

whole-transcriptome analysis of mutant and normal capitula of *Chrysanthemum morifolium*

Hua Liu¹, Chang Luo¹, Dongliang Chen¹, Yaqin Wang², Shuang Guo¹, Xiaoxi Chen¹,
Jingyi Bai¹, Mingyuan Li¹, Xinlei Huang¹, Xi Cheng¹, Conglin Huang^{1,*}

¹ Beijing Agro-Biotechnology Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing Engineering Research Center of Functional Floriculture, Beijing Key Laboratory of Agricultural Genetic Resources and Biotechnology, Beijing, 100097, China

² Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing, 100097, China

* Correspondence: conghlinh@126.com

Abstract:

Background

Chrysanthemum morifolium is one of the most economically important and popular floricultural crops in Asteraceae. Chrysanthemums have many different flower colors and shapes. However, the molecular mechanism controlling the development of chrysanthemum floral colors and shapes is still an enigma. We obtained a cut chrysanthemum variety with mutant capitula in which the ray florets became green and the inside pistils became vegetative buds, while normal capitula have many rounds of purple ray florets and few disc florets.

Results

We conducted whole-transcriptome analysis of differentially expressed genes (DEGs) between the mutant and normal capitula using third-generation and second-generation sequencing techniques. We identified DEGs between the mutant and normal capitula to reveal important regulators underlying their differential development. Regulatory genes involved in the photoperiod pathway and the control of floral organ identification as well as important functional genes in the anthocyanin synthesis pathway were also identified. Therefore, a list of candidate genes for studying flower development and anthocyanin synthesis in chrysanthemums was generated. Qualitative analysis of pigments in the florets of normal and mutant capitula revealed anthocyanins were synthesized and accumulated in the florets of normal capitula, but not in the florets of mutant capitula. It was indicated that pistils may be required for anthocyanin synthesis in chrysanthemums.

Conclusions

These results will help to elucidate the molecular mechanisms of floral organ development and will contribute to the development of techniques for studying flower shape and color regulation to promote breeding in chrysanthemum.

Key Words: *Chrysanthemum morifolium*; ray florets; pistils; flower development; mutant capitula; anthocyanin synthesis; whole-transcriptome analysis; differentially expressed genes

Introduction

Chrysanthemum morifolium is one of the most economically important and popular floricultural crops in Asteraceae, and ranks second in the cut flower trade after rose [1]. The head-like inflorescence (capitulum) resembling a single large flower is the main ornamental part of *C. morifolium*, and is also regarded as the key innovation behind the evolutionary success of the Asteraceae [2]. The typical capitulum of chrysanthemum is formed by two morphologically distinct florets: the marginal ray florets and the central disc florets. Ray florets have ligulate and zygomorphic colorful corollas (petals) and aborted stamens, which function in attracting pollinators. The disc florets have radially symmetrical colorless corollas and their fertile pollens are hermaphroditic and are used for reproduction in chrysanthemum (Additional file 1). The colors and shapes of the flowers are the most visible and amazing products of evolution, and also connect humans to nature [3].

Flowering is a key developmental process in the plant life cycle that is very complex and is controlled by endogenous factors and environmental cues. As currently understood, flower development contains three phases: flowering determination, flower evocation, and floral organ development [4]. In *Arabidopsis*, tremendous progress has been achieved toward understanding the molecular mechanisms involved in flower development [5, 6]. ABCE models have revealed that A-class together with E-class genes specify sepal identity, A-class, B-class and E-class genes specify petals, B-class, C-class and E-class genes determine stamens, and C-class and E-class genes determine carpel/gynoecium organ identity [7]. With the notable exception of A-class genes, all of these genes belong to the MADS-box family of transcription factors including the *AP1*, *AP3*, *PI*, *AG* and *SEP* genes.

The splendid colors presented mainly by flower petals have enabled plants to constantly develop new showy traits and prosper throughout millions of years of evolution. Anthocyanins and carotenoids are the two major groups of pigments generated in plant petals. Anthocyanins are accumulated in the vacuoles of petal epidermal cells and confer orange-to-violet colors in flowers [8]. Beside attracting pollinators, anthocyanins also protect against damage from UV irradiation [9]. Anthocyanins provide chrysanthemum ray florets with colorful bright colors to attract pollinators, which improves the success rate of cross pollination between different species or varieties and promotes cultivar groups with large varieties of flower types in *C. morifolium*. Anthocyanins enhance the ornamental value of chrysanthemums, and many cut flower and pot flower varieties with bright colors are on sale every year to satisfy the demand of the market. Understanding the mechanism of anthocyanin biosynthesis and its regulation will contribute to the cultivation and improvement of new color varieties of chrysanthemums.

In chrysanthemums, a few regulatory genes involved in flower development have been isolated, such as MADS-box, TCP, and WUS-like genes[10-12, 16]. Some important functional genes and transcription factors involved in the anthocyanin biosynthesis pathway have also been characterized including ANS, F3'H, F3H and MYB-like genes [13-16]. However, chrysanthemums have complex capitula containing two morphologically distinct florets and through a long period of breeding a variety of flower shapes and diverse colors have been created. The mechanism involved in chrysanthemum flower evolution and development is extremely complicated and little is known about it.

The development of RNA-seq technology has greatly improved transcriptomic studies in

chrysanthemums [1]. However, the read-length offered by second-generation high-throughput sequencing platforms is much shorter than the typical length of a eukaryotic mRNA. Additionally, the differences in transcript abundance and the presence of different isoforms make the assembly of transcriptomes from short reads extremely challenging [17]. Despite these problems, Hideki Hirakawa et al. performed de novo whole-genome assembly in *C. seticuspe* using the Illumina sequencing platform and Chi Song et al. sequenced the diploid *C. nankingense* genome using the Oxford Nanopore long-read technology; however, no more than 40% of the transcriptome sequencing reads from *C. morifolium* can be mapped to each of these two genome sequences, probably because of the extreme variation in chromosome ploidy and biological characteristics [18, 19]. Third-generation sequencing technology has dramatically increased the length of sequencing reads, which enables the precise location and sequencing of repetitive regions and isoforms within a single read.

Recently, we obtained a mutant plant of the cut chrysanthemum variety *C. morifolium* 'ZY' with both mutant and normal capitula. In the mutant capitula the ray florets became green and the inside pistils became vegetative buds, while the normal capitula had many rounds of purple ray florets and few disc florets. In this study, we analyzed a mixed sample of normal and mutant flowers, leaves, stems and roots of 'ZY' with the single-molecule long-read sequencing technology from Pacific Biosciences (PacBio). Based on the results, transcriptional sequencing and analysis of the mutant and normal capitula were performed using second-generation sequencing technology and RNA-Seq quantification analysis. Thus, we combined third-generation and second-generation sequencing to generate a

more complete/full-length *C. morifolium* transcriptome.

Based on the transcriptional sequencing and analysis, we identified differentially expressed genes (DEGs) between mutant and normal capitula to reveal important regulators controlling their differential development. Regulatory genes involved in the photoperiod pathway and the control of floral organ identification as well as important functional genes in the anthocyanin synthesis pathway were also identified to create a list of candidate genes for studying flower development and anthocyanin synthesis in chrysanthemums. These results will be helpful for elucidating the molecular mechanisms of floral organ development and will contribute to the development of techniques for studying flower shape and color regulation, as well as breeding and molecular biology in chrysanthemum.

