The transcription factor complex CmAP3-CmPI-CmUIF1 modulates carotenoid metabolism by directly regulating carotenogenic gene CmCCD4a-2 in chrysanthemum

Runing title: CmAP3-CmPI-CmUIF1 regulates the expression of CmCCD4a-2

Chenfei Lu1, Jiaping Qu1, Chengyan Deng1, Fangye Liu1, Fan Zhang1, He Huang1,* Silan Dai1,*

1Beijing Key Laboratory of Ornamental Plants Germplasm Innovation & Molecular Breeding, National Engineering Research Center for Floriculture, Beijing Laboratory of Urban and Rural Ecological Environment, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants of Education Ministry, School of Landscape Architecture, Beijing Forestry University, Beijing, 100083, China

*Correspondence should be addressed to He Huang (Email:101navy@163.com) and Silan Dai (silandai@sina.com).

© The Author(s) 2022. Published by Oxford University Press. All rights reserved. This is an Open Access article distributed under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Abstract:

Carotenoids are one of the most important pigments for the coloring in many plants, fruits and flowers. Recently, significant progress has been made in carotenoid metabolism. However, the specific understanding on transcriptional regulation controlling the expression of carotenoid metabolic genes remains extremely limited. Anemone-type chrysanthemum, as a special group of chrysanthemum cultivars, contain elongated disc florets in capitulum, which usually appear in different colors compared with the ray florets since accumulating distinct content of carotenoids. In this study, the carotenoid composition and content of the ray and disc florets of an anemone-type chrysanthemum cultivar ‘Dong Li Fen Gui’ were analyzed by high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) and the key structural gene CmCCD4a-2, of which differential expression resulted in the distinct content of carotenoids accumulated in these two types of florets, was identified. Then the promoter sequence of CmCCD4a-2 was used as bait to screen a chrysanthemum flower cDNA library and two transcription factors, CmAP3 and CmUIF1 were identified. Y2H, BiFC and Y3H experiments demonstrated that these two TFs were connected by CmPI to form CmAP3-CmPI-CmUIF1 TF complex. This TF complex regulated carotenoid metabolism through activating the expression of CmCCD4a-2 directly. Furthermore, a large number of target genes regulated directly by the CmAP3-CmPI-CmUIF1 TF complex, including carotenoid biosynthetic genes, flavonoid biosynthetic genes and flower development-related genes, were identified by DNA-affinity purification sequencing (DAP-seq), which indicated that the CmAP3-CmPI-CmUIF1 TF complex might participate in multiple processes. These findings expand our knowledge for the transcriptional regulation of carotenoid metabolism in plants and will be helpful to manipulating carotenoid accumulation in chrysanthemum.

Key words: carotenoid; chrysanthemum; CmCCD4a-2; CmUIF1-CmPI-CmAP3 transcription factor complex; transcriptional regulation

Introduction

Carotenoids, a group of important secondary metabolites that make flowers and fruits appear in diverse colors, are biosynthesized and stored in all photosynthetic organisms, including algae and plants. In the green tissues of plants, carotenoids are involved in the capture of light energy and transfer it to chlorophyll for photosynthesis. Meanwhile, they also participate in photoprotection and serve as precursors for phytohormone molecules. Besides the physiological functions in plants, carotenoids are also essential for human nutritional balance and health. It is extremely important to clarify the regulation of carotenoid metabolism for the cultivation of
horticultural crops with high nutritional value.

Generally, carotenoid metabolic genes have been identified extensively in many plants through biochemical analysis and molecular biology. Major enzymes in the carotenoid metabolic pathway include deoxyxylulose 5-phosphate synthase (DXS), deoxyxylulose 5-phosphate reductoisomerase (DXR), phytoene synthase (PSY), phytoene desaturase (PDS), ζ-carotene desaturase (ZDS), and carotene isomerase (CRTISO), which mainly participate in the synthesis of linear carotenes, such as phytoene, zeta-carotene, and lycopene. Lycopene represents the branch point of the carotenoid biosynthetic pathway and can be cyclized by lycopene β-cyclase (LCYB) and lycopene ε-cyclase (LCYE) to produce α- or β-carotene, which would be hydroxylated and epoxidated subsequently to produce various xanthophylls. Both carotenoid biosynthesis and degradation determine the steady-state level of carotenoids in plants. Thus, the catalytic function of carotenoid cleavage dioxygenases (CCDs), which cleave carotenoids to produce apocarotenoids and their derivatives, such as phytohormones (ABA and SLs) and signaling molecules, is critical in manipulating carotenoid accumulation. Previous researches have shown that the members of the CCD1 and CCD4 subfamilies could affect the carotenoid content in various plant species, such as in the fruits of peach and in the flowers of chrysanthemum and osmanthus. Many studies have declared that the expression level of these carotenoid metabolic genes is closely associated with carotenoid accumulation in plants. However, the specific transcriptional regulation that controls the expression of these structural genes remains extremely limited.

Transcription factors (TFs) act as key regulatory factors in plant growth and development by directly binding to the cis-acting elements of downstream target genes to regulate their expression. Currently, many TFs that directly modulate the carotenoid metabolic processes are characterized. For example, PHYTOCHROME-INTERACTING FACTOR 1 (PIF1) repressed the transcription of AtPSY, a key structural gene encoding a major flux-controlling enzyme in carotenoid metabolism, by directly binding to the G-box element in Arabidopsis. The transcription factor RAP2.2 belonging to the AP2/ERF family can specifically bind to the promoters of AtPSY and AtPDS to regulate their expression. Other TFs, including CpNAC1 from papaya, MtWP1 from Medicago, and AdMYB7 from kiwifruit, also were involved in the regulation of carotenoid metabolism.

MADS-box transcription factors are of ancient origin and are found in animals, fungi, and plants. They participate in various developmental processes in plants, including vegetative organ development, flowering, floral organ identity, fruit ripening, and metabolism. Recently, several MADS-box ripening regulators, including RIPENING INHIBITOR (RIN), tomato AGAMOUS-LIKE1 (TAGL1), FRUITFULL1 (FUL1), and FUL2, have been identified in tomato fruits and can form dimeric and tetrameric complexes to affect carotenoid metabolism. In
citrus, CsMADS5 was characterized as a positive regulator that interacted with CsMADS6 to synergistically promote carotenoid accumulation\textsuperscript{19}. AP3 and PI belong to the class B DEFICIENS/APETALA3 and GLOBOSA/PISTILLATA subfamilies of MADS TFs, respectively, and many researches have shown the roles of AP3 and PI homologs in floral organ identity\textsuperscript{20,21}. The GARP transcription factor family, which composed of Golden2-like, ARR-B, and Psr1, is mainly involved in hormonal signaling, nutrient response and sensing processes, chloroplast biogenesis, and plant development\textsuperscript{22}. Recently, the GARP TF UIF1 (ULT1 Interacting Factor 1), which can directly bind to the promoters of WUS and AG to exert regulatory functions on petal number in flowers, was identified in Arabidopsis\textsuperscript{23}. However, there is no report on the interaction between AP3, PI and UIF1. Besides, the exact roles of these TFs in carotenoid metabolism are still unclear.

