

# Elucidation of molecular mechanisms of flower form development in tree peony ( *Paeonia suffruticosa* ) through comparative transcriptome analysis of floral parts

**Jiuxing Lu**

Henan Agricultural University

**Yun Zheng**

Henan Agricultural University

**Haoning Wang**

Henan Agricultural University

**Zheng Wang**

Henan Agricultural University

**Yonghua Li** (✉ [lyhhnau@126.com](mailto:lyhhnau@126.com))

Henan Agricultural University

**Gary Gao**

Ohio State University

**Yan Li**

Henan Agricultural University

---

## Research article

**Keywords:** Flower shape, RNA sequencing, Flower development, MADS-box genes

**Posted Date:** January 24th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.18308/v2>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Background: Tree peony (*Paeonia suffruticosa*) is an economically, medicinally and ornamentally important woody flowering plant in East Asia. It is also a common ornamental shrub in Europe and North America. They are well known and prized for their beautiful flowers in many different shapes. Stamen petalody has been shown to be the most effective way to modify flower shapes. However, there is limited information on the molecular mechanisms of stamen petalody and flower shape formation in tree peony.

Results: In this study, RNA sequencing was used to assemble and annotate the unigenes in tree peony to identify the critical genes related to flower parts formation and verify the key genes in different flower shapes of tree peony cultivars. A total of 76,007 high quality unigenes were assembled and 30,505 were successfully annotated. A total of 1,833 transcription factors (TFs) were identified in our study, among them 16 MADS-box genes were found and characterized. Six key genes were selected to verify their functions in stamen petalody. AG and SEP showed high expression level in carpels and sepals separately both in stamen petalody and non-stamen petalody groups. PI and AP3 showed higher expression levels of inter-petals in the stamen petalody group, compared to stamens of non-stamen petalody.

Conclusion: Sixteen MADS-box genes were identified in tree peony through RNA-seq. We identified six key genes based on their differential expression levels in different flower parts. These six key genes represented all categories in the ABCDE model to verify the functions in stamen petalody. We speculate that PI and AP3 may trigger the stamen petalody in tree peony. Our study has helped establish the flower development model in tree peony, to identify key molecular mechanisms in the development of different flower shapes, and to provide valuable information for improving the genetic diversity of tree peony and many other woody plants.

## Background

Flower shapes of ornamental plants are typically controlled by multiple gene sets. The variation of flower parts such as petal, sepal, stamen and carpel are critical for the formation of flower shapes in ornamental plants. Natural mutations, human manipulations, and targeted selections have led to many attractive flower shapes in ornamental plants who are currently enjoyed by humans.

Age, ambient temperature, autonomous, gibberellin, photoperiod, and vernalization pathways have been reported to be the key pathways in flower initiation of *Arabidopsis thaliana* [1] and numerous ornamental crops [2-7]. There are many regulatory genes involved in these six pathways of flower development. The ABC model of flower development based on exhaustive studies of *Arabidopsis thaliana* and *Antirrhinum majus* is widely regarded as the basic framework of flower development in many higher plants [8]. The original ABC model developed during the 1990s, was adopted as the most comprehensive model for flower development in 2000 and is now widely known as the ABCDE model [9-13]. Nearly all genes identified in the ABCED model (A class, B class, C class, D class and E class) are members of the MADS-box gene family [13-18]. Although there are anatomical differences between herbaceous annuals and

perennial woody plants, the key regulatory genes involved in flower initiation and development have been shown to be highly conserved.

Tree peony (*Paeonia suffruticosa*), a deciduous woody shrub, has been cultivated for more than 2,000 years in China. It is a well-known for its medicinal, nutraceutical and ornamental values in East Asia and highly treasured by gardeners all over the world [19, 20]. Tree peony is well known for its large and showy flowers in various flower shapes. The flower shapes of tree peony are very diverse and have been divided into several classes including simple, lotus, chrysanthemum, rose, anemone, thousand petal crown, hydrangea globular and hundred proliferate [20]. These flower shapes have been illustrated in several published papers [21, 22] and are shown in Fig.1.

Figure 1. Flower shapes of the tree peony. A. Simple; B. Lotus; C. Chrysanthemum; D. Rose; E. Anemone; F. Thousand petal crown; G Hydrangea globular; H. Hundred proliferate

The flower shapes of tree peony vary greatly in their arrangement, number, and shape of flower parts including petal, sepal, stamen and carpel [20]. Two main reasons for these highly distinctive flower shapes include: 1) an increase in the number of naturally occurring flower petals, and (2) stamen petalody. The number of naturally occurring flower petals in tree peony ranges from 30 to 50, while stamen petalody is greater than 200 [20]. The latter leads to more cultivars of different flower shapes than the number of naturally occurring flower petals in tree peony. In order to develop highly prized cultivars with more attractive flower shapes in tree peony, it is critical to gain a deep understanding of the molecular mechanisms of flower development. A limited number of genes involved in flower development have been cloned and expressed by real-time. However, most critical genes involved in flower development in tree peony have not been identified. In this study, we aim to identify the key genes involved in the development of flower parts using by RNA-seq and verify their specific roles in controlling the presence or the absence of stamen petalody. The results from this study will help elucidate the molecular mechanisms of flower development in tree peony and could lead to the more efficient development of new cultivars for the \$14 million tree peony industry in China!

## Results

### Transcriptome sequencing and assembly

A total of 118.98 Gb of clean data from four floral parts, with three biological replications each, was acquired following a systematic quality check and data filtering. The base number of each sample was greater than 8.5 Gb (Table 1). The percentages of the GC bases of all 12 samples were greater than 44%, and the percentages of the reads with an average quality value greater than 30 were over 89%. These parameters confirmed that sequencing data were reliable and appropriate for further analysis.

A total of 76,007 unigenes were obtained with a median contig (N50) length of 1,639 bp and a mean length of 965.14 bp (Table 2). The distribution of the unigene numbers in different length ranges were highly uniform (Additional fig. 1). Hence, it is reasonable to assume that all assembled unigenes have a high degree of integrity.

Table 1. Summary of the sequencing results for four floral organs in tree peony

Samples	Number	Read Number	Base Number	GC Content	%≥Q30
Petal	PE1	30,505,090	9,088,941,390	44.52%	89.99%
	PE2	34,668,124	10,331,991,108	44.57%	90.58%
	PE3	33,988,044	10,138,469,592	45.40%	91.04%
Sepal	SE1	32,785,787	9,774,077,598	44.89%	91.20%
	SE2	31,070,742	9,253,816,890	44.76%	91.13%
	SE3	29,309,927	8,742,359,852	44.83%	90.84%
Stamen	ST1	36,870,554	10,982,297,828	45.40%	92.12%
	ST2	32,823,379	9,781,335,460	45.41%	92.15%
	ST3	35,848,484	10,696,258,334	45.46%	91.78%
Carpel	CA1	34,168,656	10,192,298,092	44.54%	91.52%
	CA2	35,251,448	10,516,365,094	44.32%	91.63%
	CA3	31,763,113	9,479,948,186	44.49%	91.36%

Note: %≥Q30 represents the percentage of bases with a quality score of 30 or higher.

