BIOLOGIA PLANTARUM 64: 504-511, 2020 DOI: 10.32615/bp.2020.068

Comprehensive transcriptome analyses of different *Crocus* flower tissues uncover genes involved in crocin biosynthesis

H.M. SHU*, S.Q. GUO, and W.C. NI*

Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, P.R. China

Abstract

The stigma of *Crocus sativus* is used in traditional Chinese medicine and has drawn attention as a rich source of crocin, a compound with a reported activity that counters various cancers, depression, and cardiovascular diseases. However, our knowledge of crocin biosynthesis in *Crocus* is still limited. To identify the genes that encode key enzymes responsible for crocin production, transcriptome analyses of *Crocus* stigma, petal, and stamen were performed. There were 109 136 unigenes in the three *Crocus* flower tissues: 10 862 unigenes were expressed explicitly in stigmas. A total of 469 and 335 down-regulated differentially expressed genes in comparison to stigmas were detected in stamens and petals, respectively. There were 290 down-regulated genes in both tissues. The gene ontology enrichment and *Kyoto Encyclopedia of Genes and Genomes* analysis revealed that most genes in the two tissues served a similar function. The results of the differences of gene expression among the three tissues showed that most of the carotenoid pathway genes in stigmas had a high expression which was a benefit for producing zeaxanthin, a substrate of crocin. Fourteen candidate genes were identified in *Crocus* from the co-down-regulated genes, and the genes were expressed explicitly in stigmas and were predicted to be involved in crocin biosynthesis. Overall, the expressions of genes in the crocin biosynthesis pathway in the three *Crocus* flower tissues had a positive relationship with their crocin content, and some new candidate genes related to crocin biosynthesis were identified. The results can help elucidate the crocin biosynthesis pathway in *Crocus*.

Additional key words: carotenoids, differentially expressed genes, petal, stamen, stigma, unigenes, zeaxanthin.

Introduction

The stigma of *Crocus sativus* has traditionally been used as one of most expensive spices and in traditional Chinese medicine for thousands of years (Fernandez 2004). Currently, *Crocus* is cultivated throughout Central Asia, Iran, Spain, India, and China. It belongs to *Iridaceae* family and it can produce apocarotenoids, such as crocin, picrocrocin, and safranal (Bouvier *et al*. 2003, D'Agostino *et al*. 2007, Baba *et al*. 2015a). These compounds provide colour, flavour, and aroma to *Crocus* plant, and have been demonstrated to improve memory significantly, possess an antidepressant effect and elicit protective effects against cardiovascular diseases (Wang *et al*. 2010, Hosseinzadeh *et al*. 2012, Rezaee and Hosseinzadeh 2013, Razavi *et al*. 2013). Additionally, the yield of stigmas is very low and stigmas manually harvested, which make it the world's most expensive plant (Gainer and Brumgard 1982).

In *Crocus*, the aforementioned compounds mainly exist in the stigma of flowers. Thus, the transcriptome of different organs has been sequenced to find some genes in crocin synthesis pathway (Baba *et al*. 2015b, Jain *et al*. 2016). The crocin biosynthesis pathway has also been identified and includes the upstream methylerythritol phosphate (MEP) pathway from pyruvate/glyceraldehyde 3-phosphate to geranylgeranyl pyrophosphate (GGPP), the midstream carotenoid pathway from GGPP to zeaxanthin, and the downstream crocin pathway from zeaxanthin to crocin (Ji *et al*. 2017). In *Crocus*, enzymes involved in crocin pathway are carotenoid cleavage dioxygenase (CCD), aldehyde dehydrogenases (ALDHs), UDP-glucosyltransferases (UGTs), and so on (Baba *et al*. 2015b,

Submitted 3 March 2020, *last revision* 4 May 2020, *accepted* 7 May 2020.

Abbreviations: ALDH - aldehyde dehydrogenase; CAP - carotenoid-associated protein; CCD - carotenoid cleavage dioxygenase; β-CH - beta-carotene hydroxylase; CRTISO - carotene *cis/trans* isomerase; DEGs - differentially expressed genes; GGPP - geranylgeranyl pyrophosphate; GO - Gene Ontology; GS - glucosidase; KEGG - Kyoto Encyclopedia of Genes and Genomes; KO - KEGG for ortholog; KOG - Eukaryotic for ortholog groups; β-LYC - lycopene beta-cyclase; MEP - methylerythritol phosphate; NCBI - National Center for Biotechnology Information; NR-NCBI - non-redundant protein sequences in NCBI; NT-NCBI - nucleotide sequences in NCBI; PDS - phytoene desaturase; PFAM - Protein family; PSY - phytoene synthase; qPCR - quantitative PCR; UGT - UDP-glucosyltransferase; ZDS - zeta-carotene desaturase; ZEP - zeaxanthin epoxidase; ZISO - *cis*-zeta-carotene isomerase.

Acknowledgements: This work was funded by the National R&D Project of Transgenic Crops of the Ministry of Science and Technology of China (2016ZX08005001).