\textit{Chrysanthemum ×morifolium}, one of the top ten well-known traditional Chinese flowers, is a leading flower with applied value worldwide\textsuperscript{24}. As the most representative ornamental flower in Asteraceae, chrysanthemum has a typical capitulum, which is composed of ray florets and disc florets. Anemone-type chrysanthemum, as a special group of chrysanthemum cultivars, contain elongated disc florets in capitulum, which have the same ability as the ray florets to accumulate pigments\textsuperscript{25}. The elongated disc florets of almost all anemone-type chrysanthemum cultivars appear yellow because they accumulate large amounts of carotenoids, while the ray florets could appear white, pink, yellow, orange, or red since different levels of carotenoids and anthocyanins accumulated simultaneously\textsuperscript{26}. Thus, these two types of florets in anemone-type chrysanthemum are excellent materials for studying the regulatory mechanism of carotenoid metabolism. In this study, the key structural gene \textit{CmCCD4a-2}, of which differential expression resulted in the distinct content of carotenoids accumulated in the ray and disc florets of anemone-type chrysanthemum cultivar ‘Dong Li Fen Gui’, was identified. Then we charactered a TF complex named CmAP3-CmPI-CmUIF1 and found that this TF complex regulated carotenoid metabolism through activating the expression of \textit{CmCCD4a-2} directly. Furthermore, other downstream target genes regulated by this TF complex were also identified by DNA-affinity purification sequencing (DAP-seq). These findings expand our knowledge for the transcriptional regulation of carotenoid metabolism in plants and will be helpful to control carotenoid accumulation in chrysanthemum.

\textbf{Materials and methods}

\textbf{Plant materials}

Chrysanthemum cultivar ‘Dong Li Fen Gui’, as the anemone-type chrysanthemum, contained elongated disc florets in capitulum, which appeared in different colors compared with the ray florets since accumulating distinct content of carotenoids (Fig. 1a), was the main material for screening the key carotenoid metabolic gene and studying the transcriptional regulatory
mechanism. Other chrysanthemum cultivars from different color groups (white group: ‘404×C34-79’, ‘404×C34-85’; pink group: ‘317’, ‘B200’; yellow group: ‘404’, ‘404×C60-10’; red group: ‘A49’, ‘D91’) (Fig. S4a) were selected to verify the expressed pattern of key structural gene in ray and disc florets. All the plant materials were grown in the chrysanthemum germplasm nursery of the Beijing Forestry University. The development of capitulum of the anemone-type chrysanthemum ‘Dong Li Fen Gui’ was divided into five stages (S1-S5) based on Lu et al. 2019\(^{27}\). The ray florets (named as R1-R5) and disc florets (named as D1-D5) of ‘Dong Li Fen Gui’ at different developmental stages (Fig. 1a) were collected for carotenoid components and contents analysis, comparative transcriptome and gene expressed profile analysis.

**Extraction and analysis of carotenoids by HPLC-MS/MS**

Carotenoid components and contents of the ray florets and disc florets of the chrysanthemum cultivar ‘Dong Li Fen Gui’ at different developmental stages were analyzed by HPLC-MS/MS. Firstly, 0.25 g of tissues were weighed and added 4 ml pigment extract. Supernatants were transferred into new 10 ml centrifuge tubes and dried using a nitrogen blowing instrument (< 30°C). The dried carotenoid extracts were dissolved in 2 ml MTBE, and then saponified in 2 ml of 10% KOH-methanol solution for 10 h away from light. After saponification, 4 ml NaCl solution and 2 ml MTBE were added to the sample, and the supernatant (MTBE-carotenoid solution) was concentrated to dryness by a nitrogen blowing instrument once again. The dried carotenoid extracts were dissolved thoroughly in 0.6-1 ml MTBE-methanol solution (1:1, v/v), then supernatants were filtered through a 0.22 μm membrane filter to prepare the sample for HPLC detection. Elution program of HPLC was as previously published by Huang et al., 2019\(^{28}\). LC-MS/MS was performed on the same HPLC as above. Carotenoid components were identified using standards (β-carotene, violaxanthin, lutein) (sigma), the specific absorption wavelength, retention time and MS data of the carotenoids reported in the published literature\(^{27,29}\).

**RNA sequencing, functional annotation and data processing**

RNA-Seq has been widely used to screen out the key structural genes and TFs in plants. In this study, the ray florets (R3) and disc florets (D3) at the S3 stage were sampled for RNA-seq. Briefly, total RNAs of R3 and D3 were extracted and the cDNA libraries was synthesized. Then obtained libraries were sequenced on the Illumina HiSeq 2500 sequencing platform (Illumina, USA). After the connector sequences, low-quality sequences were removed. Subsequently, the clean reads were assembled to unigenes using SOAPdenovo software with the parameters set to -K29, -M2, -L50. The transcript abundance of all unigenes were estimated via the FPKM method by RSEM\(^{30}\). To obtain the key carotenoid metabolic genes, the differentially expressed genes (DEGs) analysis were performed by DESeq package with the parameters that false discovery rate (FDR) < 0.01 and a fold change (FC) ≥2. Then we used Blast2GO with the cut-off of E-value <
1×10⁵ to classify DEGs into functional categories and assigned a candidate ko number by searching the KEGG pathway database³¹.

**Gene expression analysis by semi-quantitative reverse transcriptase-polymerase chain reaction and quantitative real-time PCR analysis**

Semi-quantitative reverse transcriptase-polymerase chain reaction (Semi-quantitative RT-PCR) was performed to analyze the expressed profiles of all carotenoid metabolic genes in ray and disc florets of chrysanthemum. Firstly, total RNAs of the ray (R1-R5) and disc florets (D1-D5) of chrysanthemum cultivar ‘Dong Li Fen Gui’ were extracted. Subsequently, first-strand cDNA was synthesized by the transcription kit for semi-quantitative RT-PCR. The procedure for semi-quantitative RT-PCR was following the method of previous study²⁷. Then the expression levels of carotenoid metabolic genes, including *CmDXS* (EVM0005576), *CmDXR* (EVM0031358), *CmlIPI* (EVM0051920), *CmGGPS* (EVM0014663), *CmPSY1* (EVM0027963), *CmPSY2* (EVM0033788), *CmPDS* (EVM0005023), *CmZDS* (EVM0015447), *CmLCYB* (EVM0028157), *CmLCYE* (EVM0039265), *CmCHYE* (EVM0048287), *CmCHYE* (EVM0040022), *CmVDE* (EVM0007951), *CmZEP* (EVM0021607), *CmCCD1* (EVM003018), *CmCCD4a-2* (EVM0019666), were detected. Meanwhile, we have quantified the intensity of the electrophoretic bands obtained from semi-quantitative RT-PCR by the Tanon GIS system and used these data to construct heatmap by BMKCloud for visually displaying the expression patterns of above carotenoid metabolic genes.