Table 2. Summary of the assembly for four floral organs in the tree peony

Length Range	Transcript	Unigene
200-300	23,645(13.32%)	18,966(24.95%)
300-500	26,606(14.99%)	15,324(20.16%)
500-1000	45,226(25.48%)	18,128(23.85%)
1000-2000	47,716(26.88%)	13,952(18.36%)
2000+	34,315(19.33%)	9,637(12.68%)
Total Number	177,508	76,007
Total Length	223,165,477	73,357,029
N50 Length	1,903	1,639
Mean Length	1257.21	965.14

Note: N50 Length represents the length of the unigene of N50.

## **Functional annotation and clustering of gene expression profiles of tree peony floral parts transcriptome**

BLAST was used to annotate unigenes into databases including Nr, Swiss-Prot, GO, COG, KOG, eggNOG4.5 and KEGG, resulting in 30,505 (40%) successfully annotated unigenes (Fig. 2A). Eighty eight percent of total annotated unigenes (29,872 of 30,505) were perfectly matched in Nr, while the smallest number of unigenes (8,321) were matched in KEGG. All unigenes were classified using a COG database search. In total, 8,321 unigenes were annotated into 25 COG classifications. Among them, the term “general function prediction only” (2042, 24.5%) represented the highest number of unigenes, while “extracellular structures” was represented by zero (Fig. 2B).

For the GO annotation, 17,184 annotated unigenes were classified into three principal categories and then further subdivided into 52 functional terms (Fig. 2C). Among them, “cell” (7113, 41.4%), “cell part” (7113, 41.4%) and “organelle” (5304, 30.9%) were the dominant function terms in the “cellular component” category. The term that was most represented in the “molecular function” category was “binding” (9171, 53.4%), followed by “catalytic activity” (8950, 52.1%). Within the “biological process” category, unigenes (11857, 69%) belonged to “metabolic process”.

Figure 2. A. Classification map of unigenes with COG database search; B. Classification map of unigenes with GO database search.

## **Identification of different expression unigenes and clustering of gene expression profiles of tree peony floral parts transcriptome**

Differences in gene expression profiles among four floral parts were examined by pairwise comparisons of the 12 libraries. A total of 11,801 (36.7%) unigenes were shared among all libraries, while approximately 14.3% of unigenes (4,355) were detectable in

individual floral parts. For petals, the number of unique unigenes was 328, which represented the smallest number among the four parts. Pistil had the largest number of unique unigenes, followed by stamen (1,129) and calyx (935) (Fig. 3A). Those genes expressed in certain floral parts, probably played important roles in the development of the coincident floral parts.

Based on the six comparisons (petal vs sepal, petal vs stamen, petal vs carpel, sepal vs stamen, sepal vs carpel and stamen vs carpel, we identified 9,592 different expression unigenes (DEGs) (Fig. 3C). Among them, the comparison of petal vs sepal had the smallest number of DEGs with 1,706 (17.8%), while the comparison of sepal vs stamen had the largest number of DEGs with 5,517 (57.5%) followed closely by stamen vs carpel 5,310 (55.4%) and petal vs stamen 3,505 (36.5%).

To further investigate the gene expression profiles, we performed K-Means clustering of gene expression levels of DEGs. We identified nine gene clusters with distinctive expression patterns (Fig. 3B). Cluster 7 and Cluster 9 contained 499 and 452 unigenes respectively and showed opposite patterns. While Cluster 7 showed a gradual increase from petal to carpel, Cluster 9 showed a gradual decrease, indicating different roles regarding the floral organ development. Cluster 1, 5 and 6, which included 1342, 894 and 550 unigenes, respectively, showed similar patterns with up-regulated unigenes in stamen conversely, Cluster 2, 3, 4 and 8 containing 1320, 1311, 1096 and 466 respectively, showed down-regulated unigenes in stamen. These facts indicate that stamen may be the floral part containing more complex molecular processes than the others.

Figure 3. A. Venn diagram of the unigenes of four flower parts in the tree peony; B. DEG numbers in six comparisons of four flower parts in the tree peony; C. K-means clustering map of the DEGs in four lower parts in the tree peony.

## **KEGG pathway enrichment of DEGs**

In order to identify active pathways in the four floral organs of tree peony, the DEGs were annotated in the KEGG pathways. KEGG pathway enrichment analysis was performed to search the dominant pathways in the four floral organs. Specific enrichment of genes was obtained for 102 pathways at petal vs sepal, for 113 pathways at petal vs stamen, for 102 pathways at petal vs carpel, for 116 pathways at sepal vs stamen, for 111 pathways at sepal vs carpel and for 115 pathways at stamen vs carpel. Among the six paired comparisons, a total of 5,616 unigenes were mapped in KEGG pathways including 324, 505, 361, 821, 571 and 835 DEGs, respectively, in petal vs sepal, petal vs stamen, petal vs carpel, sepal vs stamen, sepal vs carpel and stamen vs carpel. The results of each comparison are listed in Fig. 4. These pathways were divided into five groups including “cellular process,” “environmental information processing,” “genetic information processing,” “metabolism,” and “organizational systems.” A common theme was identified through these six paired comparisons and the dominant pathways were annotated with either a circle or a triangle (Fig. 4). In the “metabolism” group, the total numbers of DEGs regulating carbon metabolism, photosynthesis, biosynthesis of amino acids, starch and sucrose metabolism and phenylpropanoid biosynthesis were shown in greater abundance and these DEGs may be critical in the formation of flower parts in tree peony. In the “environmental information processing” group, the total number of DEGs for “plant hormone signal transduction” was also high.

Figure 4. KEGG analysis of six comparison sets. Circles represent dominant pathways in “metabolism group”; triangles represent dominant pathways in the “environmental information processing” group.

### **Key genes involved in regulating flowering time in tree peony**

The KEGG pathway analysis revealed a general view of the involvement of many biological processes among the floral parts development. To explore the detail expression

patterns of the key genes involving flowering time pathways, BlastT was used to search the orthologous of these key genes in *Arabidopsis thaliana*. A total of 24,163 pairs of orthologs were identified. Among them, key genes related to flowering time were selected and performed expression analysis (Fig. 5). We found that genes such as red/far-red light signaling pathways (*PHYA*, *PHYB*, *CRY1*, *ZTL*) had lower expression in stamen, which indicated that light signaling was less involved in stamen development. The expression level of genes related to the development of floral integrators and floral parts (*AP1*, *SOCI*, *FT*, *FLC*, *LHY*, *SVP*) showed expected patterns that were highly expressed in petal or sepal. The BLAST search also displayed that genes involved in autonomous pathways (*FCA*, *FY*, *FLD*, *FVE*, *FPA*) all shared similar expression patterns that reached peaks in carpel and troughs in stamen indicating that these genes are more involved in carpel development.

Figure 5. Heat map of key genes involving in flowering time pathways.