* Corresponding authors; e-mails: shuhm1982@163.com, nwchao2002@aliyun.com

Frusciante *et al*. 2014, Jain *et al*. 2016, Ji *et al*. 2017). Zeaxanthin is catalyzed by CCD to produce 3-hydroxy-cyclocitral and crocetin dialdehyde. Then the two products are catalyzed by ALDHs and UGTs to form crocin and picrocrocin (Jain *et al*. 2016). After harvesting, when the stigma is dry, picrocrocin is split to yield safranal (Jain *et al*. 2016).

Crocus is a sterile triploid that produces annual replacement corms and is propagated solely from these corms (Juan *et al*. 2009). In the autumn, saffron corms begin to blossom and produce their leaves. In the spring, the leaves of saffron turn yellow and dry up. In most planting areas, new daughter corms are under the ground and remain dormant until the growth cycle begins again (Agayev *et al*. 2007, 2009). However, in China, corms are dug up when the leaves turn yellow, and after blossoming corms are planted in the field again. Intensive differentiation of flower organs occurs when daughter corms become dormant from June to August, but this stage in China occurs indoors. This abnormal environment of flower differentiation may affect overall yield and result in different phytochemical traits of saffron (Gresta *et al*. 2009, 2010). In this study, transcriptome analyses of the Chinese *Crocus* were conducted in three different flower tissues to uncover the molecular basis of apocarotenoid biosynthesis. Moreover, the differences among various planting methods during the flower differentiation period were also compared.

Materials and methods

Tissue sampling, cDNA library construction, and sequencing: Stigmas, petals, and stamens were hand-picked from *Crocus sativus* L. flowers, which grow in a growth chamber at Jiangsu Academy of Agricultural Sciences. The growth conditions were: from June to August day/night temperatures of 25/22 °C, dark, and a relative humidity of about 60 %; from September to November a 16-h photoperiod, day/night temperatures of 18/16 °C, an irradiance of 36 μmol m^{-2} s^{-1}, and a relative humidity of about 70 %. The fresh mass of corms was greater than 25 g. The samples were frozen in liquid nitrogen and stored at -80°C for further use.

Total RNA was extracted from the stigmas, petals, and stamens using *Trizol* reagent (*Invitrogen*, Shanghai, China). RNA degradation and contamination were monitored on 1 % (m/v) agarose gels. RNA purity was checked using the *NanoPhotometer* (*Implen*, Munich, Germany). Concentration of RNA and RNA integrity were examined by *Novagene* (Beijing, China). Total RNA (10 mg) of each sample was collected for the isolation of poly (A) mRNA using beads with Oligo (dT). Then, mRNA was separated into short fragments using a fragmentation buffer. Suitable fragments were selected for the PCR amplification as templates in preparation *Illumina RNA-Seq* library. The library preparations were sequenced on an *Illumina Hiseq* platform (*Novagene*) and paired-end reads were generated. Quality control raw data (raw reads) of *fastq* format were firstly processed through *in-house perl scripts*. Clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low-quality reads from raw data. High-quality reads were used for *de novo* assembly using *Trinity* (Grabherr *et al.* 2011) software with *min_kmer_cov* set to 2.

RNA-seq by expectation maximization (*RSEM*, Li and Dewey 2011) was used to count the read numbers mapped to the transcripts. For all comparisons, the read counts were normalized to the aligned reads per kilobase of transcript, per million mapped reads (RPKM) in order to obtain the relative expression (Mortazavi *et al*. 2008). Cluster analysis of all treatments was based on the fragments per kilobase of exon per million fragments mapped (FPKM) values of the differently expressed genes (DEGs). The Venn diagrams were made using the function "Venn Diagram" in R based on the gene list for each sample.

***Gene Ontology* and *Kyoto Encyclopedia of Genes and Genomes* enrichment analysis of DEGs:** *Gene Ontology* (*GO*) is an international standardized gene functional classification system. *GO* enrichment analysis of the DEGs was implemented using the *GOseq R* packages based on Wallenius noncentral hypergeometric distribution (Young *et al*. 2010). *GO* terms with $P < 0.05$ were considered significantly enriched. The *Kyoto encyclopedia of genes and genomes* (*KEGG*) is an online database used for understanding high-level functions and utilities of the biological system (Kanehisa *et al*. 2008). In this study, *KEGG* was used to analyze the potential involvement of sequences in cellular metabolic pathways. The *KOBAS* software was used to test the statistical enrichment of DEGs in *KEGG* pathways (Mao *et al*. 2005).

Real time quantitative PCR was conducted in order to confirm the reliability of RNA-Seq data with the same RNA pools used for next-generation sequencing. First-strand cDNA was synthesized using a *PrimeScript RT* reagent kit with a gDNA eraser (*Takara*, Dalian, China) following the manufacturer's instructions. Real time quantitative PCR analyses were performed on an *ABI PRISM 7500* system (*ABI*, Foster City, CA, USA). The reaction mixture (20 mm^3) contained 10 mm^3 of *SYBR Green Premix ExTaq* (*Takara*), 1 ng of cDNA sample, 0.2 μM of each gene-specific primer (Table 1 Suppl.), and 0.4 mm^3 of *ROX Reference Dye II* (50×). Three biological replicates were used for each analysis. Relative expressions were calculated following the $2^{-\Delta\Delta T}$ method using with the *18S rRNA* gene as an internal reference gene (Livak and Schmittgen 2001, Baba *et al*. 2015b).