Results

Sequencing and assembly

As shown in Figure 1, in the mutant capitula the ray florets became green and the inside pistils became vegetative buds, while the normal capitula had many rounds of purple ray florets and few disc florets. We analyzed a mixed sample of normal and mutant capitula, leaves, stems and roots of 'ZY' using PacBio sequencing and then analyzed the normal and mutant capitula separately using the Illumina paired-end sequencing technology. The resulting sequences were assembled into 130,097 isoforms with an N50 of 3013 bp and average length of 2510 bp (Table 1).

Gene annotation and functional classification

A total of 118,589 isoforms were annotated by BLAST in at least one of the four databases searched (nr, Swiss-Prot, KOG, and KEGG), leaving 11,508 (8.85%) isoforms without annotation. In total, 118,043, 101,048, 87,630, and 54,245 isoforms were annotated in the nr, Swiss-Prot, KOG, and KEGG databases, respectively. Gene Ontology (GO) has three ontologies—molecular functions, cellular components, and biological processes—that facilitate gene annotation and analysis. A total of 36,144 isoforms were classified into 47 functional categories, including 19, 17 and 11 GO categories in the biochemical process, cellular component and molecular function ontologies, respectively. The dominant categories in the biochemical process, molecular function and cellular component ontologies were ‘metabolic process’ (20,871), ‘catalytic activity’(22,818) and ‘cell’ (11,887), respectively, indicating that numerous metabolic activities were activated during the development of chrysanthemum capitula and that this process was regulated by a wide range of genes that interacted within cells. In addition, we observed a high percentage of genes from the ‘cellular process’, ‘binding’, and ‘cell part’ categories, but few from ‘locomotion’, ‘transcription factor activity, protein binding’, and ‘extracellular matrix component’ categories in each of the three main GO classifications (Figure 2).

The EuKaryotic Orthologous Groups (KOG) database is usually used to identify orthologous and paralogous proteins, and provides a way to find JGI-predicted genes by KOG classification or ID. The annotated sequences were compared against the KOG database to assess the completeness of our transcriptome library and the reliability of our annotation process. Of 118,043 Nr hits, 87,630 sequences were assigned KOG

classifications. Among the 25 KOG categories, the cluster for 'General function prediction only' (28,904, 32.98%) represented the largest group, followed by 'Signal transduction mechanisms' (23,030, 26.68%) and 'Posttranslational modification, protein turnover, chaperones' (19,436, 22.18%). Conversely, the 'Defense mechanisms' (673, 0.77%), 'Extracellular structures' (454, 0.52%), and 'Cell motility' (170, 0.19%) clusters represented the smallest groups (Figure 3).

To further evaluate the chrysanthemum transcriptome, all the isoforms were aligned to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database using BLASTx with an E-value $<10^{-5}$. As a collection of manually drawn pathway maps, KEGG pathways show the networks of molecular interactions in cells and particular organisms. Of the 118,043 isoforms, 31,062 had significant matches with at least one KEGG pathway in the database and were assigned to 133 KEGG pathways in total (Table 2). The most represented pathways were 'Metabolic pathways' (12,473 members) and 'Biosynthesis of secondary metabolites' (6,980 members), followed by 'Biosynthesis of antibiotics' (3,268 members), 'Microbial metabolism in diverse environments' (2,777 members), and 'Carbon metabolism' (2,089 members). In addition, 1,339 isoforms were assigned to 'Plant hormone signal transduction'.

Alternatively spliced isoforms

The long reads generated by PacBio sequencing can provide extensive information about alternative splicing. In this study, 27,975 unigenes had two or more isoforms, 15,074 had three or more distinct isoforms, and 10,909 had four or more distinct isoforms (Figure 4A). Seven alternative splicing types were identified based on SUPPA analysis, including

exon skipping (938, 5.5%), alternative 5' splice site (3,044, 17.8%), alternative 3' splice site (3,336, 19.5%), mutually exclusive exon (305, 1.8%), retained intron (8,705 51%), alternative first exon (646, 3.8%), and alternative last exon (108, 0.6%). Therefore, retained intron, alternative 3' splice site and alternative 5' splice site represented the main alternative splicing types (Figure 4B).

Comparison of the transcriptomes of normal and mutant capitula

The set of isoforms common to normal and mutant capitula

In total, 124,284 isoforms were shared by normal and mutant capitula (Figure 5A). By contrast, 3269 and 955 isoforms showed specific expression in normal and mutant capitula, respectively. Therefore, more genes were expressed in normal capitula than in mutant capitula, because the pistils were mutated to vegetative buds in mutant capitula.

DEGs between mutant and normal capitula

The transcriptomes of the normal and mutant capitula were compared, and the resulting reads were mapped to the reference transcriptome. A total of 35,419 DEGs (8,232 up-regulated and 27,187 down-regulated in the mutant capitula relative to the normal capitula) were identified between the normal and mutant capitula (Figure 5B). The correlation coefficient of gene expression between the normal and mutant capitula was 0.8897, which was determined by an algorithm developed from the correlation scatter plot.

A total of 131 DEGs were specifically expressed in the mutant capitula relative to the normal capitula, including TCP1 and AP2/ERF domain-containing genes. Conversely, 2,132 genes were not expressed in the mutant capitula but were specifically expressed in the normal capitula, including some important transcription factor genes (MYB, GRAS, and

BTF3 genes), ubiquitin-conjugating enzyme genes, zinc finger protein genes and many genes without annotations. Annotation information on the DEGs specifically expressed in the mutant and normal capitula is provided in Additional files 2 and 3, respectively. These genes may play important roles during the flower development process especially during pistil determination and development in chrysanthemums. Additionally, these genes should be important candidate regulators in chrysanthemum flower development. Therefore, functional research should be conducted on these genes in the future.

GO and KEGG pathway enrichment analyses were conducted on the DEGs to identify differences in biological processes and pathways between the mutant and normal capitula. It was found that 256 genes enriched in the ‘reproduction’ term (GO:0000003) in biochemical processes were all remarkably down-regulated in the mutant capitula relative to the normal capitula, among which 11 genes were specifically expressed in the normal capitula including WD40 and UBA1-like protein-encoding genes. These genes may play important roles in the regulatory pathways related to reproduction in chrysanthemums (Additional file 4).

In total, 6,733, 7,216, and 3,879 DEGs were enriched in the biological process, molecular function and cellular component categories, respectively (Additional files 5–7). In the biological process category, the dominant terms were ‘metabolic process’ (GO:0008152, 5,128 DEGs), ‘cellular process’ (GO:0009987, 4,758 DEGs) and ‘single-organism process’ (GO:0044699, 4,017 DEGs). In the molecular function category, the most represented terms were ‘catalytic activity’ (GO: GO:0003824, 5,765), ‘binding’ (GO:0005488, 3,903 DEGs), and ‘organic cyclic compound binding’ (GO: GO:0097159,

2,386 DEGs). Finally, in the cellular component category, the most represented terms were 'cell' (GO:0005623, 2,633 DEGs), 'cell part' (GO:0044464, 2,630 DEGs), and 'intracellular' (GO:0005622, 2,490 DEGs). Thus, the physiological and biochemical activities involved in metabolic, cellular and single-organism processes differed between the mutant and normal capitula. In total, 16,342 down-regulated and 5,485 up-regulated DEGs in the mutant capitula relative to the normal capitula were enriched in many KEGG pathways (Additional files 8 and 9). Interestingly, we found that all DEGs enriched in the KEGG pathway 'Brassinosteroid biosynthesis' (ko00905) were down-regulated in the mutant capitula relative to the normal capitula. The enriched GO terms and KEGG pathways are listed in Additional files 5–9.