### **Analysis of MADS-box genes in tree peony**

A total of 1833 TFs, which represented 6% of annotated unigenes and fall into 205 TF families classified by plant transcription factor database PlantTFDB, were identified in our study. Among these TF gene families, AP2/ERF-ERF (68) was the most abundant TF family, followed by C2H2 (63) and bHLH (61). Remarkably, there were 29 MADS-box genes including 16 MADS-MIKC, were found in tree peony transcriptome data (Fig. 6A).

The MADS-box genes have been involved in many aspects of floral development and flower organs differentiation. To further examine the phylogenetic relationships among the tree peony MADS-box genes and to classified them into the established subfamilies, a phylogenetic neighbor-joining tree containing MADS-box genes in tree peony, grape, kiwifruit and arabidopsis was made (Fig. 6B). The phylogenetic analysis revealed that 16 MADS-box genes were classified into 10 subfamilies including a new clade that was not previously reported in arabidopsis. In addition, a few key MADS-box genes described in

ABCDE models were also included, such as *AP1* (*APETALA1*), *FUL* (*FRUITFULL*), *PI* (*PISTILATA*), *AP3* (*APETALA3*), *AG* (*AGAMOUS*), *STK* (*SEEDSTICK*), *SHP* (*SHATERPROOF*) and *SEP* (*SEPALLATA*). In the *AP1/FUL* subfamily (A class), c58746 was belonged to the *AP1* clade, while c87227 fell within *AGL79*. Two genes, c70349 and c80146, were assigned to the *PI/AP3* subfamily, which were functionally classified as B class. c57631, c77234 and c85161 belonged to the *AG/SHP* subfamily in C/D class, which included *AG*, *SHP* and *STK* three subclade. c77234 was classified into the *AG* clade in C class, while c57631 and c85161 were sorted into *STK* and *SHP* subclade respectively in D class. Only one gene c72216 was assigned to the *SEP* subfamily in the E functional group. In conclusion, the phylogenetic relationships showed that the tree peony contained all five classes of flower parts differentiation MADS-box genes.

To further explore how these regulators related to the flower developmental process, we examined the expression patterns of the 16 MADS-box genes (Fig. 6C). Genes including c77234 (C class) and c87227 were highly expressed in all four floral parts, while c60425, c91686 and 75076 were nearly no-expressed in all four floral parts. For stamen, c85161 (D class), c72042, c70349, and c80146 were of a highly expressions, meanwhile c70349 (B class), c80146 (B class), c87227, c58746 (A class) and c72216 (E class) were overexpressed in petal. Notably, c70164 and c69592 were specifically expressed in sepal and c57631 (D class) in carpel.

Figure 6. Analysis of 16 MADS-box genes A. Identified TFs gene families. Red points indicate MADS-box gene family. B. Phylogenetic tree map of 16 MADS-box genes. C. Expression profiles of 16 MADS-box genes.

**Closer examination of six key genes involved in controlling the development of flower shapes using stamen petalody types**

Out of the 16 MADS-box genes examined in this study, six of them were specifically selected for determining their specific roles in controlling stamen petalody based on two criteria. The criteria included genes with specifically expression in the specific flower parts and they fit for the ABCDE model. These six key genes were *AG*, *AP1-1*, *AP1-2*, *AP3*, *PI*, and *SEP*. Eight tree peony cultivars classified to stamen petalody and non-stamen petalody groups were used to verify the functions of the six candidate genes by real-time PCR (Fig. 7).

Figure. 7. The expression profiles of six key genes in stamen petalody and non-stamen petalody groups

The results showed that *SEP* had a high-level expression of sepals in both stamen petalody and non-stamen petalody cultivars. Meanwhile, *AG* showed similar expression profiles as *SEP*, but in carpels instead of sepals. *AP3* showed high level expressions in petals and much lower level expressions in stamens in the non-stamen petalody group. However, the expression profiles of *AP3* showed high level in both the own-petals and the inner-petals in the stamen petalody group. The expression profiles of *PI* are similar to those of *AP3*. *PI* showed low expression in stamen in the non-stamen petalody group and much higher expression in the stamen petalody cultivars. This indicates that *PI* likely has a similar function as *AP3* in regulating stamen petalody in tree peony. The expression profiles of *AP1-1* and *AP1-2* were inconsistent in both stamen petalody types and non-stamen petalody types.

## Discussions

Sixteen MADS-box genes have been successfully identified in tree peony through our systematic comparison of unigenes using a powerful and efficient RNA-seq method. Six exhibited unique expression in specific part(s) of tree peony flower and matched those of the ABCDE model as established in *Arabidopsis thaliana*. Our results have laid a solid foundation toward building a comprehensive model of flower parts formation in tree peony and possibly in other woody ornamental plants. Construction of a

flower development model could ultimately lead to the creation of more efficient tree peony cultivars with new floral shapes.

According to the ABCDE flower development model, A class genes determined the development of sepals; C and D class genes determined the development of carpels and stamens independently; B class genes together with A determined the development of petals; B class genes together with C determined the development of stamens, and E class genes interacted with all the other four class genes to function together. [9-13, 23]. The interactions of these five class transcription factors forming tetrameric complexes have led to the creation of diverse flower shapes in quite a few plant species. In *Matthiola incana*, the mutated *AG* alleles corresponded to the double-flower cultivars [24]. In *Magnolia stellata*, the expression changes in an *AG* orthologous gene resulted in double flowers [25]. In *Phalaenopsis orchid*, the tepal became a leaf-like organ when a *SEP*-like gene was silenced by virus-induced silencing. *SEP* played important roles in floral shape formation throughout the developmental process through the formation of various multiple protein complexes [26]. In *Prunus lannesiana*, overexpressed *AP1* transgenic plants showed carpelloid structures, and petal to stamen-like structures [27]. In our study, *AG* and *SEP* showed significantly high expression levels in carpels and sepals in both the stamen petalody and non-stamen petalody groups. This indicates that both *AG* and *SEP* may play important roles in carpels and sepals, respectively, in tree peony regardless of the presence or absence of stamen petalody. A highly significant finding in our study is that both *AP3* and *PI* showed elevated expression levels in stamen petalody cultivars compared to non-stamen cultivars. We hypothesize that the increased expression levels of *AP3* or *PI* in stamen may trigger the stamen petalody, which might be an important way to modify the flower shapes in tree peony. A previous study reported that *PsTM6* (*euAP3*) paralogs, cloned from 23 tree peony samples, displayed 24 amino acid substitutions, which might result in functional differentiation of *PsTM6* paralogs and affect stamen petalody and flower shape formation [28]. Our results indicate that *AP3* may affect stamen petalody but on the expression level. It is hard to say which of the two ways caused stamen petalody in tree peony before further functional studies. But it guides a right direction to explore the secret below the diversity flower shapes. The inconsistent expressions of *AP1-1* and *AP1-2* indicate that they may have multiple functions in regulating flower shapes that may not be tied to stamen petalody. In *Arabidopsis thaliana*, *AP1* was reported to regulate the development of sepals and petals with other MADS-box genes [9]. It is still quite possible that *AP1-1* and *AP1-2* may play important roles in regulating the formation of petals and sepals in tree peony. Additional research will be needed to characterize the specific functions of *AG*, *AP1-1*, *AP1-2*, *AP3*, *PI*, and *SEP* during flower formation for more targeted manipulation of flower shapes.

