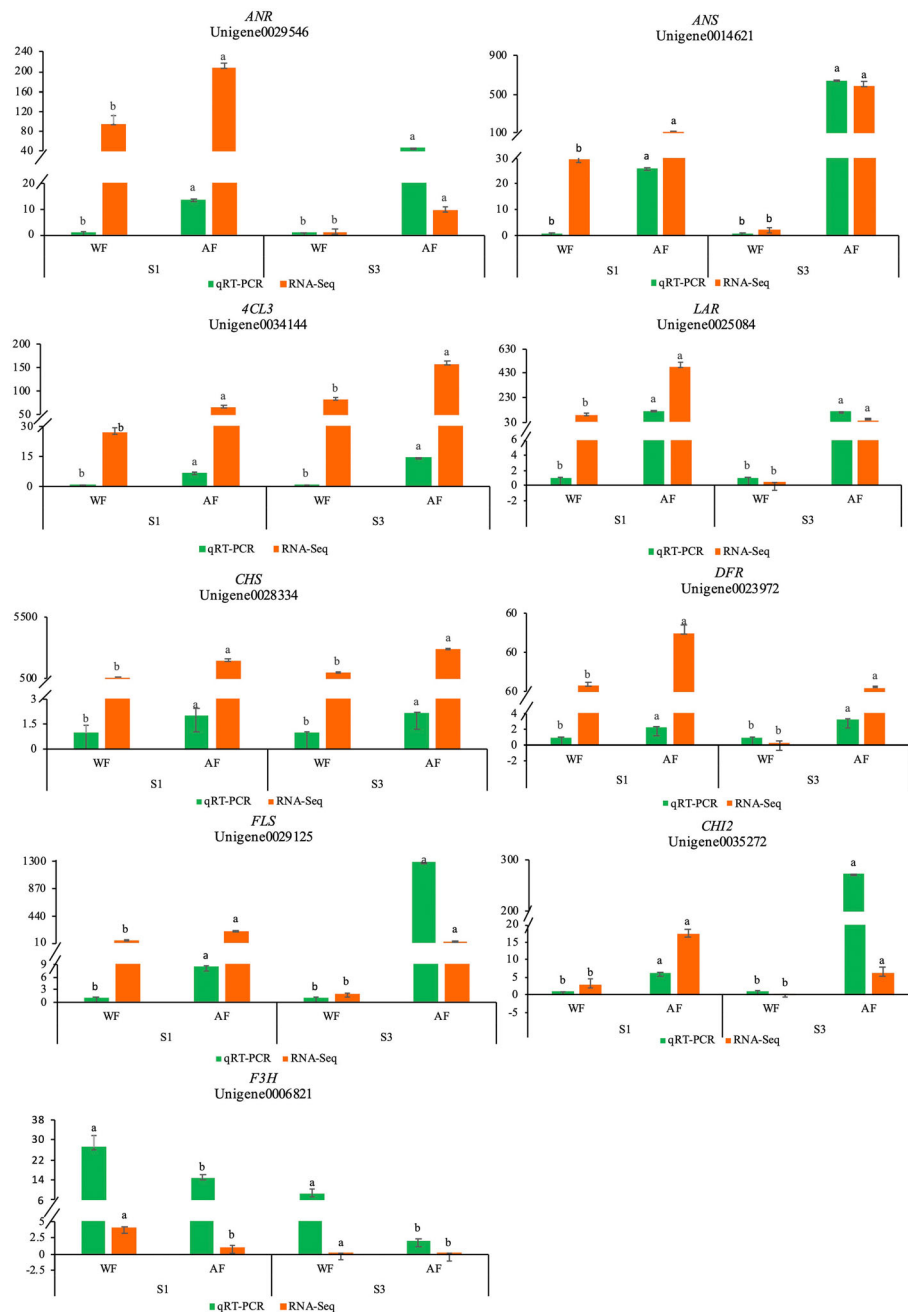


Fig. 4 Expression of flower color-related unigenes of sainfoin quantified by RNA-Seq (RNA sequencing) and RT-qPCR (quantitative real-time PCR) analysis. All RT-qPCR reactions were repeated three times for each sample, and vertical bars indicate standard errors

derivatives, petunidin derivatives, malvidin derivatives and proanthocyanidins. In this study, the total flavonoid content of AF was significantly higher than that of WF. Similar results were found in *Primula vulgaris* [35], *Fragaria ananassa* [36] and *Paeonia* [37]. This may be the main reason for the change in flower color in sainfoin. Similarly, we detected delphinidin, petunidin, and malvidin derivatives and catechin in the anthocyanin biosynthetic pathway in only the AF of sainfoin. A similar

phenomenon was discovered in the study by Lou [38], who reported that anthocyanins and their derivatives were not detected in the WF of *Muscari armeniacum* f. album. This is because the reduction in anthocyanins causes the petals to lighten in color [39].

Changes in flower phenotype in plants were related to the composition of pigments. Zhong found that the decrease in anthocyanins in *Paeonia lactiflora* resulted in an increase in the L^* value and a decrease in the a^* value

during the flowering period [40]. A similar phenomenon was discovered by Han, who reported that 14 monomolecular anthocyanins in wine were negatively correlated with L^* , b^* and h° values, and positively correlated with a^* and C^* values [41]. In our study, we found that the decrease in anthocyanin contents in WF (compared to AF) resulted in an increase in L^* , b^* and h° values and a decrease in a^* and C^* values. Our findings are consistent with previous studies. In summary, the change in flower color was closely related to the types and contents of coloration compounds in sainfoin petals, and the synthesis of those compounds was controlled by related genes.

