SpringerPlus

De novo transcriptome sequencing and comparative analysis to discover genes related to floral development in *Cymbidium faberi* Rolfe

Yuying Sun, Guangdong Wang[*], Yuxia Li, Li Jiang, Yuxia Yang and Shuangxue Guan

Abstract

Cymbidium faberi is a traditional orchid flower in China that is highly appreciated for its fragrant aroma from its zygomorphic flowers. One bottleneck of the commercial production of *C. faberi* is the long vegetative growth phase of the orchid and the difficulty of the regulation of its flowering time. Moreover, its flower size, shape and color are often targeting traits for orchid breeders. Understanding the molecular mechanisms of floral development in *C. faberi* will ultimately benefit the genetic improvement of this orchid plant. The goal of this study is to identify potential genes and regulatory networks related to the floral development in *C. faberi* by using transcriptome sequencing, de novo assembly and computational analyses. The vegetative and flower buds of *C. faberi* were sampled for such comparisons. The RNA-seq yielded about 189,300 contigs that were assembled into 172,959 unigenes. Furthermore, a total of 13,484 differentially expressed unigenes (DEGs) were identified between the vegetative and flower buds. There were 7683 down-regulated and 5801 up-regulated DEGs in the flower buds compared to those in the vegetative buds, among which 3430 and 6556 DEGs were specifically enriched in the flower or vegetative buds, respectively. A total of 173 DEGs orthologous to known genes associated with the floral organ development, floral symmetry and flowering time were identified, including 12 TCP transcription factors, 34 MADS-box genes and 28 flowering time related genes. Furthermore, expression levels of ten genes potentially involved in floral development and flowering time were verified by quantitative real-time PCR. The identified DEGs will facilitate the functional genetic studies for further understanding the flower development of *C. faberi*.

Keywords: *Cymbidium faberi* Rolfe, RNA-seq, Flower development, Flowering, Genes

Background

Orchidaceae is one of the largest families in the angiosperms with more than 25,000 species, which displays a great biodiversity resulting from adaptation to diverse habitats (Pridgeon et al. 2005). Genomic information on orchids were mainly focused on *Phalaenopsis* (Su et al. 2011; Cai et al. 2014), *Dendrobium* (Yan et al. 2015; Zhang et al. 2016), *Cymbidium ensifolium* (Li et al. 2013), *Cymbidium sinense* (Zhang et al. 2013). *Cymbidium faberi* Rolfe., common named as "Hui Lan", is one of the oldest and most popular orchids species cultivated in China, which is highly appreciated mainly because of its beautiful flower posture and fragrant aroma (Wolff, 1999). However, large scale commercial production of *C. faberi* was often hindered due to the long vegetative growth phase (usu. 5–7 years) and difficulties in flowering time control.

Plant flowering involves a transition process from vegetative growth to reproductive development with a series of conserved underlying metabolic or external phenotypic changes taking place in the shoot apical meristem. In *Arabidopsis thaliana*, there are four major pathways controlling the timing of flowering, including photoperiod, vernalization, gibberellin (GA) and autonomous

*Correspondence: gdwang@njau.edu.cn
Department of Horticulture, Nanjing Agricultural University,
Nanjing 210095, China

pathways (Mouradov et al. 2002). Major genetic elements involved in this pathways have been defined as the key switches to control flowering, such as *CONSTANS* (*CO*) in the photoperiodic pathway and *FLOWERING LOCUS C* (*FLC*) in the autonomous and vernalization pathways (Mouradov et al. 2002). The *FLOWERING LOCUS T* (*FT*) gene activates the expression of a number of flower developmental genes (Mouradov et al. 2002). In the orchid plant *Phalaenopsis aphrodite*, *PaFT1* can suppress the delayed flowering caused by *SHORT VEGATATIVE PHASE* (*SVP*) and *FRIGIDA* (*FRI*) (Jang et al. 2015). The *Cymbidium FT* orthologous gene was also cloned and ectopic expression of *CgFT* resulted in early flowering in transgenic *Arabidopsis* (Huang et al. 2012). Over expression of *DnAGL19*, a *SOC1-1/TM3*-like ortholog in *Arabidopsis* resulted in a slightly accelerated flowering time under normal growth conditions (Liu et al. 2016).

During the transition from vegetative growth to reproductive growth, MADS-box family genes play important roles in regulating floral organ specification, development and evolutionary in higher plants (Weigel and Meyerowitz 1994; Purugganan et al. 1995; Münster et al. 1997). Although conserved flowering pathways and multiple key genes were revealed in model plants, there is limited information of *C. faberi*, the very unique and important plant species featured with a 5–7 year-long vegetative growth phase. Unlike most tropical orchids (e.g. *Phalaenopsis* and *Cattleya*), *C. faberi* flowers are not brightly showy but of strong aromas to attract pollinating insects. It is of strong interest for orchid breeders to improve the appearance of its flower size, color and shape with uncompromised aromas. The flowers of *C. faberi* are zygomorphic consisting of four whorls: The first whorl is comprised of the petal-like sepals, the second whorl is of two petals and one highly specialized labellum in the middle, the third whole is of stamens, and fourth whorl is of pistils in the form a highly specialized united gynostemium (Rudall and Bateman 2002). The well-known ABC model clarify that floral development is determined by five kinds of floral organ identity genes in diverse plant groups (Coen and Meyerowitz 1991; Weigel and Meyerowitz 1994). Sepal formation is specified by the expression of A-class function genes. Expression of AB and BC determine the development of petals and stamen formation, respectively. The development of carpel is determined by C-class genes function alone, and D-class genes specify ovules. While class E function redundantly specify petals, stamens, and carpels as well as floral determinacy (Pelaz et al. 2000, 2001; Anusak et al. 2003). The floral morphology of orchid species is unique with gynostemium, labellum and resupination caused by 180° torsion of the pedicel (Rudall and Bateman 2002). 'The Perianth (P) code' clarifies the

mechanisms of sepal/petal/lip determination in *Oncidium* and *Phalaenopsis* orchids. The competition between different *APETALA3/AGAMOUS-LIKE6* (*AP3/AGL6*) homologues determines the formation of the complex perianth patterns in orchids. The formation of sepal/petal were specified by the higher-order heterotetrameric SP (sepal/petal) complex (*OAP3-1/OAGL6-1/OAGL6-1/OPI*), whereas the lip formation required the L (lip) complex (*OAP3-2/OAGL6-2/OAGL6-2/OPI*) exclusively (Hsu et al. 2015). Other MADS-box function genes participating in the sepal and petal development were also isolated from *Dendrobium madame*, *D. crumenatum*, *Oncidium* Gower Ramsey (Yu et al. 2000, 2002; Hsu and Yang 2002; Hsu et al. 2003; Xu et al. 2006). Genes involved in flower identity and floral organ specification are still unknown in *C. faberi* yet.

Besides petal shape and size, floral symmetry significantly affects the ornamental value of flowers. So far, three transcription factors (TFs) that determine the floral symmetry are identified, including *CYCLOIDEA* (*CYC*) from the TCP family (*TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR*), *DIVARICATA* (*DIV*) and *RADIALIS* (*RAD*) from the MYB family (Luo et al. 1996; Doebley et al. 1997; Almeida et al. 1997; Galego and Almeida 2002; Corley et al. 2005; Costa et al. 2005). Of these genes, *CYC* and its orthologues establish the floral monosymmetry through specifying dorsal flower identity (Luo et al. 1996, 1999; Feng et al. 2006; Wang et al. 2008). Studies on floral symmetry genes mainly focused on species with highly derived morphologies in eudicots, whereas only a few studies are available for monocots (Luo et al. 1996, 1999; Feng et al. 2006; Wang et al. 2008; Yuan et al. 2009; Bartlett and Specht 2011; Preston and Hileman 2012; Hoshino et al. 2014). Orchidaceae is characterized by highly specialized zygomorphic flowers. Studies on orchid flower symmetry is very limited (Paolo et al. 2015). Overall, the detailed molecular mechanism or the genetic elements in the regulation of floral organ specification and development in *C. faberi* remains elusive.

Transcriptome analysis is an useful tool to discriminate differences in transcript abundance among different cultivars, organs and different treatment conditions in model and non-model plants (Cheung et al. 2006; Trick et al. 2009; Li et al. 2013; Hyun et al. 2014; Zhang et al. 2014). In order to excavate genes that might regulate the floral development in *C. faberi*, we used high-throughput Illumina sequencing and bioinformatics analysis to compare the de novo transcriptomes of vegetative and flower buds from *C. faberi*. The vegetative transcriptome can be used to find genes related to the vegetative growth. The floral transcriptome was sufficiently comprehensive for gene discovery and analysis of major metabolic pathways associated

with flower traits. Genes expressed differently in the vegetative buds and flower buds may play important roles in the vegetative growth or floral development. Through transcriptome analysis, large numbers of genes related to floral organ initiation, flower symmetry patterning and flowering were identified in *C. faberi*. These results provide fundamental information for further studies on the molecular mechanism of flower development in *C. faberi*.