Results

The sequencing of *Crocus* cDNA libraries generated 60 001 800 raw reads from stigma, 61 075 690 from petal, and 42 369 356 from stamen. Approximately 5 % of the raw reads were removed after filtering of adapter sequences and low quality or short reads. Then, 57 692 616 stigma, 57 603 198 petal, and 39 930 638 stamen reads were assembled *de novo* using *Trinity* (Grabherr *et al*. 2011),

Table 1. The summary of the *Crocus* transcriptome.

Tissue	Stigma	Petal	Stamen	Total
Number of raw reads	60001800	61075690	42369356	
Number of filtered reads	57692616	57603198	39930638	
Number of transcripts	80111	109596	71105	153330
Number of unigenes	57436	79093	54603	109136
Mean length of unigenes	538	540	519	532
N50 of unigenes	748	733	665	709
Percentage of GC	43.37	44.11	49.03	

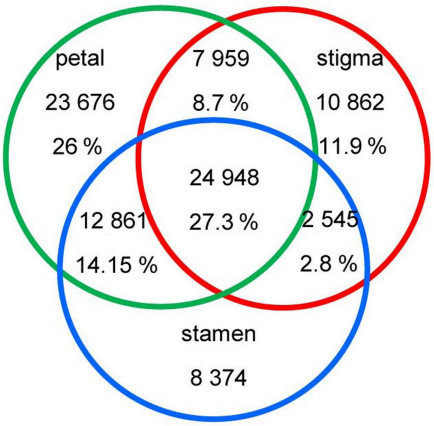

Fig. 1. The Venn diagram of the number and percentage of unigenes in three different *Crocus* tissues.

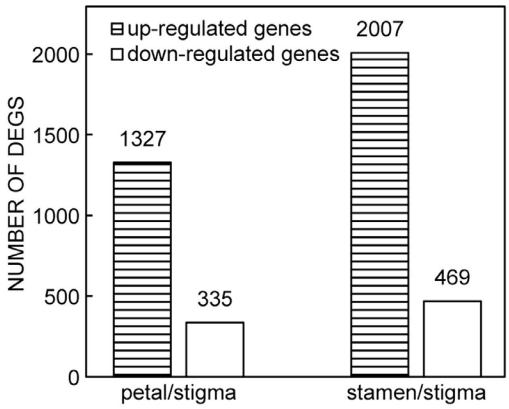

Fig. 2. Up- and down-regulated differentially expressed genes (DEGs) in petals compared to stigmas (petal/stigma) and stamens compared to stigmas (stamen/stigma).

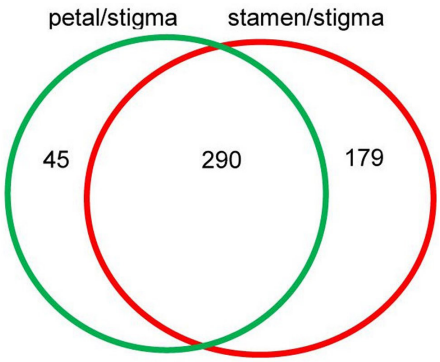

Fig. 3. The Venn diagram of down-regulated genes between petal/stigma and stamen/stigma.

which resulted in 153 330 transcripts. These transcripts were classified into 109 136 unigenes (Table 1). The average contig length was 532 bp, and N50 was 709 bp.

The functional annotation of transcripts was performed using *BLAST* (*BLASTX* tools) homology searches against various public protein databases including non-redundant protein sequences in *National Center for Biotechnology Information* (NR-*NCBI*), nucleotide (NT-*NCBI*), *KEGG for ortholog* (*KO*), *SwissProt*, *Protein family* (*PFAM*), *GO*, and *Eukaryotic for ortholog groups* (*KOG*) (Table 2 Suppl.). Of the 109 136 non-redundant unigenes, 36 805 (33.72 %) had significant similarity with known proteins in NR-NCBI database, and 28 616 (26.22 %) had hits in the *SwissProt* database.

Based on *GO* analyses, 25 606 unigenes (23.46 %) were categorized into many functional groups that belonged to biological process, cellular component, and molecular function clusters. The top 10 classes in each category were shown in Fig. 1 Suppl. In the biological process category, genes belonged to "macromolecule metabolic process" (32.9 %), "cellular macromolecule metabolic process" (30.3 %), "cellular nitrogen compound metabolic process" (24.7 %), and "organic cyclic compound metabolic process" (22.3 %). In the cellular component category, genes predominantly belonged to "intracellular" (25.6 %) and "intracellular part" (25.5 %). As for the molecular function category, genes primarily belonged to "nucleoside phosphate-binding" (13.6 %) and "anion binding" (13.6 %). *GO* term abundance results revealed high similarity with previous *Crocus* transcriptome studies (Baba *et al*. 2015b, Jain *et al*. 2016).