Genes involved in the flavonoid biosynthesis pathway are differentially regulated

The biosynthesis of flavonoids and anthocyanins has been a research hotspot in the field of plant secondary metabolism, and there is now a good understanding of the nature of related signals and how the signal transduction pathways connect biosynthetic genes [42, 43]. Previous studies have found that flavonoids are one of the most important pigments in many plant petals, and anthocyanins, the end product of the flavonoid biosynthetic pathway, make the widest range of colors, from light yellow to blue-violet [44, 45]. Our results showed that the color difference between AF and WF in sainfoin is due to the loss of malvidin, petunidin, and delphinidin derivatives in WF. This is due to the hindrance of anthocyanin biosynthesis in WF, which is largely regulated by genes [14]. Thus, the key genes for the metabolism leading to WF were identified by comparing the abundances of candidate genes in the AF and WF transcriptomes. We found 9 different genes in 2 metabolic pathways related to flower color formation, namely, the “phenylpropanoid” (*4CL3*) and “flavonoids” (*LAR*, *ANR*, *ANS*, *FLS*, *CHS*, *DFR*, *CHI2* and *F3H*) pathways. In our study, *4CL3*, *LAR*, *ANR*, *FLS*, *ANS*, *CHS*, *DFR* and *CHI2* showed much higher transcription levels in AF than in WF, but *F3H* showed the opposite expression pattern. This indicated that the change in expression of these genes might affect the color change in sainfoin.

The “phenylpropanoid biosynthesis” pathway diverts carbon flow from primary metabolism to secondary phenolic metabolism through the sequential action of *PAL*, *C4H* and *4CL* [46]. *4CL* can transform 4-coumaric acid, erucic acid as well as ferulic acid into homologous coenzyme thiol esters respectively, which is an important step in the biosynthesis of flavonoids and heteroflavonoids [47, 48]. Ehltling [49] cloned the *4CL* gene family in *Arabidopsis thaliana* and demonstrated that *At4CL3* is involved in the biosynthetic pathways of flavonoids. In our study, the RNA-Seq data revealed that sainfoin *4CL* was differentially expressed between AF and WF. Our RT-qPCR results showed that transcription in the AF

was approximately 14 times higher than that in the WF. This is the main reason why flavonoid contents in AF are higher than those in WF (Table 2). A similar result was obtained in the study of Duan [14], who reported that the transcriptional expression of *4CL* in purple flowers of alfalfa was higher than that in cream flowers. This indicated that the reduction in *4CL* expression was one of the main reasons for the appearance of white petals in sainfoin.

The “flavanone biosynthesis” pathway is directly related to the biosynthesis and accumulation of flavonoids [50, 51]. The biosynthetic pathway of anthocyanins is a branch of the phenylpropanoid and flavonoid pathways, and anthocyanins are synthesized under the catalysis of a variety of enzymes [52]. Flavanone and anthocyanins play a vital role in flower color formation and diversity in many plants [15]. *CHS* is critical for the production of chalcone, which is the precursor for the synthesis of all anthocyanins and most other flavonoid metabolites [53]. A reduction in *CHS* transcript levels led to WF lines in *Muscari botryoides*, *Petunia hybrida* and *Parrya nudicaulis* [54–56]. Our results showed that *CHS* gene expression in AF was higher than that in WF, which was a good confirmation of previous study results. *CHI* is a key enzyme involved in flavonoid synthesis, and is also one of the enzymes required in the biosynthesis of flavonoid pigments [57]. Previous studies have shown that the decreased expression or insufficient activity of *CHI* will seriously impede the flavonoid biosynthetic pathway in many plants, resulting in significant decreases in the contents of anthocyanin and flavonoid [58, 59], while *CHI* overexpression can increase flavonoid content [60]. In our study, we found that the expression of *CHI2* in AF was higher than that in WF. This is the main reason why flavonoid and anthocyanin contents in AF are higher than those in WF. *F3H* catalyzes the hydroxylation of flavonoids which is necessary for anthocyanin biosynthesis [38]. In this study, the RT-qPCR detected expression of *F3H* in WF was higher than that in AF. Studies on alfalfa and carnation have shown that low expression of *F3H* causes a deeper flower color [14, 61]. This is because the cyanidin metabolism branch is effectively limited by *F3H* gene [18]. Interestingly, proanthocyanidin was found in the white petals, suggesting that the white petals were due to the lack of anthocyanin biosynthetic pathway downstream genes. *FLS* is the main enzyme responsible for the formation of quercetin and rutin [21]. As evidenced by the qRT-PCR results, the *FLS* gene expression in the AF was higher than that in the WF. This is the main reason why quercetin and rutin contents in AF is higher than that in WF (Table 2). The *DFR* gene can reduce dihydroflavonols to colorless leucoanthocyanidins; the *ANS* gene can convert the colorless leucoanthocyanidins into

colored malvidin, pelargonidin and delphinidin [15]. The expression of *ANS* in the WF was lower than that in the AF and anthocyanins were detected in the WF in our study (Fig. 4, Table 2); therefore, the *ANS* gene might be the key factor in the inability to accumulate anthocyanins in the WF, which was similar to that of Li [18]. Many studies have reported that low expression of *DFR* hampers pigmentation in *Arabidopsis*, *Dianthus caryophyllus* and *Dendranthema morifolium* [62–64]. Therefore, it can be inferred that the *CHS*, *CHI2*, *DFR*, *ANS*, *FLS* and *F3H* genes are the key factors that lead to failure to accumulate coloration compounds in WF. Our results are consistent with those of previous studies [31, 51, 52].

