

Integrated metabolomics and transcriptome analysis on flavonoid biosynthesis in flowers of safflower (*Carthamus tinctorius* L.) during colour-transition

Chaoxiang Ren^{1,2,*}, Chao Chen^{1,2,*}, Shuai Dong³, Rui Wang^{1,2}, Bin Xian^{1,2}, Tianlei Liu⁴, Ziqing Xi^{1,2}, Jin Pei^{1,2,3} and Jiang Chen^{1,2,3}

¹ State Key Laboratory of Southwestern Chinese Medicine Resources, Chengdu, Sichuan, China

² College of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

³ The State Bank of Chinese Drug Germplasm Resources, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

⁴ College of Medical Technology, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

* These authors contributed equally to this work.

ABSTRACT

Background: Safflower (*Carthamus tinctorius* L.), well known for its flower, is widely used as a dye and traditional Chinese medicine. Flavonoids, especially flavonoid glycosides, are the main pigments and active components. However, their biosynthesis is largely unknown. Interestingly, the colour of flowers in safflower changed from yellow to red during flower development, while much of the gene and chemical bases during colour transition are unclear.

Methods: In this research, widely targeted metabolomics and transcriptomics were used to elucidate the changes in flavonoid biosynthesis from the gene and chemical points of view in flowers of safflower during colour transition. The screening of differential metabolites depended on fold change and variable importance in project (VIP) value. Differential expressed genes (DEGs) were screened by DESeq2 method. RT-PCR was used to analyse relative expressions of DEGs.

Results: A total of 212 flavonoid metabolites, including hydroxysafflor yellow A, carthamin and anthocyanins, were detected and showed a large difference. The candidate genes of glycosyltransferases and flavonoid hydroxylase that might participate in flavonoid glycoside biosynthesis were screened. Ten candidate genes were screened. Through integrated metabolomics and transcriptome analysis, a uridine diphosphate glucose glycosyltransferase gene, *CtUGT9* showed a significant correlation with flavonoid glycosides in safflower. In addition, expression analysis showed that *CtUGT9* was mainly expressed in the middle development of flowers and was significantly upregulated under MeJA treatment. Our results indicated that *CtUGT9* might play an important role in flavonoid glycoside biosynthesis during colour-transition in safflower.

Submitted 8 March 2022

Accepted 24 May 2022

Published 22 June 2022

Corresponding authors

Jin Pei, peixjin@163.com

Jiang Chen, janshen1986@163.com

Academic editor

Suresh Gawande

Additional Information and
Declarations can be found on
page 13

DOI 10.7717/peerj.13591

© Copyright

2022 Ren et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Genetics, Molecular Biology, Plant Science

Keywords Flavonoid glycosides biosynthesis, Glycosyltransferases, Colour-transition, Metabolomics, Transcriptome, Safflower

INTRODUCTION

Safflower (*Carthamus tinctorius* L.), also known as Hong Hua in China, is a member of the family Compositae or Asteraceae (Ashri et al., 1975). It is an annual, self-compatible, diploid ($2n = 2x = 24$) crop with a long cultivation history, and it is believed to have been domesticated approximately 4,000 years ago in the Fertile Crescent and has various centres of origins (Knowles, 1969). Safflower has been cultivated and used in China for more than 2,000 years, as recorded in the compendium of Materia Medica. Zhang Qian introduced safflower to China during the Han Dynasty on a diplomatic mission to the Western Regions (via the Silk Road). Safflower is known for its flowers and is used as dyes, cosmetics, and food additives worldwide (Mohammadi & Tavakoli, 2015; Azami et al., 2019). At the same time, in China and other Southeast Asian countries, dried tubular flowers are used in traditional Chinese medicine to improve cerebral blood flow and to treat various ailments, such as cerebrovascular and cardiovascular diseases (Jin et al., 2008), hypertension (Asgarpanah & Kazemivash, 2013), and coronary heart disease (Zhou et al., 2018; Chinese Pharmacopoeia Commission, 2015). Flavonoids are the main metabolite in safflower. Flavonoid glycosides (such as hydroxysafflor yellow A (HSYA)) are not only pigments of safflower (Kazuma et al., 2000; Wang et al., 2021a) but also the active component (Yang et al., 2009; Ao, Feng & Peng, 2018; Bai et al., 2020). However, their biosynthesis is largely unknown.

Currently, the core flavonoid biosynthetic pathway is well understood (Grotewold, 2006; Zhang, Butelli & Martin, 2014; Zhao & Tao, 2015). In safflower, according to a reported paper (Yue et al., 2013; Pu et al., 2021), there are mainly several flavonoid glycosides in safflower, such as HSYA, carthamin, and kaempferol-3-O- β -D-glucoside. According to their structure (Fig. 1), two kinds of glycosyltransferases might take part in their biosynthesis: O-glycosyltransferases (OGTs) (for the biosynthesis of kaempferol-3-O- β -D-glucoside) and C-glycosyltransferases (CGTs) (for the biosynthesis of HSYAs). In addition, according to the structure of HSYA, flavonoid hydroxylase, especially flavonoid four hydroxylase (Cytochrome P450, CYPs), might participate in its biosynthesis. In addition, there are many glycosyltransferases in other plants, such as C-glucosyl flavonoids in citrus plants (Ito et al., 2017), Glycyrrhiza glabra (Zhang et al., 2020), O-glucosyl flavonoids in *Medicago truncatula* (Luzia et al., 2007) and flavonoid hydroxylase in *Scutellaria baicalensis* (Zhao et al., 2018). Their work provides a reference for us to screen candidate genes involved in flavonoid glycoside biosynthesis in safflower.

Note that the flowers of safflower change their colour from yellow (Y) to red (R) during flowering (Ren et al., 2019; Pu et al., 2021). However, much of the gene and chemical bases during colour transition are still unclear. There are some reports on the chemical composition of safflower, while most research uses HPLC coupled with DAD, HPLC coupled with MS, or even UHPLC-Q-TOF-MS to analyze the components of safflower (Kazuma et al., 2000; Yue et al., 2013; Kim et al., 2020; Pu et al., 2021). However,