Based on the *KEGG* analyses, the most representative pathways were "ribosome" (474, 3.52 %), "carbon metabolism" (473, 3.51 %), and "plant-pathogen interaction" (421, 3.12 %) (Fig. 2 Suppl.). These results suggested that there was an abundance of the discovered unigenes. Additionally, the *GO* and pathway-based analyses helped elucidate the biological functions of these genes and how these genes interact.

The three different flower tissues had low similarity

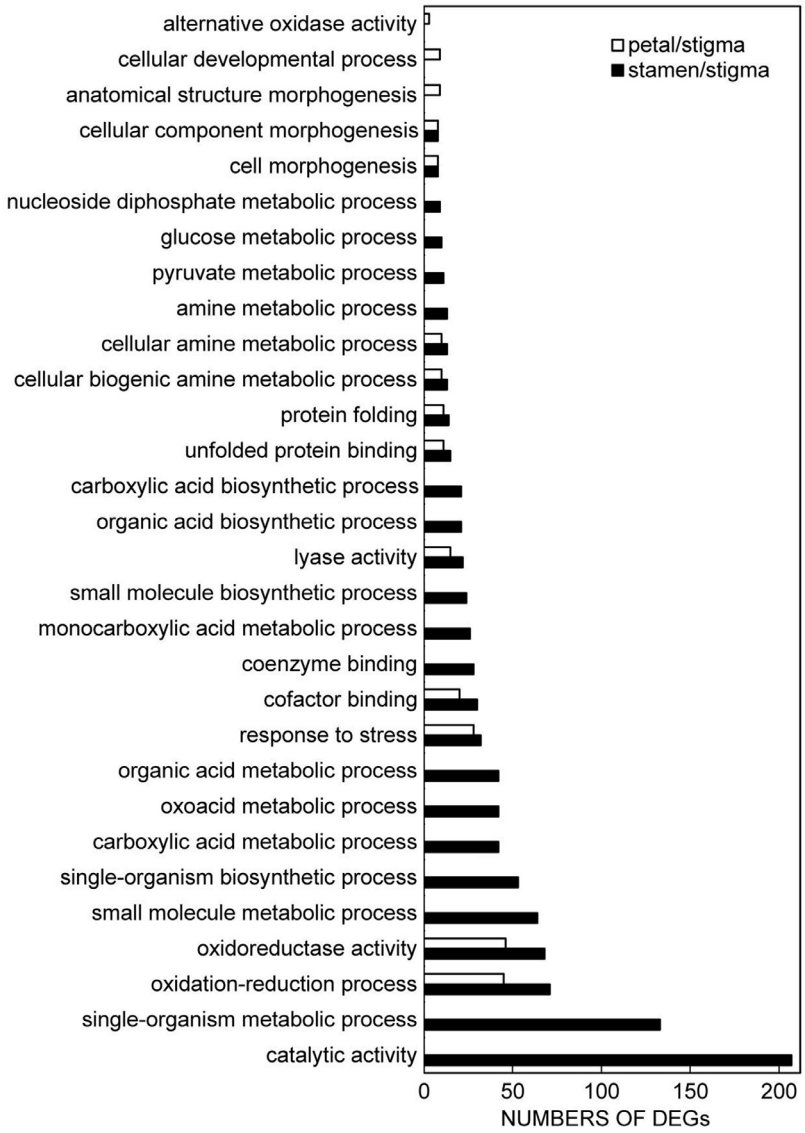

Fig. 4. Enriched *Gene ontology* (*GO*) terms of down-regulated genes between petal/stigma (*white*) and stamen/stigma (*black*). DEGs - differentially expressed genes.

in terms of expressions of unigenes (Table 3 Suppl.). The correlation coefficient of the expressions of unigenes between stigma and petal was 0.157, and that between stigma and stamen was 0.178. Among the 109 136 unigenes, there were 24 948 unigenes expressed in the three tissues, and 10 862 unigenes specifically expressed in the stigma (Fig. 1). Based on the fact that crocin mainly exist in the stigma (Jain *et al*. 2016), it is likely that some of these 10 862 unigenes may be related to crocin synthesis.

Based on q.value ≤ 0.005 (Storey *et al*. 2003) and |log2.fold change| > 1, 2 476 genes and 1 662 genes were differentially expressed in stamen and petal comparing with stigma, respectively (Fig. 2). Among the DEGs, the number of up-regulated genes was higher than the number of down-regulated genes. Cluster analyses of the DEGs among three tissues revealed similar expression patterns between petal and stamen (Fig. 3 Suppl.).

Because the crocin content of petal and stamen was significantly lower than that of stigma (Jain *et al*. 2016), the down-regulated genes in petal and stamen might be related to crocin synthesis. Compared with stigma, there were 469 down-regulated genes in stamen, and 335 down-regulated genes in petal. Among them, 290 down-regulated genes were the same (Fig. 3). Thus, it can be inferred that some genes might be related to crocin synthesis.

To understand the functions of the down-regulated DEGs, all genes were mapped to *GO* terms in the *GO* database, to identify significantly enriched terms compared to the reference gene background. Comparisons revealed similar enriched *GO* terms of the down-regulated DEGs in petal and stamen. Of the 14 *GO* categories of down-regulated DEGs from petal, 11 appeared in the enriched *GO* terms of down-regulated DEGs identified in stamen (Fig. 4). According to the *KEGG* classification analysis, enriched clusters of the down-regulated DEGs in petal and stamen were also similar, which included “protein

Table 2. Expressions of genes involved in the carotenoid/apocarotenoid pathway. FPKM: fragments per kilobase of exon per million fragments mapped.