Quantitative real-time PCR analysis (qRT-PCR) was performed to analyse the expression patterns of key carotenoid metabolic genes. According to the SYBR Premix Ex Taq kit (Japan, Takara), qRT-PCR analysis was performed on a CFX96™ real-time system (Bio-Rad Laboratories, USA). Relative expression levels were calculated using the 2⁻ΔΔCT method³². As β-actin was stably and constitutively expressed in most tissues and cells, it was widely used in rice (*Oryza sativa*)³³, maize (*Zea mays*)³⁴, and wheat (*Triticum aestivum*)³⁵ as the reference gene. In this study, we regarded β-actin as the preferred reference gene for semi-quantitative RT-PCR and qRT-PCR analysis. Besides, to further verify the accuracy of the expression patterns, we also used SAND (SAND family protein) and F-box (F-box protein), which were reported as stable expressed genes in chrysanthemum³⁶,³⁷, as new reference genes for normalizing the qRT-PCR data. The specific primer sequences of semi-quantitative RT-PCR and qRT-PCR were listed in Table S1.

**Yeast one-hybrid screening**

The promoter of the *CmCCD4a-2* gene was inserted into the pAbAi vector to form the bait construct pAbAi-*CmCCD4a-2* promoter. Subsequently, the obtained bait construct was integrated into the Y1HGold yeast strain to create bait strain including promoter sequence of *CmCCD4a-2*. This bait yeast strain was plated on the SD/-Ura medium and SD/-Ura+100/150/200/300 ng/ml
Aureobasidin A (AbA) medium to determine the minimum inhibitory concentration of AbA. Then SMART cDNA synthesis technology was used to construct a cDNA library of two types of chrysanthemum florets. This cDNA library was transformed into the bait yeast strain and inoculated on SD/-Leu+200ng/ml AbA medium for selection.

For the Y1H assay, CDS of CmUIF1, CmPI, and CmAP3 were inserted into pGADT7 plastid to produce recombinants AD-CmAP3, AD-CmPI, and AD-CmUIF1 respectively. Then above recombinants and pGADT7 were transformed into bait yeast strains integrating pAbAi-CmCCD4a-2 promoter and empty pAbAi vector (as negative control in this study) respectively. The transformed yeast strains were dotted on the SD/-Leu medium and SD/-Leu+200ng/ml AbA medium subsequently. Interactions between TFs and the bait sequence were observed after 2 days of incubation at 30 °C.

**Gene isolation and phylogenetic analysis**

The coding sequence (CDS) of CmAP3, CmPI, and CmUIF1 were amplified, and the sequences were aligned by the MUSCLE algorithm in MEGA with some amino acid sequences from Arabidopsis, gerbera, and other plants. Based on the maximum likelihood (ML) estimation method, phylogenetic trees of CmAP3, CmPI and CmUIF1 were constructed subsequently. Tree nodes were evaluated by the bootstrap method for 2,000 replicates. Evolutionary distances were computed using the p-distance method. After eliminating gaps and missing data, the phylogenetic trees of the CmAP3, CmPI and CmUIF1 were constructed.

**Subcellular localization**

The CDS of CmUIF1, CmPI, and CmAP3 were amplified and recombined into pBI121-eGFP to produce CmUIF1-GFP, CmPI-GFP, and CmAP3-GFP constructs. Agrobacterium tumefaciens strain GV3101 transformed with above recombinants or pBI121, which was regarded as negative control, were infiltrated into the leaves of Nicotiana benthamiana. The confocal laser scanning microscopy (Leica TCS SP8, Wetzlar, Germany) was used to record the expression of GFP at 48 hours after infiltration.

**Yeast two-hybrid assay**

The CDS of CmAP3, CmPI, and CmUIF1 were recombined into either pGADT7 or pGBK7 to form prey recombinants AD-CmAP3/AD-CmPI/AD-CmUIF1, or bait recombinants BD-CmAP3/BD-CmPI/BD-CmUIF1. The various prey and bait recombinants were co-transformed into Y2H strain and the transformed yeasts were dotted on the SD/-Trp/-Leu medium, SD/-Trp/-Leu/-His/-Ade+3AT medium, and SD/-Trp/-Leu/-His/-Ade+3AT+X-a-Gal medium simultaneously. The two proteins were thought to interact with each other if the transformed yeast cells could grow well on all medium and turn blue on selective medium with
Bimolecular fluorescence complementation assay

BiFC assays were performed to confirm the interaction between CmAP3, CmUIF1 and CmPI TFs. The CDS of CmPI and CmUIF1 were recombined into pSPYCE to generate YCE-CmPI and YCE-CmUIF1 constructs. Meanwhile, CDS of CmAP3 and CmUIF1 was inserted into pSPYNE173. Then GV3101 containing above recombinants or empty vectors such as pSPYCE and pSPYNE173 were co-infiltrated into the leaves of N. benthamiana. The fluorescence signals were discovered by a laser scanning confocal microscope at 48 hours after infiltration (Leica TCS SP8, Wetzlar, Germany).

Yeast three-hybrid assay

As CmPI can form heterodimers with CmAP3 and CmUIF1 respectively, Y3H assay was performed to verify whether CmAP3, CmPI and CmUIF1 would form a protein complex essential for transcriptional regulation. The procedure is as follows: firstly, CDS of CmPI was amplified and inserted into the MCS II site of the pBridge plastid expressed only in the absence of methionine to produce pBridge-CmPI construct. The CDS of CmAP3 and CmUIF1 with EcoRI and BamHI restriction sites were cloned into MCS I site of the pBridge-CmPI construct to obtain the pBridge-CmPI-CmAP3 and pBridge-CmPI-CmUIF1 constructs respectively. The construct pBridge-CmPI-CmAP3/pBridge-CmPI-CmUIF1 was co-transformed with pGADT7 or AD-CmUIF1/AD-CmAP3 into the yeast strain Y2H and incubated on SD/-Leu/-Trp medium. Finally, the transformed yeast strains were dotted at the SD/-Trp/-Leu+200ng/ml AbA medium to assess the CmAP3 and CmUIF1 interaction without the existence of CmPI, and on the SD/-Trp/-Leu/-Met+200ng/ml AbA medium to assess the CmAP3 and CmUIF1 interaction with the existence of CmPI.

Dual luciferase reporter assay

DLR assay was performed to detect transcriptional activation of AP3-PI-UIF1 TF complex on target promoters. The full-length sequences of CmAP3, CmPI, and CmUIF1 were recombined into the overexpression plastid pGreenII0029 62-SK, which were named as SK-CmAP3, SK-CmPI and SK-CmUIF1. The Promoter sequence of CmCCD4a-2 was amplified from genomic DNA, then this sequence was inserted into pGreenII 0800-LUC. GV3101 which contained SK-CmAP3 or/and SK-CmPI or/and SK-CmUIF1 were mixed with the LUC-CmCCD4a-2 promoter (10:1; v:v) and the mixture was infiltrated into the leaves of N. benthamiana. Luciferase activity was tested by the Dual-luciferase Reporter Assay System (Promega) by PerkinElmer EnVision.

