Industrial Crops & Products 134 (2019) 244–256

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

# Gene mining and identification of a flavone synthase II involved in flavones biosynthesis by transcriptomic analysis and targeted flavonoid profiling in *Chrysanthemum indicum* L.

Yanfengyang Jiang[a,1], Xiaoyu Ji[a,1], Lixin Duan[b,1], Peng Ye[a], Jinfen Yang[a], Ruoting Zhan[a], Weiwen Chen[a,*], Dongming Ma[a,*]

[a] Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine, Key Laboratory of Chinese Medicinal Resource from Lingnan (Guangzhou University of Chinese Medicine), Ministry of Education, Joint Laboratory of National Engineering Research Center for the Pharmaceutics of Traditional Chinese Medicines, Guangzhou, 510006, China

[b] International Institute for Translational Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, 510006, China

ARTICLE INFO

*Keywords:*
*Chrysanthemum indicum* L.
Flavonoid profiling
Transcriptome
Flavone synthase II
Biochemical function

ABSTRACT

*Chrysanthemum indicum* L. is a type of herb that is widely used in China, Korea, and Japan. It has been used as an ingredient in traditional medicines, tea, and functional food because of its various anti-inflammatory and anti-oxidant bioactivities. Such bioactivities have been associated with flavonoids such as apigenin, luteolin and linarin in *C. indicum*. However, the biosynthesis pathway has not been investigated. In this study, using transcriptomic analysis and targeted metabolic profiling from five different tissues, we characterize the levels of flavonoids and mine the corresponding genes involved in flavonoid biosynthesis. Transcriptomic analysis revealed that 103 unigenes are involved in flavonoid-related biosynthesis pathways. Flavone synthase (FNS) is the key enzyme responsible for flavone synthesis and provides precursors for acacetin and linarin biosynthesis. One putative FNS II gene, with the highest Reads Per Kilobase per Million mapped reads (RPKM) in flower and flower bud was cloned. Quantitative real-time polymerase chain reaction (RT-qPCR) revealed that *CiFNSII* exhibited a similar expression pattern to that in the transcriptome in terms of RPKM. In addition, a targeted metabolic profiling of three flavanones (naringenin, eriodictyol, and liquiritigenin), three flavones (apigenin, luteolin, and 7,4′-dihydroxyflavone), and two flavone derivatives (linarin and acacetin) was performed to characterize the distribution of these flavonoids in different tissues of *C. indicum*. The recombinant FNSII protein expressed in yeast was able to catalyze the conversion of three flavanones into the respective flavones. Based on the transcriptome analysis, metabolic profiling, and activity assays, a linarin biosynthesis pathway is proposed. Our study provides insight into the potential application of molecular breeding and metabolic engineering for improving the quality of cultivated *C. indicum*.

## 1. Introduction

*Chrysanthemum indicum* L. (Asteraceae) is a native plant of Asia and Northeastern Europe. In China, the dried inflorescence of *C. indicum* has been used as medicine for hundreds of years. Besides China, it also has a long history of application in Korea and Japan as an ingredient for traditional medicinal products, tea and functional foods because of its multiple bioactivities such as anti-inflammatory, anti-oxidant, anti-nociceptive, anti-bacterial, and anti-viral activities (Dong et al., 2017; Jeong et al., 2013; Kim et al., 2015; Luyen et al., 2015a; Nepali et al., 2018; Sun et al., 2016; Wu et al., 2013; Yang et al., 2017a, b; Yoo et al., 2016; Zhang et al., 2015). In 2012, the Ministry of Health of the People's Republic of China listed *C. indicum* as an item in the list of functional food, which increased its potential market demand (Zhang et al., 2007).

Many bioactive components of *C. indicum* responsible for its pharmaceutical and bioactive activities have been reported. These include flavonoids, essential oil, terpenoids and polysaccharides (Matsuda et al., 2002; Sun et al., 2015; Yoshikawa et al., 1999, 2000). Flavonoids including acacetin, linarin (acacetin-7-O-rutinoside), apigenin, luteolin were identified by HPLC analysis as the main active pharmacological components of *C. indicum* responsible for its anti-inflammatory,

* Corresponding authors.
*E-mail addresses:* chenww@gzucm.edu.cn (W. Chen), madm@gzucm.edu.cn (D. Ma).
[1] These authors contributed equally to this paper.

https://doi.org/10.1016/j.indcrop.2019.04.009
Received 8 November 2018; Received in revised form 3 April 2019; Accepted 5 April 2019

immunoregulatory and analgesic effects (Kim et al., 2016, 2013; Lee et al., 2015; Luyen et al., 2015a; Wu et al., 2013). Although major focus has been placed on identifying the bioactivities of phytochemicals in the previously-mentioned studies, genes responsible for the synthesis of the bioactive components in *C. indicum* remains unidentified.

Flavonoids, a group of valuable secondary metabolites, are extensively distributed among plants and more than 10,000 flavonoids described to date (Mathesius, 2018). These phenolic compounds play important roles in plants, such as protecting tissues from UV damage, flower coloration and signals for interspecific interactions (Dixon and Pasinetti, 2010; Mathesius, 2018; Xiao et al., 2016). All flavonoids are synthetized from the precursor naringenin (Falcone Ferreyra et al., 2012). The synthesis of flavones, one of the most abundant flavonoid subclass, is catalyzed by flavone synthase and is a rate-limiting step (Jiang et al., 2016). Based on their chemical structures, bioactive flavonoids in *C. indicum*, namely acacetin, linarin (acacetin-7-O-rutinoside), apigenin, luteolin, are classified as flavones or flavones derivatives.

FNS is responsible for the conversion of flavonones (e.g., naringenin, eriodictyol, liquiritigenin etc.) to flavones (e.g., apigenin, luteolin, 7,4′-dihydroxyflavone, etc.) (Wu et al., 2016). FNS is encoded by two gene families, namely flavone synthase I (FNSI) and flavone synthase II (FNSII) (Britsch et al., 1981; Kochs and Grisebach, 1987; Stotz and Forkmann, 1981). FNSI has been reported in rice, maize, Arabidopsis and in the Apiaceae genus (Ferreyra et al., 2015; Kim et al., 2008; Wang et al., 2018b), while the presence of FNSII is extensive in plants. FNSII enzymes are cytochrome P450 monooxygenases (CYP450s), FNSII enzymes belonging to the CYP93B subfamily are present in dicot plants, while those belonging to the CYP93 G subfamily are found in monocot plants (Akashi et al., 1998; Lam et al., 2014; Zhang et al., 2007). FNS enzymes have been characterized in *Lonicera japonica*, *Lonicera macranthoides*, and *Glycine max* to convert flavonones (e.g., naringenin, eriodictyol, Liquiritigenin, etc.) to flavones (e.g., apigenin, luteolin, 7,4′-dihydroxyflavone, etc.) (Fliegmann et al., 2010; Wu et al., 2016). However, little is known about FNS genes from *C. indicum* plants. This study aims to determine whether FNS genes encode FNSI or FNSII type enzymes, and to investigate the enzymatic reactions catalyzed by FNS enzymes from *C. indicum*.

In this study, transcriptomic sequencing and analysis of five tissue types of *C. indicum* were performed to uncover candidate genes involved in flavonoid synthesis pathway. Of the five candidate flavone synthases, one unigene with the higher RPKM in flower and flower bud was cloned and transcript levels of the unigene in different tissues were determined by RT-qPCR. The recombinant protein expressed in the WAT11 yeast strain was able to convert flavanones (naringenin, eriodictyol, and liquiritigenin) to the respective flavones (apigenin, luteolin, and 7,4′-dihydroxyflavone) in in vitro activity assays. Combined with the fact that liquiritigenin and 7,4′-dihydroxyflavone were not detected in any tissues of *C. indicum*, we provide evidence that CiFNSII might be an active flavone synthase enzyme involved in the conversion of naringenin and eriodictyol to apigenin and luteolin, respectively. These results provide a good basis for pathway elucidation, molecular breeding, and metabolic engineering for improving the quality of *C. indicum*.

## 2. Materials and methods

### 2.1. Materials

#### 2.1.1. Chemical sources

Naringenin (NAR), eriodictyol (ERI), liquiritigenin (LIQ), luteolin (LUT), apigenin (API) were purchased from Sigma (Sigma, USA). 7,4′-dihydroxyflavone (DHF), linarin and acacetin were purchased from Shanghai yuanye Bio-Technology Co., Ltd, China.

#### 2.1.2. Plant materials

*Chrysanthemum indicum* L. was planted in the nursery field of Guangzhou University of Chinese Medicine (Guangzhou City, Guangdong Province, China) in natural environment. The root, stem, leaf, flower and flower bud from healthy plants were collected from 2-year-old healthy plants in November and frozen at −80 ℃.

### 2.2. RNA extraction, reverse transcription, and sequencing

Total RNAs from the root, stem, leaf, flower, flower bud of *C. indicum* were isolated with Magen HiPure Plant RNA Mini Kit (Magen, China). We used a DNase I (GeneMark, China) treatment before cDNA synthesis to remove gDNA. After extraction, the degradation and contamination of RNA were checked in 2% agarose gels. The concentration of RNA was evaluated using an Agilent 2100 169 Bioanalyzer. RNA with $OD_{260}/OD_{280}$ at 1.8–2.2 was used for further analysis. The RNA was converted into first-strand cDNA using SuperScript III (TransGen, China) and was ligated to Illumina sequencing adapters. Paired-end sequencing was done using an Illumina HiSeq™ 2000 system by Gene Denovo Biotechnology Co. (Guangzhou, China).

### 2.3. Sequence de novo assembly and gene annotation

The de novo assembly of RNA-Seq was performed using the Trinity software (Haas et al., 2013). The expression level of the assembled contigs was examined by Bowtie version 0.12.8, and normalization using RPKM values were calculated by a program containing the RSEM algorithm (https://github.com/deweylab/RSEM), distributed by the Trinity group. The sequencing data were assessed by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), per sequence quality scores, per sequence GC content and sequence duplication levels.

### 2.4. Differentially expressed gene analysis and screening of the unigenes involved in flavonoid biosynthesis

The analysis of gene expression levels in five different tissues of *C. indicum* was performed using the edgeR program (http://www.bioconductor.org/packages/release/bioc/html/edgeR.html). The identified differentially expressed genes (DEGs) were analyzed based on a FDR (false discovery rate) of $\leq 0.05$ and fold change of $\geq 2$ to determine significant differences between the five tissues. The gene sequences were subjected to KEGG pathway analysis by using the blastx program (E value threshold set at $10^{-5}$). After annotation by the KEGG database, MEGA7 (Kumar et al., 2016) and iTOL (http://itol.embl.de/) were used for multiple sequence alignments and for preparing phylogenetic analysis.

### 2.5. Cloning of CiFNSII, sequence alignment and phylogenetic analysis

A candidate FNS gene, *CiFNSII*, was chosen for cloning. The full-length open reading frame (ORF) for *CiFNSII* was amplified using PrimerSTAR max DNA polymerase (Takara, Japan) according to the manufacturer's protocol. PCR conditions were as followed: 98 ℃, 1 min; 98 ℃, 10 s, 55 ℃, 15 s, 72 ℃, 10 s, 30 cycles; 72 ℃, 5 min. The sequences of gene-specific primers *CiFNSII*-F and *CiFNSII*-R were listed in Table S1. PCR product was cloned into pENTR/D vector (Invitrogen, USA), and the Trans1-T1 Phage resistant chemically *Escherichia coli* competent cells (TransGen, China) were consequently transformed with the recombinant plasmid pENTR/D-*CiFNSII*. Subsequently, DNA sequencing was performed. Motif analysis and sequence alignment were performed after confirming the sequence using DNAMAN version 7.0. The phylogenetic analysis was conducted by the maximum likelihood method in MEGA 7 (Kumar et al., 2016).