RNA-seq is a powerful and efficient approach for gene expression analysis at the genome level, especially in the plant species without genomic sequence information. In our study, a total of 11,801 unigenes were detected in all four floral parts. Among them, 14.3% of unigenes were found to be specifically expressed in certain individual floral parts. Among all of the flower parts examined, the unigenes that were specifically expressed in carpels were found in the greatest abundance, followed by stamen, calyx, and petal, respectively. These results suggest that the unigenes specifically expressed in certain individual floral parts might be potential candidates for the regulation of flower part formation in tree peony. Further

investigations of each set of these unigenes could lead to targeted manipulation of individual flower part development. A recent study showed that 28,199 differentially expressed genes were identified in *Paeonia rockii* by comparative transcriptome analysis. Among them, 15 candidate genes regulating the carpel quantitative variation were selected [29]. Another study showed that a set of 4876 DEGs were identified in *Asparagus officinalis*, including 43 DEGs in hormone response and biosynthesis associated with sex expression and reproduction [30]. A set of 3,7000 DEGs were identified in blueberry, among them 89% of flowering pathway genes, 86 % of MADS-box genes, and 84% of cold-regulated genes were up or down regulated in chilled flower buds compared to nonchilled flower buds [31]. A total of 1502 DEGs were detected in sterile flower buds [32]. The application of this method has yielded valuable data in tree peony. It is interesting to note that comparisons of calyx vs stamen and stamen vs carpel contained more DEGs than other comparisons. These facts indicated that the molecular process regulating the development in stamen and carpel involved more complex development events than in calyx and petal, which partially coincided with the specified unigenes results. K-Means clustering of gene expression levels of DEGs found nine gene clusters with distinctive expression patterns, of which seven of them showed up-regulated or down-regulated unigenes in stamen. It is too early to predict which sets of unigenes are more critical for the development of specific flower parts. Further investigations will reveal more information about the specific roles of these unigenes in regulating flower shape development. Our results indicated that the formation of every flower part is complex and may be regulated by hundreds of different unigenes. The large numbers of flower part-specific regulating unigenes may be linked to the evolution of higher plants due to the unique functions of flowers as the organ for pollination and seed production, and ultimately the survival of plant species. Flower colors, shapes, and sizes may play critical roles in attracting pollinators for seed production. Humans enjoy different flower shapes for aesthetic reasons. However, successful production of different flower shapes may involve the manipulation of several sets of unigenes.

## Conclusions

Sixteen MADS-box genes were identified in tree peony through the RNA-seq method. We identified six key genes based on their differential expression levels in different flower parts. These six key genes represented all categories in the ABCDE model to verify the functions in stamen petalody. We speculate that *PI* and *AP3* may play an important role in regulating stamen petalody in tree peony. *AG*, *AP1-1*, *AP1-2* and *SEP* did not show any important role in regulation stamen petalody. However, they may still be important in regulating the formation of other flower parts. Effective manipulation of *PI* and *AP3* may lead to the faster development of many exciting tree peony flower shapes. Our results provided a solid foundation for the successful establishment of the flower development model in tree peony based the ABCDE model in *Arabidopsis thaliana* and will pave the way toward the development of similar models in highly complex woody plants.

## Methods

## **Plant material and floral organs collection**

Simple form tree peony cultivar 'Fengdanbai' was selected for this study. Flower samples were collected at the early bloom stage around 10 a.m. April 6<sup>th</sup>, 2017 (Day length, 12 h 49 min 4 s; daily mean temperature, 17 °C) from 10-year old specimens planted at the experimental farm of Henan Agricultural University in Zhengzhou, Henan (34.785259°N, 113.663036°E). Three flowers were taken from three different tree peonies and were defined as biological replicates in this study. Freshly collected flowers were immediately separated into four parts: carpels, petals, sepals, and stamens. These flower parts were then frozen in liquid nitrogen and stored at -80°C for subsequent sequencing (Fig. 8). Floral parts were labeled as SE1 (sepal replication 1), SE2 (sepal replication 2), and SE3 (sepal replication 3); PE1 (petal replication 1), PE2 (petal replication 2), PE3 (petal replication 3); CA1 (carpel replication 1), CA2 (carpel replication 2), and CA3 (carpel replication 3); ST1 (stamen replication 1), ST2 (stamen replication 2), and ST3 (stamen replication 3) for subsequent RNA extraction and sequencing.

## **RNA extraction and cDNA library construction**

The total RNA was extracted from all 12 samples using an Easy-spin<sup>TM</sup> kit (Aidlab, Beijing, China) following the instructions of the manufacturer. RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was measured with NANOPhotometer Spectrophotometer (Implen, CA, USA). RNA concentration and RNA integrity were assessed separately using Qubit RNA Assay Kit in Qubit 2.0 Fluometer (Life technologies, CA, USA) and RNA Nano 6000 Assay Kit of the Agilent bioanalyzer 2100 system (Agilent technologies, CA, USA).

A total amount of 3 µg of purified RNA was used to construct the cDNA library using a NEBNext Ultra RNA library prep kit for Illumine (Nebraska, USA) following the manufacturer's instruction. The brief steps are as follows: 1) mRNA was purified from total RNA using poly-T oligo-attached magnetic beads; 2) fragmentation was carried out using divalent cations under elevated temperature in NEBNext first strand synthesis reaction buffer; 3) first strand cDNA was synthesized using random hexamer primer and M-MuLV reverse transcriptase; 4) second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H; 5) remaining overhangs were converted into blunt ends via polymerase activities and NEBnext adaptor with hairpin loop structure were ligated to prepare for hybridization; and 6) PCR was performed with fusion high-fidelity DNA polymerase, universal PCR primers and index primer. The library was checked by Qubit 2.0 and Agilent 2100 systems to examine the concentration and insert size. Sequencing was performed by Illumina HiSeq with pair-end reads of 150 bp.

## **Transcriptome assembly and annotation**

Trinity software was used to assemble the clean reads. Firstly, K-mers were obtained by breaking clean reads into short fragments [33]. Secondly, high frequency K-mers were extended from two directions with an overlap of k-1 bp until all K-mers in the library were used out. Thirdly, components were constructed by the linked K-mers (contigs) with an overlap of k-1 bp. Fourthly, De Bruijn graphs were simplified and reads were used to solve the graphs. Finally, transcripts were assembled and the longest transcripts were unigenes.

Unigenes were annotated in seven databases including Nr, COG, GO, KEGG, KOG, Pfam, Swissprot and eggNOG by BLAST with *E* value equal to or less than 1e-5 [34-36]. KOBAS2.0 was used to obtain KEGG Orthology results [37, 38]. Predicted amino acid sequences of unigenes were blasted against the Protein family (Pfam) database by HMMER with *E* value equal to or less than 1e-10 [39].