Methods

Plant materials

Mature plants of *C. faberi* with light green flowers, originally introduced from Henan province of China, were grown in the greenhouse at Department of Horticulture, Nanjing Agricultural University (Nanjing, China) under natural light and a controlled temperature of 22–28 °C. The vegetative buds were collected from lateral buds, and the flower buds (0.5–1.0 cm in length) were collected from the peduncle (Fig. 1). Organs from three individual plants were pooled as one sample. The fresh samples were frozen immediately in liquid nitrogen and stored at −80 °C.
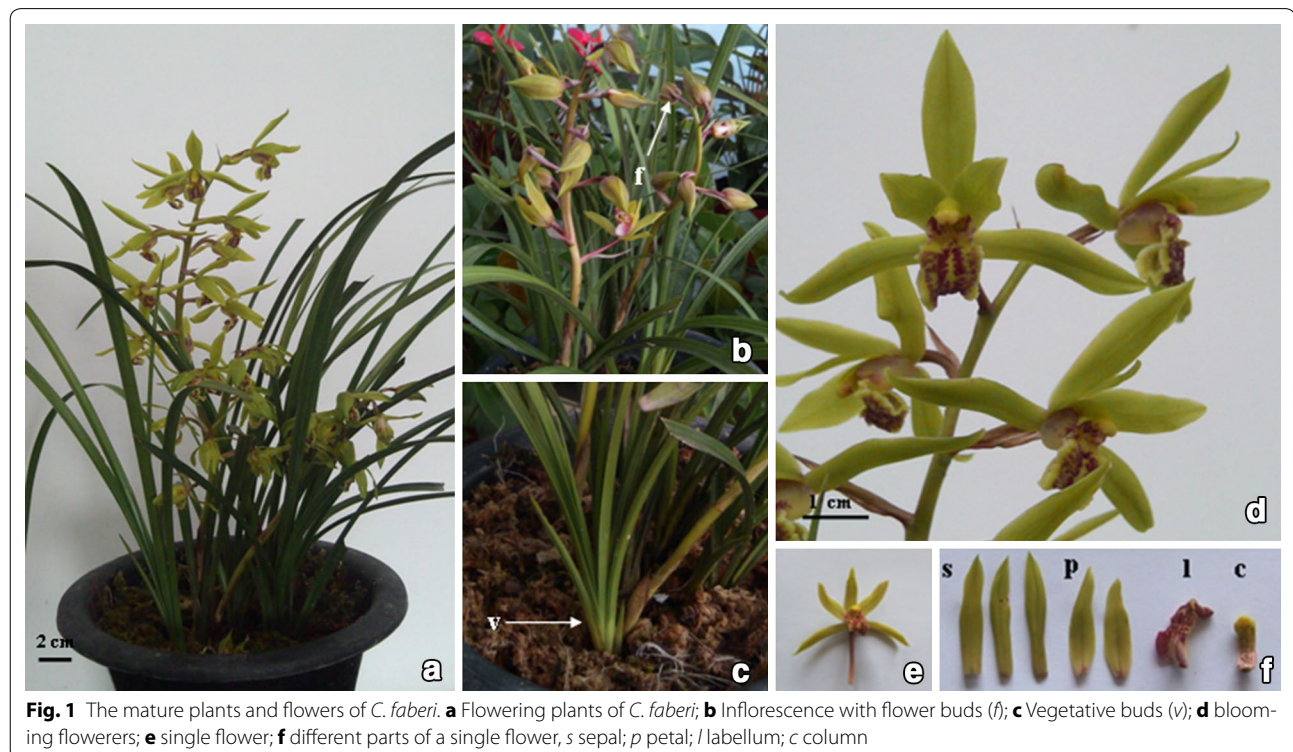
RNA extraction, cDNA library construction and de novo assembly

Total RNA of flower and vegetative buds were extracted using EASYspin plant RNA rapid extraction kit according to the manufacturer's protocol (Yuan Ping Hao Biotechnology Co. Ltd, Beijing, China). Then, the concentration was detected by a NanoDrop 2000™ micro-volume spectrophotometer (Thermo Scientific, Waltham, MA, USA) and the quality was tested by gel electrophoresis. cDNA library construction for the flower and vegetative buds and Illumina deep sequencing were performed on the HiSeq™ 2000 platform according to the manufacturer's instructions at Hangzhou Woosen Biotechnology Co. Ltd. (Hangzhou, Zhejiang, China). The Illumina reads were assembled to obtain the contigs and unigenes using Trinity software and Cap3 after removing the short raw reads and quality inspection by fastQC (Grabherr et al. 2011).

Functional annotation of unigenes

The unigenes were annotated with the National Center for Biotechnology Information non-redundant databases (NR), Gene Ontology (GO), Clusters of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases using BLASTX searches (E-value $\leq 1.0e^{-5}$). The Blast2GO software package and WEGO software were employed to compare and determine unigenes' GO annotations and obtain GO functional classifications for all annotated unigenes (Götz et al. 2008; Ye et al. 2006).

Fig. 1 The mature plants and flowers of *C. faberi*. **a** Flowering plants of *C. faberi*; **b** Inflorescence with flower buds (*f*); **c** Vegetative buds (*v*); **d** blooming flowerers; **e** single flower; **f** different parts of a single flower, *s* sepal; *p* petal; *l* labellum; *c* column

Identification of differentially expressed genes (DEGs)

Expression levels of all unigenes were calculated and then compared between the two tissue samples using the fragments per kilobase of transcript per million reads of library method (FPKM). The false discovery rate (FDR) was adopted to determine the threshold of P-values in multiple tests. DEGs were also annotated with GO assignments, COG assignments and KEGG pathways. The criteria FDR < 0.05 was used to identify DEGs and acted as a threshold of significant difference of gene expression in GO terms, COG classification and KEGG annotation.

Quantitative real-time PCR (qRT-PCR) analysis

Total RNA of flower buds and vegetative buds were extracted as above and the first strand cDNA was synthesized using PrimeScript RT reagent Kit With gDNA Eraser (Takara Bio Inc.). All primers used in this study were designed by the Primer 5 software according to the RNA-Seq data. *CfGAPDH* (JX560732) was selected as an internal reference (Additional file 1: Table S1).

The qRT-PCR analysis was performed on ABI 7500 Real-Time PCR Detection System (Applied Biosystems) using the SYBR® Premix ExTagTM reagent kit (Takara Bio Inc.) according to the manufacturer's protocol. The PCR reactions were 40 cycles (95 °C for 15 s; 55 °C for 15 s; 72 °C for 20 s) according to the instruction manual. A melting curve was generated to test the specificity of products after the qRT-PCR. The relative expression levels of the selected unigenes were normalized to *CfGAPDH* gene and calculated using the $2^{-\Delta\Delta Ct}$ method. Data were derived from three independent replicates.

Results

Transcriptome sequencing and de novo assembly

Equal amount of RNAs from the vegetative and flower buds were used to constructed cDNA libraries separately, and then sequenced. A total of 35,511,583 and 32,514,423 raw reads were obtained in the flower and vegetative buds, respectively. After removing the short raw reads and quality inspection by fastQC, the RNA-seq produced 35,510,239 and 32,513,288 clean reads for the flower and vegetative buds, respectively (Table 1). All of these reads were employed for further de novo assembly. And 189,300 contigs with an average length of 755 bp were generated using the Trinity software package. These contigs were assembled into 172,959 unigenes (200 to >10,000 bp) with an average length of 698 bp and an N50 (N50 represents weighted median length of all contigs) of 1340 bp (Table 2). Among them, 59,505 (34.4 %) unigenes were more than 500 bp. The size distribution of the assembled unigenes is shown in Additional file 2: Fig. S1. And this transcriptome shotgun assembly project has been deposited at DDBJ/EMBL/GenBank under the accession of GDHD00000000.

Functional annotation of *C. faberi* transcriptome

To validate and annotate the assembled unigenes, the 172,959 unigenes generated were subjected to BLASTX searches (E-value $\leq$ 1.0e^{-5}) against public protein databases such as NCBI NR, GO, COG and KEGG. As a result, a total of 65,577 (37.91 %) unigenes were predicted to be coding sequence and 66,000 (38.16 %), 50,161 (29.00 %), 27,443 (15.87 %), 19,715 (11.40 %) unigenes had homologous sequences respectively in NR, GO, COG and KEGG databases (Table 3).