Annotation	Abbreviation	Component number	FPKM		
			stigma	petal	stamen
Phytoene synthase	*PSY*	c58514_g2	13.04	3.50	0.34
Phytoene desaturase	*PDS*	c49360_g1	9.83	4.11	3.09
Cis-zeta-carotene isomerase	*ZISO*	c59461_g2	8.72	1.29	0.84
Zeta-carotene desaturase	*ZDS*	c56346_g1	134.90	26.54	4.68
Beta-carotene hydroxylase	*β-CH*	c55986_g1	48.70	8.53	15.75
Carotenoid isomerase protein	*CRTISO*	c62785_g1	32.69	3.26	1.15
Lycopene beta-cyclase	*β-LYC*	c48928_g1	52.01	1.08	5.68
9-cis-epoxycarotenoid dioxygenase	*NCED*	c63465_g1	2.17	25.72	6.54
Carotenoid cleavage dioxygenase 1	*CCD1*	c63178_g1	0.00	13.06	21.75
Carotenoid cleavage dioxygenase 2	*CCD2*	c40525_g1	1.58	0.00	0.00
Carotenoid cleavage dioxygenase 4a	*CCD4a*	c62543_g3	2.01	3.76	25.00
Chromoplast carotenoid cleavage dioxygenase 4b	*CCD4b*	c40281_g1	3.23	0.81	0.00
Carotenoid cleavage dioxygenase 4c	*CCD4c*	c56863_g1	1.8	3.05	0.34
Carotenoid cleavage dioxygenase 7	*CCD7*	c72778_g1	0.00	2.12	0.42
Carotenoid cleavage dioxygenase 8b	*CCD8b*	c23069_g1	0.00	0.72	2.93
Carotenoid-associated protein	*CAP*	c55919_g1	108.09	86.15	41.12
Zeaxanthin 7,8 (7',8')-cleavage dioxygenase	*ZCD*	c61748_g1	115.63	193.22	40.17
Zeaxanthin epoxidase	*ZEP*	c58932_g1	0.75	2.04	1.43
Aviolaxanthin de-epoxidase	*VDE*	c60963_g1	2.22	4.66	1.4

Table 3. Down-regulated differentially expressed genes in both petal plus stigma and stamen plus stigma that relate to the carotenoid/apocarotenoid pathway.

Gene ID	Gene name
c56346_g1	*Zeta-carotene desaturase*
c48928_g1	*Lycopene beta cyclase*
c59086_g2	*UDP-glucosyltransferase UGT85V1*
c52564_g1	*UDP-glucosyltransferase UGT85U2*
c56534_g1	*Glucan endo-1,3-beta-glucosidase GVI*
c61813_g1	*Glucan endo-1,3-beta-glucosidase 7*

Table 4. Genes specifically expressed in stigmas that relate to the carotenoid/apocarotenoid pathway.

Gene ID	Gene name
c79742_g1	*Aldehyde dehydrogenase*
c58038_g3	*Putative UDP-glycosyltransferase 86A1*
c70566_g1	*Putative UDP-glycosyltransferase 85A5*
c88212_g1	*Putative UDP-glycosyltransferase 90A1*
c64600_g1	*Putative glucan endo-1,3-beta-glucosidase*
c73629_g1	*Putative beta-glucosidase 44*
c90928_g1	*Putative glucan endo-1,3-beta-glucosidase 3*
c94736_g1	*Putative glucan endo-1,3-beta-glucosidase GVI*

processing in endoplasmic reticulum", "estrogen signaling pathway", and "MAPK signaling pathway". These results indicated that there were many down-regulated DEGs in petal had the same function with those in stamen.

In stigma, the most important biosynthesis pathway was the carotenoid/apocarotenoid pathway and 19 genes in this pathway identified in previous reports (Baba *et al*. 2015b, Jain *et al*. 2016) were found in this study. Among them, 10 genes had high expressions in stigma, including phytoene synthase (*PSY*), phytoene desaturase (*PDS*), *cis*-zeta-carotene isomerase (*ZISO*), zeta-carotene desaturase (*ZDS*), beta-carotene hydroxylase (*β-CH*), carotene *cis/trans* isomerase (*CRTISO*), lycopene beta-cyclase (*β-LYC*), carotenoid-associated protein (*CAP*), *CCD2* and *CCD4b*. Most of these genes are involved in the zeaxanthin synthesis. In the *CCD* gene family, only *CCD2* and *CCD4b* had high expression in stigma, and *CCD2* was identified only in stigma (Table 2).

In order to find potential candidate genes related to crocin synthesis, 10 862 unigenes specifically expressed in stigma and 290 co-down-regulated DEGs in stamen and petal were analyzed. Among the co-down-regulated genes in stamen and petal, 6 genes (*ZDS*, *β-LYC*, *UGT85V1*, *UGT85U2*, endo-1,3 beta-glucosidase (*β-GS*), and endo-1,3 beta-glucosidase 7 (*β-GS7*)) were identified, which might be related to crocin synthesis (Table 3). Additionally, from the genes specially expressed in stigma, 8 genes (1 *ALDH* gene, 3 *UGTs* genes and 4 *GS* genes) were identified, which might play roles in the pathway from zeaxanthin to crocin (Table 4).