### **Differential expression and unigene expression level analysis**

The threshold values of false discovery rate and fold change were less than 0.01 and more than 2, respectively [40].

### **Unigene expression level analysis**

To calculate the expression of unigenes, RSEM was used to quantify the number of reads mapped to the assembled transcriptome and read count for each gene was obtained from the mapping results [41, 42]. We used the FPKM (fragments per kilobase of gene per million mapped reads) algorithm to normalize the gene expressional abundances in each library [43].

### **Real-time quantitative PCR**

Six specifically expressed MADS-box genes were selected to perform real-time quantitative PCR. Eight tree peony cultivars with different flower shapes (classified to stamen petalody and non- stamen petalody groups) were used as the test plant materials. Flower parts in non-stamen petalody groups were separated into carpels, sepals, petals, and stamens, while stamen petalody groups were separated into carpels, sepals, outside whorl petals, and inside-whorl petals (which is produced by stamen petalody).

A total of 2 µg RNA was extracted from each of these samples. Qualified RNA was tested using the APPLIED Biosystems 7500 Real-Time PCR system with a 20 µl reaction volume containing 10 µl of SYBR Premix Ex Taq II (Takara, China), 300 pM final concentration of each primer, and 1 µl template. The reaction conditions were as follows: 30s at 95 °C, 35 cycles of 5 s at 95 °C and 34s 60 °C. The primers

are listed in Additional table 1. Ubiquitin was selected as the reference gene [44]. Each sample was performed with three replicates and the relative expressions were calculated using the  $2^{-\Delta\Delta Ct}$  method.

## List Of Abbreviations

AG: Agamous; AP1: Apetala1; AP3: Apetala3; COG: Cluster of orthologous groups of proteins; CRY1: Cryptochrome 1 gene; DEG: Different expression unigene; eggNOG: Nonsupervised orthologous groups; FCA: Arabidopsis flowering time gene; FLC: Flower locus C; FLD: Flower locus D; FPA: Autonomous pathway independent flowering time gene; FT: Flower locus T gene; FPKM: Fragments per kilobase of gene per million mapped reads; FUL: Fruitful; FVE: A classical flowering time locus of the autonomous pathway; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; KOG: Eukaryotic ortholog groups; LHY: Late elongated hypocotyl; Nr: Non-redundant protein; PFam: Protein family; PHYA: Phytochrome-directed repression A gene; PHYB: Phytochrome-directed repression B gene; Pl: Pistilata; SHP: Shatterproof; SEP: Sepallata; SOC1: Suppressor of overexpression of constans 1; STK: Seedstick; SVP: Short vegetative phase; Swiss-Prot: The manually annotated and reviewed protein sequence database; TFs: Transcription factors; ZTL: Zeitlupe.

## Declarations

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This research was supported by Key Science Research Projects of Colleges and Universities of Henan Province (19A220003) and Science Creation Projects of Henan Agricultural University (KJCX2018A05).

### Authors' contributions

JL wrote the article; YZ and HW collected samples and performed experiments; ZW reviewed the article; YL analyzed the data; YHL and GG polished the article. All authors have read and approved the manuscript.

## Acknowledgments

We thank Biomarker Technologies Company (Beijing, China) for providing technical support.

## References

1. Formar F, de Montaigu A, Coupland G. Snapshot: control of flowering in arabidopsis. *Cell* 2010; 3:550.
2. Chen J, Källman T, Ma X, Gyllenstrand N, Zaina G, Morgante M *et al.* Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in norway spruce (*Picea abies*). *Genetics* 2012; 191:865-881.
3. Burgarella C, Chantret N, Gay L, Prosperi JM, Bonhomme M, Tiffin P *et al.* Adaptation to climate through flowering phenology: a case study in *Medicago truncatula*. *Mol Ecol* 2016; 25:3397-3415.
4. Jung WY, Lee A, Moon JS, Kim Y, Cho HS. Genome-wide identification of flowering time genes associated with vernalization and the regulatory flowering networks in chinese cabbage. *Plant Biotechnol Rep* 2018; 12:347-363.
5. Li YX, Li C, Bradbury PJ, Liu X, Lu F, Romay CM *et al.* Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. *The Plant Journal* 2016; 86:391-402.
6. Wang J, Qiu Y, Cheng F, Chen X, Zhang X, Wang H *et al.* Genome-wide identification, characterization, and evolutionary analysis of flowering genes in radish (*Raphanus sativus*). *Bmc Genomics* 2017; 18:981.
7. Peng Z, Lu Y, Li L, Zhao Q, Feng Q, Gao Z *et al.* The draft genome of the fast-growing non-timber forest species moso bamboo (*Phyllostachys heterocycla*). *Nat Genet* 2013; 45:456.
8. Meyerowitz EM, Bowman JL, Brockman LL, Drews GN, Jack T, Sieburth LE *et al.* A genetic and molecular model for flower development in *Arabidopsis thaliana*. *Development* 1991; 113:157-167.
9. Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF. *Apetala1* and *sepallata3* interact to promote flower development. *Plant Journal for Cell & Molecular Biology* 2001; 26:385-394.
10. Angenent GC, Franken J, Busscher M, Van D, Went JV. A novel class of mads box genes is involved in ovule development in petunia. *Plant Cell* 1995; 7:1569-1582.
11. Bowman JL, Smyth DR, Meyerowitz EM. Genetic interactions among floral homeotic genes of arabidopsis. *Development* 1991; 112:1-20.
12. Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. B and c floral organ identity functions require sepallata MADS-box genes. *Nature* 2000; 405:200-203.