### 2.6. RT-qPCR Experiments

RNA isolation and cDNA synthesis method were described previously. RT-qPCR was carried out using EvaGreen 2× qPCR MasterMix (ABM, Canada) in Bio-Rad CFX96 Touch™ Real-Time PCR Detection System (Bio-rad, USA) according to the manufacturer's instructions. *GAPDH* (KC508619) and *EF1α* (KF305681) were used as reference genes for normalization (Gu et al., 2011; Shen et al., 2010). The relative expression levels of *CiFNSII* in different tissues were calculated using the $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak, 2008), $\Delta\Delta CT = (Ct_{target}\text{-}CT_{ref})_{target}\text{-}(Ct_{target}\text{-}CT_{ref})_{ref}$. Each reference gene was calculated separately, and the geometric mean (Gharbi et al., 2015) of the two genes was used as the standardization result of the double reference genes. The primer sequences designed by Primer Premier 5.0 which used for RT-qPCR were shown in Table S2. The analysis of each tissue type was performed in biological replicates of three and technical replicates of two.

### 2.7. Construction of expression vector and Western blot analysis

Infusion cloning (Clontech, Japan) system was used to construct yeast expression vector in this study. The pENTR/D vector with the *CiFNSII* construct was used as templates to amplify the *CiFNSII* fragment. *Eco*RI and *Spe*I were chosen as digestion sites for the linearization of expression vector of pESC-URA linearization. Then, we designed PCR amplification using internal primers pUF-*CiFNSII* and pUR-*CiFNSII*, as shown in Table S1, with 15 bp extensions that are complementary to the ends of the linearized pESC-URA vector. In addition, we added six histidines between core sequence and vector extensions as his tag for western blot detection in the reverse primer. PCR procedures were as described in Section 2.5. The PCR reaction mixture was incubated for 15 min at 50 ℃. Subsequently, *Saccharomyces cerevisiae* WAT 11 yeast strain (Urban et al., 1997) was transformed with pESC-URA-*CiFNSII* using a standard lithium acetate protocol (Gietz and Woods, 2002).

pESC-URA contains an URA3 marker for galactose-regulated expression; this system can be induced by galactose and is repressed by glucose. After incubation for 24 h in SC-U culture medium containing 2% glucose, transformant WAT 11 was isolated and resuspended in SC-U medium without sugar donor for depletion of glucose. Then the yeast cells were collected and resuspended in SC-U medium with 2% galactose for further induction of the recombinant protein expression. An equivalent of two OD600 cells from duplicate of the inductive culture were harvested at seven time points (0 h, 4 h, 8 h, 16 h, 24 h, 32 h, 48 h) over a period of 48 h (Zheng et al., 2015). Harvested cells were resuspended in 800 μL of dd $H_2O$, 200 μL of lysis buffer (2 M NaOH, 7.5% β-mercaptoethanol) was added, and the mix was incubated at 30 ℃ for 5 min. Cell suspensions were shaken for 15 min with a 0.2-g glass bead (diameter 300 μM) and added to 150 μL of 55% trichloroacetic acid before being spun at 12,000 ×*g*; the total protein of the yeats cells was subsequently extracted. *His* mouse monoclonal antibody (Transgene, China) was used to identify target proteins with a *His* epitope by Western blotting (Mahmood and Yang, 2012).

### 2.8. In vitro CiFNSII activity assays

The yeast strain WAT11 transformed with pESC-URA-*CiFNSII* was precultured as previously described and harvested after growth for 8 h, which is the time with the highest cell concentration. The cells were resuspended in pH 7.5 Tris-EDTA buffer (10 mM Tris, 1 mM EDTA) and disrupted using a high-pressure homogenizer with 20,000 psi at 4 ℃. The microsomes were precipitated with $CaCl_2$ at a final concentration of 18 mM (Liu et al., 2018). The resulting mixture turned into emulsion after 15-min of incubation on ice, which means that the microsomal fractions of the yeast cells were precipitated by $CaCl_2$. Microsomes were harvested through centrifugation at 4480 ×*g* and resuspended in 100 mM potassium phosphate buffer (pH = 7.6). The enzyme assay for CiFNSII was conducted with a 1-mL final volume containing 100 mM potassium phosphate buffer (pH = 7.6), 5 mM NADPH, 2 mM DTT, 30 μg/mL substrate (naringenin, eridictyol, or liquiritigenin), and 100 μg of total proteins. Protein concentration was determined by Bradford reagent (Tiangen, China). After incubation for 30 min at 30 ℃, each reaction was terminated by two extractions with ethyl acetate (1 mL), vacuum dried, and redissolved in HPLC-grade methanol (500 μL) for HPLC-MS analysis.

### 2.9. The determination of flavonoid contents and HPLC-MS analysis

The root, stem, leaf, flower and flower bud of *C. indicum* were ground into powder in liquid nitrogen. In total, 500 mg of powder was precisely weighed and transferred to a conical flask. 40 mL of methanol was added, and the mixture was extracted ultrasonically for 45 min at 55 ℃. Extracts were concentrated with a hypobaric drying method. 1 mL HPLC-grade methanol was used to re-dissolve the residues to obtain the extraction solution for HPLC-MS analysis. The identity of flavonoids detected in this study, including linarin, acacetin, luteolin, eriodictyol, liquiritigenin, apigenin, naringenin, and 7,4′-dihydroxyflavone, were confirmed by HPLC-MS analysis using authentic standards. The quantification of these flavonoids were performed based on by their corresponding standard curves. Each tissue type was analyzed in biological replicates of three and technical replicates of three in this experiment.

The HPLC-MS analysis was performed as follows: an Agilent 1290 Series UHPLC system was coupled online with a hybrid quadrupole time of flight (Q-TOF) mass spectrometer (6540B, Agilent Technologies, Inc., CA, USA). 2 μL of a filtered sample was applied to a reversed-phase column (SB-C18 RRHD, 2.1 × 100 mm, 1.8 μm; Agilent) with an in-line filter (1290 infinity in-line filter; Agilent). The system was operated in the negative ion mode at the flow rate of 0.4 mL/min using solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid). The gradient started from 28% B for 15 min, followed by 28% B to 100% B in 1 min, held for 4 min, and a post run time of 4 min for re-equilibration. Data were collected in the negative ESI mode separate runs on a Q-TOF (Agilent 6540B) operated in full scan mode from 100 to 1100 *m/z*. The capillary voltage was 3500 V with a scan rate of 3 scans per second; the nebulizer gas flow rate was 11 L/min; drying gas flow was 8 L/min; gas temperature was 300 ℃, and the skimmer voltage was 65 V.

## 3. Results

### 3.1. The de novo assembled transcriptome of C. indicum and annotation

To gain insight into flavone biosynthesis in *C. indicum*, next generation sequencing was performed on RNA isolated root, stem, leaf, flower, and flower bud of *C. indicum*. After performing high-throughput sequencing using the Illumina HiSeq 2000 platform, we obtained 97.69%–98.23% high quality reads from the raw reads in five different tissues, and the GC percentage ranged from 42.85% to 43.27% (Table S3). After removing the redundant sequences, 89,206 unigenes were generated with an N50 of 1138 and an average length of 714 bp (Table S4). The length distribution and reads coverage are shown in Fig. S1. It can be seen that the unigenes with length ≥500 nt account over 50% of the total unigenes, which leads to high quality of sequence assembly. The RPKM values were applied to normalize and evaluate the expression level of these assembled unigenes (Wagner et al., 2012). A total of 54,171 unigenes were annotated with a cut-off E-value of $10^{-5}$ on four databases including 53,751 (60.25%) on NCBI non-redundant (nr), 37,908 (42.49%) on Swiss-Prot, 31,973 (35.84%) on KOG (EuKaryotic Orthologous Groups) and 20,604 (23.09%) on KEGG (Kyoto Encyclopaedia of Genes and Genomes), respectively (Table S5). A venn graph of the four database annotation results is shown in Fig. S2a, indicated that 15,941 of them being annotated by the Nr, Swiss-prot, KOG

and KEGG databases simultaneously, which is approximately 29.4% of all the mapped unigenes. Based on the KOG annotation (Fig. S2b), most unigenes (10,155) were annotated as 'General function prediction only'. Additionally, 1761 unigenes were annotated as 'Secondary metabolites biosynthesis, transport and catabolism classification'. In the GO (Gene Ontology) similarity analyses shown in Fig. S3, 7473 unigenes were annotated into 3 groups: Biological Process, Cellular Component and Molecular Function. In 'Biological Process', 4332 unigenes were grouped as metabolic process, while 3615 unigenes were assigned to cellular process. Single-organism process ranked third, with 2863 unigenes. In 'Cellular component', most of the unigenes were classified to cell and cell part. For 'Molecular Function' terms, it was clearly shown that most of the unigenes were grouped as catalytic activity and binding.

### 3.2. Identification of differentially expressed genes

To determine differentially expressed genes (DEGs) from five different tissues of *C. indicum*, $|\log2FC| > 1$and FDR $\leq 0.05$ was applied as the screening thresholds. The number of DEGs in 'Leaf vs Flower bud', 'Leaf vs Flower', 'Leaf vs Stem' and 'Leaf vs Root' were 39,335, 36,589, 23,802, and 31488, respectively, with up-regulated DEGs of 22,165, 18,406, 13,170 and 20,353, and down-regulated DEGs of 17,170, 18,183 10,632 and 11,135, respectively (Fig. S4). Based on this analysis, in comparison with stems and roots, more differences in gene expression exist in flower and flower bud when compared with leaf, indicating that the number of genes under active transcription in flower buds and flowers is more abundant than in other tissues.

### 3.3. Annotated genes involved in flavonoid biosynthesis

KEGG database provides us with a basic platform for systematic study of gene function acted in the metabolic networks of gene products (Kanehisa et al., 2012). Of the 11,290 unigenes enriched in KEGG database, 103 of these belonged to the flavonoid biosynthesis related pathways (Table 1). On the other hand, 1408 unigenes were annotated by KEGG database as belonging to the secondary metabolite pathways (Fig. 1). The expression patterns of the unigenes related to flavonoid biosynthesis (Table S6) in the five different tissues (Fig. 2a) were generated by the heatmap package of the R software (Fig. 2b). In total, five unigenes were annotated and classified as part of the CYP93 family. Three unigenes were classified as part of the CYP93 A subfamily, and two unigenes—namely 0028351 and 0043683—were classified as part of the CYP93B subfamily (Table S7). Some members of the CYP93B subfamily have been previously reported as FNSII and favanone 2-hydroxylase (F2H) (Akashi et al., 1998, 1999; Martens and Forkmann, 1999; Yan et al., 2014; Zhang et al., 2007). The expression level of unigene0043683 in the flower and flower bud was much higher not only than that of unigene0028351 but also than that of other tissue types (Table S7 and Fig. 2b). Furthermore, analysis using blastx (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) programs revealed that only the unigene0043683 contains a complete ORF. Therefore, unigene0043683 was selected as a candidate gene for further cloning and biochemical characterization.