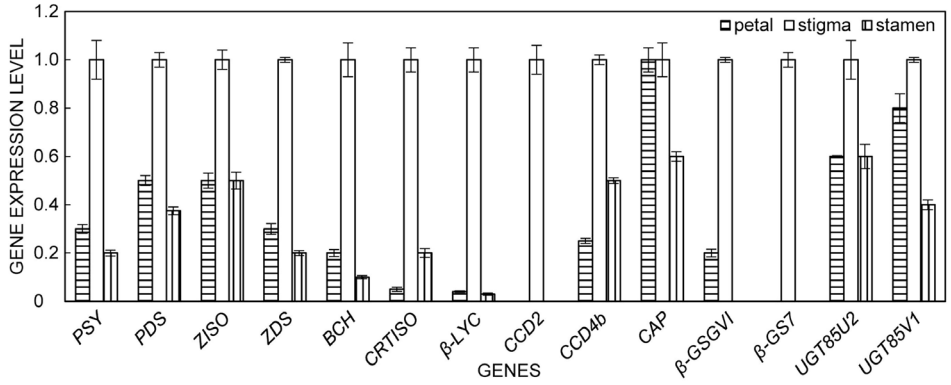

Fig. 5. Real time quantitative PCR validation of the expression patterns of 14 representative unigenes observed in *Crocus* tissues. *PSY - phytoene synthase*, *PDS - phytoene desaturase*, *ZISO - cis-zeta-carotene isomerase*, *ZDS - zeta-carotene desaturase*, *BCH - beta-carotene hydroxylase*, *CRTISO - carotene cis/trans isomerase*, *β-LYC - lycopene beta-cyclase*, *CCD2 - carotenoid cleavage dioxygenase 2*, *CCD4b - carotenoid cleavage dioxygenase 4b*, *CAP - carotenoid-associated protein*, *β-GSGVI - glucan endo-1,3-beta-glucosidase GVI*, *β-GS7 - glucan endo-1,3-beta-glucosidase 7*, *UGT85U2 - UDP-glucosyltransferase 85U2*, *UGT85V1 - UDP-glucosyltransferase 85V1*.

To confirm the reliability of the RNA-Seq data, real time quantitative PCR (qPCR) was performed with the same RNA pools used for next-generation sequencing. A total of 14 genes related to the apocarotenoid biosynthesis pathway, including *PSY*, *β-CH*, *CCD*, and UGTs, were selected for real time-qPCR analysis. The expression profiles observed in real time-qPCR matched well with those in the RNA-Seq data (Fig. 5).

Discussion

As reported in previous transcriptome studies (Baba *et al*. 2015b, Jain *et al*. 2016), the corms of *Crocus* are planted in field year-round; however, in this study, they were dug out when the leaves turned yellow. In the field condition, the flower differentiation environment of the *Crocus* cannot be controlled. Previous researches (Molina *et al*. 2005, Douglas *et al*. 2014) indicated that saffron flowering required optimal temperature for flower bud formation ranges from 23 to 27 °C during the summer and temperature for flower induction must be below 17 °C during the autumn. In this study, the growth of *Crocus* was at day/night temperatures of 25/22 °C from June to August, and of 18/16 °C from September to November, and it is very good for flower differentiation. So, we analyzed the *Crocus* tissues' transcriptome in this study to find some potential candidate genes related to crocin synthesis. The results showed that the transcriptome data in this study were as good as that in a previous study (Baba *et al*. 2015b), and some new candidate genes related to crocin synthesis were found. This indicated that the different growing environments affect the expression of some genes in *Crocus*.

Apocarotenoids (crocin, picrocrocin, and safranal) are the main active compounds found in the stigma of *Crocus*, and previous studies have identified the apocarotenoid pathway (Baba *et al*. 2015b, Jain *et al*. 2016). The crocin biosynthesis pathway in *Crocus* include the MEP, carotenoid and crocin pathways (Ji *et al*. 2017). GGPP is the product of the MEP pathway and is the substrate for many pathways. Thus, in this study, genes found in the MEP pathway were not analyzed as they were already well characterized. In the carotenoid pathway, GGPP is the substrate, and zeaxanthin is the product. Many genes involved in this pathway were previously identified in *Crocus* (Baba *et al*. 2015b, Jain *et al*. 2016). In this study, the expressions of these genes (*PSY*, *PDS*, *ZDS*, *ZISO*, *CRTISO*, *β-CH, and β-LYC*) in stigma were higher than that in stamen or petal; the higher expressions of these genes directly benefited zeaxanthin biosynthesis. β-carotene is the substrate of zeaxanthin, and it also can be catalyzed by *CCD7* and *CCD8* into strigolactones and catalyzed by *CCD4* and *CCD1* into β-cyclocitral and β-ionone. In this study, with the exception of *CCD4b*, the expressions of *CCD1*, *CCD4a*, *CCD4c*, *CCD7*, and *CCD8b* in stigma were not high. In fact, among these genes, *CCD1*, *CCD7* and *CCD8b* genes were not found in stigma at all.