Important transcription factors differentially expressed between mutant and normal capitula

A total of 3,921 important transcription factor genes from 52 classes were detected, among which 963 members were dramatically differentially expressed between the normal and mutant capitula, including members of the AP2 (14 members), ARF (35 members), B3 (41 members), BBR-BPC (3 members), BES1 (8 members), bHLH (75 members), bZIP (22 members), C2H2 (84 members), C3H (51 members), CAMTA (2 members), CO-like (5 members), DBB (11 members), Dof (9 members), E2F/DP (4 members), ERF (98 members), FAR1 (15 members), G2-like (39 members), GATA (8 members), GeBP (1 member), GRAS (29 members), GRF (1 member), HB-other (6 members), HB-PHD (1 member), HD-ZIP (51 members), HSF (20 members), LBD (7 members), LSD (1 member), MIKC (12 members), M-type (9 members), MYB (80 members), NAC (25 members), NF-

X1 (12 members), NF-YA (3 members), NF-YB (8 members), NF-YC (5 members), Nin-like (40 members), S1Fa-like (6 members), SBP (13 members), SRS (2 members), TALE (17 members), TCP (3 members), Trihelix (22 members), WRKY (56 members), YABBY (1 member), and ZF-HD (8 members) transcription factor families. Of these, the ERF, C2H2, MYB, bHLH and WRKY transcription factor families had 98, 84, 80, 75 and 56 members extremely differentially expressed between the normal and mutant capitula. Additionally, some important transcription factors had no expression in the mutant capitula and were specifically expressed in the normal capitula, including 36 C2H2 genes, 6 bZIP genes, 5 bHLH genes, 5 MYB genes, 4 HB-other genes, 2 C3H genes, 2 E2F/DP genes, 2 GATA genes, 1 ERF gene, 1 HSF gene, 1 NF-X1 gene, 1 TALE gene and 1 Trihelix gene.

These results indicated that many transcription factors play important roles during flower development including some members that have not received attention in the flower development field. The important transcription factor genes that were dramatically differentially expressed between normal and mutant capitula are presented in Additional file 10.

Identification of genes involved in the photoperiod pathway in chrysanthemum

As a typical short-day plant, chrysanthemum can flower in response to a single short day. Homologous genes of the important regulators involved in the photoperiod pathway in chrysanthemum were identified. Molecular-genetic studies have identified many genes required for the daylength response, among which some members play important roles in the regulation of flowering, while others encode components of light signal transduction pathways or pathways involved in circadian signaling, including *PHYTOCHROME (PHY)*,

CRYPTOCHROME (CRY), *LATE GIGANTEA (GI)*, *FKF1* (*Flavin binding, Kelch repeat, F-box protein 1*) and others. In this study, many genes from the transcriptome sequences were identified as homologs of photoreceptor and circadian clock components involved in the photoperiod pathway (Figure 6). Based on the protein annotations of the mutant and normal capitula transcriptome sequences, many genes were identified, including many homologs of *CRY1* and *CRY2*, as well as homologs of *PHYA*, *PHYB*, *FKF1*, *LHY*, *EFL1*, *EFL3*, *EFL4*, *TOC1*, and *GI*. *CONSTANS (CO)* plays a key role in the photoperiod response, and homologs of *CO* were identified. *FT (Flowering Locus T)* is the early target gene of *CO*, and homologs of *FT* were also identified. Additionally, many MADS box genes also play important roles in promoting floral meristem identity, including *GI*, *SHORT VEGETATIVE PHASE (SVP)*, *SUPPRESSOR OF CONSTANS1 (SOC1)*, *APETALA1 (AP1)*, *PISTILLATA (PI)* and *AGAMOUS (AG)*. We identified homologous genes of these MADS box genes. *APETALA2 (AP2)* also has important functions in promoting floral meristem identity; two homologs of *AP2* were identified. *LEAFY (LFY)* is a vital regulator of the specification of floral meristem identity, and is initially expressed very early throughout the presumptive floral meristem. We did not detect the expression of *LFY* homologs in the mutant and normal capitula of ‘ZY’, probably because the expression of *LFY* homologs had ended at the full-bloom stage of the capitula. A homolog of *SOC1* was identified in this study, which is an upstream regulatory gene of *LFY*. Interestingly, its expression level was significantly up-regulated in the mutant capitula relative to the normal capitula. As an ‘A-class’-like gene, the expression of *AP1* is directly activated by *LFY* [20, 21]. Homologs of *AP1* were identified, which were all remarkably up-regulated in the

mutant capitula relative to the normal capitula.

As another ‘A-class’ gene, *AP2* is not a MADS-box transcription factor. Two homologs of *AP2* were identified, both of which were up-regulated in the normal capitula relative to the mutant capitula. In most core eudicot species, B-class genes include three different lineages: *Pl*, *euAP3* and *TM6*; however, *TM6*-like genes seem to have been lost in *Arabidopsis* and *Antirrhinum* [22]. In this study, homologs of *Pl* and *AP3* were identified, but the expression of *TM6*-like genes was not detected. In contrast, the expression of *TM6*-like genes was detected in chrysanthemums in earlier studies [23, 24]. We also identified homologs of the C-class gene *AG* and E-like MADS-box genes in this study. We found that A-, B-, C-, and E-like genes were all expressed in the mutant capitula, although they had no normal stamens or pistils. Therefore, some other important members should also regulate the development of floral organs in chrysanthemum in the upstream or downstream pathways of A-, B-, C-, and E-like genes. Annotation information on the important genes involved in the photoperiod pathway in chrysanthemum is provided in Additional file 11.

Identification of important functional genes in the anthocyanin biosynthetic pathway and pigments in chrysanthemum

To explore the molecular basis of the difference in flower color between the normal and mutant capitula, we identified key functional genes in the anthocyanin biosynthetic pathway and analyzed their expression differences (Figure 7, Additional file 12). A series of chalcone synthase genes, chalcone flavonone isomerase genes, flavanone 3-hydroxylase genes, flavonoid 3'-hydroxylase genes, dihydroflavonol 4-reductase genes, glucosyltransferase

genes, 3-O-glucoside-6"-O-malonyltransferase genes, acyltransferase genes and one anthocyanidin synthase gene were identified in the transcriptome.

Interestingly, all four *DFR* homologs and one *CHI* homolog showed extremely low expression levels (close to 0) in the mutant capitula relative to the normal capitula. Real-time PCR showed that the *CHI* homolog was not expressed in the mutant capitula but was expressed in the normal capitula. Annotation information for these genes in the anthocyanin biosynthetic pathway is listed in Additional file 12.

The pigment types and contents in flowers determine the variety of flower colors. Qualitative analysis of pigments in the florets of normal and mutant capitula was performed by HPLC. Anthocyanins were detected in the florets of normal capitula, but not in the florets of mutant capitula (Figure 8B). As shown in Figure 8A, the detectable anthocyanins mainly included delphinidin, cyanidin, petunidin, pelargonidin, peonidin, and malvidin.