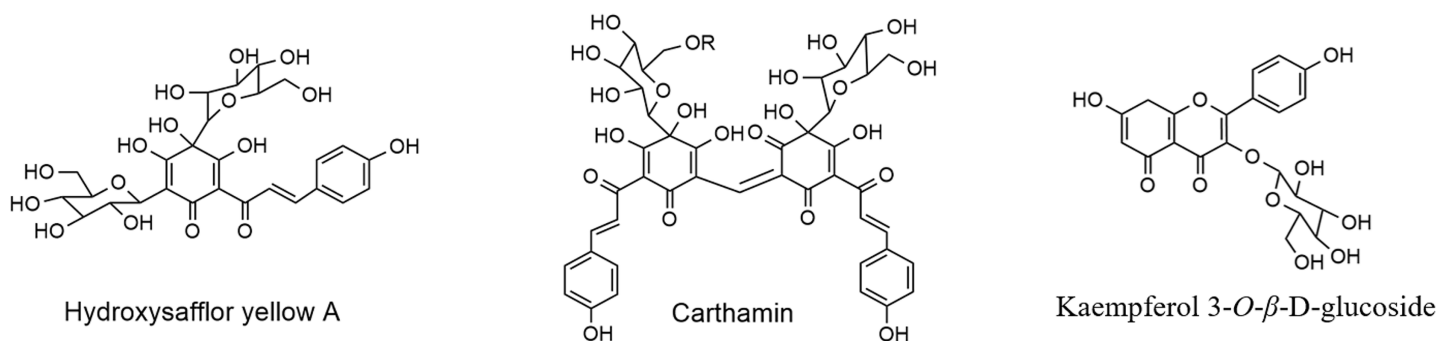


Figure 1 Structure of three flavonoid glycosides in safflower. They are Hydroxysafflor yellow A (HSYA), carthamin and kaempferol-3-O-β-D-glucoside from left to right. Full-size DOI: 10.7717/peerj.13591/fig-1

most of these reports have some shortcomings, such as a small number of compounds. In recent years, widely targeted metabolomics based on UHPLC–ESI–MS/MS has become very popular in the analysis and identification of compounds due to its advantages of high throughput, fast separation, high sensitivity, and wide coverage. At present, this method has been widely applied in plant metabolite analysis in rice (*Chen et al., 2013*), maize (*Wen et al., 2014*), tomato (*Zhu et al., 2018; Yang et al., 2021*), potato (*Cho et al., 2016*), and other plants (*Wang et al., 2021b; Xie et al., 2020*). We believe that this method can be used to detect the chemical changes during flower colour transition in safflower.

Recently, analysis of the metabolomics combined with the transcriptome had widely used in revealing the mechanisms of many important physiological processes (*Scossa, Alseikh & Fernie, 2021; Xia et al., 2021; Zou et al., 2021*). In research on *Lonicera japonica* Thunb, integrated metabolomics and transcriptome analysis of pigment accumulation in *Lonicera japonica* flower petals during colour transition revealed changes in key pigments and related biosynthesis genes associated with petal colour transitions (*Xia et al., 2021*). Besides, in our previously research, through integrated metabolomics and transcriptome methods, we analyze the molecular mechanism of flavonoid biosynthesis under MeJA treatment in safflower (*Chen et al., 2020*). Therefore, we believe that through integrated metabolomics and transcriptome analysis on flavonoid biosynthesis in safflower flowers during colour transition, some genes associated with flower colour transitions can be elucidated.

In this research, widely targeted metabolomics and transcriptomics were used to elucidate the changes from gene and chemical points of view in flavonoid biosynthesis in safflower flowers during colour transition. Expression profile of the genes that were significantly correlated with flavonoid glycosides in safflower were analysed. We aimed to elucidate the colour change from a chemical point of view and to screen candidate genes that participate in flavonoid glycoside biosynthesis through integrated metabolomics and transcriptome analysis of safflower flowers during colour transition.

MATERIALS AND METHODS

Plant materials and sampling

Safflower, “Chuanhonghua No. 1”, was used in this experiment. It was selected and bred by the Industrial Crop Research Institute, Sichuan Academy of Agricultural Sciences, and

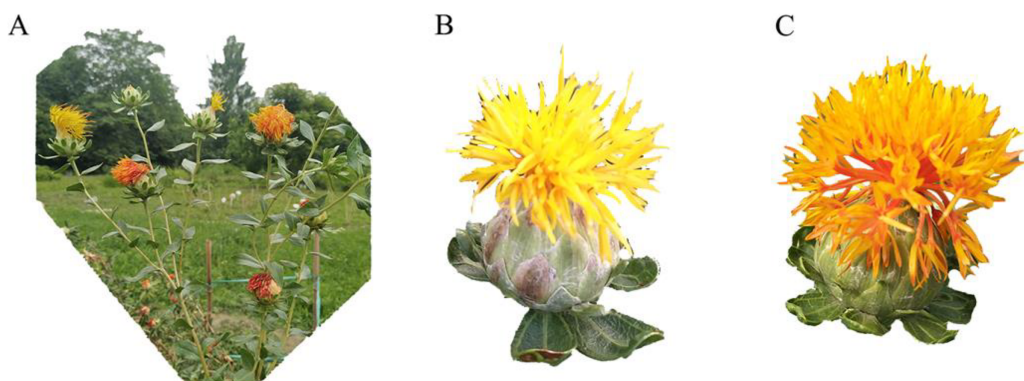


Figure 2 The colour transition of flowers in safflower. (A) Colour transition in flower at different development of flower in the same plant. (B) The flower phenotype at Day 2. (C) The flower phenotype at Day 4. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.13591/fig-2](https://doi.org/10.7717/peerj.13591/fig-2)

presented by Renchuan Yao to our lab. In this research, it was cultivated in a medicinal botanical garden at the Wenjiang campus of Chengdu University of Traditional Chinese Medicine in 2019. In our previous research, we recorded the development of safflower flowers, and the flowering period time of safflower can last 7 days, where the colour of flowers transitions from yellow (Y) to red (R) from Day 2 to Day 4. Thus, we chose the flowers of Day 2 and Day 4 as the samples in our study (Fig. 2). Samples of flowers were collected, frozen immediately in liquid nitrogen and stored in a freezer at -80°C . For RNA sequencing, five inflorescences of safflower at the same flowering stage were mixed as one sample. For metabolism analysis, 10 inflorescences of safflower plants at the same flowering stage were mixed as one sample. Three biological replicates were performed for transcriptomics metabolomics analysis.

Sample preparation and extraction for metabolomic analysis

This sample preparation and extraction were mainly consistent with our previous research (Chen *et al.*, 2020). The freeze-dried flowers were crushed using a mixer mill (MM 400; Retsch, Haan, Germany) with a zirconia bead for 1.5 min at 30 Hz. A total of 100 mg of powder was weighed and extracted overnight at 4°C with 1.0 mL of 70% aqueous methanol and then centrifuged at 10,000 g for 10 min. The extracts were absorbed (CNWBOND Carbon-GCB SPE Cartridge, 250 mg, 3 mL; ANPEL, Shanghai, China, www.anpel.com.cn/cnw) and filtered (SCAA-104, 0.22 μm pore size; ANPEL, Shanghai, China) before LC-MS analysis. The UHPLC and mass spectrometry conditions also followed those of previous research (Chen *et al.*, 2020).

Qualitative and quantitative analysis of metabolites

Flavonoid identification and quantification in our study were performed according to a previous report (Chen *et al.*, 2020). Besides, in our previous research, carthamin was not included. This time, we used the standard sample of carthamin to build a library and added them to the homemade database. The metabolites of the samples were qualitatively and quantitatively analysed by mass spectrometry. The integrated area peak of each

compound was used for principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA).

Differential metabolite analysis

The fold change and variable importance in project (VIP) value of the OPLS-DA model were combined to screen differential metabolites as described in our previous research ([Chen et al., 2020](#)). PCA was performed by using R software built-in functions (www.r-project.org/). For the parameter, scale = True. After conversion of the original data by log₂, the data were centralized (Mean Centering) and analyzed. The screening criteria were as follows: 1. Metabolites must exhibit fold changes ≥ 2 and fold changes ≤ 0.5 . If a metabolite in the experimental group exhibited a fold change more than two times or less than 0.5 times that of the metabolite in the control group, the difference was considered significant 2. If there was biological duplication in the sample group, the metabolites of VIP ≥ 1 were selected based on the above. The VIP value indicates the influence intensity of the difference between groups of corresponding metabolites in the classification and discrimination of each group of samples in the model.