**Table 1**
Distribution of unigenes involved in flavonoids related pathways.

| Pathway | All genes with pathway annotation (11290) | Pathway ID |
|---|---|---|
| Flavonoid biosynthesis | 87 (0.77%) | ko00941 |
| Anthocyanin biosynthesis | 10 (0.09%) | ko00942 |
| Isoflavonoid biosynthesis | 2 (0.02%) | ko00943 |
| Flavone and flavonol biosynthesis | 4 (0.04%) | ko00944 |

### 3.4. Isolation of full-length ORF for CiFNSII and sequence analysis

Based on sequence analysis, the unigene0043683 ORF encodes a protein with 514 amino acids and a molecular mass of 58.6 kDa. For the comparison of characteristic sequence signatures of FNSIIs, we aligned the amino acid sequences of unigene0043683 with 4 characterized FNSII enzymes: *LjFNSII-1.1 and LjFNSII-2.1* from *Lonicera japonica* (AMQ91109.1,AMQ91111.1), *LmFNSII-1.1* from *Lonicera macranthoide*s (AMQ91113.1), and CYP93B2 from *Gerbera hybrida* (AAD39549.1) by DNAMAN version 7 (Fig. 3a). It shares the highest identity (84.47%) with CYP93B2 at the amino acid level. Consequently, we named unigene0043683 as *CiFNSII* (GenBank MK419957). We found that four cytochrome P450-featured conserved motifs are present in the *CiFNSII* amino acid sequence using the 'Conserved Domains' search at the NCBI website (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The highly conserved proline-rich hinge region typically obeys the consensus LPPPPXXXP motif, and is supposed to serve as a 'hinge' that is required for proper orientation of P450 enzymes (Chapple, 1998). Furthermore, the oxygen binding pocket motif AATDTT is found to be essential for oxygen binding and activation (Hasemann et al., 1995). We also found an E-R-R triad residue that encodes a pocket locking motif that stabilizes the core structure of the meander region and the 'folding in' of the heme (Chapple, 1998; Hasemann et al., 1995). The heme-binding domain contains the signature sequence PFGXGRRXCPG is often highly conserved in all P450 polypeptides which gives these enzymes the carbon monoxide-binding ability (Chapple, 1998). A phylogenetic analysis was conducted by using the deduced amino acid sequence of *CiFNSII* and 11 FNSII proteins from 9 different plant species (Table S8) to investigate the relationship between them (Fig. 3b). Some FNSII enzymes have been reported to catalyze the direct conversion of flavanones to flavones by introducing a double bond between the C-2 and C-3 residues in flavanones, such as CYP93B2 from *G. hybrida* (Martens and Forkmann, 1999), CYP93B3 from *Antirrhinum majus*, and CYP93B4 from *Torrentia* (Akashi et al., 1999). However, F2H CYP450 enzymes catalyze the hydroxylation of flavanones generating 2-hydroxyflavanones; for example, in the Fabaceae family, it was demonstrated that CYP93B1 from *Glycyrrhiza echinata* and CYP93B10/11 from *Medicago truncatula* have F2H activity (Akashi et al., 1998; Zhang et al., 2007). *CiFNSII* is clustered closely with CYP93B2 (Martens and Forkmann, 1999) in the same clade as FNSII (Ayabe and Akashi, 2006), CYP93B1 and CYP93B10 (Akashi et al.,3 1998; Zhang et al., 2007) are involved in 2-hydroxyflavanone biosynthesis and clustered together and separated from the FNSIIs.

### 3.5. Expression analysis of the CiFNSII in different tissues of C. indicum

To determine the expression pattern of *CiFNSII*, RT-qPCR was performed on five different tissues of *C. indicum.* including root, stem, leaf, flower, and flower bud, with *GAPDH* and *EF1α* as reference genes. Higher transcript levels were detected in flower and flower bud than in other tissues (Fig. 4a–c). The expression pattern observed for *CiFNSII* is almost in agreement with the transcriptome data represented by RPKM value, which indicates that CiFNSII might be the most active flavone synthase in flower and flower bud tissues in *C. indicum* (Fig. 4d).

### 3.6. CiFNSII heterologous protein expression and in vitro enzyme activity assays

To gain insight into the functions of CiFNSII, cDNA was obtained from *C. indicum*. A pair of primers were designed to amplify the ORF of *CiFNSII* (Table S1). To investigate the catalytic activity of the enzyme encoded by the isolated putative *CiFNSII* gene, the ORF region was inserted into the pESC-URA vector and transformed into the WAT11 yeast strain. WAT11 cells expresses the *Arabidopsis* NADPH-cytochrome P450 reductase, which provides the reducing equivalents necessary for the activity of plant CYP450 s, including FNSII enzymes. Western blot

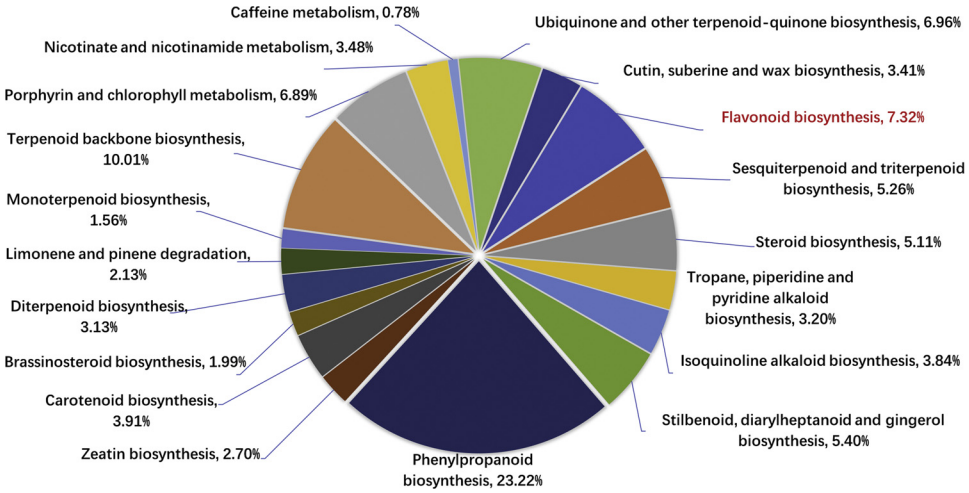


**Fig. 1.** Distribution of unigenes in different secondary metabolism categories based on KEGG classifications. Numbers follow pathway description in brackets indicate the relative percentage.

assay was performed to explore the protein expression level of CiFNSII in yeast. CiFNSII-expressing yeast cells were cultured and sampled at six different time points during a 48-h period, after induction with Galactose. CiFNSII was successfully expressed during the 48-h induction period, with a maximum expression level at 8 h, as shown by Western blot analysis (Fig. S5). Consequently, induction with galactose for 8 h was selected for the collection of cells used for microsomes preparation.

Microsomes were prepared from harvested cells and tested in FNSII enzyme assays with three flavanone substrates, namely naringenin, eridictyol, and liquiritigenin. Reaction products were analyzed by HPLC-MS analysis. A new candidate product peak was detected from the products of the enzyme assays.

When narigenin was used as a substrate, the product showed a molecular ion of $m/z$ 269 in the negative ion mode, which is close to naringenin's molecular ion of $m/z$ 271 (Fig. 5a). This product exhibits

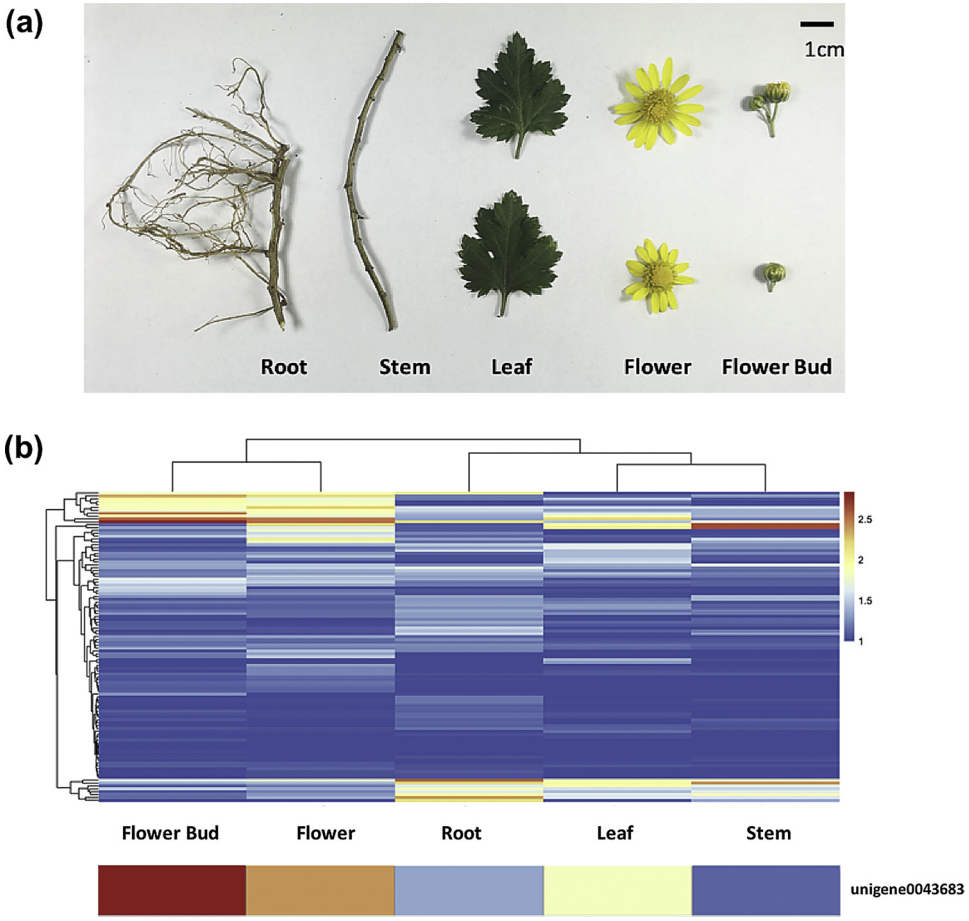


**Fig. 2.** (a) Representative images of *C. indicum* tissues used for transcriptome and metabolic profiling analysis. (b) Expression pattern analysis of unigenes involved in flavonoid-related biosynthesis pathway.