According to the NR annotation, 66,000 (38.16 %) unigenes had homologous sequence in the database,

Table 1 Statistical summary of *C. faberi* transcriptome sequencing data

Statistics of data production	Flower buds	Vegetative buds
Raw reads	35,511,583	32,514,423
Raw bases	5,326,737,450	4,877,163,450
Clean reads	35,510,239	32,513,288
Clean bases	5,280,378,528	4,841,503,927
Clean data rate (%)	99.125	99.27

Table 2 Summary of assembly quality from transcriptome in *C. faberi*

	Contigs	Unigenes
Total number	189,300	172,959
Average length(bp)	755.549	698.016
Max length(bp)	11,797	11,797
Min length(bp)	201	201
GC percentage (%)	42.1	42.0
Length of N50 (bp)	1452	1340

Table 3 Summary statistics of functional annotation for *C. faberi* vegetative and flower buds unigenes in public databases

Public protein database	No. of unigene hits	Percentage (%)
Predict protein number	65,577	37.91
COG	27,443	15.87
NR (E-value < 10^{-5})	66,000	38.16
GO	50,161	29.00
KEGG	19,715	11.40

of which 14,365 (21.77 %), 4309 (6.53 %), 4173 (6.32 %) unigenes were annotated with homologous genes from *Vitis vinifera*, *Oryza sativa* Japonica Group and *Prunus persica*, respectively (Fig. 2a). The similarity distribution indicated that 54.87 % of the unigenes showed a similarity higher than 60, and 40.04 % unigenes had similarity between 60 and 80 % (Fig. 2b). For the E value distribution, 53.79 % of the top hits had high homology with the E-value <$1.0e^{-50}$ (Fig. 2c).
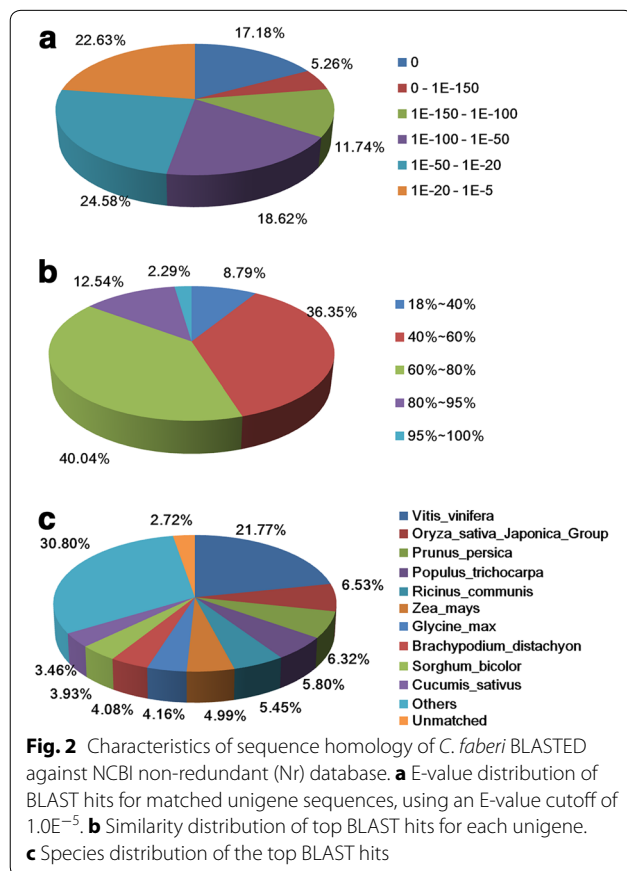
As an international standardized gene functional classification system, GO describes properties of genes and their products in many organisms and provides a comprehensive description of gene properties across species and databases (Hao et al. 2011). Based on the BlastX results, a total of 50,161 (29.00 %) unigenes were assigned into 49 GO term annotations, including three categories, i.e., molecular function, biological process, and cellular component (Fig. 3). Within the molecular function category (70,280 GO terms), "binding" (32,021, 45.6 %) and "catalytic activity" (26,742, 38.0 %) were the most highly represented terms. For biological process (108,214 GO terms), "cellular process" (28,941, 26.7 %) and "metabolic process" (29,813, 27.6 %) were the highly represented terms. Among the cellular component category (87,253

GO terms), "cell" (28,076, 32.2 %), "cell part" (13,723, 15.7 %) and "organelle" (28,076, 32.2 %) were the three main categories.

A total of 27,443 unigenes were classified into 24 COG functional categories (Fig. 4). And the largest category was "general function prediction only" (6177, 22.5 %), followed by "signal transduction mechanisms" (2220, 8.1 %) and "replication, recombination and repair" (2196, 8.0 %). Unigenes annotated as the "signal transduction mechanisms" in our study may allow for the identification of novel genes involved in signal transduction pathways. Approximately 16.9 % of the unigenes were associated with biochemical synthesis and metabolism, such as "carbohydrate transport and metabolism" (1242, 4.5 %), "amino acid transport and metabolism" (1184, 4.3 %) and "secondary metabolites biosynthesis, transport and catabolism" (544, 2.0 %). A total of 1225 unigenes (4.46 %) was "function unknown".

Furthermore, we classified 19,715 (11.40 %) unigenes into 274 KEGG pathways by searching against the KEGG database (Additional file 3: Table S2). Among these pathways, "metabolic pathways"(3684, 18.7 %) represented the largest group, followed by "biosynthesis of secondary metabolites"(1616, 8.2 %). The other highly represented pathways included "microbial metabolism in diverse environment (803, 4.1 %)", "ribosome (509, 2.6 %)" and "protein processing in endoplasmic reticulum (507, 2.6 %)" (Table 4). A total of 301 unigenes involved in plant hormone signal transduction were found, including 9 pathways controlling the signal transduction of auxin, cytokinin, gibberellin acid, abscisic acid, ethylene, brassinosteroid, jasmonic acid and salicylic acid (Table 5). Besides, through the KEGG pathway annotation, a circadian rhythm pathway including 58 unigenes was also found, which can be used for further studies on the flowering-related genes of the photoperiod pathway.

Based on the annotation of unigenes, 118 MADS-box genes were also discovered (Additional file 4: Table S3), these including the sepal development related genes such as *APETALA1* (*AP1*), *CAULOFLOWER* (*CAL*); petal development related genes *DEFICIENS* (*DEF*)and *PISTILLATA* (*PI*); the C/D/E class function genes *AGAMOUS* (*AG*), *SEEDSTICK* (*STK*) and *FLORAL BINDING PROTEIN*-LIKE (*FBP*-like) were also identified. TCP gene family play important role in the development of plants and can be divided into two classes: PCF class and TCP-C class (which involes the CYC/TB1 clade and CIN clade) (Howarth and Donoghue 2006; Navaud et al. 2007). The CYC/TB1 clade includes genes mainly involved in the development of axillary meristems and floral bilateral symmetry (Luo et al. 1996; Doebley et al. 1997). To determine the TCP gene family in *C. faberi*, we analyzed the transcriptome database generated by
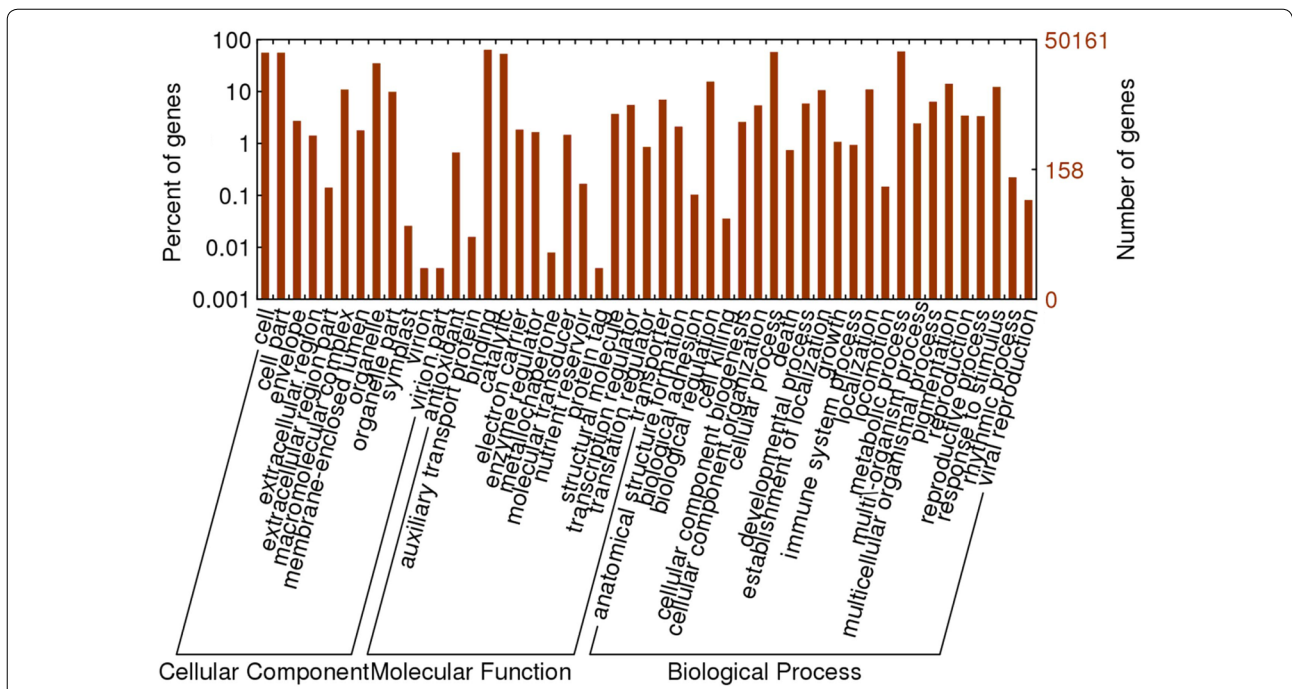