In plants, the production of carotenoids is essential for photosynthesis and photoprotection. However, in the stigma of *Crocus*, carotenoids are not the end products. Zeaxanthin is the substrate of many different pathways, and it can be catalyzed by zeaxanthin epoxidase (ZEP) and violaxanthin de-epoxidase (VDE) to neoxanthin and catalyzed by ZEP and 9-*cis*-epoxycarotenoid dioxygenase (NCED) to ABA (Baba *et al*. 2015b). In this study, the expressions of these 3 genes were the highest in petal. These results revealed that when genes catalyzed the substrate for other products, not for apocarotenoids, their expressions were not the highest in stigma.

Apocarotenoid compounds are derived by oxidative cleavage of zeaxanthin in several steps after being catalyzed by CCDs, ALDHs, and UGTs (Baba *et al*. 2015b). Then, picrocrocin is converted to safranal through the addition of heat and β-GS (Baba *et al*. 2015b). This pathway specifically occurs in *Crocus*, thus, few papers have been published on the genes involved in the apocarotenoid synthesis pathway. *CCD2* was identified, and it was

found to be the key gene in this pathway (Frusciante *et al*. 2014). In this study, *CCD2* gene was only expressed in stigma and was not observed in the other flower tissues. Six *CsALDH* genes were identified by Demurtas *et al.* (2018), and *CsALDH3I1* converts crocetin dialdehyde into crocetin. *CsUGT74AD1* converts crocetin to crocin 1 and 2' (Demurtas *et al.* 2018). In this study, 14 potential candidate unigenes which might be related with crocin biosynthesis were identified, and they included *ZDS*, *β-LYC*, *UGT85V1*, *UGT85U2*, 3 *UGT* genes, 1 *ALDH* gene, and 6 *GS* genes. *ZDS* and *β-LYC* were identified in the tissues of this study and played important roles in the carotenoid pathway, and their expressions in stigma were higher than that in stamen or petal. *UGT85V1* and *UGT85U2* were also identified in the three tissues. Previous research demonstrated that the two genes in the stigma of *Crocus* were regulated by stress, and overexpression of them in *Arabidopsis* could enhance resistance to salt and oxidative stress (Ahrazem *et al*. 2015). In this study, the two genes were down-regulated in petal and stamen compared with stigma. These revealed that the two genes might be involved in apocarotenoid synthesis, and they might have other functions. The *ALDH* gene identified in this study was different from the six *CsALDH* genes identified by Demurtas *et al*. (2018). So, the function of the *ALDH* gene identified in this study is also unknown. The other 9 genes identified in this study were also different from those reported in previous studies. *UGT* genes, and *GS* genes involved in crocin synthesis, but the actual function of these candidate genes are needed to be identified by other experiments.

To summarize, *Crocus* transcriptome analyses were performed in this study. The differences in gene expressions among the three flower tissues were associated with their crocin content. Some potential candidate genes involved in crocin synthesis were found.

References

Agayev, Y.M., Fernandez, J.A., Zarifi, E.: Clonal selection of saffron (*Crocus sativus* L.): the first optimistic experimental results. - Euphytica **169**: 81-99, 2009.

Agayev, Y.M., Shakib, A.M., Soheilivand, S., Fathim M.: Breeding of saffron (*Crocus sativus*): possibilities and problems. - Acta Hort. **739**: 203-207, 2007.

Ahrazem, O., Rubio-Moraga, A., Trapero-Mozos, A., Climent, M.F.L., Gomez-Cadenas, A., Gomez-Gomez, L.: Ectopic expression of a stress-inducible glycosyltransferase from saffron enhances salt and oxidative stress tolerance in *Arabidopsis* while alters anchor root formation. - Plant Sci. **234**: 60-73, 2015.

Baba, S.A., Malik, A.H., Wani, Z.A., Mohiuddin, T., Shah, Z., Abbas, N., Ashraf, N.: Phytochemical analysis and antioxidant activity of different tissue types of *Crocus sativus* and oxidative stress alleviating potential of saffron extract in plants, bacteria, and yeast. - S. Afr. J. Bot. **99**: 80-87, 2015a.

Baba, S.A., Mohiuddin, T., Basu, S., Swarnkar, M.K., Malik, A.H., Wani, Z.A., Abbas, N., Singh, A.K., and Ashraf, N.: Comprehensive transcriptome analysis of *Crocus sativus* for discovery and expression of genes involved in apocarotenoid biosynthesis. - BMC Genomics **16**: 698, 2015b.

Bouvier, F., Suire, C., Mutterer, J., Camara, B.: Oxidative remodeling of chromoplast carotenoids: identification of the carotenoid dioxygenase *CsCCD* and *CsZCD* genes involved in *Crocus* secondary metabolite biogenesis. - Plant Cell **15**: 47-62, 2003.

D'Agostino, N., Pizzichini, D., Chiusano, M.L., Giuliano, G.: An EST database from saffron stigmas. - BMC Plant Biol. **7**: 53, 2007.