DNA-affinity purification sequencing

DAP-seq was performed as the procedure which was published by O’Malley et al\textsuperscript{38} and Bartlett et al\textsuperscript{39}. Firstly, genomic DNA of the leaves and flowers of Chrysanthemum nankingense,
one of the key progenitors of domesticated chrysanthemum\textsuperscript{40}, was extracted as the procedure of the DNeasy plant mini kit (Qiagen, Germany). The gDNA was broken to fragments with an average of 200 bp and attached a short DNA sequencing adaptor. Meanwhile, the coding sequences of \textit{CmAP3}, \textit{CmPI}, \textit{CmUIF1} were cloned into Halo Tag expression vector to generate the recombinants respectively. \textit{Escherichia coli} strain containing the recombinants were cultivated and the fusion proteins were purified by Magne HaloTag Beads (Promega, USA). Mixture of Halo-CmAP3/CmPI/CmUIF1 and the Magne HaloTag Beads was incubated with 200 ng of fragmented gDNA on a rotator for 1h at room temperature. After incubation, gDNA fragments enriched from beads were amplified by PCR and sequenced on the Illumina Novaseq 6000 platform. DAP-seq peaks were found by MACS2 accompanied with default parameters. BEDtools was performed to detect the distribution of DAP-seq peaks at the whole genome level. Motif discovery was performed by the MEME-ChIP suite 5.0.5\textsuperscript{41}.

Statistical analysis

All data were presented as the mean ± SD (error bars indicated the standard deviations of the means) from at least three biological replicates. Statistical analysis was performed by Student’s t-test and significant differences between means were defined with $P<0.01$ (**) and $P<0.05$ (*).

Results

Carotenoid composition and content analysis of ray and disc florets of ‘Dong Li Fen Gui’

Carotenoids are a group of important pigments affecting the flower coloration of chrysanthemum\textsuperscript{36}. HPLC-MS/MS was used to clarify whether the different flower coloration of ray and disc florets of ‘Dong Li Fen Gui’ was caused by accumulating variable levels of carotenoids. We found that both ray florets and disc florets were able to accumulate carotenoids with similar content and composition, mainly including violaxanthin, lutein, and β-carotene, at the early developmental stage of the capitulum (S1) (Fig. 1b). As the capitulum developed, the total carotenoids of the ray florets gradually decreased, only trace amounts of carotenoids accumulated at the S5 stage (0.025 μg/g). While the total content of carotenoids increased significantly in disc florets, reaching a maximum of 48.974 μg/g at the S4 stage and 41.295 μg/g at the S5 stage, which was 4.960- and 4.182-fold higher than that at the S1 stage. Furthermore, lutein and its isomers such as (9Z)-lutein, (9’Z)-lutein, and (9’Z)-lutein-5,6-epoxide were the main carotenoid components in the disc florets at the S4 and S5 stages, accounting for 76.41% and 75.17% of the total carotenoids, respectively (Fig. 1b).
Fig. 1 Characteristics of capitulum developmental stages of ‘Dong Li Fen Gui’ and carotenoid components and contents analysis. a. The development of chrysanthemum capitulum of ‘Dong Li Fen Gui’ was divided into five stages (S1-S5). R1-R5 and D1-D5 represented ray florets and disc florets of ‘Dong Li Fen Gui’ at different development stages. Bar, 1 cm. b. Carotenoid components and contents analysis of ray florets (R1-R5) and disc florets (D1-D5) of ‘Dong Li Fen Gui’ at different development stages. The data were presented as the mean ± SD from at least three biological replicates. (*, P < 0.05; **, P < 0.01; Student’s t-test).

Screening for candidate carotenoid metabolic genes by comparative transcriptomics and gene expression analysis

To screen the key carotenoid metabolic genes of which differential expression resulted in the
distinct content of carotenoids accumulated in the ray florets and disc florets, RNA samples derived from these two types of florets of anemone-type chrysanthemum cultivar ‘Dong Li Fen Gui’ at the S3 stage were analyzed by RNA sequencing (Fig. 2a-b). A total of 54,603 unigenes were assembled and annotated against public databases: Nr (38,527/54,603; 70.55%), COG (17,315/54,603; 31.71%), Pfam (41,697/54,603; 76.36%), KOG (29,690/54,603; 54.37%), Swiss-Prot (38,285/54,603; 70.11%), KEGG (19,807/54,603; 36.27%), and GO (29,534/54,603; 54.09%). Then the transcriptome data of these two types of florets were compared (R3 vs. D3), and we have identified 4,329 differentially expressed genes (DEGs) eventually (Fig. S1a). Gene ontology (GO) project was carried out to clarify the function of the DEGs. 4,329 DEGs were classified into 20, 16, and 15 categories and assigned to biological processes, cellular components, and molecular functions (Fig. S1b-c). KEGG enrichment analysis was carried out to further understand the biological functions of these DEGs through pathway classification and enrichment. The pathways that starch and sucrose metabolism (ko00500), carbon metabolism (ko01200), and the biosynthesis of amino acids (ko01230) were the most DEG enrichment. In addition, some DEGs, such as CmDXS (EVM0005576), CmVDE (EVM0007951), and CmCCD4a-2 (EVM0019666), were also enriched in carotenoid biosynthesis (ko00906) (Fig. S1d-e).

To validate the RNA-Seq data and identify the key structural genes, all carotenoid metabolic genes were obtained in this study (Fig. 2c), and semi-quantitative RT-PCR was performed to measure the expressed patterns of these genes during capitulum development. The results showed that only CmCCD4a-2, encoding carotenoid cleavage dioxygenase in chrysanthemum, was expressed differently between ray florets and disc florets at all developmental stages. This key structural gene highly expressed in ray florets of ‘Dong Li Fen Gui’, while it was almost not expressed in the disc florets, showing a negative correlation with carotenoid accumulation (Fig. 3a; Fig. S2), and qRT-PCR verified this result (Fig. 3b; Fig. S3). Sequence alignment and phylogenetic analysis showed that CmCCD4a-2 had high homology with CmCCD4a, which made the ray florets of chrysanthemum appear white by cleaving carotenoids into colorless compounds (Fig. S4). Thus, we regarded CmCCD4a-2 as the key carotenoid metabolic gene causing distinct carotenoid contents to accumulate in the ray florets and disc florets of ‘Dong Li Fen Gui’.

Furthermore, chrysanthemum cultivars from different color groups (Fig. S5a) were selected to verify whether CmCCD4a-2 had the same expression pattern in the ray and disc florets of other chrysanthemum cultivars. Through carotenoid content and gene expression analysis, we found large amount of carotenoids were accumulated in the disc florets of white- and pink-colored chrysanthemum cultivars, while only low transcript levels of CmCCD4a-2 were detected. In contrast, CmCCD4a-2 was found that highly expressed in ray florets, which contained only trace amounts of carotenoids (Fig. S5b-c; Fig. S3). These results indicated that CmCCD4a-2, of which
differential expression caused the distinct content of carotenoids to accumulate in the ray and disc florets, was the most important carotenoid metabolic gene in chrysanthemum.