Fig. 2 Characteristics of sequence homology of *C. faberi* BLASTED against NCBI non-redundant (Nr) database. **a** E-value distribution of BLAST hits for matched unigene sequences, using an E-value cutoff of $1.0E^{-5}$. **b** Similarity distribution of top BLAST hits for each unigene. **c** Species distribution of the top BLAST hits

Fig. 3 GO classification of all annotated unigenes. GO classification of all annotated 50,161 unigenes. All terms belonged to the three main GO categories: biological process, cellular component and molecular function. The *x*-axis indicated the subcategories, the right *y*-axis indicated the number of genes in each category, the *left y*-axis indicated the percentage of a specific category of genes in the corresponding GO category. *Red column* indicated all annotated unigenes
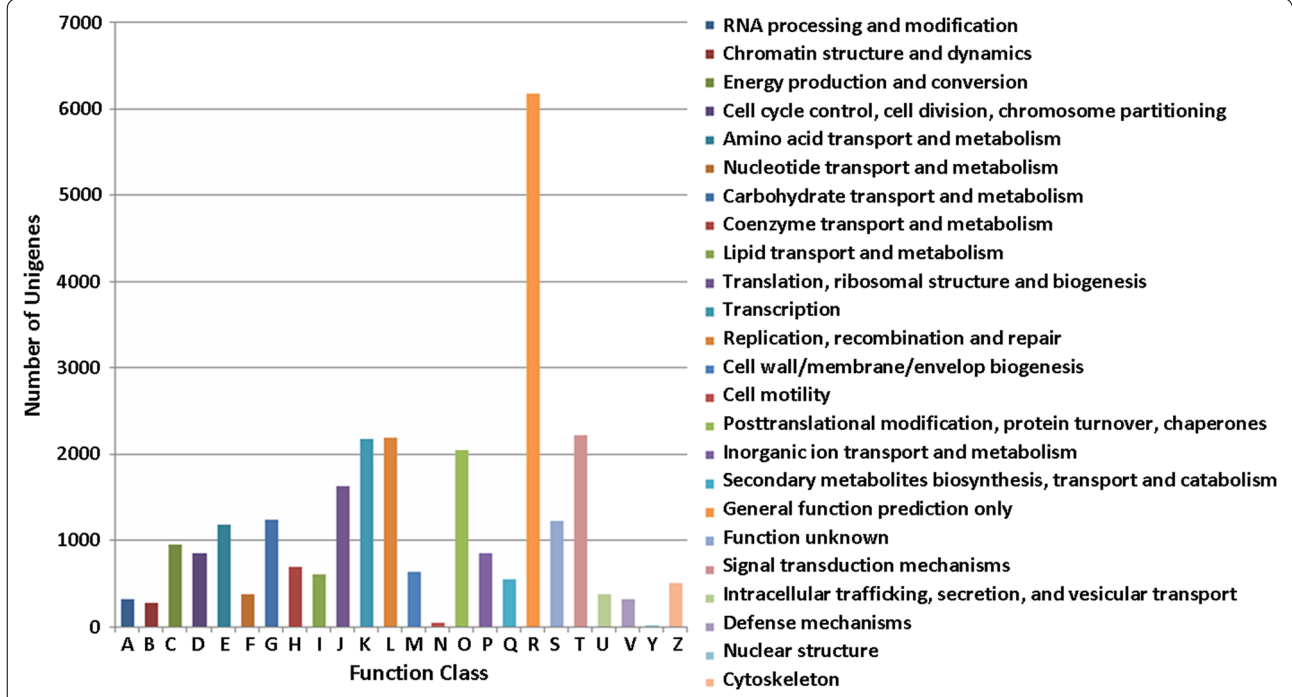

Fig. 4 COG annotations of putative proteins. All putative proteins were aligned to the COG database and were functionally classified into at least 24 molecular families. The capital letters in *x*-axis indicates the COG categories as listed on the right of the histogram and the *y*-axis indicates the number of unigenes in the corresponding COG category

Table 4 Significantly enriched pathways of all unigenes and DEGs in *C. faberi*

Pathway	No. of all genes with pathway annotation (19,715)	No. of DEGs with pathway annotation (66)		Pathway ID
		Up-regulated genes	Down-regulat-ed genes	
Metabolic pathways	3681	60	240	ko01100
Biosynthesis of secondary metabolites	1616	66	186	ko01110
Microbial metabolism in diverse environments	803	46	96	ko01120
Ribosome	509	18	71	ko03010
Protein processing in endoplasmic reticulum	507	14	48	ko04141
Spliceosome	469	27	17	ko03040
RNA transport	457	21	21	ko03013
Starch and sucrose metabolism	403	12	48	ko00500
Cell cycle	355	25	8	ko04110
Tuberculosis	337	11	26	ko05152
Pyrimidine metabolism	330	19	13	ko00240
RNA degradation	327	19	7	ko03018

Table 5 The pathways involved in plant hormone biosynthesis

| Pathway	Product	Pathway ID	Unigene number
Cysteine and methionine metabolism	Ethylene	ko00270	125
Phenylalanine metabolism	Salicylic acid	ko00360	103
Tryptophan metabolism	Auxin	ko00380	90
alpha-Linolenic acid metabolism	Jasmonic acid	ko00592	59
Diterpenoid biosynthesis	Gibberellin	ko00904	47
Carotenoid biosynthesis	Abscisic acid	ko00906	28
Zeatin biosynthesis	Cytokinin	ko00908	18
Brassinosteroid biosynthesis	Brassinosteroid	ko00905	17
Indole alkaloid biosynthesis	Indole-acetic acid	ko00901	2

this study and found 35 unigenes annotated as TCP TFs, including 14 *CIN*-like genes, 20 *PCF*-like genes and 1 *CYC*/*TB1*-like genes. As shown in Additional file 5: Table S4, they showed homology with 13 *Arabidopsis* TCP genes.

A total of 240 genes associated with flowering time were obtained (Additional file 6: Table S5). These include floral meristem identity genes *LEAFY* (*LFY*) and *APETALA1* (*AP1*); autonomous pathways genes *FCA*, *FPA*, *FLOWERING LOCUS* (*FLD*), *FY*, *FVE*, *FLOWER-ING LATE KH MOTIF* (*FLK*); vernalization pathways genes *FRI* and *VERNALIZATION INSENSITIVE* (*VIN*); photoperiod pathway genes such as *FT*, *Phytochrome A* (*PHYA*), *Phytochrome B* (*PHYB*), *PIF3*, *ELF3*, *LHY*, *SUP-PRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) and *CO*; Gibberellin (GA) pathway genes such as

GIBBERELLIC ACID INSENSITIVE (*GAI*). All these unigenes provide important resources for future study of floral organ development, floral bilateral symmetry and flowering time.