RNA sequencing and annotation

RNA isolation and sequencing were performed as previously described ([Love, Wolfgang & Simon, 2014](#)). The samples were ground on dry ice, and total RNA was prepared by using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). To remove DNA, an aliquot of total RNA was treated with DNase (Takara, Dalian, China). RNA was determined by using a NanoDrop 2000 spectrophotometer (NanoDrop, Wilmington, DE, USA) and an Agilent 2100 bioanalyzer (Agilent, Santa Clara, CA, USA). mRNA was isolated from total RNA using magnetic beads with oligo (dT) primers; cDNA was synthesized by using a cDNA synthesis kit (Takara, Dalian, China).

The library preparations were sequenced on an Illumina HiSeq platform with three biological replicates in each group. In this project, Trinity v2.0.6 ([Grabherr et al., 2011](#)) was used to splice the filtered high-quality sequencing data to obtain the transcriptome, which was used as the reference sequence for subsequent differential expression analysis.

After the transcriptome was assembled, clean reads were compared with the transcriptome. The average sequencing depth (counts * 150/gene_len) was about 424.

A power analysis was performed by RNASeqpower in R package ([Hart et al., 2013](#)). The statistical power of this experimental design, calculated in RNASeqpower was 0.9210878 (depth = 424, $n = 3$, CV = 0.24, effect = 2, alpha = 0.05).

The UniGene sequences obtained were annotated through six databases: KEGG, GO, NR, Swiss-Prot, trEMBL and KOG. The whole set of transcript data can be found in the National Center for Biotechnology Information (NCBI) database (BioProject ID: [PRJNA774916](#)).

Screening of differential genes

The screening of differential expressed gene was mainly consistent with our previous research ([Wang et al., 2021a](#)). The gene expression levels were calculated through RSEM

(RNAseq by Expectation-Maximization) v1.2.12 (Li & Colin, 2011). DESeq2 (Love, Wolfgang & Simon, 2014) was used for differential expression analysis. Read counts of genes are the expected count outputs calculated using RSEM. After discrepancy analysis, multiple hypothesis tests are needed to correct the hypothesis test probability (P value) with the Benjamini-Hochberg procedure to obtain the false discovery rate (FDR) when $|\log_2\text{fold change}| \geq 1$ and $\text{FDR} < 0.05$.

Expression profile in the development of different flower tissues

According to previous research (Ren *et al.*, 2019), the development of safflower can be listed for 7 days. However, in the last 2 days, the flowers are almost dry. Thus, we chose the beginning of the first 5 days to check gene expression. In addition, the root, stem and leaf at anthesis were used to detect gene expression in different tissues. RT-PCR was used to measure their expression, as previously described (Ren *et al.*, 2019). Total RNA was isolated using an RNA extraction kit (Invitrogen, CA, USA), and reverse transcription was carried out using the Prime Script Reagent Kit (Takara, Dalian, China). The primers used to amplify the screened gene *CtUGT9* by real-time PCR were designed by Primer 5.0, with the sequences ACTCAAAGGCTTCTCAAT (Forward) and GGTGTAGTGTTAAA GGGTAAAT (Reverse), and parts of the safflower 28S coding region were used as internal reference genes, with the sequences GGGTCCTTTCACGTTTCTGA (Forward) and GGCTGACTTATCGGTAGCA (Reverse). RT-PCR was performed using SYBR Premix Ex Taq II (Takara, Shiga, Japan) with three replicates, and the cycling conditions were set according to the manual. The Bio-Rad CFX96 real-time system (Hercules, CA, USA) was used in this experiment.

MeJA treatment of the flower

The MeJA treatment of the flowers was mainly performed according to our previous research (Chen *et al.*, 2020). A 100 μM MeJA (Sigma-Aldrich, St. Gallen, Switzerland) was sprayed onto healthy safflower flowers 3 days after anthesis (DAA). The flowers sprayed with the same solution without MeJA were used as the control group. The flowers were then enclosed in clear plastic bags. After treatment for 6 h, flowers were collected, immediately frozen in liquid nitrogen for further analysis.

RESULTS

Metabolic profiling and differential flavonoid analysis in flowers of safflower during colour transition

The flavonoids in flowers of safflower during colour transition from yellow to red were detected based on UHPLC ESI-MS/MS (including HSYA and carthamin). A total of 212 flavonoids were detected, including 64 flavones, 41 flavonols, 40 flavone C-glycosides, 22 flavonones, 10 isoflavones, 10 catechin derivatives, 19 anthocyanins, two quinone chalcones, two flavonolignans, one alkaloid, and one proanthocyanidin (Table S1).

PCA shows that the two samples (Y and R) can be clearly clarified based on flavonoids (Fig. S1). The fold change and VIP value of the OPLS-DA model were combined to screen differential flavonoids. There were 41 significantly differential flavonoids in the