Verification of gene expression profiles using qRT-PCR

To further verify the expression profiles of the isoforms in the Illumina sequencing analyses, 16 isoforms were selected for qRT-PCR using the mutant and normal capitula originally used for RNA-seq. The selected isoforms comprised four *MYB*-like genes (*MYB8CZ-1*, *MYB8CZ-2*, *MYB8CZ-3* and *MYB8CZ-4*), one *CHI* gene (*CHICZ-1*), one *COP1* gene (*COP1CZ*) and one *EF2* gene (*EF2CZ*) specifically expressed in the normal capitula, one histone *H2B* gene (*HIS2BCZ*), one catalase gene (*CAT3CZ*), one photosystem II 5 kDa protein gene (*PSBTCZ*), one annexin D8 gene (*ANN2CZ*), one arginine kinase gene (*ARGKCZ*) and four unigenes encoding uncharacterized proteins (*UnknownCZ1*, *UnknownCZ2*, *UnknownCZ3* and *UnknownCZ4*). All 16 genes were

consistent with the sequencing data (Additional file 13, Figure 9).

Discussion

After the pistils mutated to become vegetative buds, anthocyanins were not synthesized, which indicated pistils may be required for anthocyanin synthesis in chrysanthemums

As one of the major flower pigment groups in higher plants, anthocyanin synthesis and accumulation is an integral part of flower development in most plant species and is tightly linked with petal cell expansion. It is thought that the activation of the anthocyanin synthesis pathway during petal development requires both environmental and endogenous signals. In *Petunia hybrida*, gibberellins (GAs), sugars and light were revealed to be required to induce the transcription of anthocyanin synthesis genes and accumulate pigment in the developing corolla. Research results have also indicated that GAs, sugars and light are involved in the regulation of various pathways to complete the entire flower development process [25]. David Weiss presented evidence that in the early stages of *Petunia hybrida* flower development, GAs produced by the anthers controlled anthocyanin synthesis and accumulation in the corolla by activating the transcription of related genes in the anthocyanin synthesis pathway, and the removal of stamens at the early stage of flower development inhibited anthocyanin synthesis in corollas [26]. However, in *C. morifolium*, the capitula of many varieties, such as those with the 'Pingpong' shape, have only ray florets, because of the long period of double flower breeding. Interestingly, in *C. morifolium* 'ZY', most capitula have ray florets without disc florets. Therefore, in chrysanthemums, anthocyanin synthesis may be associated with pistils.

In the mutant capitula, the pistils of ray florets mutated to become vegetative buds and the corollas of ray florets became green because of a lack of anthocyanins. When the pistils of ray florets mutated into vegetative buds, anthocyanin synthesis in the corolla may have been blocked. Therefore, it was hypothesized that pistils may be required for anthocyanin synthesis in chrysanthemums.

Many transcription factors are implicated in the developmental regulation of florets in chrysanthemums

In this study, 963 of 3,921 detected important transcription factors genes were dramatically differentially expressed between the normal and mutant capitula, including members of the ERF, C2H2, MYB, bHLH and WRKY transcription factor families. This indicated that these members participated in the process of floret development in chrysanthemums, including the determination and growth regulation of pistils and anthocyanin synthesis and accumulation in ray floret corollas. Both pistil development and anthocyanin synthesis in chrysanthemum florets may be regulated by common genes.

Interestingly, some transcription factors showed no expression in the mutant capitula and were specifically expressed in the normal capitula, including 35 C2H2 genes, 7 MYB genes, 6 bZIP genes, 6 bHLH genes, 4 HB-other genes, 2 C3H genes, 2 E2F/DP genes, 2 GATA genes, 1 ERF gene, 1 HSF gene, 1 NF-X1 gene, 1 TALE gene and 1 Trihelix gene, among which 4 MYB genes (*MYB8CZ-1*, *MYB8CZ-2*, *MYB8CZ-3* and *MYB8CZ-4*) were confirmed to have no expression in the mutant capitula by real-time PCR.

C2H2 zinc finger protein genes functioning as transcription factors play important roles in many biological processes related to plant growth and development, hormone signaling

and stress responses [27]. In many plants, C2H2 zinc finger protein genes are involved in salt, cold, drought and oxidative tolerance, light stress and pathogen defense [28]. Additionally, some members participate in the developmental regulation of flowers. For example, *SIZF2* controls flower and leaf shape in *A. thaliana*. *SUPERMAN (SUP)* determines the boundary between the stamen and carpel whorls, suppresses class B gene expression and promotes stem cell termination in the fourth whorl of *A. thaliana* flowers [28-30]. In this study, 35 C2H2 genes showed no expression in the mutant capitula, which indicated that these C2H2 zinc finger protein genes may have important functions in the determination and growth regulation of pistils or anthocyanin synthesis and accumulation in the ray floret corolla in chrysanthemums.

As one of the largest transcription factor families, basic leucine zipper (bZIP) family transcription factors play key roles in controlling plant development and stress responses [31]. bZIP genes have been reported to be involved in various flower developmental processes in plants, including pollen development, the floral transition and flower initiation [32, 33]. In this study, the expression deficiency of six bZIP genes in mutant capitula suggested that these bZIPS are also important regulators of ray floret development in chrysanthemums.

MYB transcription factors comprise one of the largest and most diverse transcription factor families in the plant kingdom, exist widely in eukaryotes, and play essential roles in a wide range of physiological and biochemical processes controlling plant growth and development [34, 35]. bHLH transcription factors possess the highly conserved bHLH domain including a basic region and a HLH region, and play important roles in plant growth

and development, metabolic regulation, and responses to environmental changes. The regulatory functions of bHLH transcription factors in active secondary metabolism especially anthocyanin synthesis have been a focus of research [36].

It was revealed that R2R3-MYB, bHLH, and WD40 proteins form a ternary complex called the MBW complex that regulates the anthocyanin biosynthesis pathway in model plants [37]. We found that five bHLH genes and seven MYB genes showed no expression in the mutant capitula, which indicated that these genes are essential regulators of normal ray floret development in chrysanthemums and the functions of these genes should be further explored. Hua Li et al. suggested that MdMYB8 was involved in the regulation of flavonoid biosynthesis by binding to the MdFLS promoter, and the overexpression of MdMYB8 promoted flavonol biosynthesis in crabapple (MdMYB8 is associated with flavonol biosynthesis via the activation of the MdFLS promoter in the crabapple fruits). In this study, four MYB8 like genes (*MYB8CZ-1*, *MYB8CZ-2*, *MYB8CZ-3* and *MYB8CZ-4*) were not expressed in the mutant capitula lacking anthocyanins. Flavonol is an upstream substrate in anthocyanin biosynthesis. Therefore, this result indicated these four *MYB8*-like genes may be important regulators in anthocyanin biosynthesis in chrysanthemum capitula. The other transcription factors showing no expression in the mutant capitula should also be important regulators of ray floret growth and development in chrysanthemums, including HB genes, 2 C3H genes, 2 E2F/DP genes, 2 GATA genes, 1 ERF gene, 1 HSF gene, 1 NF-X1 gene, 1 TALE gene and 1 Trihelix gene.