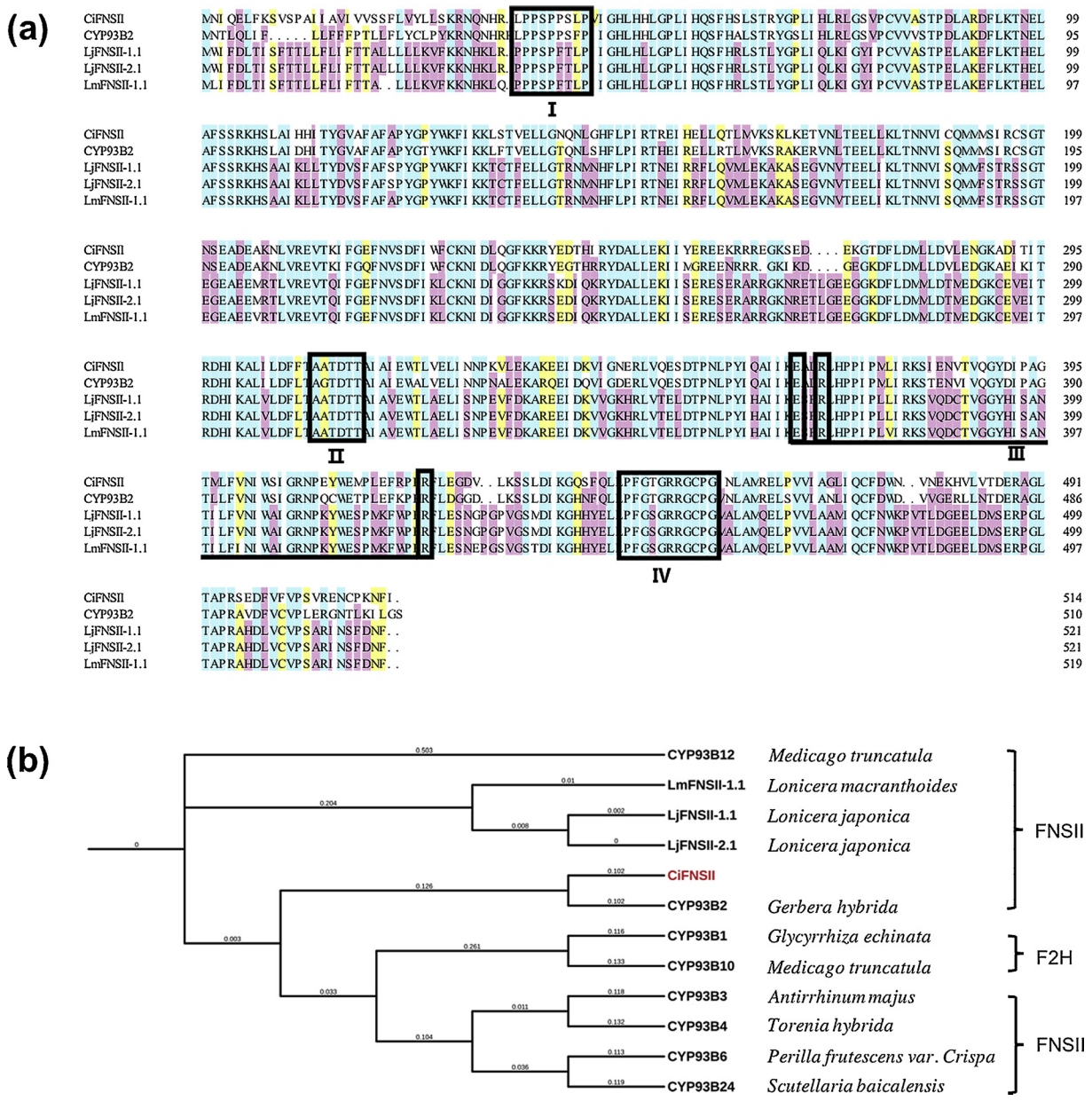


**Fig. 3.** (a) Multiple sequences alignment of the *CiFNSII* protein and four other flavone synthases, *Gerbera hybrida* CYP93B2 (AAD39549.1), *Lonicera japonica* LjFNSII-1.1&LjFNSII-2.1 (AMQ91109.1,AMQ91111.1), *Lonicera macranthoides* LmFNSII-1.1 (AMQ91113.1). Alignment was carried out with gap penalty of 7 and K-tupe of 2 by DNAMAN version 7. Light-blue, yellow, and pink coloration reflect 100%, 75% and 50% amino acid residues conservation. CYP450-specific conserved motifs including proline-rich membrane hinge LPPPPXXXP (I), oxygen binding pocket motif AATDTT (II), E-R-R triade motif (III), and heme-binding domain PFGXGRR-XCPG (IV) are shown with black frame. (b) Maximum likelihood tree illustrating the phylogenetic relationship of *CiFNSII* with flavone synthases from several plant species. Phylogenetic analysis was performed by MEGA 7.0 and test of phylogeny was carried out by bootstrap method with 500 replications. Substitution model was built by poisson correction data subset to use pairwise deletion. *CiFNSII* is marked in red.

the same retention time and mass fragmentation pattern as an authentic standard of apigenin. Therefore, we propose that the reaction product is apigenin, which has a *m/z* ratio of 2 less than that of naringenin. Similarly, eriodictyol (*m/z* 287) was converted luteolin (*m/z* 285) following the enzyme reaction (Fig. 5b), while liquiritigenin (*m/z* 255) was converted to 7,4′-dihydroxyflavone (*m/z* 253) (Fig. 5c). The MS/MS fragmentation profiles were shown in Figure S6-S11. Therefore, CiFNSII-expressing microsomes appear to metabolize eriodictyol, naringenin, and liquiritigenin to luteolin, apigenin, and 7,4′-dihydroxyflavone, respectively. In our test, moreover, 2-hydroxynaringenin (*m/z* 287) and 2-hydroxyliquiritigenin (*m/z* 271) could be identified through the extracted ion peak (Fig. 5a,c), although the respective standard samples were inaccessible to us. This supports the hypothesis that *CiFNSII* encodes a flavone synthase, catalyzing the conversion of flavanones to flavones, probably through 2-hydroxyflavanone intermediates. Thus, a first step seem to involve the hydroxylation of flavanones, followed by a second step of dehydration, as both 2-hydroxyflavanones and flavones could be identified without acid treatment.

### 3.7. Profiling of Flavonoid Metabolites in different tissues of C. indicum

The biosynthesis of flavones (apigenin, luteolin and 7,4′-dihydroxyflavone) from flavanones (naringenin, eridictyol, and liquiritigenin, respectively) is catalyzed by FNSII enzymes (Fliegmann et al., 2010; Wu et al., 2016). Flavone derivatives (linarin and acacetin) are important

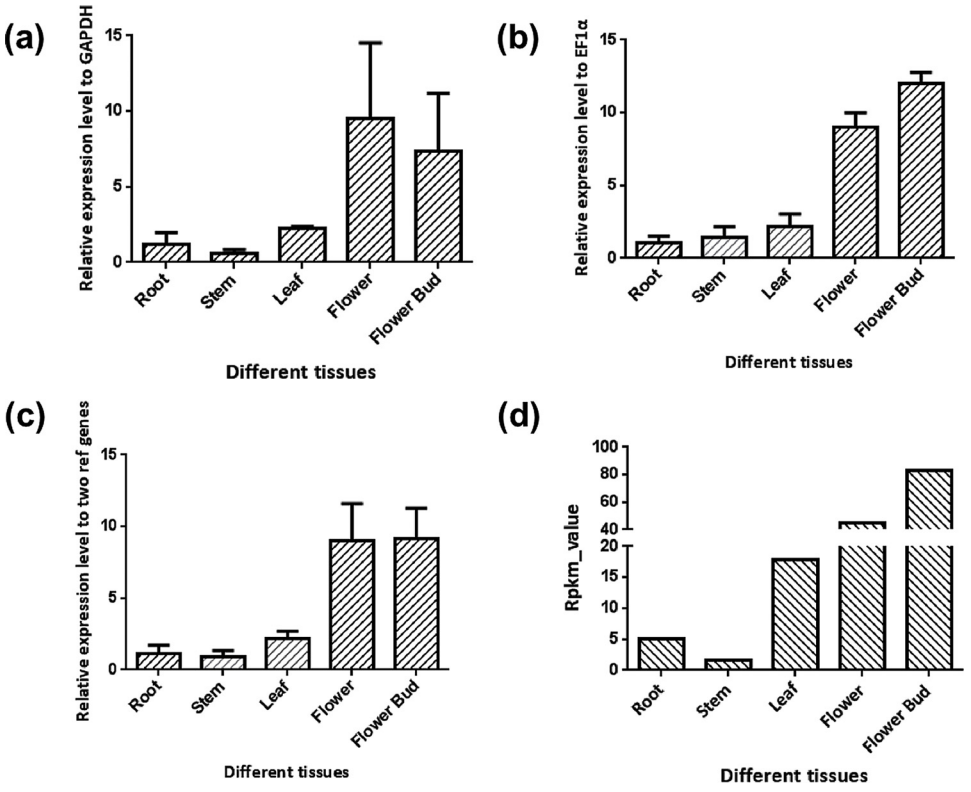


**Fig. 4.** (a,b,c) RT-qPCR analysis of *CiFNSII* expression in different tissues. The relative mRNA expression levels were calculated using the $2^{-\Delta\Delta CT}$ method and were normalized to *GAPDH*, *EF1α* and the double references of *GAPDH* and *EF1α*. Each tissue type was performed in three biological replicates. Error bars represent the standard deviations between biological replicates. (d) RPKM values of *CiFNSII* expression in different tissues from the transcriptome data.

pharmaceutical integredients. To explore the distribution of FNSII-related flavones and flavones derivatives such as linarin and acacetin accumulated in *C. indicum*, we gathered methanolic extracts from root, stem, leaf, flower, and flower bud for HPLC-MS analysis. The results are shown in Fig. 6, Fig. S12 and Table 2. From eight targeted compounds (naringenin, apigenin, eridictyol, luteolin, liquiritigenin, 7,4′-dihydroxyflavone, linarin, and acacetin), 6 flavonoids except liquiritigenin and 7,4′-dihydroxyflavone were found among the different tissues, and the composition was very different from each other. The phytochemical results showed that some common constituents, such as linarin, luteolin, and apigenin, were present in all the tissues. Flower represent the organ with the most abundant flavonoid variety, i.e., six flavonoids detected in our experimentt. Notably, acacetin was not detected in flower bud. The total contents of flavonoids detected in flower and flower bud were higher than those of the other organs such as root, stem and leaf. It was obvious that luteolin and apigenin were the most abundant metabolites among all the tissues, and they also exhibited much higher concentration in flower and flower bud sections as compared with other tissues. The flower had the highest luteolin (27.635 μg/mg FW) and apigenin (2.353 μg/mg FW) levels. As luteolin and apigenin are produced by CiFNSII, we hypothesize that the higher contents of luteolin and apigenin are possibly related to higher FNSII transcript levels in *C. indicum*. This hypothesis is consistent with our transcriptomic data. It is interesting that the root but not flower has the highest level of linarin (0.364 μg/mg FW), which are about 6-fold higher than those in flower.

## 4. Discussion

In recent years, the integration of both gene transcription and metabolism analyses has been widely used to reveal the biosynthesis pathways of the bioactive components, such as flavonoids, caffeoylquinic acids, and bornyl acetate in medicinal plants (Wang et al., 2018a; Yang et al., 2018). In the present study, bioinformatics analysis was carried out based on the acquired unigenes data from five different tissues of *C. indicum*. The genes of the main flavonoid pathways were identified by KEGG, covering the majority of enzymes in key processes such as flavonoid biosynthesis.

Based on data from our transcriptomic and enzyme activity assays, a putative linarin pathway was proposed. As illustrated in Fig. 7, enzymes involved in the proposed pathway encompass phenylalanine ammonia-lyase (PAL), trans-cinnamate 4-monooxygenase (C4H), 4-coumarate-CoA ligase (4CL), chalcone synthase (CHS), chalcone reductase (CHR), chalcone isomerase (CHI), FNSII, flavonoid 3′-hydroxylase (F3′H), methyltransferase (MT), and UDP-glycosyltransferase (UGT). Eight candidate unigenes of methyltransferases were found as putative ones involved in the catalysis of apigenin to acacetin. Acacetin and linarin (acacetin-β-rutinoside) have been detected in *C. indicum* and *Chrysanthemum boreale* (Gao et al., 2008; Nugroho et al., 2013). Furthermore, acacetin-7-O-glucoside has been observed in *Chrysanthemum morifolium* cv. Hangju and *Chrysanthemum morifolium* cv. Kotobuki (Nishina et al., 2019; Zhang et al., 2019a). Therefore, UGTs might be involved in acacetin-7-O-glucoside biosynthesis, and this glucoside derivate might be further rhamnosylated by another UGT (or the same UGT) to form linarin. In addition, apigenin-7-O-glucoside and luteolin-7-O-glucoside have been identified in *C. indicum* (Feng et al., 2010; Gao et al., 2008; Wu et al., 2013). Therefore, UGTs, not yet identified, are probably involved in flavone glycosides biosynthetic pathways. Therefore, the 53 flavonoid O-glucosyltransferase unigenes reported in our transcriptome have to be functionally characterized to demonstrate their involvement in acacetin-7-O-glucoside, linarin, apigenin-7-O-glucoside, or luteolin-7-O-glucoside biosynthesis.