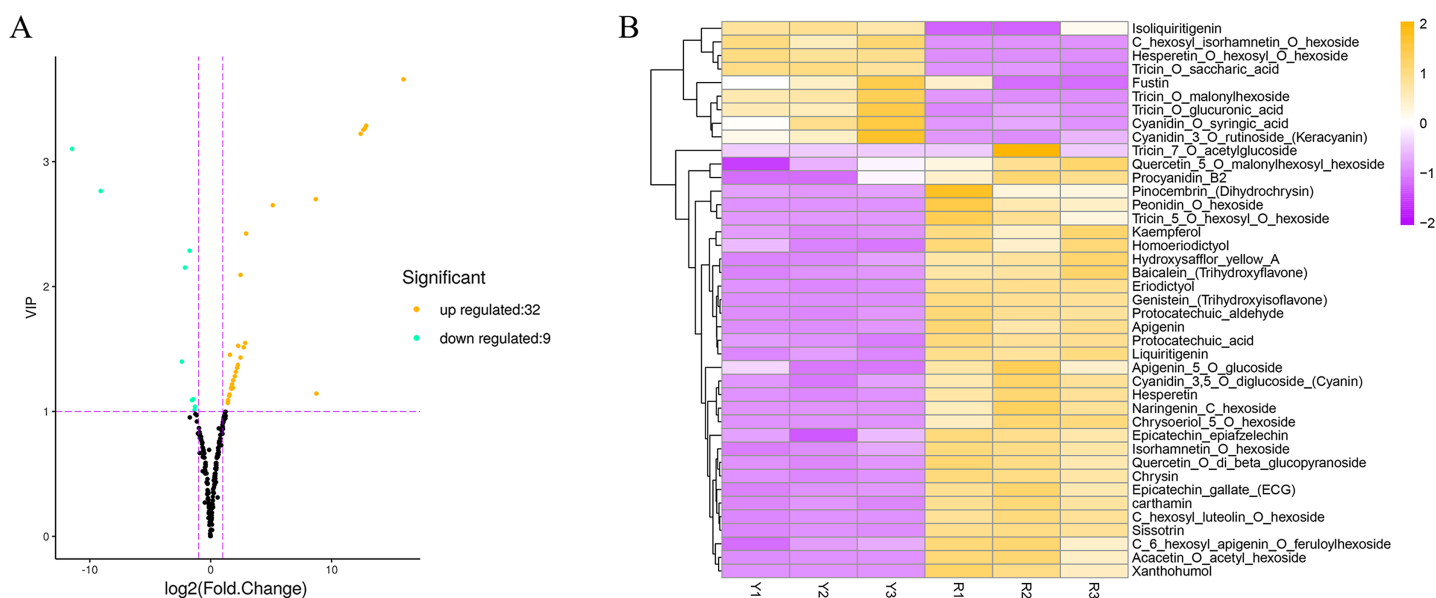


Figure 3 Differential flavonoid analysis in flowers of safflower during colour transition. (A) Volcanic map of differential flavonoids, red dot means the up-regulated flavonoid, while green dot means the down-regulated flavonoid. (B) Heatmap of the differential flavonoids.

Full-size DOI: 10.7717/peerj.13591/fig-3

flowers of safflower during the colour transition from Y to R. Among them, 32 flavonoids were upregulated, and nine flavonoids were downregulated (Fig. 3A). A heatmap of the differential flavonoids is drawn (Fig. 3B). It shows that 4 flavone C-glycosides are differently regulated, among which three flavone C-glycosides are significantly upregulated from Y to R. It also shows that four anthocyanins are differently regulated, and two anthocyanins (peonidin *O*-hexoside and cyanidin 3,5-*O*-diglucoside) are significantly upregulated from Y to R. All the details of the differential flavonoids can be viewed in Table S2.

Transcriptome sequencing and differential transcript analysis in flowers of safflower during colour transition

The transcriptome of flowers during colour transition was sequenced: 47.49 G clean bases were obtained, including 320,594,102 raw reads and 318,575,696 clean reads. The Q30 value was above 93.3%. The details can be viewed in Table S3. Trinity (Grabherr *et al.*, 2011) was used to splice the filtered high-quality sequencing data to obtain the transcriptome as a reference sequence for subsequent differential expression analysis. A total of 437,060 transcripts and 270,890 UniGenes were extracted (Table S4). Six databases (KEGG, GO, NR, Swiss-Prot, trEMBL and KOG) were used for protein annotation, and 90,961 genes were annotated by at least one database (Table S5).

The differentially expressed genes were analyzed with DESeq2 (Love, Wolfgang & Simon, 2014). A total of 4,820 genes were differentially expressed in the flowers of safflower during colour transition, among which 2,018 genes were downregulated and 2,802 genes were upregulated (Fig. 4). The details can be viewed in Table S6. The differentially expressed genes were annotated through the KEGG database. The differential genes mainly

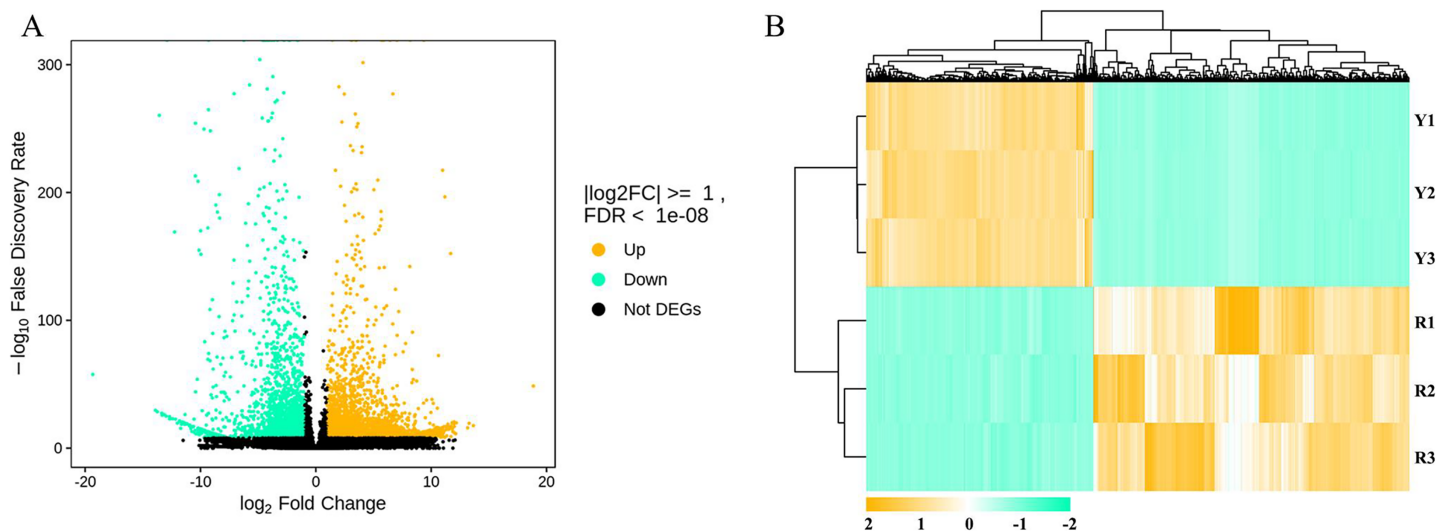


Figure 4 Differential transcript analysis in flowers of safflower during colour transition. (A) Volcanic map of differentially expressed genes, red dot means the up-regulated genes, while green dot means the down-regulated genes. (B) Heatmap of differentially expressed genes. Cluster analysis was performed according to gene expression levels. [Full-size !\[\]\(5fd6ef84f97f42d7f8b34275f1b65312_img.jpg\) DOI: 10.7717/peerj.13591/fig-4](https://doi.org/10.7717/peerj.13591/fig-4)

participated in metabolic pathways (648), biosynthesis of secondary metabolites (377), ribosomes (276), and MAPK signalling (171). Among them, 67 genes participated in phenylpropanoid biosynthesis, including 18 genes that participated in flavonoid biosynthesis (KO00941), three genes that participated in anthocyanin biosynthesis (KO00942), and two genes that participated in flavone and flavonol biosynthesis (KO00944). The details can be viewed in [Table S7](#).

Screening and expression analysis of glycosyltransferases and flavonoid hydroxylase participating in flavonoid glycoside biosynthesis

It can be inferred from the structure of the flavonoid glycosides in safflower ([Yue et al., 2013](#); [Pu et al., 2021](#)) ([Fig. 1](#)) that two kinds of UGT might participate in flavonoid glycosides, O-glycosyltransferases (OGTs) and C-glycosyltransferases (CGTs). As all of the UGT family members have the domain PF00201 (UDP-glucuronosyl and UDP-glucosyl transferase domains), the UGT gene sequences were first extracted from the transcriptome of safflower flowers based on PF00201. A total of 79 genes were screened. Phylogenetic analysis was performed with the reported UGTs ([Luzia et al., 2007](#); [Ito et al., 2017](#); [Zhang et al., 2020](#)), including 11 CGTs and 6 OGTs. Nine candidate UGTs were screened, including 2 CGTs and 7 OGTs, which were named *CtUGT1-CtUGT9* ([Fig. S2](#)). The genes used for BLAST in the research are listed in [Table S8](#). In addition, according to the structure of HSYA, flavonoid hydroxylases, isoforms such as CYP82D2 in *Scutellaria baicalensis* ([Zhao et al., 2018](#)), might participate in flavonoid hydroxylase in safflower. All flavonoid hydroxylases contain the domain PF00067 (cytochrome P450 domain). Therefore, all flavonoid hydroxylase gene sequences were extracted from the transcriptome of flowers of safflower on PF00067. A total of 271 genes were screened.