Fig. 2 Key carotenoid metabolic genes identified based on comparative transcriptomics analysis. a. The ray and disc florets of ‘Dong Li Fen Gui’ at the S3 stage were used for RNA sequencing. b. Heatmap analysis of all obtained genes based on the transcriptome data. C. Transcripts of all carotenoid metabolic genes in ray and disc florets of ‘Dong Li Fen Gui’ at the S3 stage based on the transcriptome data.
Fig. 3 Gene expression analysis of *CmCCD4a*-2 in the ray and disc florets of ‘Dong Li Fen Gui’ at different developmental stages. **a.** Identifying the key carotenoid metabolic genes that caused the distinct content of carotenoids to accumulate in the ray and disc florets of ‘Dong Li Fen Gui’ by heatmap analysis using the semi-quantitative RT-PCR data. **b.** The expression profile of *CmCCD4a*-2 was analyzed by qRT-PCR. All data were presented as the mean ± SD from at least three biological replicates. (*, P < 0.05; **, P < 0.01; Student’s t-test).

**Yeast one-hybrid screening and candidate TFs identifying**

To identify TFs that directly regulated the expression of *CmCCD4a*-2, Y1H screening was performed in this study. Firstly, the upstream sequence with a total length of 1352 bp, which prior to the ‘ATG’ of *CmCCD4a*-2, was isolated from ‘Dong Li Fen Gui’ by genome walking method. As TATA-box was a core promoter element and it was close to transcription initiation site, the sequence ranged from 5’-end (‘ACCCTTCAA’) to TATA-box was defined as the promoter of *CmCCD4a*-2 (Fig S6a). Possible cis-acting elements, including ABA-responsive elements ABRE motif, MYB-binding motif, MYC motif, GARP binding motif and MADS-box binding motif, were predicted by PlantCare and PlantPAN 3.0 databases (Fig. S6a-b; Table S2). Then this promoter sequence was integrated into the Y1HGgold yeast strain to create bait yeast strain. We have plated the bait yeast strain on SD/-Ura medium with 100/150/200/300 ng/ml AbA respectively and found 200 ng/ml AbA was the minimum inhibitory concentration for the pAbAi-*CmCCD4a*-2 promoter bait strain (Fig. S7). Subsequently, cDNA library was transformed into this bait yeast strain and inoculated on SD/-Leu+200ng/ml AbA medium for selection. Thirteen positive clones were obtained successfully through Y1H screening, and two proteins (CmAP3 and CmUIF1) were identified and regarded as candidate TFs that might directly bind to the promoter of *CmCCD4a*-2 (Table S3). Phylogenetic analysis revealed that CmAP3 belonged to the class B DEFICIENS/APETALA3 subfamily of MADS TFs and was closely linked with the *C. lavandulifolium* AP3 protein (Fig. S8a). Through multiple sequence alignment, we found that
CmAP3 contained highly conserved domains (MADS domain and K-box region) that were typical of the MADS TFs (Fig. S8c). In addition, CmUIF1 that belonged to the Golden 2-like subfamily of GARP TFs, was highly homologous with AtUIF1 (Fig. S8b, S8d), which modulated flower development by binding to the WUS and AG promoters directly in Arabidopsis\textsuperscript{23}. Then the subcellular localization of GFP fusions with CmAP3 and CmUIF1 was examined. All GFP fluorescence was detected in the nucleus (Fig. 4a), indicating that the CmAP3 and CmUIF1 proteins, as typical transcription factors, localize in the nucleus. Gene expression analysis by qRT-PCR have showed that CmAP3 expressed higher levels in ray florets of ‘Dong Li Fen Gui’ than in disc florets, while the expressed level of CmUIF1 was not significantly different in these two types of florets (Fig. 4b; Fig. S3).

Y1H assays were performed to detect the interaction between the CmAP3/CmUIF1 and the CmCCD4a-2 promoter. We found that bait yeast cells containing either AD or the recombinant vector AD-CmAP3/AD-CmUIF1 could grow well on synthetic dropout medium without leucine (SD/-Leu). Only the bait yeast cells containing recombinant vector AD-CmAP3 or AD-CmUIF1 could survive on selective medium supplemented with 200 ng/ml Aureobasidin A (SD/-Leu+200 ng/ml AbA) (Fig. 4c). Meanwhile, AD-CmAP3 and AD-CmUIF1 were also transformed into Y1HGOLD yeast strain integrating empty pAbAi vector (as negative control in this study) and we found these transformed yeast cells cannot grow well on SD/-Leu+200 ng/ml AbA medium (Fig. S9), which indicated that the interaction between CmAP3 and CmUIF1 with the CmCCD4a-2 promoter was specific and functional. To identify the binding sites of CmAP3 and CmUIF1, all the MADS-box and GARP binding motifs predicted by PlantCare and PlantPAN 3.0 (Fig. S6a-b; Table S2) were compared with the binding sites of AtAP3 (Arabidopsis thaliana), AqAP3-3 (Aquilegia coerulescens) and AtUIF1 identified by ChIP-seq data. These reported CHIP-seq data indicated that CA-repeat boxes, including C(A/T)\textsubscript{3}G or CC(A/T)\textsubscript{5}GG, could recruit AP3 TF in Arabidopsis\textsuperscript{42} and clematis\textsuperscript{43}, while UIF1 preferred to bind the motif (AGA(A/T)TC) in Arabidopsis\textsuperscript{23}. Thus, we have speculated that CmAP3 and CmUIF1 TFs could interact with the CmCCD4a-2 promoter by binding to the CA-repeat box (‘CTTAAAAAG’; ‘CTTAAATAAG’) and UIF1-binding motif (‘AGAATC’; ‘AGATTC’) respectively. However, dual luciferase assays showed that single infiltration of CmAP3, or CmUIF1 can not induce any CmCCD4a-2 promoter activity (Fig. 4d). Transcription factors often perform regulatory functions by forming protein complexes\textsuperscript{7,19}. Y2H assay was carried out to test the protein-protein interaction (PPI) between CmAP3 and CmUIF1. Yeast cells containing both CmAP3 and CmUIF1 recombinant vectors was not able to grow on the selective medium supplemented with 3AT and X-\textalpha-gal, as shown in Fig. S10a, indicating that CmAP3 could not interact with CmUIF1 directly. BiFC assay verified the above result (Fig. S10b).
Fig. 4 Candidate upstream transcription factors CmAP3 and CmUIF1 identified. a. Subcellular localization of CmAP3 and CmUIF1 in tobacco leaves. The green GFP fluorescence signal was imaged 48 hours after injection by a laser scanning confocal microscope. b. The expression profile of CmAP3 and CmUIF1 were analyzed by qRT-PCR. c. Y1H showed the binding of CmAP3 and CmUIF1 to the CmCCD4a-2 promoter. d. Effect of CmAP3 and CmUIF1 on the activity of the CmCCD4a-2 promoter by dual-luciferase assay. All data were presented as the mean ± SD from at least three biological replicates. (*, P < 0.05; **, P < 0.01; Student’s t-test).