13. Colombo L, Franken J, Koetje E, Went JV, Dons HJ, Angenent GC *et al.* The petunia mads box gene *fbp11* determines ovule identity. *Plant Cell* 1995; 7:1859-1868.
14. Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF. The *sep4* gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Curr Biol* 2004; 14:1935-1940.
15. Jofuku KD, den Boer BG, Van MM, Okamuro JK. Control of arabidopsis flower and seed development by the homeotic gene *apetala2*. *Plant Cell* 1994; 6:1211-1225.
16. Qin Q, Yin T, Chen J, Xie M, Zhang S. *Apetala3/deficiens* and *pistillata/globosa* genes with floral development of plant. *Chinese Journal of Cell Biology* 2006; 28:571-576.
17. Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM. The protein encoded by the arabidopsis homeotic gene *agamous* resembles transcription factors. *Nature* 1990; 346:35-39.
18. Honma T, Goto K. Complexes of mads-box proteins are sufficient to convert leaves into floral organs. *Nature* 2001; 409:525-529.
19. Cheng F, Li J, Chen D: Chinese flare mudan, 1 edn. Beijing: Forestry publishing house of China; 2005.
20. Li J, Zhang X, Zhao X: Tree peony of china. Beijing: Encyclopedia of China publishing house; 2011.
21. Yan Han X, Wang L, Shu Q, An Liu Z, Xia Xu S, Tetsumura T. Molecular characterization of tree peony germplasm using sequence-related amplified polymorphism markers. *Biochem Genet* 2008; 46:162-179.
22. Han XY, Wang LS, Shu QY, Liu ZA, Xu SX, Tetsumura T. Molecular characterization of tree peony germplasm using sequence-related amplified polymorphism markers. *Biochem Genet* 2008; 46:162-179.
23. Coen ES, Meyerowitz EM. The war of the whorls: genetic interactions controlling flower development. *Nature* 1991; 353:31-37.
24. Nakatsuka T, Koishi K. Molecular characterization of a double-flower mutation in *Matthiola incana*. *Plant Sci* 2018; 268:39-46.
25. Zhang B, Liu Z, Ma J, Song Y, Chen F. Alternative splicing of the *agamous* orthologous gene in double flower of *Magnolia stellata* (Magnoliaceae). *Plant Sci* 2015; 241:277-285.
26. Pan ZJ, Chen YY, Du JS, Chen YY, Chung MC, Tsai WC *et al.* Flower development of phalaenopsis orchid involves functionally divergent *sepallata*-like genes. *New Phytol* 2014; 202:1024-1042.
27. Chen M, Lin I, Yang C. Functional analysis of three lily (*Lilium longiflorum*) *apetala1*-like mads box genes in regulating floral transition and formation. *Plant Cell Physiol* 2008; 49:704-717.
28. Shu Q, Wang L, Wu J. Analysis of the formation of flower shapes in wild species and cultivars of tree peony using the MADS-box subfamily gene. *Gene* 2012; 493:113-123.
29. Liu N, Cheng F, Zhong Y, Guo X. Comparative transcriptome and co-expression network analysis of carpel quantitative variation in *Paeonia rockii*. *Bmc Genomics* 2019; 20:1-18.
30. Li S, Zhang G, Zhang X, Yuan J, Deng C, Gao W. Comparative transcriptome analysis reveals differentially expressed genes associated with sex expression in garden asparagus (*Asparagus officinalis*). *Bmc Plant Biol* 2017; 17:143.

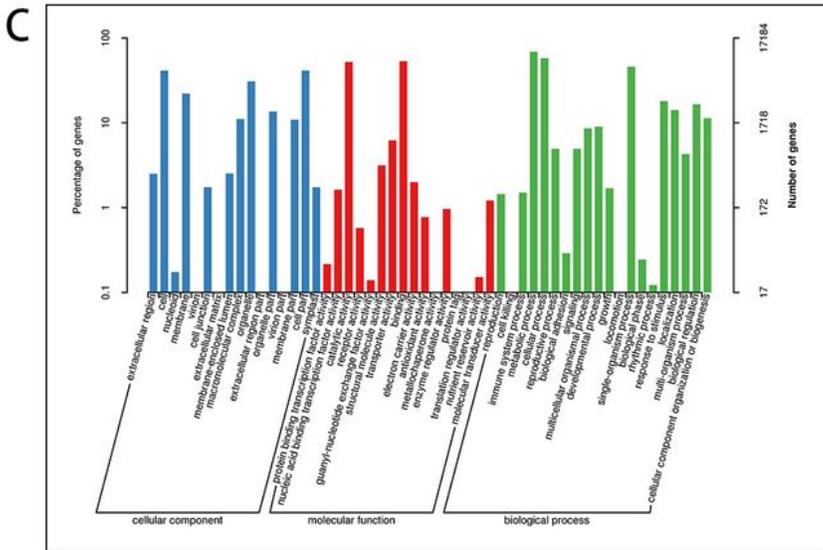
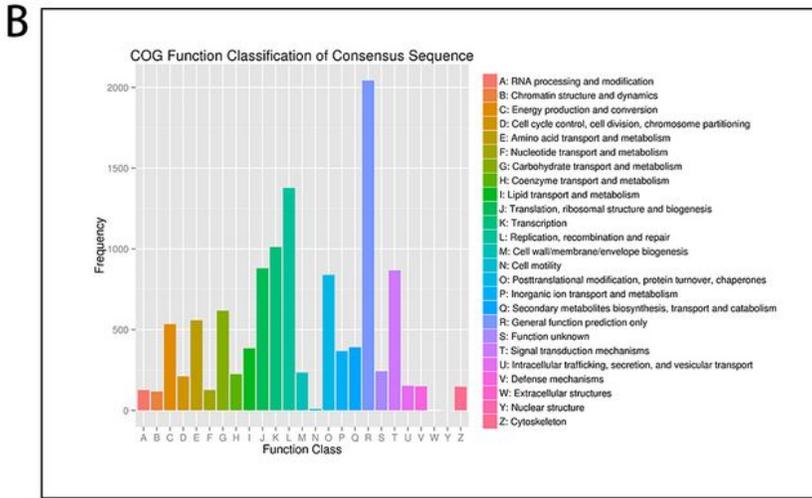
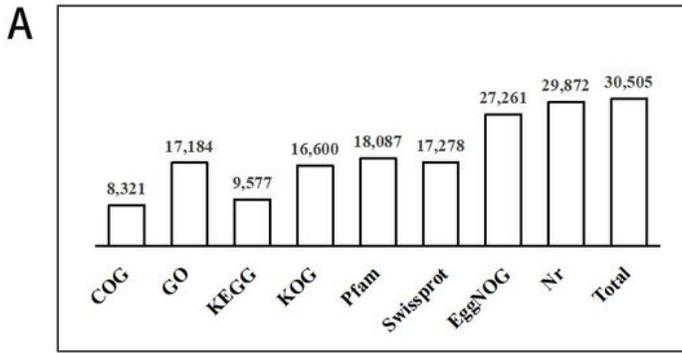
31. Song G, Chen Q. Comparative transcriptome analysis of nonchilled, chilled, and late-pink bud reveals flowering pathway genes involved in chilling-mediated flowering in blueberry. *Bmc Plant Biol* 2018; 18:98.
32. Liu H, Tan M, Yu H, Li L, Zhou F, Yang M *et al.* Comparative transcriptome profiling of the fertile and sterile flower buds of a dominant genic male sterile line in sesame (*Sesamum indicum*). *Bmc Plant Biol* 2016; 16:250.
33. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I *et al.* Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol* 2011; 29:644.
34. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W *et al.* Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Res* 1997; 25:3389-3402.
35. Apweiler R, Bairoch A, Wu C, Barker W, Boeckmann B, Ferro B *et al.* Uniprot: the universal protein knowledgebase. *Nucleic Acids Res* 2004; 32:115-119.
36. Koonin E, Fedorova N, Jackson J, Jacobs A, Krylov D, Makarova K *et al.* A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol* 2003; 5:60.
37. Kanehisa BM, Goto S, Kawashima S. *Et al.*: the KEGG resource for deciphering the genome. *Nucleic Acids Res* 2013; 32:D277-D280.
38. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S *et al.* KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 2011; 39:W316-W322.
39. Finn RD. Pfam: the protein families database. *Nucleic Acids Res* 2014; 42:D222-D230.
40. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010; 11:R106.
41. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short dna sequences to the human genome. *Genome Biol* 2009; 10:1-10.
42. Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BM *et al.* Ebseq: an empirical bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 2013; 29:1035-1043.
43. Trapnell C, Williams B, Pertea G, Mortazavi A, Kwan G, van Baren M *et al.* Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010; 28:511-515.
44. Liu CJ, Wang SL, Xue JQ, Zhu F, Ren X, Li M *et al.* Molecular cloning of ubiquitin protein gene and study on this gene as reference gene in tree peony. *Acta Horti Sin* 2015; 42:1983-1992.