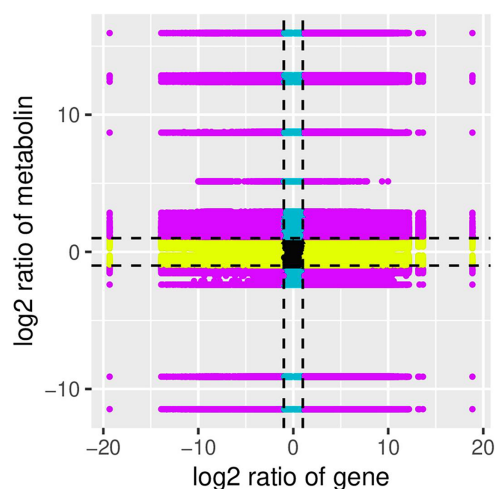


Figure 5 The nine quadrants constructed based on the Pearson correlation coefficient. Black dashed lines were used to divide the figure into 1–9 quadrants. Five means that genes and flavonoids are not differently expressed. Three and seven mean differential expression patterns of genes and flavonoids are consistent, and genes and flavonoids that are positively correlated in the two quadrants. Metabolite changes may be positively regulated by genes. One, two and four mean that the expression abundance of metabolite is higher than that of genes. Genes and flavonoids are negatively correlated to each other, which means that flavonoids are up-regulated, while genes are unchanged or down-regulated. Six, eight and nine mean that the expression abundance of metabolite is lower than that of genes. Genes are negatively correlated to flavonoids. When genes are up-regulated, flavonoids remain unchanged or down-regulated.

Full-size  DOI: [10.7717/peerj.13591/fig-5](https://doi.org/10.7717/peerj.13591/fig-5)

Then, phylogenetic analysis was performed with CYP82D2. One homologue of the CYP82D gene was screened and named *CtCYP1* (Fig. S3). All the sequences of the screened genes are listed in Table S9.

Integrated metabolomics and transcriptome analysis in flowers of safflower during colour transition

In this study, nine quadrants were constructed to perform the integrated metabolomics and transcriptome analyses. Correlation analysis of the genes and metabolites detected in flowers of safflower during colour transition was performed using the *cor* program in R to calculate the Pearson correlation coefficients between the genes and metabolites. The nine quadrants were constructed based on the Pearson correlation coefficient (Fig. 5). The results with Pearson correlation coefficients greater than 0.8 were selected, and the clustering heatmap based on the correlation coefficient is shown in Fig. 6. The nine genes screened through phylogenetic analysis were analysed, and the results showed that only the *CtUGT9* gene (TRINITY_DN101065_c9_g1) was positively correlated (PPC above 90%) to 25 flavonoids, including one flavonoid glycoside. Details are shown in Table S10.

Expression profile of *CtUGT9* and its response to MeJA treatment

Integrated metabolomics and transcriptome analysis in flowers of safflower during colour transition indicated that *CtUGT9* is significantly correlated with many flavonoids, including one flavonoid glycoside, indicating that it might play an important role in the

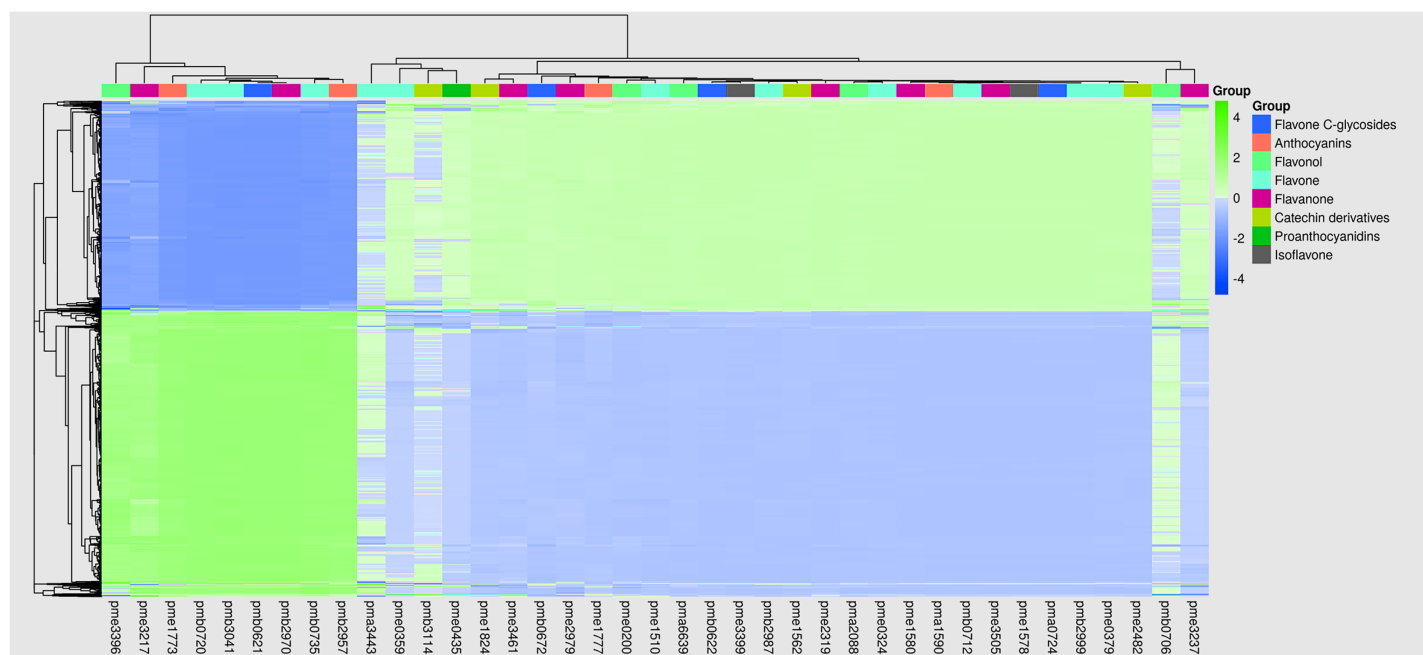


Figure 6 The clustering heat map of genes and flavonoids. The result is showed based on the pearson correlation coefficient greater than 0.8.