Since their active dissemination and breeding around the world, a large cultivar group for chrysanthemums has been built. Differing from traditional model plants,

chrysanthemums have a pseudanthium named the capitulum composed of two morphologically distinct florets (outer-layer ray florets and inner-layer disc florets). Interestingly, the outer-layer ray florets and disc florets are different in corolla symmetry, flower color and stamen fertility. The zygomorphous ligulate ray florets have various colors and abundant corolla types including flat petals, tube petals, spoon-shaped petals and abnormal petals. The changing permutations and combinations of ray florets and disc florets constitute daedal flower head types. Additionally, different corolla types exist in certain flower head types in chrysanthemums [38]. Therefore, the molecular mechanisms involved in chrysanthemum flower development are very complicated and go beyond the basic ABCE model. Besides the MADS-box genes and other genes revealed in previous studies, many other regulators should also participate in chrysanthemum flower developmental regulation. The transcription factors specifically expressed in the normal capitula in this study are important candidate genes to explore the regulatory molecular mechanisms in chrysanthemum flower development.

Carpels may originate from vegetative buds

The female reproductive organ, named the carpel, is specific to the angiosperms, or flowering plants. The carpel is probably the major factor contributing to the success of angiosperms, because it allowed angiosperms to diversify from an unconfirmed, possibly gymnosperm-like ancestor to form the more than 300,000 species alive today [39]. As in most species, the chrysanthemum carpel is differentiated into stigma, style, and ovary tissues. Flower developmental mechanism studies show that flowers develop from a floral apex. On the flanks of the floral apex, floral organ development starts in a centripetal

sequence: the outermost organs are initiated first and the innermost organs are initiated last. However, the evolutionary sequence of floral organs is reversed. The ovules in the center of the flower are the evolutionarily oldest organs and the perianth organs at the periphery are the youngest [40]. Each organ has its own evolutionary history.

In this study, the pistils in the mutant capitula were mutated and became vegetative buds. In the early stages, the vegetative buds mutated from the pistils looked very much like normal pistils (Figure 1). As the mutant florets grew, the two seemingly bifurcate stigma became two tender leaves and the center part between the two seemingly bifurcate stigma grew into a shoot apex. Here we present a bold hypothesis; that the pistil is derived from a vegetative bud. The predecessor of the stigma may be tender leaves developed from the phyllopedium of the vegetative bud; thus, the ovule would have evolved out of the apical meristem of the vegetative bud. This hypothesis supports the inference that the ovule did not originate from a phyllome, and each ovule represents a short shoot [41, 42]. Additionally, the dorsiventral gene expression in the outer integument of the ovule indicates that the outer integument of ovules has a leaf-derived origin, which is consistent with the inference that the stigma is derived from leaves [43, 44].

Conclusions

In our study, comparative transcriptome analysis revealed significant differences in gene expression and signaling pathways between the mutant and normal capitula. We identified DEGs between the mutant and normal capitula to reveal important regulators underlying their differential development. The transcription factors specifically expressed in the normal capitula in this study are important candidate genes to explore the regulatory molecular

mechanisms in chrysanthemum flower development. Regulatory genes involved in the photoperiod pathway and the control of floral organ identification as well as important functional genes in the anthocyanin synthesis pathway were also identified. Qualitative analysis of pigments in the florets of normal and mutant capitula revealed anthocyanins were synthesized and accumulated in the florets of normal capitula, but not in the florets of mutant capitula. It was indicated that pistils may be required for anthocyanin synthesis in chrysanthemums. This study represents the first step in exploring the molecular mechanism of floral organ development and will contribute to the development of techniques for studying flower shape and color regulation to promote breeding in chrysanthemum.

Methods

Plant materials and RNA extraction

The tissues (normal and mutant capitula) used in this study were obtained from a cut-flower chrysanthemum variety (*C. morifolium* 'ZY', a hybrid of chrysanthemum varieties) cultivated in a greenhouse under an 8-h light/16-h dark cycle at 23 °C in the Beijing Academy of Agriculture and Forestry Sciences (116.3°E, 39.9°N). After the florets of the capitula were fully formed, about 3–6 normal capitula, 3–6 mutant capitula, 3–4 fully expanded leaves, 1–2 g roots and 1–2 g stems were collected between 9:00 and 12:00 a.m. Three biological duplicates were collected for each sample. After collection, the samples were placed immediately in liquid N₂ and stored at –80 °C until RNA extraction. The extraction of total RNA was performed using the RNeasy Plant Mini Kit (Qiagen, China). RNA quantity and quality assessment was performed using a NanoDropND2000

instrument (Thermo Scientific).

Library construction, PacBio sequencing and Data processing

Library construction and PacBio sequencing were conducted with the PacBio Sequel system (PacBio, CA, USA). Briefly, 1 µg of total RNA extracted from each of the five tissues was equally pooled together and the mRNA was enriched by Oligo (dt) magnetic beads. The enriched mRNA was reverse transcribed into cDNA with Clontech AMSRTer PCR cDNA Synthesis Kits. After amplification, size selection was performed using the BluePippin Size Selection System with two bins: 0.5–2 and 2–6 kb. Then, large-scale PCR was performed and the cDNA products were used to construct next SMRTbell template libraries using the SMRTbell Template Prep Kit. The SMRTbell template was annealed to the sequencing primers and bound to polymerase, and sequenced on the PacBio Sequel platform using P6-C4 chemistry with 10 h movies by Gene Denovo Biotechnology Co.

The raw sequencing reads of the cDNA libraries were classified and clustered into consensus transcripts using the PacBio IsoSeq pipeline (https://github.com/PacificBiosciences/IsoSeq_SA3nUP) and the SMRT Analysis software suite (<https://www.pacb.com/support/software-downloads/>). Briefly, CCS (circular consensus sequence) reads were extracted out of the subreads BAM file. Then, the CCS reads were classified into four categories: full-length (FL) non-chimeric, full-length chimeric, non-FL and short reads based on the cDNA primers and polyA tail signals. Short reads were discarded. Subsequently, the FL non-chimeric reads were clustered using the Iterative Clustering for Error Correction (ICE) software to generate cluster consensus isoforms. Two strategies were employed to improve the accuracy of the PacBio reads. First,

the non-FL reads were used to polish the cluster consensus isoforms with the Quiver software to obtain FL polished high-quality consensus sequences (accuracy \geq 99%). Second, the low-quality isoforms were further corrected using Illumina short reads obtained from the same samples with the LoRDEC tool (version 0.8) [45]. Then, the final transcriptome isoform sequences were filtered by removing the redundant sequences with the CD-HIT-v4.6.7 software using a threshold of 0.99 identities.

Illumina sequencing

The total RNA isolated from the normal and mutant capitula was used for Illumina sequencing on an Illumina HiSeq™ 2000 system (Illumina, San Diego, CA, USA). We purified the poly (A) mRNAs, fragmented them into small pieces, and then synthesized the first- and second-strand cDNAs.

After the double-stranded cDNAs were purified and resolved for end reparation and poly (A) tail addition, the short fragments were connected with sequencing adapters. Briefly, a cDNA library with average insert sizes of 300–500 bp was created and cDNA sequencing was performed using the Illumina HiSeq™ 2000 system, with paired end 2 \times 100 nt multiplexes.