The specific production pathways of proanthocyanidins include *LAR* and *ANR*, both of which are key enzymes for their synthesis [65]. *LAR* can convert leucoanthocyanidin into catechin [66]. The *ANR* gene is an anthocyanin reductase that can transform anthocyanidins to flavan-3-ols needed for the proanthocyanidins produced in the flavonoid pathway [67]. In our study, catechin was detected in AF but not WF of sainfoin, and proanthocyanidins showed the opposite pattern. In addition, the RNA-Seq data and RT-qPCR revealed that the expression of the *ANR* and *LAR* genes in AF was higher than that in WF. Overall, the high expression of the *LAR* gene in AF compared with WF led to a higher catechin content in AF. This result was similar to that of Wang [68]. However, our result was opposite to that of Xie [69], who reported that the overexpression of *ANR* in *Arabidopsis thaliana* leaves resulted in anthocyanin loss and proanthocyanidin accumulation. Therefore, the role of the *ANR* gene in flower petals in sainfoin needs further study.

In summary, compared with that of AF, the flavonoid biosynthetic pathway of WF was blocked upstream by *4CL3*. Furthermore, the downregulated expression of *4CL3*, *FLS*, *ANS*, *CHS*, *DFR*, and *CHI2* resulted in a decrease in flavonoid and flavone compounds, such as rutin, kaempferol and its derivatives, reducing anthocyanin synthesis. At the same time, the high expression level of *F3H* might disrupt anthocyanins synthesis, leading to the formation of WF.

Conclusion

The contents and metabolic pathways of flavonoids and anthocyanins in amaranth and white petals in sainfoin were compared by UPLC, RNA-Seq and RT-qPCR. The main anthocyanins in AF in sainfoin were malvidin, petunidin and delphinidin derivatives, but these anthocyanins were not detected in WF. The main reason for the appearance of WF in sainfoin was the differential expression of multiple genes

related to flavonoid and anthocyanin biosynthesis, resulting in the differences in the types and contents of flavonoids and anthocyanins. Our RNA-Seq data greatly enrich sainfoin genomic research. Our results will provide valuable molecular information for genetic breeding and provide a reference for the future study of flower color polymorphisms in sainfoin.

Methods

Plant material

The sainfoin (*Onobrychis viciifolia* Scop ‘Mengnong’) material used in this experiment was provided by Inner Mongolia Agricultural University. This variety was approved by The Chinese Herbage Varietal Resources Registration Board in 1994 and registered as a new variety (Variety registration No.: 151). At the same time, it was also put on record in the Department of Animal Husbandry and Veterinary Medicine, Ministry of Agriculture. The EMS (Sigma Co.) concentration that yielded 50% sainfoin seed lethality (LD50) was 0.9% (v/v) after 18 h. Seeds treated with LD50 were transferred into the field, a WF mutant was found in 2013, and its seeds were collected individually. After mixing and planted for another three generations, WF from F4 generation were confirmed in 2017. In May 2018, AF and WF plants were planted at the experimental base of Inner Mongolia Agricultural University, located in Hohhot, Inner Mongolia, North China (latitude: 40°80'N, longitude: 111°69'E, elevation: 1058 m). The four developmental stages were defined according to the petal changes: S1, calyx higher than petal; S2, calyx and petal equal in height; S3, calyx lower than petal; S4, floret in full bloom (Fig. 5).

Petal color measurements

Petal color was measured at the S3 stage by a colorimeter (NH300, 3nh, China) with D65 illuminant. The colors are expressed as CIELAB [70] values (L^* , a^* , b^* , C^* and h°), and the average of six measurements per flower and ten flowers per treatment was used. L^* represents the lightness of the color with a range from black (0 = black) to white (100 = white). The color parameters a^* and b^* vary from -60 to 60; a^* describes redness and greenness, and b^* describes yellowness and blueness. C^* is used to denote the saturation of the color, and the higher the C^* value was, the more saturated the color was. h° expresses the hue of the color, where 0° = red and 270° = blue [71].

Extraction and qualitative and quantitative analyses of flavonoids

Freeze-dried petals at stage S3 were ground in liquid nitrogen and powder (20 mg). The 1 mL of 0.1% acetic