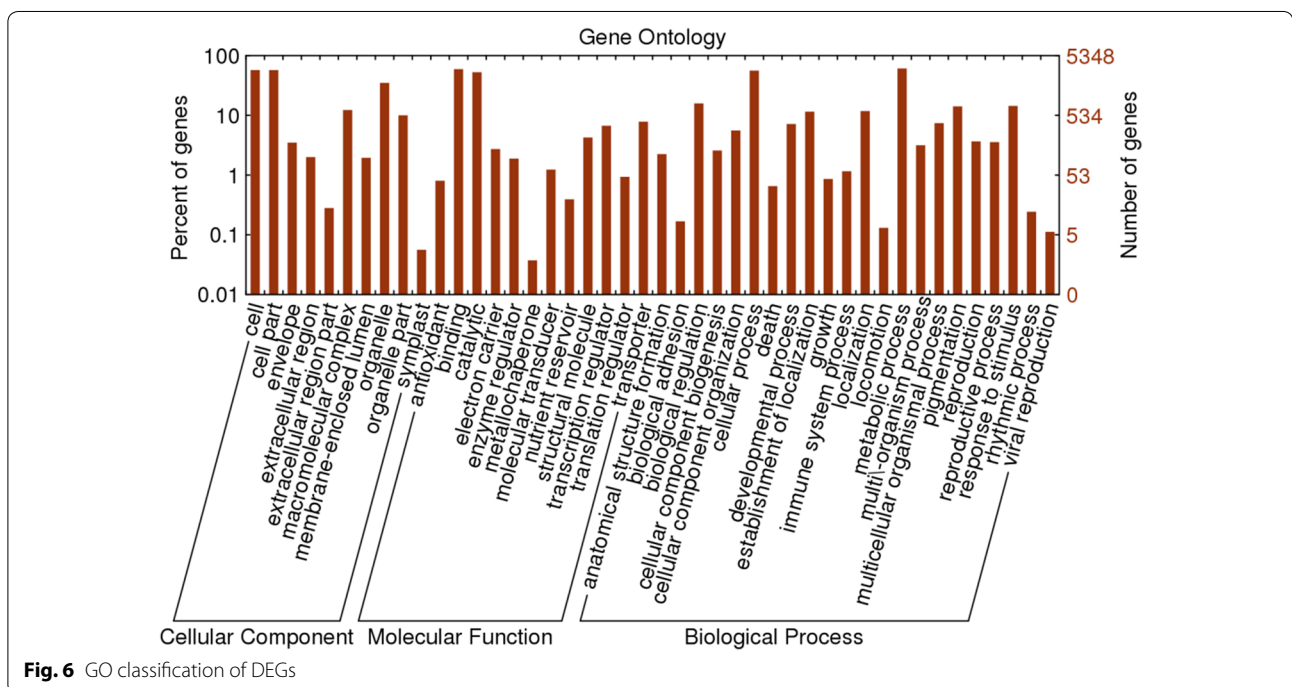**DEG analysis of vegetative and flower buds**

The FPKM methods were used to analyze the gene expression in the two libraries: 123,128 and 125,862 unigenes were obtained in the flower buds and vegetative buds, of which 39,549 and 42,283 genes expressed specifically in the flower buds and vegetative buds, respectively (Fig. 5). To analyze different gene expression in the two libraries, 13,484 DEGs were identified using FPKM methods. Among them, 7683 were down-regulated and 5801 were up-regulated in the flower buds when compared to those in the vegetative buds, including 3430 and 6556 genes specifically enriched in flower
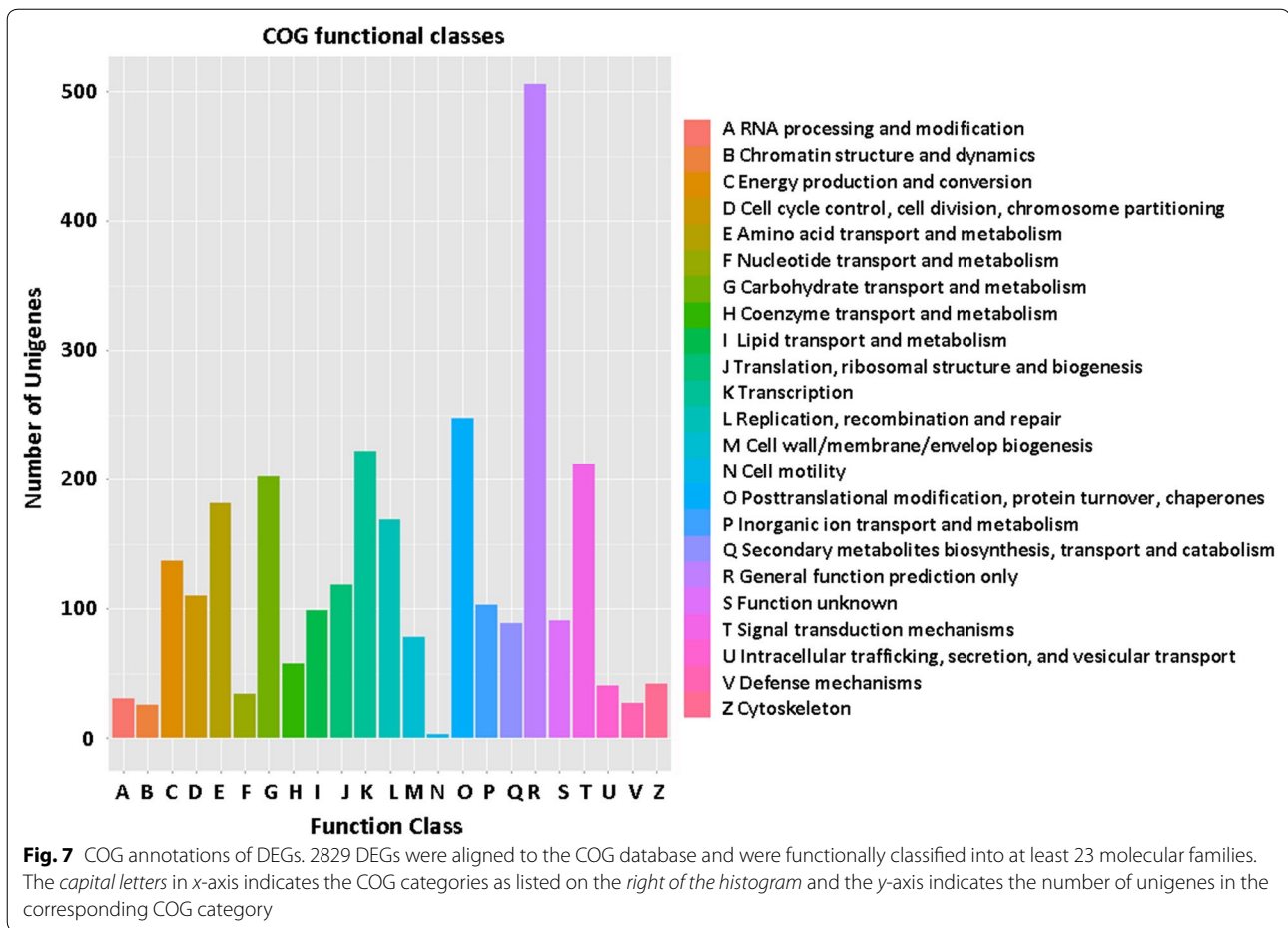
Fig. 5 The numbers of specific genes and shared genes between the flower and vegetative buds of *C. faberi*
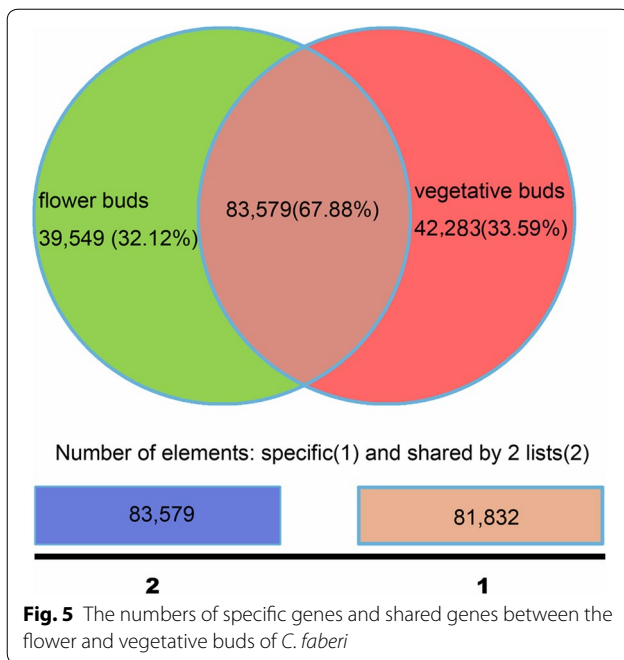
and vegetative buds, respectively. To validate and annotate the assembled DEGs, the 13,484 DEGs were subjected to BLASTX comparisons (*E*-value $\leq 10^{-5}$) against GO, COG and KEGG database to identify putative functions of these unigenes. As a result, 5348 (39.7 %), 2829 (21.0 %) and 3927 (29.1 %) DEGs had homologous sequences in GO, COG and KEGG databases,

respectively. In addition, 5348 DEGs were classified into 44 GO term annotations, including cellular component (1399), biological process (3921) and molecular function (7006), with multiple terms assigned to the same transcript (Fig. 6). A total of 2829 DEGs were classified into 23 functional COG categories (Fig. 7). "general function prediction only" (506) was the most, followed by "posttranslational modification, protein turnover, chaperones" (248) and "transcription" (222). To further understand the function in biological processes, 3927 DEGs were classified into 245 KEGG pathways. These results showed that most of the DEGs were involved in the "metabolic pathways"(300), "biosynthesis of secondary metabolites"(252), "microbial metabolism in diverse environment (142)" and "ribosome"(89) (Table 4). After transcriptome analysis, TFs involved in the floral differentiation were revealed and their expression levels in the flower and vegetative buds were calculated and compared using the FPKM method.