## Figures



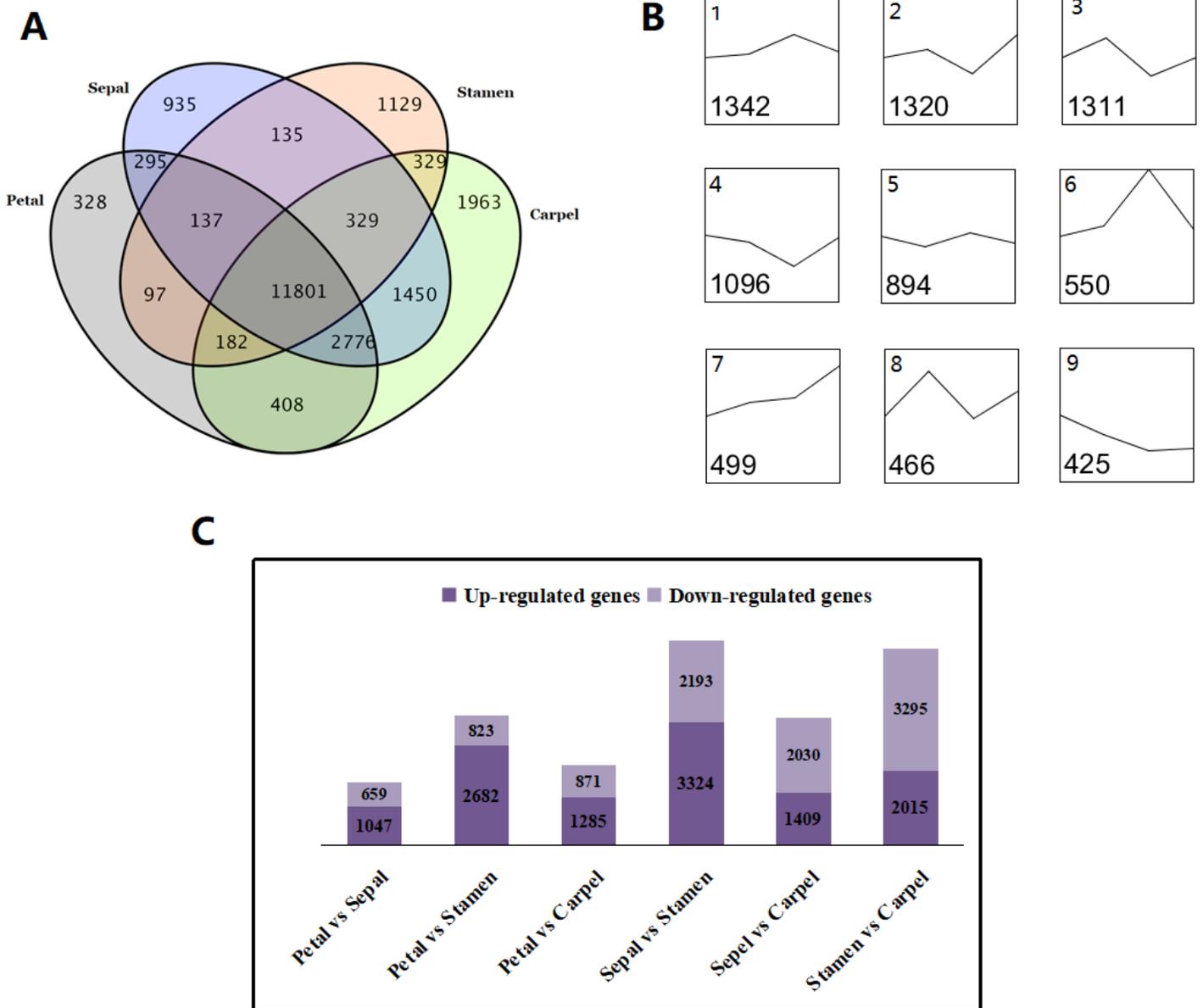
**Figure 1**

Flower shapes of the tree peony. A. Simple; B. Lotus; C. Chrysanthemum; D. Rose; E. Anemone; F. Thousand petal crown; G Hydrangea globular; H. Hundred proliferate



**Figure 2**

A. Classification map of unigenes with COG database search; B. Classification map of unigenes with GO database search.



**Figure 3**

A. Venn diagram of the unigenes of four flower parts in the tree peony; B. DEG numbers in six comparisons of four flower parts in the tree peony; C. K-means clustering map of the DEGs in four lower parts in the tree peony.