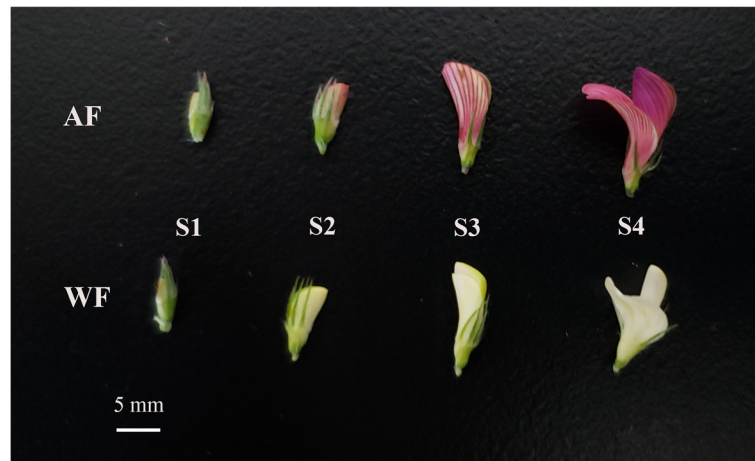


Fig. 5 Different flower colors of sainfoin (white flower 'WF' and amaranth flower 'AF') at four developmental stages

acid/methanol was used to extract sample at 4 °C overnight. Extracts were centrifuged at 10000 rpm with 10 min. The identification and quantification of flavonoids and anthocyanin compounds were conducted with an ultrahigh-performance liquid chromatograph-mass spectrometer coupled to a triple-quadrupole mass spectrometer (XEVO®-TQ, Waters, Milford, MA, USA) with ESI [34]. The relative anthocyanin and flavonoid contents were computed from the peak areas of the ion peaks of the characteristic Mass spectrometry daughter based on the strength of the corresponding standard compounds. MassLynx™ (V 4.1, SCN 846, Waters Corp., Manchester, UK) was used for Mass spectrometry data acquisition and data analysis. SDs were obtained from three biological replicates.

RNA extraction, cDNA library construction and sequencing

Petals from the AF and WF of sainfoin at the S1 and S3 stages were sampled, for a total of twelve samples, including three biological replicates. All samples were preserved at -80 °C for RNA extraction. The total RNA was extracted following the instruction manual of Qiagen RNeasy Plant Mini Kit. The RNA concentration and quality were determined by NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). Then the mRNA was enriched after removing rRNA by using Ribo-Zero™ Magnetic Kit (Epicentre). The cDNA library was built based on enriched mRNA with NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® followed the manufacturer's instructions. The cDNA was stored at -80 °C for sequencing and RT-qPCR experiment. The RNA-Seq library was sequenced on an Illumina HiSeq4000 instrument by Gene Denovo Biotechnology Co., Guangzhou, China.

De novo transcriptome assembly, unigene annotation, and DEG analysis

Our transcriptome datasets were deposited in the NCBI database under a BioProject ID: PRJNA643568. Transcriptome de novo assembly was performed with clean reads filtered from the raw reads by removing adapters, unknown nucleotides (> 10%), and low-quality reads (Q-values ≤ 10). Then, FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used to verify sequence quality, including the Q20, Q30 and GC content of clean reads. Since sainfoin (*O. viciifolia* Scop) genome information was not available, the clean reads of all twelve samples were combined for de novo assembly of the transcriptome using the reference genome independent Trinity method [72]. Trinity software package was used to combine the Inchworm, Chrysalis and Butterfly components [73]. First, short clean reads of a certain length were combined with overlap to form longer contigs by inchworm. Second, based on their paired-end information, clean reads were mapped back to the corresponding contigs by Chrysalis. At last, the path that were taken by reads and pairs of reads were analyzed by Butterfly. The finished sequences of the transcripts were defined as unigenes. BLASTx program were used to annotate all assembled unigenes (<http://www.ncbi.nlm.nih.gov/BLAST/>) with a threshold of E-value < 0.00001 to the Nr database (<http://www.ncbi.nlm.nih.gov>), Swiss-Prot protein database (<http://www.expasy.ch/sprot/>), KEGG database (<http://www.genome.jp/kegg>), KOG database and GO (<http://www.geneontology.org>) [74]. Unigene expression was normalized to RPKM values and differentially expressed genes were identified among samples or groups by edgeR software with a criteria of |fold change| ≥ 2, and FDR < 0.05 (<http://www.r-project.org/>). Next, GO and KEGG enrichment analyses were carried out for all DEGs, and hypergeometric tests with

$p \leq 0.05$ as a threshold were used to determine the significant enrichment of GO terms and KEGG pathways.

RT-qPCR analysis

Eleven selected DEGs involved in flavonoid synthesis were determined by one-step RT-qPCR. The experiment was performed on an ABI 7500 system (Applied Biosystems, USA) using SYBR (TaKaRa). The primers for the DEGs were designed by Primer Premier 5.0 (Premier, Canada) and the reference gene was JZ818469 (Additional file 8) [75]. The relative expression levels of DEGs were analyzed using the $2^{-\Delta\Delta CT}$ method [76]. Each sample (including three biological replicates) was quantified in triplicate.

All the data were subjected to statistical analysis using the t-test (SPSS ver.19.0), and they are presented as the mean \pm SD. The effect was considered significant when $P < 0.05$.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-02827-8>.

- Additional file 1.** Assembly of sainfoin transcriptome.
- Additional file 2.** The size distribution of sainfoin unigenes.
- Additional file 3.** GO classification of assembled unigenes.
- Additional file 4.** DEGs generated from WF and AF of sainfoin in S1 and S3 stages.
- Additional file 5.** GO functional annotations and the number of DEG statistics.
- Additional file 6.** KEGG pathway annotation of DEGs between WF and AF of sainfoin in S1 and S3 stages.
- Additional file 7.** KEGG pathway showing phenylpropanoid biosynthesis, flavonoid biosynthesis, flavone and flavonol biosynthesis in sainfoin.
- Additional file 8.** Primers used for RT-qPCR analysis.