A total of 173 DEGs related to the floral development were discovered in our transcriptome data (Additional file 7: Table S6). These TFs were attributed to different gene families, including MADS (34 DEGs), TCP (12 DEGs), MYB (34 DEGs), ARF (9 DEGs), NAC (21 DEGs), b-ZIP (17 DEGs), WRKY (18 DEGs) (Table 6). Twenty-eight DEGs related to the flowering time were identified (Additional file 8: Table S7) and some shown in Table 7. Most of these TFs showed higher expression in the flower buds indicating that these genes might involve in the flower development.

Fig. 6 GO classification of DEGs

Fig. 7 COG annotations of DEGs. 2829 DEGs were aligned to the COG database and were functionally classified into at least 23 molecular families. The *capital letters* in *x*-axis indicates the COG categories as listed on the *right of the histogram* and the *y*-axis indicates the number of unigenes in the corresponding COG category

Table 6 DEGs related to floral differentiation in the flower and vegetative buds

Genes related to floral differentiation	Total No. of DEGs	No. of DEGs in flower buds	No. of up-regulated DEGs in flower buds	No. of DEGs in vegetative buds	No. of up-regulated DEGs in vegetative buds
MADS	34	33	29	8	5
TCP	12	7	0	12	12
MYB	34	30	29	11	5
ARF	9	7	1	9	8
NAC	21	20	20	8	1
b-ZIP	17	14	13	13	4
WRKY	18	17	15	10	3

qRT-PCR validation of selected genes

To verify the reliability of RNA-Seq data and explicit the expression patterns of genes related to floral development, ten genes mainly related to the floral organ development, floral symmetry and flowering time associated genes were selected to perform qRT-PCR analysis. Our results indicated that *DEF, PI, AG, AP1, BRANCHED2* (*BRC2*) and *CCA1* had higher expression and *TCP4, CO,*

SOC1 and *LFY* displayed lower expression in the flower buds (Fig. 8a-j). *PI* and *DEF*, the two B class function genes controlling the petal development, showed significantly higher expression in the flower buds than the vegetative buds. While *CCA1* participating in the circadian rhythm showed significantly higher expression in the flower buds than the vegetative buds. Overall, all of the 10 genes performed the same expression patterns in our

Table 7 Some DEGs related to flowering time in the flower and vegetative buds

Genes related to flowering	Total No. of DEGs	No. of DEGs in flower buds	No. of up-regulated DEGs in flower buds	No. of DEGs in vegetative buds	No. of up-regulated DEGs in vegetative buds
PHYA	1	1	0	1	1
CCA1	3	2	2	3	1
LHY	1	1	1	1	0
FCA	2	2	1	1	1
FY	1	1	0	1	1
FRI	3	1	1	2	2
CO	2	1	1	2	1

qRT-PCR results as in the RNA-Seq data, verifying the reliability of the Illumina sequencing data.
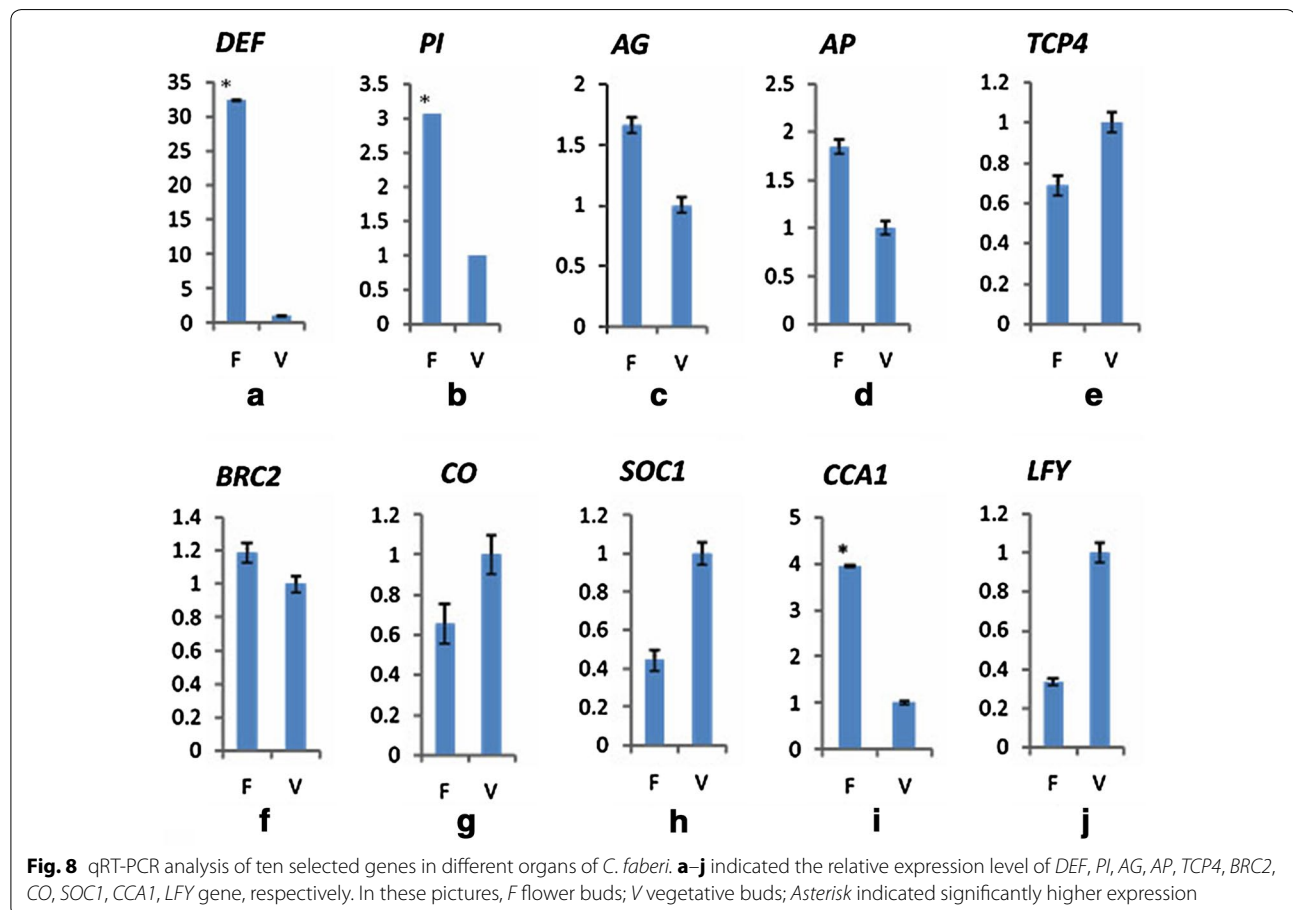
Discussion

Capacity of the transcriptome database

In recent years, transcriptome analysis based on deep RNA sequencing has been applied to many plants to identify differences in gene expression levels among different cultivars, organs and different treatment conditions (Chen et al. 2014; Yates et al. 2014). In this study, we obtained 189,300 contigs and identified 172,959 unigenes by de novo assembly through RNA-Seq technology in the vegetative and flower buds of *C. faberi*. According to the Nr protein database, 66,000 (38.16 %) unigenes were successfully annotated. Furthermore, 173 DEGs involved in floral organ development, floral zygomorphy and flowering time were found. These results supported that plant conservative genes, *C. faberi*-specific genes, and *C. faberi* tissue-specific genes all were identified in our transcriptomic analysis. Our research will provide valuable information for future study to inquire the flower development mechanisms of *C. faberi*.

Genes related to the floral organ development

MADS-box genes are known for their roles in the flower organ development in *Arabidopsis* and *Antirrhinum* (Coen and Meyerowitz, 1991; Weigel and Meyerowitz, 1994). While some monocot species like the orchid family possess distinct floral structures, thus floral patterning in *Arabidopsis* may not be comparable to such flowering

Fig. 8 qRT-PCR analysis of ten selected genes in different organs of *C. faberi*. **a**–**j** indicated the relative expression level of *DEF*, *PI*, *AG*, *AP*, *TCP4*, *BRC2*, *CO*, *SOC1*, *CCA1*, *LFY* gene, respectively. In these pictures, *F* flower buds; *V* vegetative buds; *Asterisk* indicated significantly higher expression

plants. In recent years, a few MADS-box genes have been identified and characterized in *D. thyrsiflorum* (Reichb. f.) (Martin et al. 2006), *Oncidium* Gower Ramsey (Hsu and Yang 2002), *Phalaenopsis* (Tsai et al. 2004) and *C. faberi* (Xiang et al., 2011). According to the "Orchid code" theory, the identity of the lateral petals and floral lip were determined by four different *AP3/DEF*-like genes, whereas the *PI/GLO*-like genes retained the function unchanged like class A, C, D and E genes (Mondragón-Palomino and Theissen 2008, 2009, 2011; Aceto and Gaudio 2011). In this study, 34 DEGs of MADS-like genes were found. Meanwhile, 29 DEGs were up-regulated in the flower buds (26 DEGs specifically expressed in the flower buds and most of them were class B and class C genes) and only 5 DEGs were up-regulated in the vegetative buds. Meanwhile, the expression level of *PI* and *DEF* showed significantly higher in the flower buds than the vegetative buds, proving that *PI* and *DEF* played a pivotal role in the flower development.