Basic annotation of isoforms

Isoform functional annotation was performed by BLASTX alignment (E-value of 1.00E $^{-5}$) between the isoforms and protein databases. The Nr (non-redundant protein database, NCBI), Swiss-Prot (<http://www.expasy.ch/sprot>), KEGG (<http://www.genome.jp/kegg>), and KOG (<http://www.ncbi.nlm.nih.gov/KOG>) databases were used, and the optimal alignment results determined the sequence direction of the isoforms. The Blast2GO software was

employed to get GO annotations for the isoforms [38]. GO annotation was also performed with the Blast2GO software [46]. Isoforms ranking in the 20 highest scores and no shorter than 33 HSP (High-scoring Segment Pair) hits were selected for Blast2GO analysis. Then, functional classification of the isoforms was performed using the WEGO software [47].

Analysis of chrysanthemum transcriptome sequencing results

RNA-Seq quantification analysis using the number of reads per kilobase of exon model per million mapped reads (RPKM) was performed to calculate the expression level of each isoform [39]. The chrysanthemum transcriptome of the samples was used as a reference for the screening and analysis of DEGs. A rigorous algorithm was created based on the method of Audic et al. to screen the DEGs [48]. The false discovery rate (FDR) was used to affirm the threshold of the *P*-value in multiple tests and analyses [41]. The absolute value of log₂ (ratio) ≥ 2 and FDR < 0.05 were used as the thresholds to determine significant differences in gene expression [49]. Only the DEGs with a minimum of a twofold change in expression level were used in the differential gene expression analysis.

Alternative splicing detection

To analyze alternative splicing events in the transcript isoforms, the Coding Genome reconstruction Tool (Cogent) was employed to partition the transcripts into gene families based on *k*-mer similarity and then each family was reconstructed into a coding reference genome with De Bruijn graph methods [50]. The alternative splicing events of the transcript isoforms were analyzed using the SUPPA tool [51].

Gene expression analysis based on qRT-PCR

Total RNA was extracted from the normal and mutant capitula as described above. First,

we treated the total RNA with DNase (Promega, USA), and then subjected it to reverse transcription into cDNA using a reverse transcription system (Tiangen, China). Real-time RT-PCR was done using the PikoReal real-time PCR system (Thermo Fisher Scientific, Germany). Each reaction was carried out in a total volume of 20 µL with 2 µL first-strand cDNA as a template (amplification program: 30 s at 95 °C, followed by 40 cycles of 5 s at 95 °C and 30 s at 60 °C). The gene-specific primers shown in Additional file 13 were used to evaluate the relative quantification of each gene. All real-time RT-PCR experiments were performed in three biological replicates. Each replicate was standardized in triplicate. The relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method with the *protein phosphatase 2A (PP2Acs)* gene from *C. morifolium* as the reference gene [52].

Qualitative analysis of pigments in chrysanthemum flowers

The analysis of anthocyanin profiles in the normal and mutant capitula was performed using high pressure liquid chromatography. A 1.0–10.0 g sample was ground into fine powder in liquid N₂ and homogenized into 50 ml anthocyanin extracts [ethyl alcohol: distilled water: hydrochloric acid (2:1:1, v/v/v)] assisted by sonication at 20 °C for 30 min [53]. Then, the mixture was heated in boiling water for 1 h using a water bath and centrifuged at 16000 ×g for 10 min at 20 °C. The supernate was passed through a 0.45 µm reinforced nylon membrane filter before injection. An X Bridge BEH C18 column (250 × 4.6 mm × 5 µm) was used to separate the anthocyanins and flavonols. The column was maintained at 35 °C and water containing 1% (v/v) formic acid (A) and 1% (v/v) acetonitrile (B) was used as the mobile phase. Gradient elution was applied at a flow rate of 0.8 mL/min with the following conditions: 92% A + 8% B, 0 min; 88% A + 12% B, 2 min; 82% A + 18%

B, 5 min; 80% A + 20% B, 10 min; 75% A + 25% B, 12 min; 70% A + 30% B, 15 min; 55% A + 45% B, 18 min; 20% A + 80% B, 20 min; 92% A + 8% B, 22 min; 92% A + 8% B, 30 min. The injection volume was 20 μ L and the photodiode array detector was set at 530 nm for anthocyanins [54]. Three biological replicates of each sample type were analyzed.

Competing interests

The authors declare that they have no competing interests.

Funding

Research was supported by Beijing Natural Science Foundation (6194033); Natural Science Foundation of Beijing Academy of Agriculture and Forestry Sciences (QNJJ201817); Innovation Foundation of Beijing Academy of Agriculture and Forestry Sciences (KJCX20180204); National Natural Science Foundation of China (31901354); Beijing Municipal Committee of Science and Technology (Z181100002518003)

Authors' contributions

Hua Liu, Chang Luo, Dongliang Chen and Yaqin Wang performed the research; Hua Liu analyzed the data and prepared the manuscript; Conglin Huang and Mingyuan Li guided the research. Shuang Guo, Xiaoxi Chen, Jingyi Bai, Xinlei Huang and Xi Cheng provided assistance for the research. **All authors read and approved the final manuscript.**

Acknowledgments

We thank Robbie Lewis, MSc, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing a draft of this manuscript.

References:

- [1]. Su, J., et al., Current achievements and future prospects in the genetic breeding of chrysanthemum: a review. *Horticulture Research*, 2019. 6: p. 109.
- [2]. Ja, T.D.S., Chrysanthemum: advances in tissue culture, cryopreservation, postharvest technology, genetics and transgenic biotechnology. 2003. 21(8): p. 715-766.
- [3]. Roberts, W.R. and E.H. Roalson, Comparative transcriptome analyses of flower development in four species of Achimenes (Gesneriaceae). *Bmc Genomics*. 2017. 18(1): p. 240.
- [4]. O'Maoileidigh, D.T.E.A., Gene network analysis of *Arabidopsis thaliana* flower development through dynamic gene perturbations. *Plant Journal*, 2015. 83(2): p. 344-358.
- [5]. Irish, V.F., The flowering of *Arabidopsis* flower development. *Plant Journal for Cell & Molecular Biology*. 2010. 61(6): p. 1014-1028.
- [6]. Lyu, J., et al., The Floral Repressor GmFLC-like Is Involved in Regulating Flowering Time Mediated by Low Temperature in Soybean. *International Journal of Molecular Sciences*. 2020.
- [7]. Smaczniak, C., et al., Developmental and evolutionary diversity of plant MADS-domain factors: Insights from recent studies. *Development (Cambridge, England)*, 2012. 139: p. 3081-98.
- [8]. Weiss, D., Regulation of flower pigmentation and growth: Multiple signaling pathways control anthocyanin synthesis in expanding petals. *Physiologia Plantarum*, 2000. 110(2): p. 152-157.
- [9]. Elomaa, P., et al., Activation of Anthocyanin Biosynthesis in *Gerbera hybrida* (Asteraceae) Suggests Conserved Protein-Protein and Protein-Promoter Interactions between the Anciently Diverged Monocots and Eudicots. *Plant Physiology*. 133(4): p. 1831-1842.
- [10]. Zhang, Y., et al., Altered expression of CmNRRa changes flowering time of *Chrysanthemum morifolium*. *Plant biotechnology journal*, 2013. 11(3): p. 373-379.
- [11]. Wei, Q., et al., Control of chrysanthemum flowering through integration with an aging pathway. *Nature Communications*, 2017. 8(1): p. 829.
- [12]. Cui-Hui, et al., The MADS transcription factor CmANR1 positively modulates root system development by directly regulating CmPIN2 in chrysanthemum. [J]. *Horticulture research*, 2018.
- [13]. Noda, N., et al., Generation of blue chrysanthemums by anthocyanin B-ring hydroxylation and glucosylation and its coloration mechanism. *Science Advances*. 2017. 3(7): p. e1602785.
- [14]. Akemi and Ohmiya, Molecular mechanisms underlying the diverse array of petal colors in chrysanthemum flowers. *Breed Sci*. 2018. 68(1):119–127.
- [15]. Junyang, Y., et al., Transcriptome analysis of differentially expressed unigenes involved in flavonoid biosynthesis during flower development of *Chrysanthemum morifolium* ‘Chuju’ . *Scientific Reports*. 2018. 8(1): p. 13414-.