Abbreviations

DEGs: Differentially expressed genes; nt: nucleotides; BLAST: Basic Local Alignment Search Tool; Nr: NCBI nonredundant protein; KEGG: Kyoto Encyclopedia of Genes and Genomes; KOG: Clusters of Orthologous Groups of proteins; Swiss-Prot: A manually annotated and reviewed protein sequence database; GO: Gene Ontology; RPKM: Reads per kb per million reads; RT-qPCR: Real-time quantitative PCR; ESI: Electrospray ionization; SDs: Mean values and standard deviations; 4CL: 4-coumarate: CoA ligase; ANS: Anthocyanidin synthase; CHI: Chalcone isomerase; CHS: Chalcone synthase; DFR: Dihydroflavonol 4-reductase; F3H: Flavanone-3-hydroxylase; FLS: Flavonol synthase; LAR: Anthocyanidin reductase; ANR: Leucoanthocyanidin reductase; FG3: Flavonol-3-O-glucoside; PMAT1: Isoflavone 7-O-glucoside-6"-O-malonyltransferase; UPLC: Ultra-high performance liquid chromatography–mass; WF: White flower; AF: Amaranth flowers

Acknowledgements

We thank the Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences and Gene Denovo Biotechnology Co. (Guangzhou, China) for providing technical support.

Authors' contributions

FLS conceived and designed the experiments. YQ and YTZ performed the experiments. QY and QMC wrote the manuscript. YQ, WY and FYY carried

out the data analysis. All authors reviewed and considered the final manuscript. The author(s) read and approved the final manuscript.

Funding

This study was supported by the Inner Mongolia Autonomous Region Applied Technology Research and Development Fund Project (2019GG244). The funding body was not involved in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

All raw sequence data are available at NCBI project PRJNA643568 and Sequence Read Archive (SRA) with accession number SRR12130587, SRR12130586, SRR12130585, SRR12130584, SRR12130583, SRR12130582, SRR12130581, SRR12130580, SRR12130579, SRR12130578, SRR12130577 and SRR12130576. The addresses are as follows: <https://submit.ncbi.nlm.nih.gov/subs/sra>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Grassland Resources and Environment, Key Laboratory of Forage Cultivation, Processing and High Efficient Utilization of the Ministry of Agriculture and Key Laboratory of Grassland Resources of the Ministry of Education, Inner Mongolia Agricultural University, Hohhot 010011, China. ²Inner Mongolia Academy of Agricultural and Animal Husbandry Sciences, Hohhot, China.

Received: 2 July 2020 Accepted: 5 January 2021

Published online: 22 January 2021

References

- Kliker R, Kempf K, Malisch CS, Lüscher, Andreas. Promising options for improving performance and proanthocyanidins of the forage legume sainfoin (*Onobrychis viciifolia* Scop.). *Euphytica*. 2017;213(8):179.
- Bhattarai S, Coulman B, Biligetü B. Sainfoin (*Onobrychis viciifolia* Scop.): renewed interest as a forage legume for western Canada. *Can J Plant Sci*. 2016;96(5):748–56.
- Wilman D, Asiedu FHK. Growth, nutritive value and selection by sheep of sainfoin, red clover, lucerne and hybrid ryegrass. *J Agr Sci*. 1983; 100(1):115–26.
- Brinkhaus AG, Wyss U, Arrigo Y, Girard M, Bee G, Zeitz JO, Kreuzer M, Dohmemeier F. In vitro ruminal fermentation characteristics and utilisable CP supply of sainfoin and birdsfoot trefoil silages and their mixtures with other legumes. *Animal*. 2017;11(4):580–90.
- Hatev B, Hayot Carbonero C, Stringano E, Sales LF, Smith LM, Muellerharvey I, Hendriks WH, Pellikaan WF. Diversity of condensed tannin structures affects rumen in vitro methane production in sainfoin (*Onobrychis viciifolia*) accessions. *Grass Forage Sci*. 2015;70(3):474–90.
- Girard M, Dohmemeier F, Wechsler D, Goy D, Kreuzer M, Bee G. Ability of 3 tanniferous forage legumes to modify quality of milk and Gruyère-type cheese. *J Dairy Sci*. 2016;99(1):205–20.
- Richards KW, Edwards PD. Density, diversity, and efficiency of pollinators of sainfoin, *Onobrychis viciaefolia* Scop. *Can Entomol*. 1988;120(12):1085–100.
- Kells AR. Sainfoin: an alternative forage crop for bees. *Bee World*. 2001;82(4): 192–4.
- Zhang YY, Zhou TH, Dai ZW, Dai XY, Li W, Cao MX, Li CR, Tsai WC, Wu XQ, Zhai JW, Liu ZJ, Wu SS. Comparative Transcriptomics provides insight into floral color polymorphism in a *Pleione limprichtii* orchid population. *Int J Mol Sci*. 2020;21(1):247.
- Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L. Carotenoid metabolism in plants: the role of plastids. *Mol Plant*. 2018;11(1):58–74.
- Davies KM, Albert NW, Schwinn KE. From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Funct Plant Biol*. 2012;39(8):619.