Genes related to the floral zygomorphy

The ornamental value of *C. faberi* is determined by many factors, especially the novel flower color, shape and fragrance. Floral zygomorphy have evolved multiple times from radially symmetrical (actinomorphic; polysymmetric) ancestors in different angiosperm lineages (Endress 1999; Stebbins 1974). The mechanism of TCP model for bilateral flower symmetry has been well established in model species, such as snapdragon (*A. majus*), *Lotus japonicus* and rice (Luo et al. 1996; Feng et al. 2006; Yuan et al. 2009). In *A. majus*, *CYC* and *DICH* control the dorsoventral asymmetry (Luo et al. 1999). In *L. japonicus*, floral dorsoventral asymmetry is regulated by three *LjCYC* genes (*LjCYC*1, *LjCYC*2, *LjCYC*3), especially the *LjCYC*2. In monocot, *RETARDED PALEA1* (*REP1*) controls palea development and floral zygomorphy in rice, whose flower is different from that of *A. majus* or *L. japonicus*. In our *C. faberi* transcriptome data, we identified 35 *TCP* genes, of which 10 *CIN*-like genes and one *PCF*-like gene showed significantly higher expression levels in the vegetative buds compared to the flower buds. qRT-PCR analysis revealed that the expression level of *BRC2* was higher in the flower buds than the vegetative buds (Fig. 8e), suggesting that it might play important role in the regulation of floral zygomorphy of *C. faberi*.

Genes related to the regulation of flowering

Previous studies showed that the flowering of *Arabidopsis* was regulated by four pathways, including autonomous pathway, vernalization pathway, gibberellic acid (GA)-dependent pathway and the photoperiod pathway. And *CCA1* influenced the circadian rhythm, overexpression of *CCA1* resulted in long hypocotyls and late flowering (Wang and Tobing 1998). *CO* involved in the

photoperiod pathway in *Arabidopsis* and acted as a floral promoter, which was regulated by the circadian clock (Yano et al. 2000). The temporal and spatial regulation of *CO* turned out to be important to the photoperiod-dependent induction of flowering (An et al. 2004). Early in a day, the expression of *CO* was low, then increased rapidly 10 h after dawn and peaked at around 15 h (Suárez-López et al. 2001). *LFY* is a major floral meristem regulator, its overexpression caused early flowering in transgenic *Arabidopsis* (Benlloch et al. 2007). A single *PHYA* gene, *LHY* gene and three *CCA1* genes involved in the photoperiod pathways showed more highly expressed in the vegetative buds than in the flower buds. Two *VRN* genes playing important roles in the vernalization pathway were detected in our study, and were found to be more highly expressed in the vegetative buds than in the flower buds. Besides, DEGs related to the autonomous pathway were also detected, including two *FCA* genes, a *FLD* gene and a *FY* gene. qRT-PCR revealed that expression of *LFY* was higher in the vegetative buds and expression of *CO*, *SOC1*, *CCA1* was higher in the flower buds, which was consistent with the RNA-Seq data (Fig. 8g–j). Expression of *CCA1* was significantly higher in the flower buds than the vegetative buds. This observation suggesting that the mechanism of flowering regulation in *C. faberi* might be similar to *Arabidopsis*.

In conclusion, the current RNA-seq and DEG analysis revealed 393 genes associated to the floral development, including 35 TCP transcription factors (TFs), 118 MADS-box genes and 240 flowering time related genes. A total of 173 DEGs were identified, including 12 TCP genes, 34 MADS-box genes and 28 flowering time related genes. The transcriptome database in the present study will be a valuable supplement to the genomic sequence dataset of *C. faberi*, and will serve as an important public information platform for further studies on the flower development mechanism in *C. faberi*.

Additional files

Additional file 1: Table S1. Gene-specific primers for qRT-PCR.

Additional file 2: Fig. S1. Length distribution of unigenes.

Additional file 3: Table S2. Result of metabolic pathway analysis for unigenes by KEGG annotation.

Additional file 4: Table S3. Unigenes that share homology with MADS-box genes.

Additional file 5: Table S4. Unigenes that share homology with TCP family genes in *C. faberi*.

Additional file 6: Table S5. Unigenes that related to the flowering time in *C. faberi*.

Additional file 7: Table S6. The functional annotation for the DEGs related with flower development in *C. faberi*.

Additional file 8: Table S7. DEGs related to the flowering time in *C. faberi*.

References

Aceto S, Gaudio L (2011) The mads and the beauty: genes involved in the development of orchid flowers. Curr Genomics 12:342–356

Almeida J, Rocheta M, Galego L (1997) Genetic control of flower shape in *Antirrhinum majus*. Development 124:1387–1392

An H, Roussot C, Suárez-López P, Corbesier L, Vincent C, Piñeiro M, Hepworth S, Mouradov A, Justin S, Turnbull C, Coupland G (2004) *CONSTANS* acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of *Arabidopsis*. Development 131:3615–3626

Anusak P, Ditta DS, Beth S, Liljergren SJ, Elvira B, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424:85–88

Bartlett ME, Specht CD (2011) Changes in expression pattern of the *teosinte branched1*-like genes in the *Zingiberales* provide a mechanism for evolutionary shifts in symmetry across the order. Am J Bot 98:227–243

Benlloch R, Berbel A, Serrano-Mislata A, Madueño F (2007) Floral initiation and inflorescence architecture: a comparative view. Ann Bot 100:659–676

Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, Chen LJ, He Y, Xu Q, Bian C, Zheng Z, Sun F, Liu W, Hsiao YY, Pan ZJ, Hsu CC, Yang YP, Hsu YC, Chuang YC, Dievart A, Dufayard JF, Xu X, Wang JY, Wang J, Xiao XJ, Zhao XM, Du R, Zhang GQ, Wang M, Su YY, Xie GC, Liu GH, Li LQ, Huang LQ, Luo YB, Chen HH, Van de Peer Y, Liu ZJ (2014) The genome sequence of the orchid *Phalaenopsis equestris*. Nat Genet 47:65–72

Chen X, Zhang J, Liu QZ, Guo W, Zhao TT, Ma QH, Wang GX (2014) Transcriptome sequencing and identification of cold tolerance genes in Hardy *Corylus* species (*C. heterophylla* Fisch) floral buds. PLoS One 9(9):e108604

Cheung F, Haas BJ, Goldberg SM, May GD, Xiao Y, Town CD (2006) Sequencing *Medicago truncatula* expressed sequenced tags using 454 Life Sciences technology. BMC Genom 7:272

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. Nature 353:31–37

Corley SB, Carpenter R, Copsey L, Coen E (2005) Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. Proc Natl Acad Sci 102:5068–5073

Costa MMR, Fox S, Ai Hanna, Baxter C, Coen E (2005) Evolution of regulatory interactions controlling floral asymmetry. Development 132:5093–5101

Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386:485–488

Endress PK (1999) Symmetry in flowers: diversity and evolution. Int J Plant Sci 160:3–23

Feng X, Zhao Z, Tian Z, Xu S, Luo Y, Cai Z, Wang Y, Yang J, Wang Z, Weng L, Chen J, Zheng L, Guo X, Luo J, Sato S, Tabata S, Ma W, Cao X, Hu X, Sun C, Luo D (2006) Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proc Natl Acad Sci 103:4970–4975

Galego L, Almeida J (2002) Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. Genes Dev 16:880–891

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420–3435

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, Palma FD, Birren BW, Nusbaum C, Lindblad-TohK Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652

Hao QN, Zhou XA, Sha AH, Wang C, Zhou R, Chen SL (2011) Identification of genes associated with nitrogenuse efficiency by genome-wide transcriptional analysis of two soybean genotypes. BMC Genom 12:1–15

Hoshino Y, Igarashi T, Ohshima M, Shinoda K, Murata N, Kanno A, Nakano M (2014) Characterization of *CYCLOIDEA*-like genes in controlling floral zygomorphy in the monocotyledon *Alstroemeria*. Sci Hortic 169:6–13

Howarth DG, Donoghue MJ (2006) Phylogenetic analysis of the 'ECE' (CYC/TB1) clade reveals duplications predating the core eudicots. Proc Natl Acad Sci USA 103:9101–9106