Flavone synthase is a rate-limiting enzyme that catalyzes the biosynthesis of flavones (Martens and Mithofer, 2005). The metabolic flux from flavanones to the biosynthesis of flavones controlled by rice FNSII has been demonstrated (Lam et al., 2014). Given the importance of FNS in flavonoid biosynthesis, RT-qPCR of a unigene 0043683 encoding FNSII was performed with the highest transcript levels observed in flower and flower bud. This result is consistent with the expression pattern of flavonoid transcriptome (Fig. 4). In other words, transcriptomic analysis has provided insights for elucidating the molecular signatures of flavonoid accumulation in *C. indicum*.

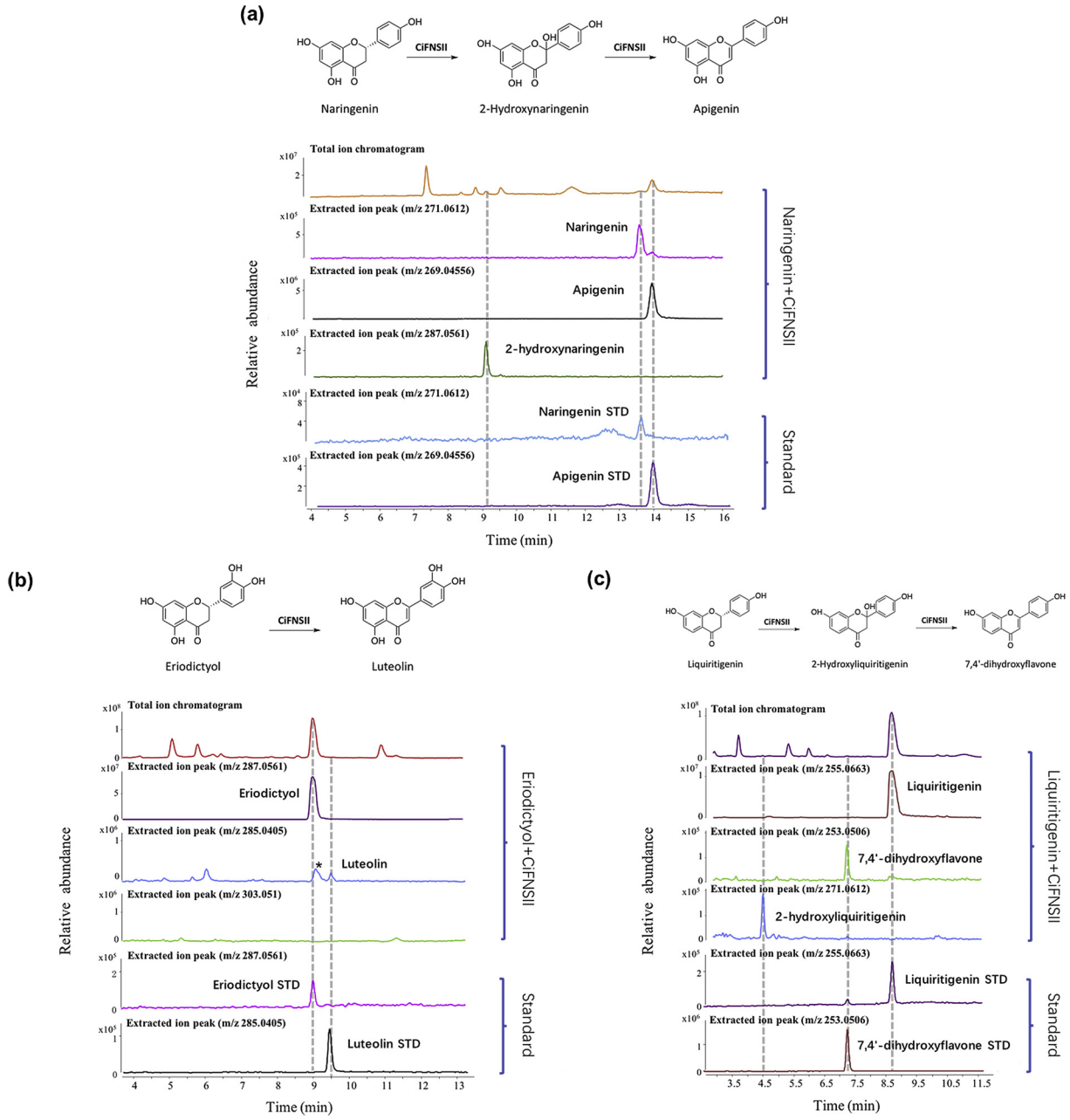


**Fig. 5.** Analysis of reaction products generated by CiFNSII recombinant protein by HPLC-MS. Chromatographic profiles show the conversion of (a) naringenin (NAR) to apigenin (API), (b) eriodictyol (ERI) to luteolin (LUT), and (c) liquiritigenin (LIQ) to and 7,4′-dihydroxyflavone (DHF). * represents a contamination peak.

Two FNS enzymes, FNSI ans FNSII, have been shown to be responsible for the conversion of flavanones to flavones by catalyzing a double bond formation between C2 and C3 of flavanones. At present, most identified FNSII enzymes are classified as CYP450 s, and are widely distributed in dicot plants. On the other hand, FNSI has been reported in *Arabidopsis thaliana*, rice (*Oryza sativa*) and several species of Apiaceae (Ferreyra et al., 2015; Jiang et al., 2016; Wang et al., 2018b; Wu et al., 2016; Zhao et al., 2016). In our study, a *C. indicum* FNS with a higher expression in flower and flower bud was clustered with FNSIIs from the other plants (Fig. 3b). *C. indicum* FNS is closely related with FNSIIs and distantly related with F2Hs. Thus, we predicted that FNS in *C. indicum* is a FNS II, and named it as *CiFNSII*.

For enzyme assays, CiFNSII was expressed in WAT11 yeast cells, and the recombinant protein was demonstrated to convert flavanones (naringenin, eriodictyol, and liquiritigenin) into the corresponding flavones (apigenin, luteolin, and 7,4′-dihydroxyflavone) in vitro (Fig. 5). When low amounts of substrates and total proteins were used for activity assays (10 μg and 50 μg, respectively), negative results were achieved (Fig. S13). The CYP93B1 (F2H) enzyme from *G. echinata* has been reported to catalyze the NADPH-dependent hydroxylation of naringenin and liquiritigenin to the corresponding 2−OH flavanones. Furthermore, eriodictyol is converted into luteolin by microsome of yeast expressing CYP93B1 only after dehydration by acid treatment (Akashi et al., 1998). Currently, it have been reported that CYP450 *AFNS2* and *TFNS5* from snapdragon and torenia, respectively, are FNSII enzymes whose reaction mechanisms for flavone biosynthesis involve 2-hydroxyflavanone intermediates (Akashi et al., 1999). The second type of reaction mechanism of FNSII introducing a C2 = C3 double bond without 2-hydroxyflavanones formation has been observed in *Gerbera hydrida*, *G. max*, *L. japonica*, *and L. macranthoides* (Fliegmann et al., 2010; Martens and Forkmann, 1999; Wu et al., 2016). In our assays, 2-hydroxynaringenin (Fig. 5a) and 2-hydroxyliquiritigenin

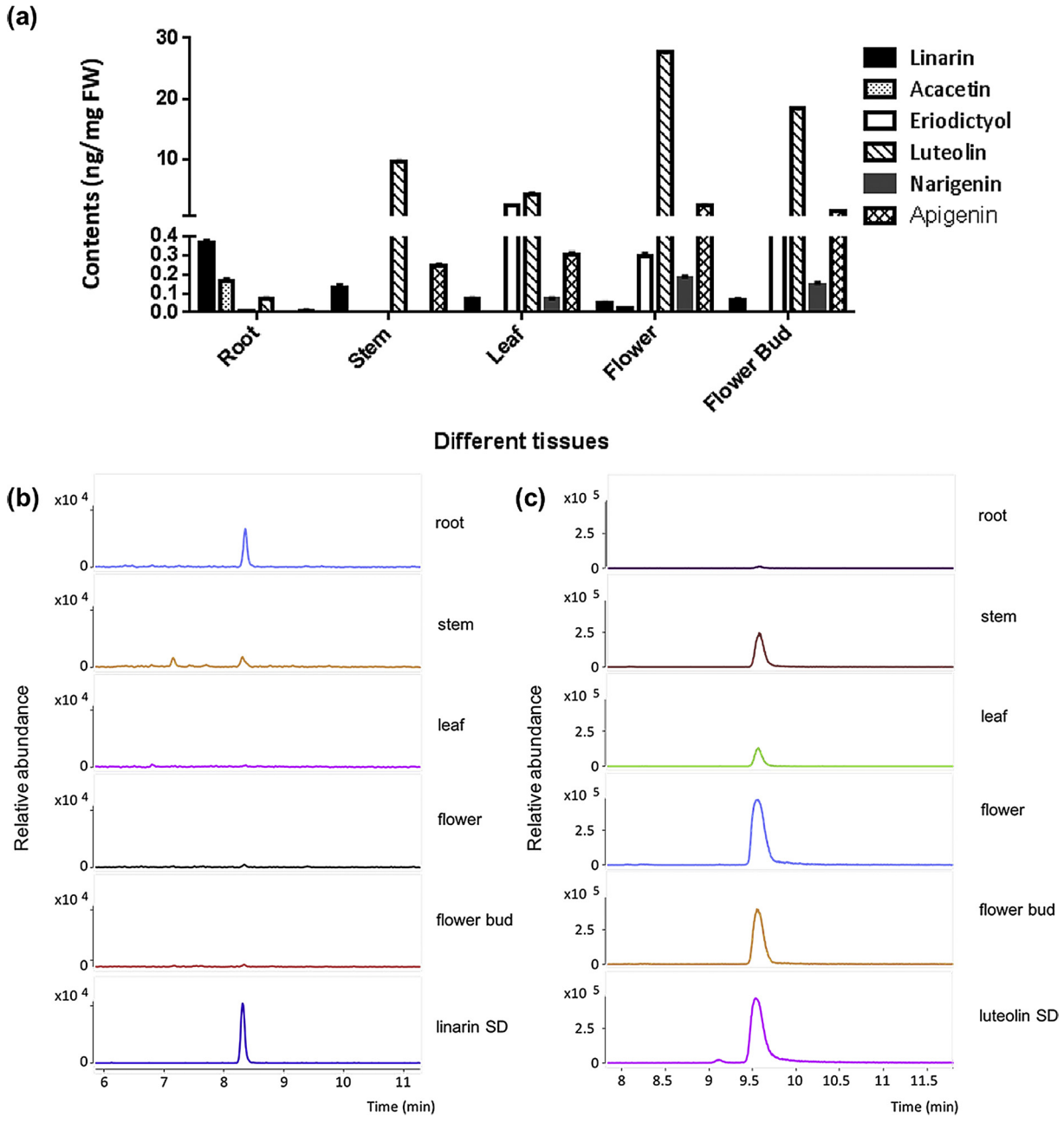


**Fig. 6.** Flavonoid contents (a) and HPLC-MS chromatographic profiles of (b) linarin and (c) luteolin in different tissues of *C. indicum*.