12. Schiest FP, Johnson SD. Pollinator-mediated evolution of floral signals. *Trends Ecol Evol.* 2013;28(5):307–15.
13. Hanumappa M, Choi G, Ryu S, Choi G. Modulation of flower colour by rationally designed dominant-negative chalcone synthase. *J Exp Bot.* 2007;58(10):2471–8.
14. Duan HR, Wang LR, Cui GX, Zhou XH, Duan XR, Yang HS. Identification of the regulatory networks and hub genes controlling alfalfa floral pigmentation variation using RNA-sequencing analysis. *BMC Plant Biol.* 2020;20(1):1–17.
15. Tanaka Y, Sasaki N, Ohmiya A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 2008;54(4):733–49.
16. Tripathi AM, Niranjan A, Roya S. Global gene expression and pigment analysis of two contrasting flower color cultivars of canna. *Plant Physiol Bioch.* 2018;27:1–10.
17. Grotewold E. The genetics and biochemistry of floral pigments. *Annu Rev Plant Biol.* 2006;57(1):761–80.
18. Li X, Wang J, Zhao J, Zheng Y, Wang HF, Wu X, Xian C, Lei JJ, Zhong CF, Zhang YT. Study on cyanidin metabolism in petals of pink-flowered strawberry based on transcriptome sequencing and metabolite analysis. *BMC Plant Biol.* 2019;19(1):1–16.
19. Li Y, Ma R M YJ, Zhang F, Duan HR, Yang HS, Tian FP, Zhou XH, Wang CM. Transcriptomic analysis of *Lycium ruthenicum* Murr. During fruit ripening provides insight into structural and regulatory genes in the anthocyanin biosynthetic pathway. *PLoS One.* 2018;13(12):1–12.
20. Holton Timothy A, Cornish EC. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell.* 1995;7(7):1071–83.
21. Li L, Zhai YB, Luo XB, Zhang YC, Shi QB. Comparative transcriptome analyses reveal genes related to pigmentation in the petals of red and white *Primula vulgaris* cultivars. *Physiol Mol Biol Pla.* 2019;25(4):1029–41.
22. Zhang H, Tian H, Chen M, Xiong J, Cai H, Liu Y. Transcriptome analysis reveals potential genes involved in flower pigmentation in a red-flowered mutant of white clover (*Trifolium repens* L.). *Genomics.* 2018;110(3):191–200.
23. Li X, Liu S, Yuan G, Zhao P, Yang W, Jia J, Cheng L, Qi D, Chen S, Liu G. Comparative transcriptome analysis provides insights into the distinct germination in sheepgrass (*Leymus chinensis*) during seed development. *Plant Physiol Bioch.* 2019;139:446–58.
24. Xie WG, Zhang JC, Zhao XH, Zhang ZY, Wang YR. Transcriptome profiling of *Elymus sibiricus*, an important forage grass in Qinghai-Tibet plateau, reveals novel insights into candidate genes that potentially connected to seed shattering. *BMC Plant Biol.* 2017;17(1):78.
25. Huang B, Rong H, Ye YJ, Ni ZX, Xu M, Zhang WX, Xu LA. Transcriptomic analysis of flower color variation in the ornamental crabapple (*Malus* spp.) half-sib family through Illumina and PacBio sequel sequencing. *Plant Physiol Bioch.* 2020;149:27–35.
26. Lu C, Pu Y, Liu Y, Li Y, Qu J, Huang H, Dai S. Comparative transcriptomics and weighted gene co-expression correlation network analysis (WGCNA) reveal potential regulation mechanism of carotenoid accumulation in *Chrysanthemum x morifolium*. *Plant Physiol Bioch.* 2019;142:415–28.
27. Hiroyasu Y. Mutation breeding of ornamental plants using ion beams. *Breeding Sci.* 2018;68(1):71–8.
28. Akhtar S, Sikder S, Biswas P, Hazra P, D'Souza SF. Induction of mutation in tomato (*Solanum Lycopersicum* L.) by gamma irradiation and EMS. *Indian J Genet Plant Breed.* 2013;73(4):392–9.
29. Shah SNM, Gong ZH, Arisha MH, Khan A, Tian SL. Effect of ethyl methyl sulfonate concentration and different treatment conditions on germination and seedling growth of the cucumber cultivar chinese long (9930). *Genet Mol Res.* 2015;14(1):2440–9.
30. Awais A, Nualsri C, Soonswun W. Induced mutagenesis for creating variability in Thailand's upland Rice (cv. Dawk pa-yawm and Dawk Kha 50) using ethyl methane Sulphonate (EMS). *Sarhad J Agr.* 2019;35(1):293–301.
31. Amin R, Wani MR, Raina A, Khurshheed S, Khan S. Induced morphological and chromosomal diversity in the mutagenized population of black cumin (*Nigella sativa* L.) using single and combination treatments of gamma rays and ethyl methane Sulfonate. *Jordan J Bio Sci.* 2019;12(1):23–30.
32. Malisch C, Luscher A, Baert N, Engstrom MT, Studer B, Frygasas C, Suter D, Mueller-Harvey I, Salminen J. Large variability of proanthocyanidin content and composition in sainfoin (*Onobrychis viciifolia*). *J Agr Food Chem.* 2015;63(47):10234–42.
33. Nigel CV, Ionela R, Geoffrey CK, Dieter T. Acylated flavonol glycosides from the forage legume, *Onobrychis viciifolia* (sainfoin). *Phytochemistry.* 2011;72(4–5):423–9.