Full-size DOI: 10.7717/peerj.13591/fig-6

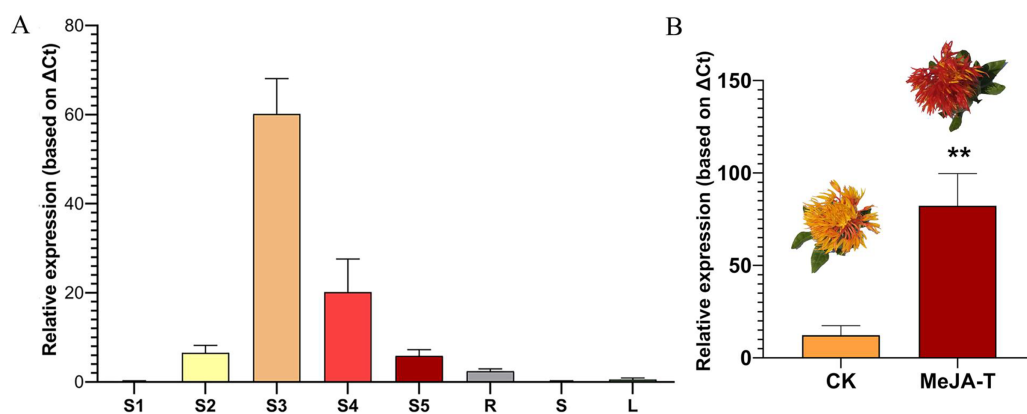


Figure 7 The expression profile of the *CtUGT9* and its response to MeJA treatment. (A) The expression profile of the *CtUGT9*. S1–S5 mean five different development stages of flower in safflower. R means Root, S means Stem, L mean leaf. (B) *CtUGT9* response to MeJA treatment. CK means untreated of MeJA. MeJA-T mean the treatment of MeJA. Two asterisks (**) mean significance at 0.01%. The flower phenotypes of untreated and treatment of MeJA on safflower were shown in the figure.

Full-size DOI: 10.7717/peerj.13591/fig-7

biosynthesis of flavonoid glycosides. Gene expression is closely related to its function; thus, we further detected the expression profile of *CtUGT9*. Its expression in different tissues (root, stem, leaf and flower tissue) and in the development of flowers was detected. The results showed that *CtUGT9* was mainly expressed in flowers, and along with the development of flowers, the expression of *CtUGT9* showed a trend of rising first and then decreasing, with the highest expression at Day 3 of flowering (Fig. 7A).

Our previous study reported that flavonoids of safflower, including flavonoid glycosides, can be upregulated after MeJA treatment (Chen *et al.*, 2020). Therefore, we speculate that *CtUGT9* should also be upregulated by MeJA. Thus, its expression under MeJA treatment was detected. The results showed that there was a significant increase after MeJA treatment ($P = 0.01$) (Fig. 7B). Overall, the results indicated that *CtUGT9* might play an important role in the biosynthesis of flavonoids and flavonoid glycosides in the flowers of safflower.

DISCUSSION

The flowers of safflower have multipurpose usage, not only in dyes and cosmetics but also in traditional Chinese medicine. There are two well-known pigments in safflower: yellow pigments, such as HSYA, which is also an active component in traditional Chinese medicine, and red pigments, such as carthamin (Kazuma *et al.*, 2000; Pu *et al.*, 2021). Our results showed that during the colour transition of safflower flowers, HSYA and carthamin showed a substantial difference, with a significant increase from Y to R, which is the same as a previous report (Pu *et al.*, 2021). However, anthocyanins were also detected and showed a significant difference during color transition in safflower. Four anthocyanins were differentially regulated, among which 2 anthocyanins (peonidin *O*-hexoside and cyanidin 3,5-*O*-diglucoside) were significantly upregulated during the colour transition from Y to R. It is possible that anthocyanins might play a role in the colour transition of safflower.

A total of 4,820 genes were differentially expressed in the flowers of safflower during colour transition. Among these genes, 67 participate in phenylpropanoid biosynthesis, including 18 genes that participate in flavonoid biosynthesis (KO00941), three genes that participate in anthocyanin biosynthesis (KO00942), and two genes that participate in flavone and flavonol biosynthesis (KO00944). We analysed gene expression in the KEGG map (Fig. S4). Most genes downstream of the flavonoid biosynthesis pathway were downregulated during the colour transition of flowers. Only several genes upstream of the flavonoid biosynthesis pathway, such as *HCT* and *CHI*, were unregulated during colour transition. One possible reason is that the flavonoid components in safflower, especially flavonoid glycosides, branch upstream from the flavonoid biosynthesis pathway, which is not the same as the known pathway of flavonoid biosynthesis. In our previous report, we also found that after treatment with MeJA, the contents of flavonoid glycosides, especially HSYA components, increased, but gene expression downstream of the flavonoid biosynthesis pathway decreased (Chen *et al.*, 2020). These results also reveal that the flavonoid biosynthesis pathway in safflower shows a substantial difference from the known flavonoid biosynthesis pathway in model plants, such as *Arabidopsis thaliana*.

In our experiment, 10 genes were screened as candidate genes that might participate in flavonoid glycoside biosynthesis in safflower, including *CtCYP1* and *CtUGT1-CtUGT9*. However, only *CtUGT9* showed a significant correlation with most flavonoids, including flavonoid glycosides. All the other 10 candidate genes were detected in our experiment, and the results showed that only *UGT9* was highly expressed in safflower and had a

substantial change during the colour transition of flowers compared with the other nine genes (Fig. S5).

Flavonoid glycosides, the main components of the pigment in safflower, are mainly detected in flowers (Kazuma *et al.*, 2000; Kim *et al.*, 2020; Wang *et al.*, 2021a; Pu *et al.*, 2021). Gene expression is closely related to its function. We further detected the expression of *CtUGT9*, and the results showed that it was mainly expressed in flowers. During flower development, the expression of *CtUGT9* showed a trend of rising first and then decreasing (Fig. 7A). In addition, *CtUGT9* showed a strong response to MeJA treatment. The expression analysis indicated that *CtUGT9* might participate in flavonoid glycoside biosynthesis (Fig. 7B).

In our research, we planned to find CGTs and CYPs that might participate in flavonoid biosynthesis. However, the screened CYPs or CGTs did not seem to participate in flavonoid glycoside biosynthesis in safflower. The UGTs in safflower may differ from the reported UGTs. From the reported UGTs (Ito *et al.*, 2017; Luzia *et al.*, 2007; Zhang *et al.*, 2020), most of the substrate contains flavonoids; however, most of the C-glycosyl in safflower is synthesized based on chalcone (Fig. 1). In addition, for HSYA, there is a –OH group at the four position. Therefore, in the future, to identify candidate genes that participate in flavonoid glycoside biosynthesis in safflower, especially flavonoid C-glycoside biosynthesis, some other methods might be used, such as GWAS (Chen *et al.*, 2014) or mGWAS (Fang & Luo, 2019).