**Table 2**
Flavonoids in different tissues of *C. indicum*.

| Name | Formula | Mass $[M-H]^-$ | RT | Contents of flavonoids (μg/mg FW) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Root | Stem | Leaf | Flower | Flower bud |
| Luteolin | $C_{15}H_{10}O_6$ | 285.0405 | 9.57 | 0.069 ± 0.005 | 9.535 ± 0.072 | 4.230 ± 0.087 | 27.635 ± 0.055 | 18.324 ± 0.093 |
| Eriodictyol | $C_{15}H_{12}O_6$ | 287.0561 | 9.14 | 0.007 ± 0.001 | 0.002 ± 0.001 | 2.274 ± 0.03 | 0.294 ± 0.014 | 0.57 ± 0.002 |
| Apigenin | $C_{15}H_{10}O_5$ | 269.0455 | 14.09 | 0.010 ± 0.002 | 0.244 ± 0.008 | 0.301 ± 0.012 | 2.353 ± 0.020 | 1.307 ± 0.025 |
| Naringenin | $C_{15}H_{12}O_5$ | 271.0612 | 13.75 | n.d. | n.d. | 0.070 ± 0.005 | 0.179 ± 0.010 | 0.146 ± 0.009 |
| 7,4′-Dihydroxy flavone | $C_{15}H_{10}O_4$ | 253.0506 | 8.92 | n.d. | n.d. | n.d. | n.d. | n.d. |
| Liquiritigenin | $C_{15}H_{12}O_4$ | 255.0663 | 7.36 | n.d. | n.d. | n.d. | n.d. | n.d. |
| Acacetin | $C_{16}H_{12}O_5$ | 283.0612 | 17.08 | 0.163 ± 0.012 | n.d. | n.d. | 0.019 ± 0.001 | n.d. |
| Linarin | $C_{28}H_{32}O_{14}$ | 591.1719 | 8.33 | 0.364 ± 0.014 | 0.133 ± 0.012 | 0.070 ± 0.005 | 0.052 ± 0.001 | 0.064 ± 0.007 |

n.d., not detected.

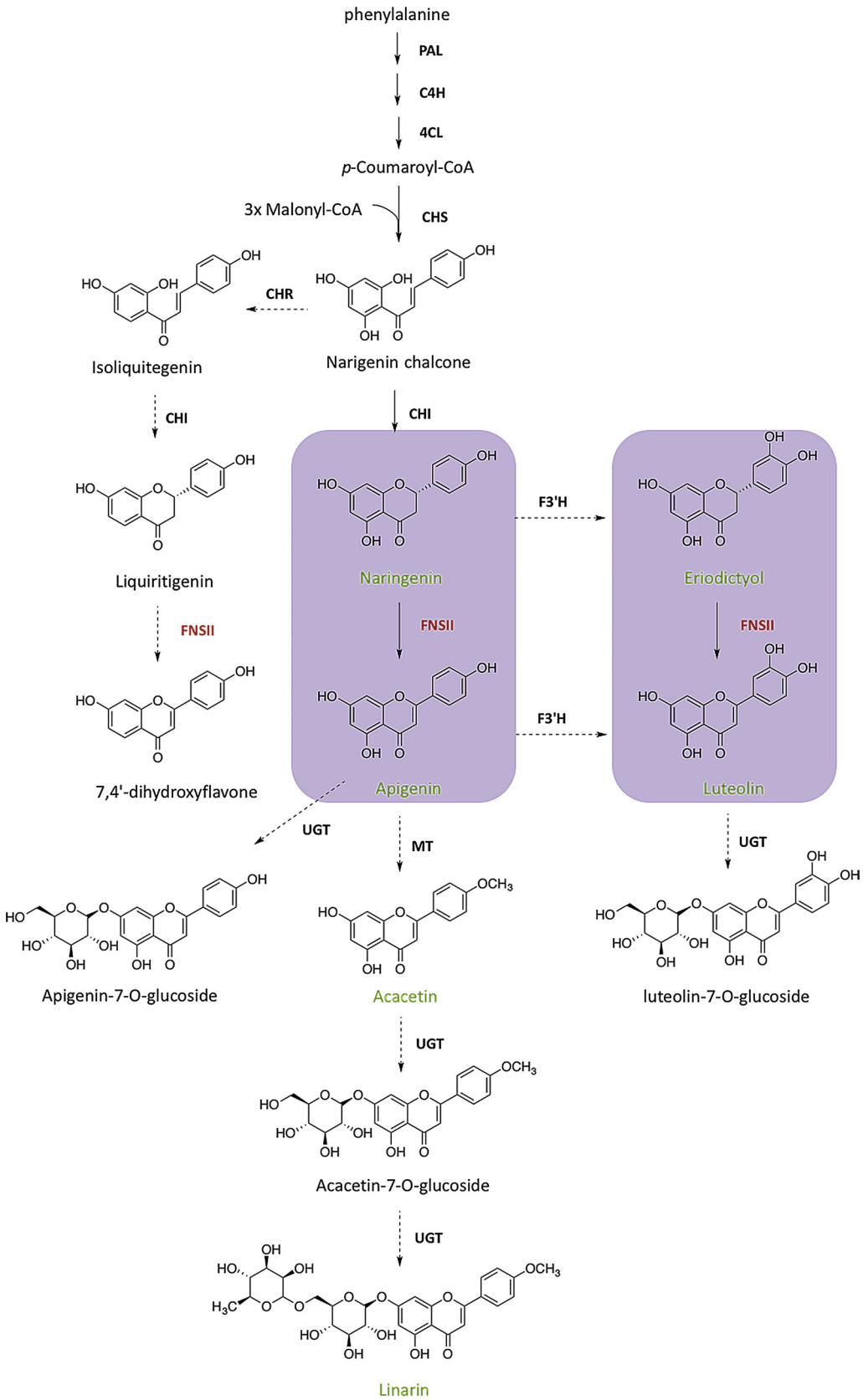


**Fig. 7.** A schematic outline of flavonoid pathway proposed for *C. indicum*. PAL, phenylalanine ammonia-lyase; C4H, trans-cinnamate 4-mono-oxygenase; 4CL, 4-coumarate-CoA ligase; CHS, chalcone synthase; CHR, chalcone reductase; CHI, chalcone isomerase; F3′H, flavonoid 3′-hydroxylase; FNSII, flavone synthase II; MT, methyltransferase; UGT, UDP-glycosyltransferase. Enzymes highlighted in red represents the FNSII that has been biochemically characterized, while compounds highlighted in green represents the compounds detected in our study. The purple colored range shows the proposed pathway of flavanone to flavone identified in *C. indicum* according to the biochemical and phytochemical results. The solid arrows represent the common upstream pathway of flavonoid biosynthesis in many plants (Artigot et al., 2013; Casas et al., 2014; Koes et al., 1994), the dashed arrows represent proposed conversion in *C. indicum* (Nugroho et al., 2013).

(Fig. 5c) could also be identified through the extracted ion peak. Our results suggest that CiFNSII catalyzes the conversion of flavanones to flavones, probably through the generation of 2-hydroxyflavanone intermediates. Additional experiments are needed to investigate the catalytic mechanism of CiFNSII.

In our enzyme reaction assay, the extracted ion peak (271) in the second plot (from top to bottom) of the panel (Fig. 5a) exhibited two peaks. The reason why two peaks were present when the *m/z* search

corresponds to the specific product could be the $O^{18}$ isotopic peak leading to the double peaks. The ratio of the peak areas in 271 to 269 is 0.2483% (the red arrow in Fig. S14a,b). The ratio of the intensity of the ions in 271 to 269 (Fig. S14c) is 0.2476%. The two values are similar to the natural abundance of $O^{18}$ (0.20%). Consequently, we concluded that the confused peak of naringenin in Fig. 5b is the isotopic peak.

Traditionally, the dried flowers of *C. indicum* have been used for the extraction of metabolites to perform phytochemical and pharmacological analysis (Cha et al., 2018; Cheon et al., 2009; Matsuda et al., 2002; Seo et al., 2013; Zhang et al., 2019b). Different drying methods can affect the chemical composition and abundance of metabolites (Abdul-Hamid et al., 2015; Farag et al., 2017; Pariyani et al., 2017). In our study, the fresh root, stem, leaf, flower, and flower bud of *C. indicum* were used to determine the levels of flavonones, flavones and their derivatives in these different tissues and elucidate the metabolic pathway with the help of transcriptome. As a result, two flavonones (narigenin and eriodictyol), two flavones (apigenin and luteolin) and two flavone derivatives (acacetin and linarin) were identified in *C. indicum*. Unfortunately, liquiritigenin and 7,4′-dihydroxyflavone were not detected in any tissues of *C. indicum* in our study. Recently, it has been identified in *C. indicum* flavonols and their derivatives such as quercetin, isorhamnetin, quercimeritroside and astroside (Cha et al., 2018; Luyen et al., 2015b). Future studies will be performed to compare the levels of flavone and their derivatives with other accumulated flavonoids, such as flavonols, anthocyanins and isoflavonoids in order to learn more about flavonoids accumulated in *C. indicum*.

In Pharmacopoeia of the People's Republic of China, linarin level is used as the standard for determining whether certain *C. indicum* product is qualified for medicinal use (Zhang et al., 2007). Notably, a precursor of linarin (acacetin-7-O-rutinoside), acacetin has been reported to have anti-inflammatory, antioxidative and cardioprotective activities (Kim et al., 2016; Liu et al., 2016; Sun et al., 2017). In our study, both acacetin and linarin were highly accumulated in root, with concentrations of over 6-times higher than those in flower (Fig. 6b, Figs. S12d, S15). It is interesting that linarin levels are lower in leaf, flower, and flower bud than those in root and stem. Our hypothesis is that *C. indicum* is a perennial plant, which means it dies during winter and grows back from its rootstock. If linarin synthesis occurs all year round, this may explain higher accumulation levels of linarin (acacetin-β-rutinoside) observed in root and stem than flower, flower bud, and leaf. One possible explanation is that acacetin and linarin may use apigenin as a substrate or precursor, resulting in low accumulation of apigenin in root and stem. Likelywise, luteolin levels are higher in flower and flower bud than those of the other tissues. It has also been reported higher levels of flavonoid aglycones (luteolin and apigenin) than glycosides (linarin glycosylated acacetin) in *Chrysanthemum morifolium* with yellow petals and *Chrysanthemum morifolium* cv. Noble Wine with purple striped white petals (Han et al., 2017). The reason why flavone aglycones are higher than linarin in *C. indicum* is worth further exploring. Luteolin, one of the most common flavones, exists in vegetables and fruits such as celery, parsley, and peppermint. It has been reported to exhibit antitumor, anti-inflammatory, and anti-obesity effects (Kwon et al., 2018; Negro et al., 2012). These observations are consistent with luteolin-rich artichoke, an important vegetable in the Mediterranean (e.g., Italy, Spain, France, and North Africa) and America diet (Kwon et al., 2018; Negro et al., 2012).

Recently, quercitrin present in *C. indicum* flowers has been reported to reduce high-fat diet-induced obesity in mice (Cha et al., 2018; Nepali et al., 2018). Quercitrin is a flavonol glycoside consisting in the flavonol quercetin and the deoxy sugar rhamnose. Quercetin is formed from eriodictyol by flavanone 3-hydroxylase (F3H) and flavonol synthase (FLS) enzymes (Winkel-Shirley, 2001). In our transcriptome analysis, genes of five families of enzymes, namely F3H, FLS, dihydroflavonol 4-reductase (DFR), and glucosyltransferase (GT) involved in flavonol and anthocyanidin biosynthesis were also identified, respectively (Table S6). These unigenes displayed a cluster pattern (Fig. 2b) similar to the common intermediates of flavonoid pathway, such as naringenin, apigenin, luteolin, and eriodictyol in the different tissues of *C. indicum* (Fig. 6a). If gene expression and metabolite accumulation are consistent, it could be hypothesized that F3H, FLS activities and quercetrin might have similar accumulation patterns in *C. indicum*. Further studies to explore this hypothesis are warranted.