34. Regos I, Urbanella A, Treutter D. Identification and quantification of phenolic compounds from the forage legume Sainfoin (*Onobrychis viciifolia*). *J Agr Food Chem.* 2009;57(13):5843–52.
35. Li LA, Zhai YB, Luo XB, Zhang YC, Shi QB. Comparative transcriptome analyses reveal genes related to pigmentation in the petals of red and white *Primula vulgaris* cultivars. *Physiol Mol Biol Pla.* 2019;25(4):1029–41.
36. Li X, Wang J, Zhao J, Zheng Y, Zhang YT. Study on cyanidin metabolism in petals of pink-flowered strawberry based on transcriptome sequencing and metabolite analysis. *BMC Plant Biol.* 2019;19(1).
37. Guo L, Wang Y, Silva JA, Fan Y, Yu X. Transcriptome and chemical analysis reveal putative genes involved in flower color change in *Paeonia 'coral sunset'*. *PPB.* 2019;138:130–9.
38. 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):3157–64.
39. Yang Q, Yuan HH, Sun XB. Preliminary studies on the changes of flower color during the flowering period in two tree peony cultivars. *Acta Horti Sin.* 2015;42 (5).
40. Zhong PX, Wang LS, Li SS, Xu YJ, Zhu ML. The changes of floral color and pigments composition during the flowering period in *Paeonia officinalis*. *Acta Horticulturae Sinica.* 2012;11.
41. Han FL, Zhang WN, Pan QH, Zheng CR, Chen HY, Duan CQ. Principal component regression analysis of the relation between CIELAB color and monomeric anthocyanins in young cabernet sauvignon wines. *Molecules.* 2008;13(11):2859–70.
42. Cheyrier V, Comte G, Davies KM, Lattanzio V, Martens S. Plant phenolics: recent advances on their biosynthesis, genetics, and ecophysiology. *Plant Physiol Bioch.* 2013;72(1):1–20.
43. Hao ZD, Liu SQ, Hu LF, S JS, Chen JH. Transcriptome analysis and metabolic profiling reveal the key role of carotenoids in the petal coloration of *Liriodendron tulipifera*. *Horticulture Research.* 2020;70(1).
44. Zhou Y, Wu XX, Zhang Z. Comparative proteomic analysis of floral color variegation in peach. *Biochem Bioph Res Co.* 2015;464(4):1101–6.
45. Wang YL, Wang YQ, Song ZQ. Repression of MYB2 by both microRNA585a and HY5 leads to the activation of anthocyanin biosynthetic pathway in *Arabidopsis*. *Mol Plant.* 2016;9(10):1395–405.
46. Douglas CJ. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends Plant Sci.* 1996;1(6):171–8.
47. Gang DR, Lavid N, Zubieta C, Chen F, Beuerle T, Lewinsohn E, Noel JP, Pichersky E. Characterization of Phenylpropene O-Methyltransferases from sweet basil: facile change of substrate specificity and convergent evolution within a plant O-methyltransferase family. *Plant Cell.* 2002;14(2):505–19.
48. Cukovic D, Ehlting J, VanZiffle JA, Douglas CJ. Structure and evolution of 4-Coumarate: coenzyme a ligase (4CL) gene families. *Biol Chem.* 2005;382(4):645–54.
49. Ehlting J, Büttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E. Three 4-coumarate: coenzyme a ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* 1999;19(1):9–20.
50. Zhao DQ, Tao J. Recent advances on the development and regulation of flower color in ornamental plants. *Front Plant Sci.* 2015;6:261.
51. Nakatsuka T, Nishihara MM, Mishiba K, Yamamura S. Temporal expression of flavonoid biosynthesis-related genes regulates flower pigmentation in gentian plants. *Plant Sci.* 2005;168(5):1309–18.
52. Fang ZW, Hou ZH, Wang SP, Liu ZX, Wei SD, Zhang YX, Song JH, Yin JL. Transcriptome Analysis Reveals the Accumulation Mechanism of Anthocyanins in Buckwheat (*Fagopyrum esculentum* Moench.) Cotyledons and Flowers. *Int J Mol Sci.* 2019;20(6):1493.
53. Achilonu CC, Maleka FM. Characterization and Expression Analyses of Chalcone Synthase (CHS) and Anthocyanidin Synthase (ANS) Genes in *Clivia miniata*. OMICS Publishing Group. 2016;4(2).
54. Katsumoto Y, Fukuchimizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekurasakakibara K, Togami J, Pigeaire A, Tao GQ, Nehra NS, Lu CY, Dyson BK, Tsuda S, Ashikari T, Kusumi T, Mason JG, Tanaka Y. Engineering of the rose flavonoid biosynthetic pathway successfully generated blue-hued flowers accumulating Delphinidin. *Plant & Cell Physiology.* 2007;48(11):1589–600.
55. Spitzer B, Zvi MM, Ovadis M, Marheva E, Barkai O, Edelman O, Marton I, Masci T, Alon M, Morin S, Rogachev I, Aharoni A, Vainstein A. Reverse genetics of floral scent: application of tobacco rattle virus-based gene silencing in *Petunia*. *Plant Physiol.* 2007;145(4):1241–50.