CONCLUSIONS

Safflower (*Carthamus tinctorius* L) is widely used as a dye and traditional Chinese medicine. Flavonoids, especially flavonoid glycosides, are pigments and active components in safflower. Interestingly, the colour of safflower flowers changes from yellow to red during development. In this research, widely targeted metabolomics and transcriptomics were adapted to elucidate the changes from a gene and chemical point of view in flavonoid biosynthesis in flowers of safflower during colour transition. A total of 212 differential flavonoid metabolites and 4,820 differentially expressed genes were detected in the flowers of safflower during colour transition. Expression profiles of the genes that were significantly correlated with flavonoid glycosides in safflower were analysed. We found that *CtUGT9* was significantly associated with flavonoids and flavonoid glycosides. It was mainly expressed in flowers and showed a trend of rising first and then decreasing, with the highest expression at Day 3 of flowering. In addition, *CtUGT9* was significantly upregulated by MeJA treatment. Our results indicated that *CtUGT9* played an important role in flavonoid glycoside biosynthesis, which might affect the flower colour transition in safflower.

ACKNOWLEDGEMENTS

We acknowledge pro. Cathie Martin (JIC) for the help on the screening of UGT in safflower, Yuanyuan Qin and Wenhan Hu (Metware Biotechnology Co., Ltd., Wuhan, China) for the data analysis, American Journal Experts for the language editing.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This project is funded by grants from the National Natural Science Foundation of China (81803669, U19A2010), the Key R&D Plan of Science and Technology Department of Sichuan Province (2021YFYZ0012-5, 2020YFN0152), the Sichuan Provincial central guiding local science and technology development special project (2020ZYD058), and the Xinglin Talent Program of Chengdu University of TCM (0300510007). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 81803669, U19A2010.

Key R&D Plan of Science and Technology Department of Sichuan Province: 2021YFYZ0012-5, 2020YFN0152.

Technology Development Special Project: 2020ZYD058.

Chengdu University of TCM: 0300510007.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Chaoxiang Ren conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Chao Chen performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Shuai Dong analyzed the data, prepared figures and/or tables, and approved the final draft.
- Rui Wang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Bin Xian performed the experiments, prepared figures and/or tables, and approved the final draft.
- Tianlei Liu performed the experiments, prepared figures and/or tables, and approved the final draft.
- Ziqing Xi performed the experiments, prepared figures and/or tables, and approved the final draft.
- Jin Pei conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Jiang Chen conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The transcriptome data is available at NCBI: [PRJNA774916](https://pubmed.ncbi.nlm.nih.gov/3574916/).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.13591#supplemental-information>.

REFERENCES

- Ao H, Feng WW, Peng C. 2018.** Hydroxysafflor yellow A: a promising therapeutic agent for a broad spectrum of diseases. *Evidence-Based Complementary and Alternative Medicine* **2018**:8259280 DOI [10.1155/2018/8259280](https://doi.org/10.1155/2018/8259280).
- Asgarpanah J, Kazemivash N. 2013.** Phytochemistry, pharmacology and medicinal properties of *Carthamus tinctorius* L. *Chinese Journal of Integrative Medicine* **19**(2):153–159 DOI [10.1007/s11655-013-1354-5](https://doi.org/10.1007/s11655-013-1354-5).
- Ashri A, Zimmer DE, Urie AL, Knowles PF. 1975.** Evaluation of the germ plasm collection of safflower *Carthamus tinctorius* L. VI. Length of planting to flowering period and plant height in Israel, Utah and Washington. *Theoretical and Applied Genetics* **46**(7):359–364 DOI [10.1007/BF00281678](https://doi.org/10.1007/BF00281678).
- Azami K, Hayashi T, Kusumi T, Ohmori K, Suzuki K. 2019.** Total synthesis of carthamin, a traditional natural red pigment. *Angewandte Chemie International Edition* **58**(16):5321–5326 DOI [10.1002/anie.201900454](https://doi.org/10.1002/anie.201900454).
- Bai X, Wang W, Fu R, Yue S, Gao H, Chen Y, Tang Y. 2020.** Therapeutic potential of hydroxysafflor yellow A on cardio-cerebrovascular diseases. *Frontiers in Pharmacology* **11**:01265 DOI [10.3389/fphar.2020.01265](https://doi.org/10.3389/fphar.2020.01265).
- Chen W, Gao Y, Xie W, Gong L, Lu K, Wang W, Li Y, Liu X, Zhang L, Dong H, Zhang W, Zhang L, Yu S, Wang G, Lian X, Luo J. 2014.** Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nature Genetics* **46**(7):714–721 DOI [10.1038/ng.3007](https://doi.org/10.1038/ng.3007).
- Chen W, Gong L, Guo Z, Wang W, Zhang H, Liu X, Yu S, Xiong L, Luo J. 2013.** A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Molecular Plant* **6**(6):1769–1780 DOI [10.1093/mp/sst080](https://doi.org/10.1093/mp/sst080).
- Chen J, Wang J, Wang R, Xian B, Ren C, Liu Q, Wu Q, Pei J. 2020.** Integrated metabolomics and transcriptome analysis on flavonoid biosynthesis in safflower (*Carthamus tinctorius* L.) under MeJA treatment. *BMC Plant Biology* **20**(1):353 DOI [10.1186/s12870-020-02554-6](https://doi.org/10.1186/s12870-020-02554-6).
- Chinese Pharmacopoeia Commission. 2015.** *Pharmacopoeia of the people's Republic of China*. Beijing: China Medical Science Press.
- Cho K, Cho K, Sohn H, Ha I, Hong S, Lee H, Kim Y, Nam M. 2016.** Network analysis of the metabolome and transcriptome reveals novel regulation of potato pigmentation. *Journal of Experimental Botany* **67**(5):1519–1533 DOI [10.1093/jxb/erv549](https://doi.org/10.1093/jxb/erv549).
- Fang C, Luo J. 2019.** Metabolic GWAS-based dissection of genetic bases underlying the diversity of plant metabolism. *The Plant Journal* **97**(1):91–100 DOI [10.1111/tpj.14097](https://doi.org/10.1111/tpj.14097).
- Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren B, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011.** Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**(7):644–652 DOI [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883).