In our phytochemical analysis, liquiritigenin and 7,4′-dihydroxyflavone were not detected in any tissues of *C. indicum*. To date, liquiritigenin and 7,4′-dihydroxyflavone have been found only in a few plants, such as licorice and soybean (Akashi et al., 1998; Fliegmann et al., 2010). Both F2H from licorice and soybean are able to catalyze the conversion of liquiritigenin to 7,4′-dihydroxyflavone (Akashi et al., 1998; Fliegmann et al., 2010): Although no unigene for CHR has been found in the transcriptome analysis of *C. indicum*, here, we demonstrated the enzymatic conversion of liquiritigenin to 7,4′-dihydroxyflavone by CiFNSII. Similarly, FNSII proteins from *L. japonica* and *L. macranthoides* have been reported to catalyze the conversion of liquiritigenin to 7,4′-dihydroxyflavone, although neither liquiritigenin nor 7,4′-dihydroxyflavone were detected in these plants (Wu et al., 2016). In the root of *Scutellaria baicalensis*, FNSII-2 uses pinocembrin rather than naringenin as an intermediate to produce chrysin (Zhao et al., 2016). In our studies, both naringenin and eriodictyol were detected in *C. indicum* by phytochemical analysis, and their conversions into apigenin and luteolin, respectively, by in vitro activity assays were confirmed. As a result, our results suggest that CiFNSII is an active flavone synthase in planta involved in the conversion of naringenin and eriodictyol to apigenin and luteolin, respectively, *in planta*. Future studies with knock-out and knock-down plants are needed to verify this hypothesis. Finally, we propose a linarin biosynthesis pathway (Fig. 7) based on transcriptomic analysis, metabolic profiling, and enzyme activities of CiFNSII. These results provide genomic resources for future enhancement of linarin production using genetic strategies and metabolic engineering in *C. indicum*.

## 5. Conclusion

In this study, we report the use of transcriptomic analysis and integrative metabolomics to reveal genes involved in flavone biosynthesis and flavonoid accumulation levels in *C. indicum*. It is clearly demonstrated that flavonoid distribution differs considerably among different tissues of *C. indicum*. Based on transcriptomic data, CiFNSII was identified as the key enzyme responsible for flavone biosynthesis in *C. indicum*. We further confirmed the biochemical function of CiFNSII in converting naringenin, eriodictyol and liquiritigenin directly to their respective flavones. Our study provides insight into the potential application of molecular breeding and metabolic engineering for improving the quality of cultivated *C. indicum*.

## Conflict of interest

The authors declare no conflict of interest.

## Funding

We are grateful for the financial support of the start-up fund of Guangzhou University of Chinese Medicine, Guangdong Pearl River Talents Plan (2017GC010368), key platform construction project of Department of Education of Guangdong Province (Grant number 2014KTSPT016) & Guangdong Science & Technology (Grant number 2013B090600115).

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.indcrop.2019.04.009.

## References


Abdul-Hamid, N.A., Abas, F., Ismail, I.S., Shaari, K., Lajis, N.H., 2015. Influence of different drying treatments and extraction solvents on the metabolite profile and nitric oxide inhibitory activity of ajwa dates. J. Food Sci. 80, H2603–2611.

Akashi, T., Aoki, T., Ayabe, S., 1998. Identification of a cytochrome P450 cDNA encoding (2S)-flavanone 2-hydroxylase of licorice (*Glycyrrhiza echinata* L.; Fabaceae) which represents licodione synthase and flavone synthase II. FEBS Lett. 431, 287–290.

Akashi, T., Fukuchi-Mizutani, M., Aoki, T., Ueyama, Y., Yonekura-Sakakibara, K., Tanaka, Y., Kusumi, T., Ayabe, S., 1999. Molecular cloning and biochemical characterization of a novel cytochrome P450, flavone synthase II, that catalyzes direct conversion of flavanones to flavones. Plant Cell Physiol. 40, 1182–1186.

Artigot, M.-P., Daydé, J., Berger, M., 2013. Expression of key genes of the isoflavonoid pathway in hypocotyls and cotyledons during soybean seed maturation. Crop Sci. 53, 1096–1108.

Ayabe, S.-i., Akashi, T., 2006. Cytochrome P450s in flavonoid metabolism. Phytochem. Rev. 5, 271–282.

Britsch, L., Heller, W., Grisebach, H., 1981. Conversion of flavanone to flavone, dihydroflavonol and flavonol with an enzyme system from cell cultures of parsley. Zeitschrift für Naturforschung C 36, 742–750.

Casas, M.I., Duarte, S., Doseff, A.I., Grotewold, E., 2014. Flavone-rich maize: an opportunity to improve the nutritional value of an important commodity crop. Front. Plant Sci. 5, 440.

Cha, J.Y., Nepali, S., Lee, H.Y., Hwang, S.W., Choi, S.Y., Yeon, J.M., Song, B.J., Kim, D.K., Lee, Y.M., 2018. *Chrysanthemum indicum* L. ethanol extract reduces high-fat diet-induced obesity in mice. Exp. Ther. Med. 15, 5070–5076.

Chapple, C., 1998. Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. Annu. Rev. Plant Biol. 49, 311–343.

Cheon, M.S., Yoon, T., Lee, D.Y., Choi, G., Moon, B.C., Lee, A.Y., Choo, B.K., Kim, H.K., 2009. *Chrysanthemum indicum* Linne extract inhibits the inflammatory response by suppressing NF-kappaB and MAPKs activation in lipopolysaccharide-induced RAW 264.7 macrophages. J. Ethnopharmacol. 122, 473–477.

Dixon, R.A., Pasinetti, G.M., 2010. Flavonoids and isoflavonoids: from plant biology to agriculture and neuroscience. Plant Physiol. 154, 453–457.

Dong, M., Yu, D., Abdullah Al-Dhabi, N., Duraipandiyan, V., 2017. The impacts of *Chrysanthemum indicum* extract on oxidative stress and inflammatory responses in adjuvant-induced arthritic rats. Evid. Based Complement. Altern. Med. 2017.

Falcone Ferreyra, M.L., Rius, S.P., Casati, P., 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front. Plant Sci. 3, 222.

Farag, M.A., Ali, S.E., Hodaya, R.H., El-Seedi, H.R., Sultani, H.N., Laub, A., Eissa, T.F., Abou-Zaid, F.O.F., Wessjohann, L.A., 2017. Phytochemical profiles and antimicrobial activities of *Allium cepa* Red cv. and *A. sativum* subjected to different drying methods: a comparative MS-based metabolomics. Molecules 22.

Feng, Z., Yang, Y., Jiang, J., Zhang, P., 2010. Chemical constituents from flowers of *Chrysanthemum indicum*. China J. Chinese Matera Med. 35, 3302–3305.

Ferreyra, M.L.F., Emiliani, J., Rodriguez, E.J., Campos-Bermudez, V.A., Grotewold, E., Casati, P., 2015. The identification of maize and Arabidopsis type I flavone synthases links flavones with hormones and biotic interactions. Plant Physiol. 169, 1090–1107.

Fliegmann, J., Furtwangler, K., Malterer, G., Cantarello, C., Schuler, G., Ebel, J., Mithofer, A., 2010. Flavone synthase II (CYP93B16) from soybean (*Glycine max* L.). Phytochemistry 71, 508–514.

Gao, M.H., Li, H., Zhang, L., Xiao, S.X., 2008. Studies on chemical constituents from flowers of Chrysanthemum indicum. J. Chin. Med. Mater. 31, 682–684.

Gharbi, S., Shamsara, M., Khateri, S., Soroush, M.R., Ghorbanmehr, N., Tavallaei, M., Nourani, M.R., Mowla, S.J., 2015. Identification of reliable reference genes for quantification of microRNAs in serum samples of sulfur mustard-exposed veterans. Cell J. (Yakhteh) 17, 494.

Gietz, R.D., Woods, R.A., 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol 350, 87–96.

Gu, C., Chen, S., Liu, Z., Shan, H., Luo, H., Guan, Z., Chen, F., 2011. Reference gene selection for quantitative real-time PCR in Chrysanthemum subjected to biotic and abiotic stress. Mol. Biotechnol. 49, 192.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494.

Han, A.R., Kim, H.Y., So, Y., Nam, B., Lee, I.S., Nam, J.W., Jo, Y.D., Kim, S.H., Kim, J.B., Kang, S.Y., Jin, C.H., 2017. Quantification of antioxidant phenolic compounds in a new *Chrysanthemum Cultivar* by high-performance liquid chromatography with diode array detection and electrospray ionization mass spectrometry. Int. J. Anal. Chem. 2017, 1254721.

Hasemann, C.A., Kurumbail, R.G., Boddupalli, S.S., Peterson, J.A., Deisenhofer, J., 1995. Structure and function of cytochromes P450: a comparative analysis of three crystal structures. Structure 3, 41–62.

Jeong, S.C., Kim, S.M., Jeong, Y.T., Song, C.H., 2013. Hepatoprotective effect of water extract from *Chrysanthemum indicum* L. flower. Chin. Med. 8, 7.

Jiang, N., Doseff, A.I., Grotewold, E., 2016. Flavones: from biosynthesis to health benefits. Plants (Basel) 5.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 40, D109–114.

Kim, J.H., Cheon, Y.M., Kim, B.-G., Ahn, J.-H., 2008. Analysis of flavonoids and characterization of the OsFNS gene involved in flavone biosynthesis in Rice. J. Plant Biol. 51, 97.

Kim, C., Kim, M.C., Kim, S.M., Nam, D., Choi, S.H., Kim, S.H., Ahn, K.S., Lee, E.H., Jung, S.H., 2013. *Chrysanthemum indicum* L. extract induces apoptosis through suppression of constitutive STAT3 activation in human prostate cancer DU145 cells. Phytother. Res. 27, 30–38.

Kim, T.W., Kim, Y.J., Park, S.R., Seo, C.S., Ha, H., Shin, H.K., Jung, J.Y., 2015. *Chrysanthemum indicum* attenuates cisplatin-induced nephrotoxicity both in vivo and in vitro. Nat. Prod. Commun. 10, 397–402.

Kim, B., Lee, J.H., Seo, M.J., Eom, S.H., Kim, W., 2016. Linarin down-regulates phagocytosis, pro-inflammatory cytokine production, and activation marker expression in RAW264.7 macrophages. Food Sci. Biotechnol. 25, 1437–1442.

Kochs, G., Grisebach, H., 1987. Induction and characterization of a NADPH-dependent flavone synthase from cell cultures of soybean. Zeitschrift für Naturforschung C 42, 343–348.

Koes, R.E., Quattrocchio, F., Mol, J.N., 1994. The flavonoid biosynthetic pathway in plants: function and evolution. BioEssays 16, 123–132.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Kwon, E.Y., Kim, S.Y., Choi, M.S., 2018. Luteolin-enriched artichoke leaf extract alleviates the metabolic syndrome in mice with high-fat diet-induced obesity. Nutrients 10.

Lam, P.Y., Zhu, F.Y., Chan, W.L., Liu, H., Lo, C., 2014. Cytochrome P450 93G1 is a flavone synthase II that channels flavanones to the biosynthesis of tricin O-Linked conjugates in rice. Plant Physiol. 165, 1315–1327.

Lee, H.J., Seo, H.S., Ryu, J., Yoon, Y.P., Park, S.H., Lee, C.J., 2015. Luteolin inhibited the gene expression, production and secretion of MUC5AC mucin via regulation of nuclear factor kappa B signaling pathway in human airway epithelial cells. Pulm. Pharmacol. Ther. 31, 117–122.

Liu, H., Yang, L., Wu, H.J., Chen, K.H., Lin, F., Li, G., Sun, H.Y., Xiao, G.S., Wang, Y., Li, G.R., 2016. Water-soluble acacetin prodrug confers significant cardioprotection against ischemia/reperfusion injury. Sci. Rep. 6, 36435.