Hsu HF, Yang CH (2002) An orchid (*Oncidium* 'Gower Ramsey') *AP3*-like MADS genes regulate floral formation and initiation. Plant Cell Physiol 43:1198–1209

Hsu HF, Huang CH, Chou LT, Yang CH (2003) Ectopic expression of an orchid (*Oncidium* Gower Ramsey) *AGL6*-like gene promotes flowering by activating flowering time genes in *Arabidopsis thaliana*. Plant Cell Physiol 44(8):783–794

Hsu HF, Hsu WH, Lee YI, Mao WT, Yang JY, Li JY, Yang CH (2015) Model for perianth formation in orchids. Nat Plants 46:1–8

Huang WT, Fang ZM, Zeng SJ, Zhang JX, Wu KL, Chen ZC, Silva TS, Duan J (2012) Molecular cloning and functional analysis of three *FLOWERING LOCUS T* (*FT*) homologous genes from Chinese *Cymbidium*. Int J Mol Sci 13:11385–11398

Hyun TK, Lee S, Kumar D, Rim Y, Kumar R, Lee SY, Lee CH, Kim JY (2014) RNA-seq analysis of *Rubus idaeuscv*. Nova: transcriptome sequencing and *de novo* assembly for subsequent functional genomics approaches. Plant Cell Rep 33:1617–1628

Jang S, Choi SC, Li HY, An G, Schmelzer E (2015) Functional characterization of *Phalaenopsis aphrodite* flowering genes *PaFT1* and *PaFD*. PLoS One 10(8):e0134987

Li XB, Luo J, Yan TL, Xiang L, Jin F, Qin DH, Sun CB, Xie M (2013) Deep sequencing-based analysis of the *Cymbidium ensifolium* floral transcriptome. PLoS One 8(12):e85480

Liu XR, Pan T, Liang WQ, Gao L, Wang XJ, Li HQ, Liang S (2016) Over expression of an orchid (*Dendrobium nobile*) SOC1/TM3-like ortholog, *DnAGL19*, in *Arabidopsis* regulates *HOS1-FT* expression. Front Plant Sci 7:99

Luo D, Carpenter R, Vincent C, Copsey L, Coen E (1996) Origin of floral asymmetry in *Antirrhinum*. Nature 383:794–799

Luo D, Carpenter R, Copsey L, Vincent C, Clark J, Coen E (1999) Control of organ asymmetry in flowers of *Antirrhinum*. Cell 99:367–376

Martin S, Johansen LB, Pederson KB, Signe F, Bo BJ (2006) Cloning and transcription analysis of an *AGAMOUS*- and *SEEDSTICK* ortholog in the orchid *Dendrobium thyrsiflorum* (Reichb. f.). Gene 366:266–274

Mondragón-Palomino M, Theissen G (2008) MADS about the evolution of orchid flowers. Trends Plant Sci 13:51–59

Mondragón-Palomino M, Theissen G (2009) Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. Ann Bot 104:583–594

Mondragón-Palomino M, Theissen G (2011) Conserved differential expression of paralogous *DEFICIENS*- and *GLOBOSA*-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. Plant J 66:1008–1019

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: interacting pathways as a basis for diversity. Plant Cell 14(Suppl):111–130

Münster T, Pahnke J, Di Rosa A, Kim JT, Martin W, Saedler H, Theissen G (1997) Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. Proc Natl Acad Sci 94:2415–2420

Navaud O, Dabos P, Carnus E, Tremousaygue D, Hervé C (2007) TCP transcription factors predate the emergence of land plants. J Mol Evol 65:23–33

Paolo SD, Gaudio L, Aceto S (2015) Analysis of the TCP genes expressed in the inflorescence of the orchid *Orchis italica*. Sci Rep 5:16265

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require *SEPALLATA* MADS-box genes. Nature 405:200–203

Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF (2001) *APETALA1* and *SEPALLATA3* interact to promote flower development. Plant J 26:385–394

Preston JC, Hileman LC (2012) Parallel evolution of TCP and B-class genes in Commelinaceae flower bilateral symmetry. Evodevo 3:6

Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN (2005) Genera Orchidacearum: Epidendroideae (part one). Oxford University Press, Oxford

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics 140:345–356

Rudall PJ, Bateman RM (2002) Roles of synorganization, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. Biol Rev 77:403–441

Stebbins GL (1974) Flowering plants: evolution above the species level. Harvard University Press, Cambridge

Su CL, Chao YT, Yao-Chien AC (2011) *De Novo* assembly of expressed transcripts and global analysis of the *Phalaenopsis aphrodite* transcriptome. Plant Cell Physiol 52(9):1501–1514

Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature 410:1116–1120

Trick M, Long Y, Meng JL, Bancroft I (2009) Single nucleotide polymorphism (SNP) discovery in the polyploidy *Brassica napus* using Solexa transcriptome sequencing. Plant Biotechnol J 7:334–346

Tsai WC, Kuoh CS, Chuang MH, Chen WH, Chen HH (2004) Four *DEF*-like MADS box genes displayed distinct floral morphogenetic roles in *Phalaenopsis* orchid. Plant Cell Physiol 45:831–844

Wang ZY, Tobing EM (1998) Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED I* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. Cell 93:1207–1217

Wang Z, Luo Y, Li X, Wang L, Xu S, Yang J, Weng L, Sato S, Tabata S, Ambrose M, Rameau C, Feng X, Hu X, Luo D (2008) Genetic control of floral zygomorphy in pea (*Pisum sativum* L.). PNAS 105(30):10414–10419

Weigel D, Meyerowitz EM (1994) The ABCs of floral homeotic genes. Cell 78:203–209

Wolff E (1999) Chinese *Cymbidium* species. Orchids: the Magazine of the American Orchid Society 68(7):682–693

Xiang L, Qin DH, Li XB, Li BJ, Guo FQ, Wu C, Sun CB (2011) Cloning and expression analysis of B Class MADS-box genes from *Cymbidium faberi*. Acta Hortic Sin 38:143–147

Xu YF, Teo LL, Zhou J, Prakash PK, Yu H (2006) Floral organ identity genes in the orchid *Dendrobium crumenatum*. Plant J 46:54–68

Yan L, Wang X, Liu H, Tian Y, Lian J, Yang R, Hao S, Wang X, Yang S, Li Q, Qi S, Kui L, Okpekum M, Ma X, Zhang J, Ding Z, Zhang G, Wang W, Dong Y, Sheng S (2015) The genome of *Dendrobium officinale* illuminates the biology of the important traditional Chinese orchid herb. Mol Plant 8(6):922–934

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. Plant Cell 12:2473–2484

Yates SA, Swain MT, Hegarty MJ, Chernukin I, Lowe M, Allison GG, Ruttink T, Abberton MT, Jenkins G, Skøt Leif (2014) *De novo* assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC Genom 15:453

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34:293–297

Yu H, Yang SH, Goh CJ (2000) *DOH1*, a class 1 knox gene, is required for maintenance of the basic plant architecture and floral transition in orchid. Plant Cell 12(11):2143–2159

Yu H, Yang SH, Goh CJ (2002) Spatial and temporal expression of the orchid floral homeotic gene *DOMADS1* is mediated by its upstream regulatory regions. Plant Mol Biol 49(2):225–237

Yuan Z, Gao S, Xue DW, Luo D, Li LT, Ding SY, Yao X, Wilson ZA, Qian Q, Zhang DB (2009) *RETARDED PALEA1* controls palea development and floral zygomorphy in rice. Plant Physiol 149:235–244

Zhang JX, Wu KL, Zeng SJ, Silva JT, Silva JA, Xia HQ, Tian CE, Duan J (2013) Transcriptome analysis of *Cymbidium sinense* and its application to the identification of genes associated with floral development. BMC Genom 14:1–17

Zhang HN, Wei YZ, Shen JY, Lai B, Huang XM, Ding F, Su ZX, Chen HB (2014) Transcriptomic analysis of floral initiation in litchi (*Litchi chinensis* Sonn.) based on *de novo* RNA sequencing. Plant Cell Rep 33:1723–1735

Zhang GQ, Xu Q, Bian C, Tsai WC, Yeh CM, Liu KW, Yoshida K, Zhang LS, Chang SB, Chen F, Shi Y, Su YY, Zhang YQ, Chen LJ, Yin Y, Lin M, Huang H, Deng H, Wang ZW, Zhu SL, Zhao X, Deng C, Niu SC, Huang J, Wang M, Liu GH, Yang HJ, Xiao XJ, Hsiao YY, Wu WL, Chen YY, Mitsuda N, Ohme-Takagi M, Luo YB, Van de Peer Y, Liu ZJ (2016) The *Dendrobium catenatum* Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Sci Rep 6:19029