- Grotewold E. 2006.** The genetics and biochemistry of floral pigments. *Annual Review of Plant Biology* 57(1):761–780 DOI 10.1146/annurev.arplant.57.032905.105248.
- Hart S, Therneau T, Zhang Y, Poland G, Kocher J. 2013.** Calculating sample size estimates for RNA sequencing data. *Journal of Computational Biology* 20(12):970–978 DOI 10.1089/cmb.2012.0283.
- Ito T, Fujimoto S, Suito F, Shimosaka M, Taguchi G. 2017.** C-Glycosyltransferases catalyzing the formation of di-C-glucosyl flavonoids in citrus plants. *The Plant Journal* 91(2):187–198 DOI 10.1111/tpj.13555.
- Jin Y, Xiao Y, Zhang F, Xue X, Xu Q, Liang X. 2008.** Systematic screening and characterization of flavonoid glycosides in *Carthamus tinctorius* L. by liquid chromatography/UV diode-array detection/electrospray ionization tandem mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis* 46(3):418–430 DOI 10.1016/j.jpba.2007.10.036.
- Kazuma K, Takahashi T, Sato K, Takeuchi H, Matsumoto T, Okuno T. 2000.** Quinochalcones and flavonoids from fresh florets in different cultivars of *Carthamus tinctorius* L. *Bioscience, Biotechnology and Biochemistry* 64(8):1588–1599 DOI 10.1271/bbb.64.1588.
- Kim J, Assefa A, Song J, Mani V, Park S, Lee S, Lee K, Kim D, Hahn B. 2020.** Assessment of metabolic profiles in florets of carthamus species using ultra-performance liquid chromatography-mass spectrometry. *Metabolites* 10(11):440 DOI 10.3390/metabo10110440.
- Knowles P. 1969.** Centers of plant diversity and conservation of crop germplasm: safflower. *Economic Botany* 23(4):324–329 DOI 10.2307/4253083.
- Li B, Colin N. 2011.** RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* 12:323 DOI 10.1186/1471-2105-12-3231.
- Love M, Wolfgang H, Simon A. 2014.** Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12):550 DOI 10.1186/s13059-014-0550-8.
- Luzia V, Jack W, Lahoucine A, Marina A, Wang X, Richard A. 2007.** A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*. *Plant Molecular Biology* 64(5):499–518 DOI 10.1007/s11103-007-9167-6.
- Mohammadi M, Tavakoli A. 2015.** Effect of harvest time of spring safflower (*Carthamus tinctorius* L.) florets on the production of red and yellow pigments. *Quality Assurance and Safety of Crops & Foods* 7(5):581–588 DOI 10.3920/QAS2014.0427.
- Pu Z, Zhang S, Tang Y, Shi X, Tao H, Yan H, Chen J, Yue S, Chen Y, Zhu Z, Zhou G, Su S, Duan J. 2021.** Study on changes in pigment composition during the blooming period of safflower based on plant metabolomics and semi-quantitative analysis. *Journal of Separation Science* 44(22):4082–4091 DOI 10.1002/jssc.202100439.
- Ren C, Tang X, Chen C, Chen J, Pei J, Wu Y, Wu Q. 2019.** Cloning and expression analysis of a new chalcone isomerase gene during flowering in safflower. *Turkish Journal of Botany* 43(2):143–150 DOI 10.3906/bot-1809-25.
- Scossa F, Alseekh S, Fernie A. 2021.** Integrating multi-omics data for crop improvement. *Journal of Plant Physiology* 257(2):153352 DOI 10.1016/j.jplph.2020.153352.
- Wang R, Ren C, Dong S, Chen C, Xian B, Wu Q, Wang J, Pei J, Chen J. 2021a.** Integrated metabolomics and transcriptome analysis of flavonoid biosynthesis in safflower (*Carthamus tinctorius* L.) with different colors. *Frontiers in Plant Science* 12:712038 DOI 10.3389/fpls.2021.712038.
- Wang R, Shu P, Zhang C, Zhang J, Chen Y, Zhang Y, Du K, Xie Y, Li M, Ma T, Zhang Y, Li Z, Grierson D, Pirrello J, Chen K, Bouzayen M, Zhang B, Liu M. 2021b.** Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor

formation in kiwifruit (*Actinidia chinensis*). *The New Phytologist* **233**(1):373–389
DOI 10.1111/nph.17618.

- Wen W, Li D, Li X, Gao Y, Li W, Li H, Liu J, Liu H, Chen W, Luo J, Yan J. 2014. Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nature Communications* **5**(1):3438–3447 DOI 10.1038/ncomms4438.
- Xia Y, Chen W, Xiang W, Wang D, Xue B, Liu X, Xing L, Wu D, Wang S, Guo Q, Liang G. 2021. Integrated metabolic profiling and transcriptome analysis of pigment accumulation in *Lonicera japonica* flower petals during colour-transition. *BMC Plant Biology* **21**(1):98 DOI 10.1186/s12870-021-02877-y.
- Xie Z, Wang J, Wang W, Wang Y, Xu J, Li Z, Zhao X, Fu B. 2020. Integrated analysis of the transcriptome and metabolome revealed the molecular mechanisms underlying the enhanced salt tolerance of rice due to the application of exogenous melatonin. *Frontiers in Plant Science* **11**:618680 DOI 10.3389/fpls.2020.618680.
- Yang Z, Yang J, Jia Y, Tian Y, Wen A. 2009. Pharmacokinetic properties of hydroxysafflor yellow A in healthy Chinese female volunteers. *Journal of Ethnopharmacology* **124**(3):635–638 DOI 10.1016/j.jep.2009.02.026.
- Yang X, Zhang X, Yang Y, Zhang H, Zhu W, Nie W. 2021. The histone variant Sl_H2A.Z regulates carotenoid biosynthesis and gene expression during tomato fruit ripening. *Horticulture Research* **8**(1):85 DOI 10.1038/s41438-021-00520-3.
- Yue S, Tang Y, Li S, Duan J. 2013. Chemical and biological properties of quinochalcone C-glycosides from the florets of *Carthamus tinctorius*. *Molecules* **18**(12):15220–15254 DOI 10.3390/molecules181215220.
- Zhang Y, Butelli E, Martin C. 2014. Engineering anthocyanin biosynthesis in plants. *Current Opinion in Plant Biology* **19**:81–90 DOI 10.1016/j.pbi.2014.05.011.
- Zhang M, Li F, Li K, Wang Z, Wang Y, He J, Su H, Zhang Z, Chi C, Shi X, Yun C, Zhang Z, Liu Z, Zhang L, Yang D, Ma M, Qiao X, Ye M. 2020. Functional characterization and structural basis of an efficient di-C glycosyltransferase from *Glycyrrhiza glabra*. *Journal of the American Chemical Society* **142**(7):3506–3512 DOI 10.1021/jacs.9b12211.
- Zhao Q, Cui M, Levsh O, Yang D, Liu J, Li J, Hill L, Yang L, Hu Y, Weng J, Chen X, Martin C. 2018. Two CYP82D enzymes function as flavone hydroxylases in the biosynthesis of root-specific 4'-Deoxyflavones in *Scutellaria baicalensis*. *Molecular Plant* **11**(1):135–148 DOI 10.1016/j.molp.2017.08.009.
- Zhao D, Tao J. 2015. Recent advances on the development and regulation of flower color in ornamental plants. *Frontiers in Plant Science* **6**:261 DOI 10.3389/fpls.2015.00261.
- Zhou D, Qu Z, Wang H, Su Y, Wang Y, Zhang W, Wang Z, Xu Q. 2018. The effect of hydroxysafflower yellow A on coronary heart disease through Bcl-2/Bax and PPAR- γ . *Experimental and Therapeutic Medicine* **15**(1):520–526 DOI 10.3892/etm.2017.5414.
- Zhu G, Wang S, Huang Z, Zhang S, Liao Q, Zhang C, Lin T, Qin M, Peng M, Yang C, Cao X, Han X, Wang X, Knaap E, Zhang Z, Cui X, Klee H, Fernie A, Luo J, Huang S. 2018. Rewiring of the fruit metabolome in tomato breeding. *Cell* **172**(1–2):249–261 DOI 10.1016/j.cell.2017.12.019.
- Zou Y, Han C, Wang F, Tan Y, Yang S, Huang C, Xie S, Xiao X. 2021. Integrated metabolome and transcriptome analysis reveal complex molecular mechanisms underlying selenium response of *Aloe vera* L. *Journal of Plant Biology* **64**(2):135–143 DOI 10.1007/s12374-020-09285-z.