Liu, X., Cheng, J., Zhang, G., Ding, W., Duan, L., Yang, J., Kui, L., Cheng, X., Ruan, J., Fan, W., Chen, J., Long, G., Zhao, Y., Cai, J., Wang, W., Ma, Y., Dong, Y., Yang, S., Jiang, H., 2018. Engineering yeast for the production of breviscapine by genomic analysis and synthetic biology approaches. Nat. Commun. 9, 448.

Luyen, B.T., Tai, B.H., Thao, N.P., Cha, J.Y., Lee, H.Y., Lee, Y.M., Kim, Y.H., 2015a. Anti-inflammatory components of *Chrysanthemum indicum* flowers. Bioorg. Med. Chem. Lett. 25, 266–269.

Luyen, B.T., Tai, B.H., Thao, N.P., Lee, Y.M., Lee, S.H., Jang, H.D., Kim, Y.H., 2015b. The anti-osteoporosis and antioxidant activities of chemical constituents from *Chrysanthemum indicum* flowers. Phytother. Res. 29, 540–548.

Mahmood, T., Yang, P.-C., 2012. Western blot: technique, theory, and trouble shooting. N. Am. J. Med. Sci. 4, 429.

Martens, S., Forkmann, G., 1999. Cloning and expression of flavone synthase II from Gerbera hybrids. Plant J. 20, 611–618.

Martens, S., Mithofer, A., 2005. Flavones and flavone synthases. Phytochemistry 66, 2399–2407.

Mathesius, U., 2018. Flavonoid Functions in Plants and Their Interactions with Other Organisms. Multidisciplinary Digital Publishing Institute.

Matsuda, H., Morikawa, T., Toguchida, I., Harima, S., Yoshikawa, M., 2002. Medicinal flowers. VI. Absolute stereostructures of two new flavanone glycosides and a phenylbutanoid glycoside from the flowers of *Chrysanthemum indicum* L.: their inhibitory activities for rat lens aldose reductase. Chem. Pharm. Bull. 50, 972–975.

Negro, D., Montesano, V., Grieco, S., Crupi, P., Sarli, G., De Lisi, A., Sonnante, G., 2012. Polyphenol compounds in artichoke plant tissues and varieties. J. Food Sci. 77, C244–252.

Nepali, S., Cha, J.Y., Ki, H.H., Lee, H.Y., Kim, Y.H., Kim, D.K., Song, B.J., Lee, Y.M., 2018. *Chrysanthemum indicum* inhibits adipogenesis and activates the AMPK pathway in high-fat-diet-induced obese mice. Am. J. Chin. Med. 46, 119–136.

Nishina, A., Sato, D., Yamamoto, J., Kobayashi-Hattori, K., Hirai, Y., Kimura, H., 2019. Antidiabetic-like effects of naringenin-7-O-glucoside from edible *Chrysanthemum*' Kotobuki' and Naringenin by activation of the PI3K/Akt pathway and PPARγ. Chem. Biodivers. 16, e1800434.

Nugroho, A., Lim, S.C., Choi, J., Park, H.J., 2013. Identification and quantification of the sedative and anticonvulsant flavone glycoside from Chrysanthemum boreale. Arch. Pharm. Res. 36, 51–60.

Pariyani, R., Ismail, I.S., Ahmad Azam, A., Abas, F., Shaari, K., 2017. Identification of the compositional changes in *Orthosiphon stamineus* leaves triggered by different drying techniques using $^{1}$H NMR metabolomics. J. Sci. Food Agric. 97, 4169–4179.

Schmittgen, T.D., Livak, K.J., 2008. Analyzing real-time PCR data by the comparative $C_T$ method. Nat. Protoc. 3, 1101–1108.

Seo, D.W., Cho, Y.R., Kim, W., Eom, S.H., 2013. Phytochemical linarin enriched in the flower of *Chrysanthemum indicum* inhibits proliferation of A549 human alveolar basal epithelial cells through suppression of the Akt-dependent signaling pathway. J. Med. Food 16, 1086–1094.

Shen, G.-M., Jiang, H.-B., Wang, X.-N., Wang, J.-J., 2010. Evaluation of endogenous references for gene expression profiling in different tissues of the oriental fruit fly *Bactrocera dorsalis* (Diptera: Tephritidae). BMC Mol. Biol. 11, 76.

Stotz, G., Forkmann, G., 1981. Oxidation of flavanones to flavones with flower extracts of *Antirrhinum majus* (snapdragon). Zeitschrift für Naturforschung C 36, 737–741.

Sun, H., Zhang, T., Fan, Q., Qi, X., Zhang, F., Fang, W., Jiang, J., Chen, F., Chen, S., 2015. Identification of floral scent in *chrysanthemum cultivars* and wild relatives by gas chromatography-mass spectrometry. Molecules 20, 5346–5359.

Sun, S., Jiang, P., Su, W., Xiang, Y., Li, J., Zeng, L., Yang, S., 2016. Wild chrysanthemum extract prevents UVB radiation-induced acute cell death and photoaging. Cytotechnology 68, 229–240.

Sun, L.C., Zhang, H.B., Gu, C.D., Guo, S.D., Li, G., Lian, R., Yao, Y., Zhang, G.Q., 2017. Protective effect of acacetin on sepsis-induced acute lung injury via its anti-inflammatory and antioxidative activity. Arch. Pharm. Res. 1–12.

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., Pompon, D., 1997. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J. Biol. Chem. 272, 19176–19186.

Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285.

Wang, H., Ma, D., Yang, J., Deng, K., Li, M., Ji, X., Zhong, L., Zhao, H., 2018a. An integrative volatile terpenoid profiling and transcriptomics analysis for gene mining and functional characterization of AvBPPS and AvPS involved in the monoterpenoid biosynthesis in *Amomum villosum*. Front. Plant Sci. 9, 846.

Wang, Q.Z., Downie, S.R., Chen, Z.X., 2018b. Genome-wide searches and molecular analyses highlight the unique evolutionary path of flavone synthase I (FNSI) in Apiaceae. Genome 61, 103–109.

Winkel-Shirley, B., 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 126, 485–493.

Wu, X.L., Li, C.W., Chen, H.M., Su, Z.Q., Zhao, X.N., Chen, J.N., Lai, X.P., Zhang, X.J., Su, Z.R., 2013. Anti-inflammatory effect of supercritical-carbon dioxide fluid extract from flowers and buds of *Chrysanthemum indicum* Linnen. Evid. Complement. Alternat. Med. 2013, 413237.

Wu, J., Wang, X.C., Liu, Y., Du, H., Shu, Q.Y., Su, S., Wang, L.J., Li, S.S., Wang, L.S., 2016. Flavone synthases from *Lonicera japonica* and *L. macranthoides* reveal differential flavone accumulation. Sci. Rep. 6, 19245.

Xiao, J., Capanoglu, E., Jassbi, A.R., Miron, A., 2016. Advance on the flavonoid C-glycosides and health benefits. Crit. Rev. Food Sci. Nutr. 56, S29–S45.

Yan, J., Wang, B., Jiang, Y., Cheng, L., Wu, T., 2014. GmFNSII-controlled soybean flavone metabolism responds to abiotic stresses and regulates plant salt tolerance. Plant Cell Physiol. 55, 74–86.

Yang, H.M., Sun, C.Y., Liang, J.L., Xu, L.Q., Zhang, Z.B., Luo, D.D., Chen, H.B., Huang, Y.Z., Wang, Q., Lee, D.Y., Yuan, J., Li, Y.C., 2017a. Supercritical-carbon dioxide fluid extract from *Chrysanthemum indicum* enhances anti-tumor effect and reduces toxicity of bleomycin in tumor-bearing mice. Int. J. Mol. Sci. 18, 465.

Yang, W.S., Kim, D., Yi, Y.S., Kim, J.H., Jeong, H.Y., Hwang, K., Kim, J.H., Park, J., Cho, J.Y., 2017b. AKT-targeted anti-inflammatory activity of the methanol extract of *Chrysanthemum indicum* var. albescens. J. Ethnopharmacol. 201, 82–90.

Yang, Y., Jiang, Z., Guo, J., Yang, X., Xu, N., Chen, Z., Hao, J., Li, J., Pang, J., Shen, C., 2018. Transcriptomic analyses of *Chrysanthemum morifolium* Ramat under UV-B radiation treatment reveal variations in the metabolisms associated with bioactive components. Ind. Crops Prod. 124, 475–486.

Yoo, K.Y., Kim, I.H., Cho, J.H., Ahn, J.H., Park, J.H., Lee, J.C., Tae, H.J., Kim, D.W., Kim, J.D., Hong, S., Won, M.H., Kang, I.J., 2016. Neuroprotection of *Chrysanthemum indicum* Linne against cerebral ischemia/reperfusion injury by anti-inflammatory effect in gerbils. Neural Regen. Res. 11, 270–277.

Yoshikawa, M., Morikawa, T., Murakami, T., Toguchida, I., Harima, S., Matsuda, H., 1999. Medicinal flowers. I. Aldose reductase inhibitors and three new eudesmane-type sesquiterpenes, kikkanols A, B, and C, from the flowers of *Chrysanthemum indicum* L. Chem. Pharm. Bull. 47, 340–345.

Yoshikawa, M., Morikawa, T., Toguchida, I., Harima, S., Matsuda, H., 2000. Medicinal flowers. II. Inhibitors of nitric oxide production and absolute stereostructures of five new germacrane-type sesquiterpenes, kikkanols D, D monoacetate, E, F, and F monoacetate from the flowers of *Chrysanthemum indicum* L. Chem. Pharm. Bull. 48, 651–656.

Zhang, J., Subramanian, S., Zhang, Y., Yu, O., 2007. Flavone synthases from *Medicago truncatula* are flavanone-2-hydroxylases and are important for nodulation. Plant Physiol. 144, 741–751.

Zhang, X., Xie, Y.L., Yu, X.T., Su, Z.Q., Yuan, J., Li, Y.C., Su, Z.R., Zhan, J.Y., Lai, X.P., 2015. Protective effect of super-critical carbon dioxide fluid extract from flowers and buds of *Chrysanthemum indicum* Linnen against ultraviolet-induced photo-aging in mice. Rejuvenation Res. 18, 437–448.

Zhang, N., He, Z., He, S., Jing, P., 2019a. Insights into the importance of dietary chrysanthemum flower (*Chrysanthemum morifolium* cv. Hangju)-wolfberry (*Lycium barbarum* fruit) combination in antioxidant and anti-inflammatory properties. Food Res. Int. 116, 810–818.

Zhang, X., Wu, J.Z., Lin, Z.X., Yuan, Q.J., Li, Y.C., Liang, J.L., Zhan, J.Y., Xie, Y.L., Su, Z.R., Liu, Y.H., 2019b. Ameliorative effect of supercritical fluid extract of *Chrysanthemum indicum* Linnen against D-galactose induced brain and liver injury in senescent mice via suppression of oxidative stress, inflammation and apoptosis. J. Ethnopharmacol. 234, 44–56.

Zhao, Q., Zhang, Y., Wang, G., Hill, L., Weng, J.K., Chen, X.Y., Xue, H., Martin, C., 2016. A specialized flavone biosynthetic pathway has evolved in the medicinal plant, Scutellaria baicalensis. Sci. Adv. 2, e1501780.

Zheng, X., Luo, X., Ye, G., Chen, Y., Ji, X., Wen, L., Xu, Y., Xu, H., Zhan, R., Chen, W., 2015. Characterisation of two oxidosqualene cyclases responsible for triterpenoid biosynthesis in Ilex asprella. Int. J. Mol. Sci. 16, 3564–3578.