CmPI, an intermediate “bridge” protein connecting CmAP3 and CmUIF1

As one of the most important TFs determining floral organ identity, PI protein tends to interact with AP3 to form heterodimers to perform their regulatory roles. Meanwhile, PI, AP3 and UIF1, all of which can regulate the petal development in plants. We have speculated that PI could act as an intermediate protein that connects CmAP3 and CmUIF1. To verify the above hypothesis, here, the full-length sequence of CmPI was cloned and phylogenetic analysis revealed that CmPI belonged to the class B GLOBOSA/PISTILLATA subfamily of MADS TFs and was closely linked with the C. lavandulifolium PI protein (Fig. S11a). Through multiple sequence alignment, we have found that CmPI contained highly conserved domains that were typical of the MADS TFs (Fig. S11b). qRT-PCR showed that CmPI expressed higher levels in ray florets of ‘Dong Li Fen Gui’ than that in disc florets, which was similar as the expressed pattern of CmAP3 (Fig. 5a; Fig. S3). However, unlike CmAP3 and CmUIF1, CmPI was not able to interact with the
CmCCD4a-2 promoter directly (Fig. 5b). Y2H assays showed that yeast cells cotransformed with both CmPI and CmAP3 constructs, CmPI and CmUIF1 constructs grew well on the selective medium supplemented with 3AT and X-α-gal and turned blue (Fig. 5c), suggesting that CmPI could interact with CmAP3 and CmUIF1 separately. BiFC assays verified the above results (Fig. 5d). It was declared that CmPI would be an intermediate “bridge” protein connecting CmAP3 and CmUIF1 in chrysanthemum.

![Fig. 5 Intermediate “bridge” protein CmPI identified.](image)

**Fig. 5 Intermediate “bridge” protein CmPI identified.** a. The expression profile of CmPI was analyzed by qRT-PCR. The data were presented as the mean ± SD from at least three biological replicates. (*, P < 0.05; **, P < 0.01; Student’s t-test). b. Y1H assay showing CmPI was not able to interact with the CmCCD4a-2 promoter. c. Protein interaction analysis between CmPI and CmAP3/CmUIF1 by Y2H assay. d. BiFC assay confirmed the interaction between CmPI and CmAP3/CmUIF1 in Nicotiana benthamiana leaves. YFP fluorescence signal was imaged 48 hours after injection by a laser scanning confocal microscope.

**CmAP3-CmPI-CmUIF1 TF complex activates the promoter of CmCCD4a-2**

Y3H assay was used to further verify the interaction between CmAP3, CmPI and CmUIF1. Yeast cells containing both Bridge-CmPI-CmUIF1 and AD-CmAP3 constructs and Bridge-CmPI-CmAP3 and AD-CmUIF1 constructs grew well on SD/-Trp/-Leu/-Met medium.
added with 200 ng/ml AbA (SD/-Trp/-Leu/-Met+200 ng/ml AbA), while they could not grow on SD/-Trp/-Leu medium added with 200 ng/ml AbA (SD/-Trp/-Leu+200 ng/ml AbA) (Fig. 6a), which indicated that CmAP3 could interact with CmUIF1 indirectly only if CmPI existed. CmPI was regarded as a physical “bridge” that connected CmAP3 and CmUIF1 to form the CmAP3-CmPI-CmUIF1 TF complex.

Dual luciferase assays were conducted to investigate the effect of CmAP3-CmPI-CmUIF1 TF complex on the activity of the CmCCD4a-2 promoter. The relative luciferase expression (LUC/REN ratios) was significantly higher in the presence of CmAP3-CmPI-CmUIF1 TF complex than in the control (Fig. 6b), indicating that CmAP3-CmPI-CmUIF1 heterotrimer could effectively activate the promoter of CmCCD4a-2 in chrysanthemum. Thus, we proposed that the CmAP3-CmPI-CmUIF1 TF complex existed abundantly in the ray florets of anemone-type chrysanthemum cultivar ‘Dong Li Fen Gui’ could bind to CmCCD4a-2 promoter directly and activate its transcription, which resulted in only trace amounts of carotenoids was accumulated in the ray florets.

Fig. 6 Effect of CmAP3-CmPI-CmUIF1 TF complex on the activity of the CmCCD4a-2 promoter. a. Y3H was carried out to detect the interaction among CmAP3, CmPI and CmUIF1. b. Dual-luciferase assay showed CmAP3-CmPI-CmUIF1 heterotrimer effectively activated the promoter of CmCCD4a-2. All data were presented as the mean ±SD from at least three biological replicates. (*, *P < 0.05; **, *P < 0.01; Student’s t-test).

**Genome-wide binding site analysis of CmAP3-CmPI-CmUIF1 TF complex**

TF complexes usually modulate plant growth and development, fruit ripening, and secondary metabolism by controlling the transcription levels of multiple downstream structural or regulatory genes simultaneously. In this study, DNA affinity purification sequencing (DAP-seq), an in vitro DNA-TF binding assay that has been widely applied in plants for identifying downstream target
genes\textsuperscript{44,45}, was used to investigate putative downstream target genes regulated by the CmAP3-CmPI-CmUIF1 TF complex. Firstly, recombinantly expressed CmAP3, CmPI, and CmUIF1 were incubated with gDNA libraries of \textit{C. nankingense}, one of the key progenitors of domesticated chrysanthemum, and CmAP3/CmPI/CmUIF1-binding DNA fragments were enriched and sequenced using next-generation sequencing. Then, a total of 4,734/4,945/43,115 (CmAP3/CmPI/CmUIF1) peaks were obtained, with a total length of 1,027,630/988,744/13,662,116 bp and an average length of 217/199/316 bp. We analyzed the distribution of these peaks located in annotated genic regions and found that 78.31%/77.37%/67.41% (CmAP3/CmPI/CmUIF1) were enriched in intergenic regions, 8.91%/8.88%/8.23% within the gene promoter regions (2 kb upstream of TSS), 4.14%/4.85%/5.26% within the 0.5 kb downstream regions, 1.63%/1.40%/0.83% within exon regions, and 6.55%/7.14%/9.98% within intron regions (Fig. 7a). To identify the DNA motifs of CmAP3, CmPI, and CmUIF1, we applied MEME-Chip and performed de novo motif discovery in the sequences of the DAP-seq peaks. CarG boxes, including C(A/T)$_8$G and CC(A/T)$_6$GG, were most centrally enriched for both CmAP3 and CmPI (Fig. 7b-c), which was consistent with previous reports\textsuperscript{42,43}. Furthermore, significant putative DNA binding motifs AGA(A/T)TC(A/C/T), which were quite similar to that of AtUIF1\textsuperscript{23}, were identified for CmUIF1 (Fig. 7d).