56. Cynthia AD, Jason B, Timothy B, Matthew LC, Daniel JK, Justen BW. Arctic mustard flower color polymorphism controlled by petal-specific downregulation at the threshold of the anthocyanin biosynthetic pathway. *PLoS One*. 2011;16(4):1–10.
57. Wu YQ, Zhu MY, Jiang Y. Molecular characterization of chalcone isomerase (*CHI*) regulating flower color in herbaceous peony (*Paeonia lactiflora* pall.). *J Integr Agr*. 2018;17(1):122–9.
58. Van Tunen AJ, Koes RE, Spelt CE, Van DKAR, Stuitje AR, Mol JN. Cloning of the two chalcone flavanone isomerase genes from *Petunia hybrida*: coordinate, light-regulated and differential expression of flavonoid genes. *EMBO J*. 1988;7(5):1257–63.
59. Kim S, Jones R, Yoo KS, Pike LM. Gold color in onions (*Allium cepa*): a natural mutation of the chalcone isomerase gene resulting in a premature stop codon. *Mol Gen Genomics*. 2004;272(4):411–9.
60. Muir SR, Collins GJ, Robinson S, Hughes S, Bovy A, Ric DVCH, Van Tunen AJ, Verhoeven ME. Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nat Biotechnol*. 2001; 19(5):470–4.
61. Masami M, Takashi O, Yoshihiro O, Daisuke H, Yoshio I, Tamotsu H, Hiroyuki Y, Michio S. Flavonoid Biosynthesis in Pink-flowered Cultivars Derived from 'William Sim' Carnation (*Dianthus caryophyllus*). *J Jpn Soc Horticult Sci*. 2001;70 (3).
62. Chen SM, Li CH, Zhu XR, Deng YM, Sun W, Wang LS, Chen FD, Zhang Z. The identification of flavonoids and the expression of genes of anthocyanin biosynthesis in the chrysanthemum flowers. *Biol Plantarum*. 2012;56(3):458–64.
63. Stich K, Eidenberger T, Wurst F, Forkmann G. Enzymatic conversion of dihydroflavonols to flavan-3,4-diols using flower extracts of *Dianthus caryophyllus* L. (carnation). *Planta*. 1992;187(1):103–8.
64. Feyissa DN, Lvdal T, Olsen KM, Slimestad R, Lillo C. The endogenous *GL3*, but not *EGL3*, gene is necessary for anthocyanin accumulation as induced by nitrogen depletion in *Arabidopsis* rosette stage leaves. *Planta*. 2009;230(4):747–54.
65. Zhu Q, Sui S, Lei X, Yang Z, Lu K, Liu G, Liu YG, Li M, Han Y. Ectopic expression of the *Coleus R2R3* MYB-type Proanthocyanidin regulator gene *SsMYB3* alters the flower color in transgenic tobacco. *PLoS One*. 2015;10(10):e139392.
66. Han YP, Vimolmangkang S, Soria-Guerra R E, Korban S S. Introduction of apple *ANR* genes into tobacco inhibits expression of both *CHI* and *DFR* genes in flowers, leading to loss of anthocyanin. *J Exp Bot*. 2012;(7):2437–2447.
67. Sverine G, Soizic L, Olivier C, Laurence G. Leucoanthocyanidin reductase and anthocyanidin reductase gene expression and activity in flowers, young berries and skins of *Vitis vinifera* L. cv. Cabernet-sauvignon during development. *Plant Physiol Bioch*. 2009;47(4):282–90.
68. Wang PQ, Liu YJ, Zhang LJ, Wang WZ, Hou H, Zhao Y, Jiang XL, Yu J, Tan HR, Wang YS, Xie DY, Gao LP, Xia T. Functional demonstration of plant flavonoid carbocations proposed to be involved in the biosynthesis of proanthocyanidins. *Plant J*. 2020;101(1):18–36.
69. Xie DY, Sharma B, Paiva NL, Ferreira D, Dixon RA. Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science (New York, N.Y.)*. 2003;299(5605):396–399.
70. Zhang YY, Zhou TH, Dai ZW, Dai XY, Wei L, Cao MX, Li CR, Tsai WC, Wu XQ, Zhai JW, Liu ZJ, Wu SS. Comparative Transcriptomics provides insight into floral color polymorphism in a *Pleione limprichtii* orchid population. *Int J Mol Sci*. 2020;21(N1):247.
71. Cui HL, Zhang YN, Shi XL, Gong FF, Xiong X, Kang XP, Xing GM, Li S. The numerical classification and grading standards of daylily (*Hemerocallis*) flower color. *PLoS One*. 2019;14(6):1–16.
72. Grabherr M, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Xian A, Lin F, Raktima R, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52.
73. Cheng QM, Bai SQ, Ge GT, Li P, Liu LY, Zhang CD, Jia YS. Study on differentially expressed genes related to defoliation traits in two alfalfa varieties based on RNA-Seq. *BMC Genomics*. 2018;19(1):1–8.
74. Chen JJ, Duan YJ, Hu YL, Li WM, Sun DQ, Hu HG, Xie JH. Transcriptome analysis of *atemoya* pericarp elucidates the role of polysaccharide metabolism in fruit ripening and cracking after harvest. *BMC Plant Biology*. 2019;19(1).
75. Wang JJ, Zhao Y, Ray I, Song MZ. Transcriptome responses in alfalfa associated with tolerance to intensive animal grazing. *Scientific reports*. 2016;6(1):19438.
76. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods*. 2001;25(4):402.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