- [16]. Yang, Y., et al., Interactions between WUSCHEL- and CYC2-like Transcription Factors in Regulating the Development of Reproductive Organs in *Chrysanthemum morifolium*, in International Journal of Molecular Sciences. 2019.
- [17]. Shaokui, et al., Full-length transcriptome of *Misgurnus anguillicaudatus* provides insights into evolution of genus *Misgurnus* .[J]. Scientific reports, 2018.
- [18]. Song, C., et al., The *Chrysanthemum nankingense* Genome Provides Insights into the Evolution and Diversification of *Chrysanthemum* Flowers and Medicinal Traits. Molecular Plant, 2018. 11(12): p. 1482-1491.
- [19]. Hirakawa, H., et al., De novo whole-genome assembly in *Chrysanthemum seticuspe*, a model species of Chrysanthemums, and its application to genetic and gene discovery analysis. DNA research. 2019. 26: p. 195-203.
- [20]. Litt, A. and E.M. Kramer, The ABC model and the diversification of floral organ identity. 2010. Semin Cell Dev Biol. 2010. 21(1): 129-37.
- [21]. Wang, H., et al., Next-Generation Sequencing of the *Chrysanthemum nankingense* (Asteraceae) Transcriptome Permits Large-Scale Unigene Assembly and SSR Marker Discovery. Plos One, 2013. 8(4): p. e62293.
- [22]. Ren, L., et al., A transcriptomic analysis of *Chrysanthemum nankingense* provides insights into the basis of low temperature tolerance. [J]. BMC Genomics, 2014, 15(1):844.
- [23]. Liu, H., et al., Whole-transcriptome analysis of differentially expressed genes in the ray florets and disc florets of *Chrysanthemum morifolium*. BMC Genomics, 2016. 17(1): p. 398.
- [24]. Liu, H., et al., Whole-Transcriptome Analysis of Differentially Expressed Genes in the Vegetative Buds, Floral Buds and Buds of *Chrysanthemum morifolium*. PLOS ONE, 2015. 10: p. e0128009.
- [25]. Weiss, D., Regulation of flower pigmentation and growth: Multiple signaling pathways control anthocyanin synthesis in expanding petals. Physiologia Plantarum, 2000. 110(2): p. 152-157.
- [26]. Weiss, D., et al., Stamens and Gibberellic Acid in the Regulation of Flavonoid Gene Expression in the Corolla of *Petunia hybrida*. Plant Physiology. 94(2): p. 511-515.
- [27]. Alam, I., et al., Comprehensive genomic survey, structural classification and expression analysis of C2H2 zinc finger protein gene family in *Brassica rapa* L. PLOS ONE, 2019. 14(5): p. e0216071.
- [28]. Alam, I., et al., Comprehensive genomic survey, structural classification and expression analysis of C2H2 zinc finger protein gene family in *Brassica rapa* L. PLOS ONE, 2019. 14(5): p. e0216071.
- [29]. Dathan and N., The *Arabidopsis* SUPERMAN protein is able to specifically bind DNA through its single Cys2-His2 zinc finger motif. Nucleic Acids Research. 2002. 30(22): p. 4945-4951.
- [30]. Prunet, N., et al., SUPERMAN prevents class B gene expression and promotes stem cell termination in the fourth whorl of *Arabidopsis thaliana* flowers. Proceedings of the National Academy of Sciences of the United States of America, 2017. 114(27): p. 201705977.
- [31]. Yang, Y., et al., The bZIP gene family in watermelon: genome-wide identification and expression analysis under cold stress and root-knot nematode infection. Peer J, 2019. 7: p. e7878.
- [32]. Gibalová, A., et al., Characterization of pollen-expressed bZIP protein interactions and the role of ATbZIP18 in the male gametophyte. Plant Reproduction, 2017. 30(1): p. 1-17.
- [33]. Wang, Y., et al., The inhibitory effect of ABA on floral transition is mediated by ABI5 in *Arabidopsis*. Journal of Experimental Botany, 2013(2): p. 2.
- [34]. Chen, C., et al., MYB Transcription Repressors Regulate Plant Secondary Metabolism. Critical Reviews in Plant Sciences, 2019. 38(3): p. 159-170.

- [35]. Li, J., Y. An and L. Wang, Transcriptomic Analysis of *Ficus carica* Peels with a Focus on the Key Genes for Anthocyanin Biosynthesis, *International Journal of Molecular Sciences*. 2020.
- [36]. Amírez, C.I.O., et al., Expression and function of the bHLH genes ALCATRAZ and SPATULA in selected Solanaceae species. *The Plant Journal*, 2019. 99(4).
- [37]. The MYB5 - driven MBW complex recruits a WRKY factor to enhance the expression of targets involved in vacuolar hyper - acidification and trafficking in grapevine. *The Plant Journal*, 2019.
- [38]. Yuan ZHANG, et al., A classification study for chrysanthemum (*Chrysanthemum*×*grandiflorum* Tzvelv.) cultivars based on multivariate statistical analyses. *Journal of Systematics & Evolution*, 2014. 52(5): p. 612-628.
- [39]. Fourquin, C., et al., Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proceedings of the National Academy of Sciences of the United States of America*. 2005. 102(12): p. 4649-4654.
- [40]. Endress, P.K., Angiosperm Floral Evolution: Morphological Developmental Framework. *Advances in Botanical Research*, 2006. 44: p. 1-61.
- [41]. Endress, P.K., Angiosperm Floral Evolution: Morphological Developmental Framework. *Advances in Botanical Research*, 2006. 44: p. 1-61.
- [42]. Taylor, D.W. and L.J. Hickey, Evidence for and Implications of an Herbaceous Origin for Angiosperms. 2007: Springer US.
- [43]. Meister, R.J., L.M. Kotow and C.S. Gasser, SUPERMAN attenuates positive INNER NO OUTER autoregulation to maintain polar development of *Arabidopsis* ovule outer integuments. *Development*, 2002. 129(18): p. 4281-4289.
- [44]. Yamada, T., M. Ito and M. Kato, Expression pattern of INNER NO OUTER homologue in *Nymphaea* (water lily family, Nymphaeaceae). *Development Genes & Evolution*. 2003. 213(10): p. 510-513.
- [45]. Schunter, C., et al., Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics. *Bmc Genomics*. 2014. 15(1): p. 167.
- [46]. Conesa, A., et al., Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005. 21(18): p. 3674-3676.
- [47]. Ye, J., WEGO: a web tool for plotting GO annotations. 2006. 34(Web Server issue): p. W293-W297.
- [48]. Samarskiĭ, A., Claverie JM: The significance of digital gene expression profiles. *Genome Res*. 1997, 7(10):986-995.
- [49]. Mortazavi, A., et al., Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 2008. 5(7): p. 621-628.
- [50]. Li, J., et al., Long read reference genome-free reconstruction of a full-length transcriptome from *Astragalus membranaceus* reveals transcript variants involved in bioactive compound biosynthesis. 2017. 3: p. 17031.
- [51]. Alamancos, G.P., et al., Leveraging transcript quantification for fast computation of alternative splicing profiles. *RNA*, 2015. 21(9): p. 1521 -1531.
- [52]. Xue, C.M., et al., Reference Gene Selection for Quantitative Real-time PCR in *Pandora neoaphidis*. *Journal of Agricultural Biotechnology*, 2014.
- [53]. Du, H., et al., Methylation mediated by an anthocyanin, O-methyltransferase, is involved in purple flower coloration in *Paeonia*. *Journal of Experimental Botany*, 2015. 66(21).