The putative target genes related to CmAP3/CmPI/CmUIF1-binding peaks were integrated, and a total of 13,428 genes were eventually identified (Fig. 7e). KEGG analysis was performed to further understand the biological functions of these target genes (Fig. 7f). \textit{PSY1} (CHR00034017), \textit{PSY2} (CHR00091590), \textit{CRTISO} (CHR00019807), \textit{LCYB} (CHR00027913), \textit{CHYE} (CHR00027078), \textit{VDE} (CHR00000429), and \textit{ZEP} (CHR00026800), enriched in the carotenoid biosynthesis pathway (ko00906), were identified in this study (Fig. S12a; Table S4), which indicated that the CmAP3-CmPI-CmUIF1 TF complex might be involved in carotenoid biosynthesis in chrysanthemum. \textit{CHS1} (CHR00038317), \textit{CHS2} (CHR00047213), \textit{CHS3} (CHR0008801), \textit{CHI} (CHR00036904), \textit{F3H} (CHR00037184), \textit{FLS} (CHR00000448), \textit{F3′H} (CHR00050525), and \textit{DFR} (CHR00058078), participating in flavonoid biosynthesis, were also directly bound by the CmAP3-CmPI-CmUIF1 TF complex (Fig. S12b; Table S4). Furthermore, we also found that this TF complex might be involved in petal development in chrysanthemum by directly binding flower symmetry \textit{CYC2}-like genes such as \textit{CYC2a} (CHR00035692, CHR00048990), \textit{CYC2b} (CHR00065022), \textit{CYC2d} (CHR00093761), and \textit{CYC2f} (CHR00074698) (Fig. S13; Table S4).
Fig. 7 Downstream target genes regulated by the CmAP3-CmPI-CmUIF1 TF complex were identified by DAP-seq. a. Genome-wide analysis of the CmAP3/CmPI/CmUIF1-binding peaks. Conserved motifs of CmAP3 (b), CmPI (c), and CmUIF1 (d) identified from the DAP-seq data. e. Venn diagram showing peak overlap between the putative target genes related to CmAP3, CmPI, and CmUIF1-binding peaks. f. KEGG enrichment analysis of all putative target genes regulated by the CmAP3-CmPI-CmUIF1 TF complex.

Discussion

*CmCCD4a-2* is the most important carotenoid metabolic gene causing distinct carotenoid accumulation in ray and disc florets

Carotenoid biosynthetic genes in plants have been studied for more than 30 years, however, the role of carotenoid degradation has only been addressed more recently. Carotenoids are
cleaved by arotenoid cleavage dioxygenases (CCDs), and these enzymes are generically divided into five subfamilies, including NCED (nine-cis-epoxycarotenoid dioxygenase), CCD1, CCD4, CCD7, and CCD8. Currently, a great number of studies have declared that the differential expression of CCD1s and CCD4s results in distinct contents of carotenoids accumulated in various plant species and cultivars, such as in the different colored fruits of peach (PpCCD), summer squash (CpCCD and CpCCD), and in the different colored flowers of chrysanthemum (CmCCD), osmanthus (OfCCD), oncidium (OgCCD), and daffodil (NpCCD).

Our previous study found that CmCCD4a-2, which had high homology with CmCCD4a, was accompanied by CmLCYE and CmPAPs to simultaneously affect carotenoid accumulation in the ray florets of chrysanthemum cultivars of different colors. Here, we found that of all the carotenoid metabolic genes, CmCCD4a-2 was the only gene that was highly expressed in the ray florets of ‘Dong Li Fen Gui’ at all developmental stages, and it was almost not expressed in the disc florets, showing a negative correlation with carotenoid accumulation (Fig. 2c; Fig. 3a-b; Fig. S3). Furthermore, CmCCD4a-2 was also highly expressed in the ray florets of other white- and pink-colored chrysanthemum cultivars, while only low transcript levels of CmCCD4a-2 were detected in the disc florets, which accumulated a great number of carotenoids (Fig. S5a-c; Fig. S3). These results indicated that the expression pattern of CmCCD4a-2 was conserved and consistent in different colored chrysanthemum cultivars. Thus, we proposed that significantly differential expression levels of CmCCD4a-2 resulted in distinct carotenoid contents accumulated in the ray and disc florets of chrysanthemum. However, it was not clear why CmCCD4a-2 was differentially expressed in these two types of florets.

Trans-acting factors, particularly upstream TFs, affect the transcription level of genes by binding to cis-acting regulatory regions. Currently, only a few TFs were found to directly regulate the expression of CCD4. OfWRKY3 and OfERF61 in osmanthus activate OfCCD4 expression, resulting in arotenoid cleavage in petals. VvMADS4 in grapes negatively regulated the expression of VvCCD4b. However, until now, the underlying mechanism that TFs or TF complexes transcriptionally regulate the expression of CmCCD4a-2 in chrysanthemum remained poorly understood.

**CmA3-Cmp1-CmUIF1 TF complex modulates carotenoid metabolism by directly regulating the expression of CmCCD4a-2**

Through yeast one-hybrid (Y1H) screening and Y1H assays, we first demonstrated that CmA3 and CmUIF1 TFs could directly bind to CmCCD4a-2 promoter to perform regulatory functions (Fig. 4c; Table S3). Phylogenetic analysis revealed that CmA3 belonged to the class B DEFICIENS/APETALA3 subfamily of the MADS-box TF family, and CmUIF1 was regarded as a
member of the GARP TF family (Fig. S8a-b). Currently, many MADS TFs that directly regulate the carotenoid metabolic process have been characterized. For example, RIN could affect carotenoid biosynthesis by directly regulating the transcription levels of several carotenoid metabolic genes during tomato fruit ripening. Tomato FRUITFULL1 (FUL1) has also been shown to promote the expression of SLPSY1 by directly binding to its promoter. In citrus, CsMADS6 modulated the transcription levels of LCYB1 together with PSY and PDS to affect carotenoid accumulation. While the members of the GARP TF family have been widely reported to mainly participate in hormonal signaling, nutrient response and sensing processes, chloroplast biogenesis and plant development, particularly in leaves and flowers, few studies have published on GARP TFs regulating carotenoid metabolism.