Demurtas, O.C., Frusciante, S., Ferrante, P., Diretto, G., Azad, N.H., Pietrella, M., Aprea, G., Taddei, A.R., Romano, E., Mi, J., Al-Babili, A., Frigerio, L., Giuliano, G.: Candidate enzymes for saffron crocin biosynthesis are localized in multiple cellular compartments. - Plant Physiol. **177**: 990-1006, 2018.

Douglas, M.H., Smallfield, B.M., Wallace, A.R., McGimpsey, J.A.: Saffron (*Crocus sativus* L.): The effect of mother corm size on progeny multiplication, flower and stigma production. - Sci Hort. **166**: 50-58, 2014.

Fernandez, J.: Biology, biotechnology and biomedicine of saffron. - Recent Res. Dev. Plant Sci. **2**: 127-159, 2004.

Frusciante, S., Diretto, G., Bruno, M., Ferrante, P., Pietrella, M., Prado-Cabrero, A., Beyer, P., Gomez-Gom, L., Al-Babili, S., Giuliano, G.: Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. - Proc. nat. Acad. Sci. USA **111**: 12246-12251, 2014.

Gainer, J.L., Brumgard, F.B. Using excess volume of mixing to correlate diffusivities in liquids. - Chem. Eng. Commun. **15**: 323-329, 1982.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli. E., Hacohen, N., Gnirke, A., Rhind, N., Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A.: Full-length transcriptome assembly from RNA-Seq data without a reference genome. - Nat. Biotechnol. **29**: 644-652, 2011.

Gresta, F., Avola, G., Lombardo, G.M., Siracusa, L., Ruberto, G.: Analysis of flowering, stigmas yield and qualitative traits of saffron (*Crocus sativus* L.) as affected by environmental conditions. - Sci Hort. **119**: 320-324, 2009.

Gresta, F., Lombardo, G.M., Siracusa, L., Ruberto, G.: Effect of mother corm dimension and sowing time on stigma yield, daughter corms and qualitative aspects of saffron (*Crocus sativus* L.) in a Mediterranean environment. - J. Sci. Food Agr. **88**: 1144-1150, 2010.

Hosseinzadeh, H., Sadeghnia, H.R., Ghaeni, F.A., Motamedshariaty, V.S., Mohajeri, S.A.: Effects of saffron (*Crocus sativus* L.) and its active constituent, crocin, on recognition and spatial memory after chronic cerebral hypoperfusion in rats. - Phytotheraphy Res. **26**: 381-386, 2012.

Jain, M., Srivastava, P.L., Verma, M., Ghangal, R., Garg, R.: *De novo* transcriptome assembly and comprehensive expression profiling in *Crocus sativus* to gain insights into apocarotenoid biosynthesis. - Sci. Rep. **6**: 22456, 2016.

Ji, A., Jia, J., Xu, Z., Li, Y., Bi, W., Ren, F., He, C., Liu, J., Hu, K., Song, J.: Transcriptome-guided mining of genes involved in crocin biosynthesis. - Front. Plant Sci. **8**: 518, 2017.

Juan, J.A., Corcoles, H.L., Munoz, R.M., Picornell, M.R.: Yield and yield components of saffron under different cropping systems. - Ind. Crop Prod. **30**: 212-219, 2009.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y.: KEGG for linking genomes to life and the environment. - Nucl. Acids Res. **36**(S1): 480-484, 2008.

Li, B., Dewey, C.: RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. - BMC

Bioinformatics **12**: 323, 2011.

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.

Mao, X., Cai, T., Olyarchuk, J.G., Wei, L.: Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. - Bioinformatics **21**: 3787-3793, 2005.

Molina, R.V., Valero, M., Navarro, Y., Guardiola, J.L., Garcia-Luis, A.: Temperature effects on flower formation in saffron (*Crocus sativus* L.). - Sci. Hort. **103**: 361-379, 2005.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, A.: Mapping and quantifying mammalian transcriptomes by RNA-Seq. - Nat. Methods **5**: 621-628, 2008.

Razavi, B.M., Hosseinzadeh, H., Movassaghi, A.R., Imenshahidi, M., Abnous, K.: Protective effect of crocin on diazinon induced cardiotoxicity in rats in subchronic exposure. - Chemicobiol. Interact. **203**: 547-555, 2013.

Rezaee, R., Hosseinzadeh, H.: Safranal: from an aromatic natural product to a rewarding pharmacological agent. - Iran J. basic med. Sci. **16**: 12-26, 2013.

Storey, J.D., Tibshirani, R.: Statistical significance for genome wide studies. - Proc. nat. Acad. Sci. USA **100**: 9440-9445, 2003.

Wang. Y., Han. T., Zhu, Y., Zheng, C.J., Ming, Q.L., Rahman, K., Qin, L.P.: Antidepressant properties of bioactive fractions from the extract of *Crocus sativus* L. - J. Natur. Med. **64**: 24-30, 2010.

Young, M.D., Wakefield, M.J., Smyth, G.K., Oshlack, A.: Gene ontology analysis for RNA-seq: accounting for selection bias. - Genome Biol. **11**: R14, 2010.