[54]. Revilla, E. and J.M. Ryan, Analysis of several phenolic compounds with potential antioxidant properties in grape extracts and wines by high-performance liquid chromatography-photodiode array detection without sample preparation. 2000. 881(1): p. 461-469.

Figures Legends

Figure 1. The mutant plant of *C. morifolium* 'ZY'. A The mutant and normal capitula. B The normal capitulum. C The mutant capitulum. D The vegetative buds in the mutant capitula (the left one) and the pistils in the normal capitula (the right one). E The normal ray florets (the left one) and the mutant ray florets (the right one). F The new shoots from the mutant capitulum.

Figure 2. Histogram presentation of Gene Ontology classifications. The results are summarized in three main categories: biological processes, cellular components, and molecular functions. The y-axis on the left side indicates the percentage of genes in a category, and the y-axis on the right side shows the number of genes.

Figure 3. EuKaryotic Orthologous Groups (KOG) classifications in chrysanthemum. A total of 87,630 sequences with KOG classifications within the 25 categories are shown.

Figure 4. The alternatively spliced genes analysis. A The alternatively spliced isoforms analysis. B The alternative splicing types.

Figure 5. The Venn diagram of the number of expressed genes. A. Venn diagram of the number of expressed genes in the normal capitula (NorC) and the mutant capitula (MutC) B. The number of up-regulated and down-regulated genes between the normal and mutant capitula

Figure 6. A schematic of flowering regulatory networks involved in the photoperiod pathway in *Chrysanthemum morifolium*. Arrows indicate activation. All homologs of the regulators involved in the photoperiod pathway are listed in Additional file 11.

Figure 7. The anthocyanin biosynthetic pathway and the regulatory genes identified in chrysanthemum transcriptome. CHS: chalcone synthase, CHI: chalcone isomerase, F3H: flavanone 3-hydroxylase, F3'H: flavonoid 3'-hydroxylase, DFR: dihydroflavonol 4-reductase, ANS: anthocyanidin synthase, OT: 3-O-glucoside-6"-O-malonyltransferase, GT: glucosyltransferase, AT: acyltransferase. The genes in the anthocyanin biosynthetic pathway is listed in Additional file 12.

Figure 8. The qualitative analysis of Pigments in the normal capitula. (A) and the mutant capitula (B). 1. Delphinidin 2. Cyanidin 3. Petunidin 4. Pelargonidin 5. Peonidin 6. Malvidin

Figure 9. The expression profiles of 16 transcripts in *Chrysanthemum morifolium* by qRT-PCR.

Additional files Legends

Additional file 1 Flowers of chrysanthemums. A The capitulum. B The disc floret. C The ray floret.

Additional file 2 The DEGs specifically expressed in the mutant capitula

Additional file 3 The DEGs specifically expressed in the normal capitula

Additional file 4 The DEGs enriched in the “reproduction” term (GO0000003)

Additional file 5 The DEGs enriched in biological process

Additional file 6 The DEGs enriched in molecular function

Additional file 7 The DEGs enriched in cellular component

Additional file 8 KEGG enrichment of DEGs down-regulated in the mutant capitula

Additional file 9 KEGG enrichment of DEGs up-regulated in the mutant capitula

Additional file 10 The important transcription factors genes that dramatically differentially expressed between normal and mutant capitula

Additional file 11 The important genes involved in the photoperiod pathway in chrysanthemum

Additional file 12 The important function genes in anthocyanin biosynthetic pathway in chrysanthemum

Additional file 13 Primers used in real-time quantitative PCR of *Chrysanthemum morifolium*

Tables

Table 1 Results of de novo assembly.

Isoform	N50 (bp)	Max length (bp)	Min length (bp)	Average length (bp)	Total assembled bases	GC%	Annotation counts	Annotation ratio
130,097	3013	14273	57	2510	494377327	40.06	118,043	90.73%

Table 2. Categorization of Chrysanthemum unigenes to KEGG biochemical pathways

KEGG Categories	Unigene number	Ratio of no.	Pathway ID
Metabolic pathways	12473	40.16%	ko01100

Biosynthesis of secondary metabolites	6980	22.47%	ko01110
Biosynthesis of antibiotics	3268	10.52%	ko01130
Microbial metabolism in diverse environments	2777	8.94%	ko01120
Carbon metabolism	2089	6.73%	ko01200
Protein processing in endoplasmic reticulum	1858	5.98%	ko04141
Biosynthesis of amino acids	1766	5.69%	ko01230
Spliceosome	1725	5.55%	ko03040
Endocytosis	1546	4.98%	ko04144
Starch and sucrose metabolism	1511	4.86%	ko00500
RNA transport	1350	4.35%	ko03013
Plant hormone signal transduction	1339	4.31%	ko04075
Ubiquitin mediated proteolysis	1205	3.88%	ko04120
Plant-pathogen interaction	1202	3.87%	ko04626
mRNA surveillance pathway	1085	3.49%	ko03015
Purine metabolism	1080	3.48%	ko00230
RNA degradation	1045	3.36%	ko03018
Ribosome	1016	3.27%	ko03010
Oxidative phosphorylation	1013	3.26%	ko00190
Aminoacyl-tRNA biosynthesis	933	3.00%	ko00970
Amino sugar and nucleotide sugar metabolism	898	2.89%	ko00520
Glycolysis / Gluconeogenesis	811	2.61%	ko00010
Pyruvate metabolism	799	2.57%	ko00620

Glycerophospholipid metabolism	793	2.55%	ko00564
Pyrimidine metabolism	790	2.54%	ko00240
Glyoxylate and dicarboxylate metabolism	781	2.51%	ko00630
Peroxisome	711	2.29%	ko04146
Fatty acid metabolism	708	2.28%	ko01212
Phagosome	699	2.25%	ko04145
alpha-Linolenic acid metabolism	678	2.18%	ko00592
Cysteine and methionine metabolism	667	2.15%	ko00270
Carbon fixation in photosynthetic organisms	614	1.98%	ko00710
Glycine, serine and threonine metabolism	583	1.88%	ko00260
Phosphatidylinositol signaling system	556	1.79%	ko04070
Phenylpropanoid biosynthesis	525	1.69%	ko00940