To test the effect of CmAP3 and CmUIF1 on the activity of the CmCCD4a-2 promoter, dual luciferase assays were carried out in this study. However, the results showed that the CmCCD4a-2 promoter was not activated when CmAP3 or CmUIF1 existed alone (Fig. 4d). TFs usually regulate the secondary metabolism, including carotenoid metabolism, by forming the TF complexes. For example, WHITE PETAL1 (MtWP1) in Medicago truncatula interacted with MtTT8 and MtWD40-1 physically to promote carotenoid-derived flower pigmentation. CsMADS5, as a positive regulator, could interact with CsMADS6 to synergistically promote carotenoid accumulation. Therefore, we analyzed the relationship among CmAP3, CmUIF1, and CmPI, which belonged to the class B GLOBOSA/PISTILLATA subfamily of MADS-box TFs (Fig. S11) and were reported to interact with AP3 to determine the organ identity of petals and stamens. The results showed that CmAP3 was not able to interact with CmUIF1 (Fig. S10a-b), while CmPI, regarded as a physical “bridge”, could connect CmAP3 and CmUIF1 to form the CmAP3-CmPI-CmUIF1 TF complex in chrysanthemum (Fig. 5c-d; Fig. 6a). Furthermore, dual luciferase assays showed that the CmAP3-CmPI-CmUIF1 TF complex could effectively activate the promoter of CmCCD4a-2 (Fig. 6b). In summary, we proposed a working model: the CmAP3-CmPI-CmUIF1 TF complex existed abundantly in the ray florets of anemone-type chrysanthemum cultivar ‘Dong Li Fen Gui’ and could bind to CmCCD4a-2 promoter directly and activate its transcription effectively, which resulted in the accumulation of only trace or low amounts of carotenoids. While low-abundance transcripts of CmAP3 and CmPI, as important components of CmAP3-CmPI-CmUIF1 TF complex, were detected in the disc florets of ‘Dong Li Fen Gui’ (Fig. 4b; Fig. 5a). So there was not enough TF complex existed in the disc florets to effectively activate the expression of CmCCD4a-2, thereby accumulating a large number of carotenoids (Fig. 8).
**Fig. 8** The regulation model of the differential carotenoids accumulated in the ray and disc florets of *chrysanthemum*. CmAP3-CmPI-CmUIF1 TF complex existed abundantly in the ray florets could directly bind to the CmCCD4a-2 promoter and activate its transcription effectively, which resulted in the accumulation of only trace or low amounts of carotenoids. While low-abundance transcripts of CmAP3 and CmPI, as important components of CmAP3-CmPI-CmUIF1 TF complex, were detected in the disc florets. So there was not enough TF complex to effectively activate the expression of CmCCD4a-2, thereby accumulating a large number of carotenoids. The color saturation of the circles (deep/light) represented the expression levels of CmAP3, CmPI and CmUIF1 (high/low).

**CmAP3-CmPI-CmUIF1 TF complex might participate in multiple processes in chrysanthemum**

Currently, a number of TFs or TF complexes that simultaneously regulate multiple secondary metabolic processes have been characterized\(^{14,15,62}\). For example, AdMYB7, an R2R3-MYB TF in kiwifruit, regulated carotenoid accumulation via the transcriptional activation of the promoter of AdLCYB. Overexpression of the AdMYB7 gene in *N. benthamiana* altered chlorophyll accumulation and increased the expression of the chlorophyll biosynthesis genes NbGGR and NbSGR1\(^\text{15}\). WHITE PETAL1 (MtWP1) in *Medicago truncatula* interacted with MtTT8 and MtWD40-1 to activate the expression of both carotenoid and anthocyanin biosynthetic genes\(^\text{14}\). SlMYB72 in tomato interacted with the auxin response factor SIARF4 to regulate chlorophyll,
carotenoid, and flavonoid biosynthesis simultaneously. In this study, a large number of metabolic genes directly bound by the CmAP3-CmPi-CmUIF1 TF complex, including carotenoid biosynthetic genes such as PSY1, PSY2, CRTISO, LCYB, CHYE, VDE, and ZEP and flavonoid biosynthetic genes such as CHS1, CHS2, CHS3, CHI, F3H, FLS, F3′H, and DFR, were identified through DAP-seq (Fig. S12a-b; Fig. S13; Table S4). These results indicated that the CmAP3-CmPi-CmUIF1 TF complex might be involved not only in carotenoid degradation (as mentioned above, this TF complex could directly regulate the expression of CmCCD4a-2, one of the most important structural genes in carotenoid degradation) but also in the biosynthesis of carotenoids and flavonoids. However, the specific regulatory mechanism needs further investigation.

Furthermore, some transcription factors, which belonged to the CYCLOIDEA/TEOSINTE BRANCHED1 (CYC/TB1) or MADS-box TF families, were reported to participate in pigment metabolism, still play a very important role in flower and fruit development. In Torenia fournieri, TfCYC2 could directly bind to the regulatory regions of TfMYB1 to establish an asymmetric pigmentation pattern. When TfCYC2 was up- or downregulated, dorsal petal identity, petal shape, and corolla pigmentation pattern were simultaneously changed. CsMADS6 in citrus activated the transcription of CsLCYB1 along with CsPSY, CsPDS, and CsCCD1 to modulate carotenoid metabolism. Overexpression of CsMADS6 in tomato dramatically altered carotenoid accumulation, accompanied by changes in sepal morphology. Here, some flower development-related genes, particularly flower symmetry CYC2-like genes such as CYC2a (CHR00035692, CHR00048990), CYC2b (CHR00065022), CYC2d (CHR00093761), and CYC2f (CHR00074698) (Fig. S12; Table S4), were identified and regarded as potential target genes regulated by the CmAP3-CmPi-CmUIF1 TF complex based on DAP-seq data. Developmental genetic studies have found that CYC2-like genes are the key genes which participate in the regulation of organ symmetry in many eudicot plant groups. These CYC2-like members are also involved in capitulum architecture, allometric growth and fusion of corolla lobes, and stamen regression in Asteraceae. Among them, CYC2b and CYC2e were more likely to participate in regulating the petal ligule length of ray florets. CYC2c and CYC2g determined the position and zygomorphy of ray florets. CYC2d was proven to repress the growth of stamens and dorsal corolla lobes. Generally, the regulatory mechanism of CYC2-like genes is mainly focused on protein-protein interaction, while the regulatory networks upstream of these genes are poorly understood. Our data showed that the CmAP3-CmPi-CmUIF1 TF complex might serve as a potential upstream TF complex that directly bound to the regulatory regions of CYC2-like genes to affect the development of ray and disc florets in chrysanthemum. However, the specific regulatory mechanism needs further analysis.
Data availability statement
All data supporting the findings of this study are available within the paper and within its supplementary materials published online.

Acknowledgements

This research was supported by the National Key Research and Development Program (No. 2018YFD1000405) and the Beijing Natural Science Foundation of China (grant no. 6212022).

Conflict of interests

The authors declare no competing interests.

Contributions

HH and SLD conceived and designed this study. CFL and JPQ performed the experiments. CFL, CYD, and FZ carried out the data analysis. CFL, FYL and HH wrote this manuscript. All authors read and approved the final manuscript.

Supplementary data

Supplementary information accompanies the manuscript on the Horticulture Research website http://www.nature.com/hortres.

References

7. Sun, T., Li, L. Toward the ‘golden’era: the status in uncovering the regulatory control of carotenoid accumulation in plants. Plant Science, 290, 110331 (2020).


53. Han, Y. et al. Mechanism of floral scent production in *Osmanthus fragrans* and the production and regulation of its key floral constituents, β-ionone and linalool. *Horticulture research*, **6**(1), 1-12 (2019).