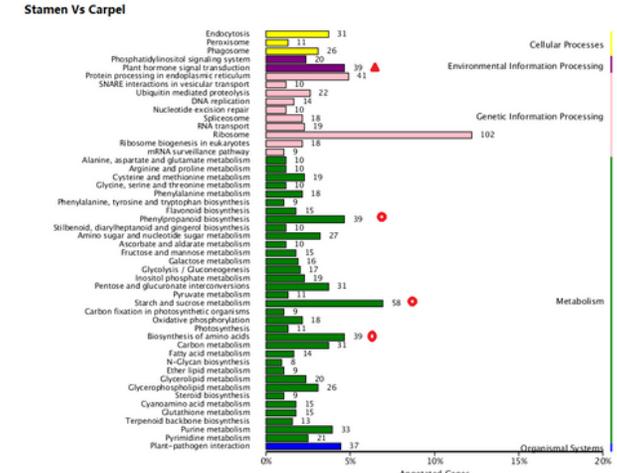
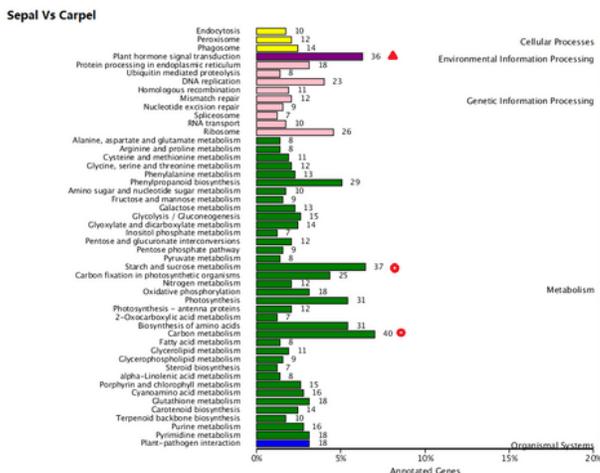
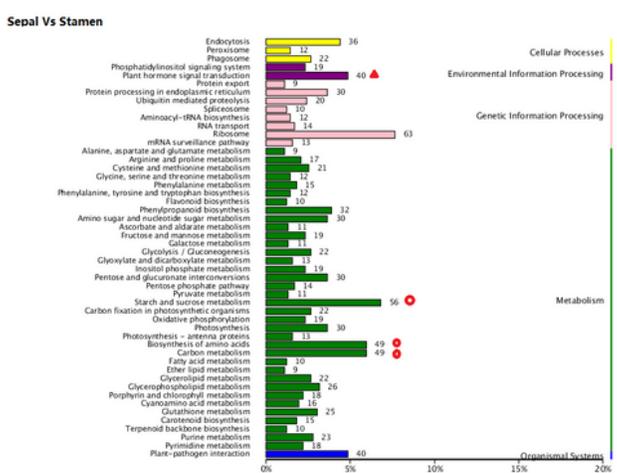
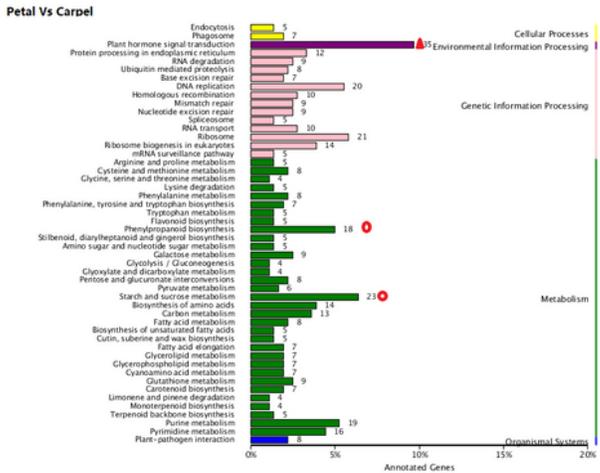
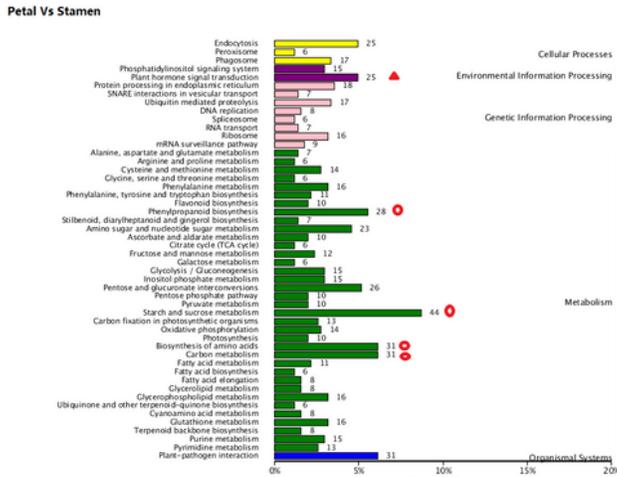
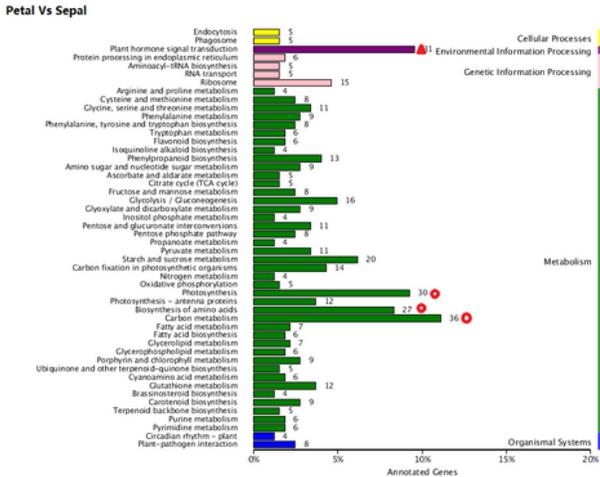


Figure 4

KEGG analysis of six comparison sets. Circles represent dominant pathways in "metabolism group"; triangles represent dominant pathways in the "environmental information processing" group.

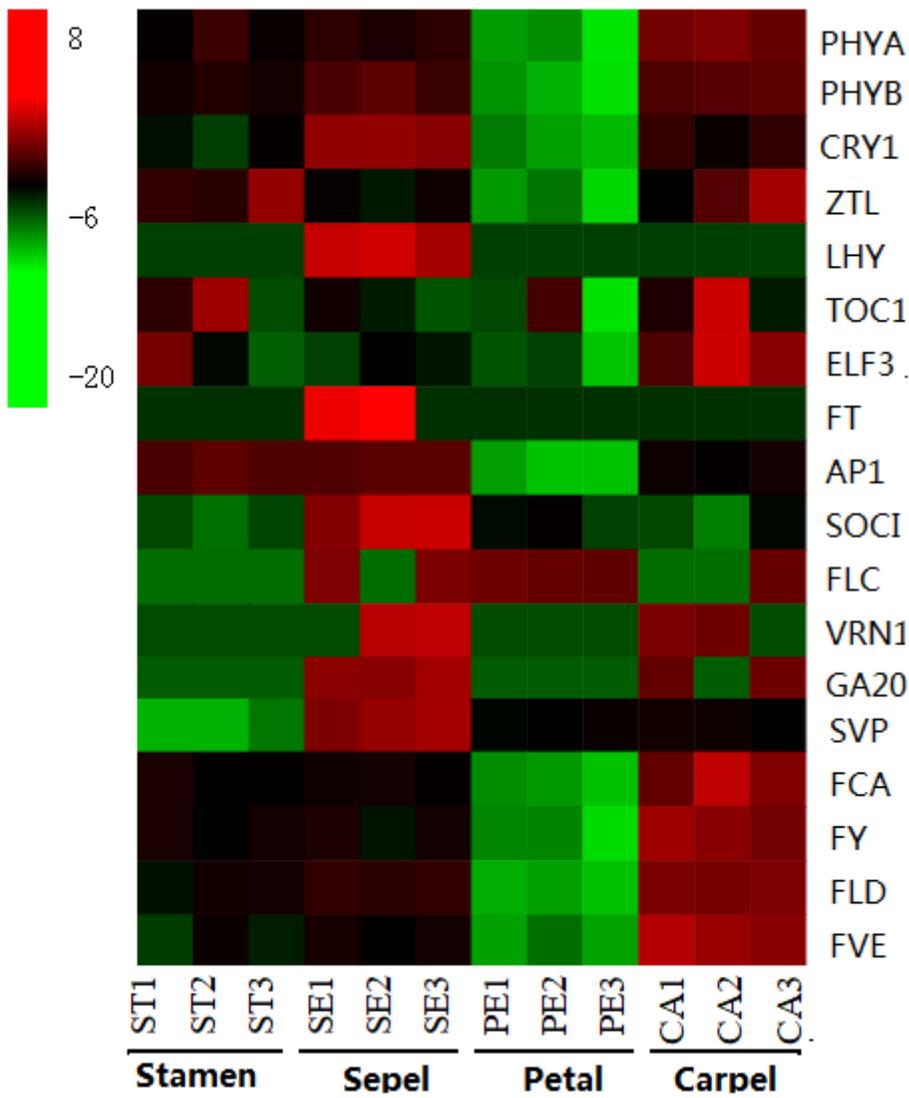
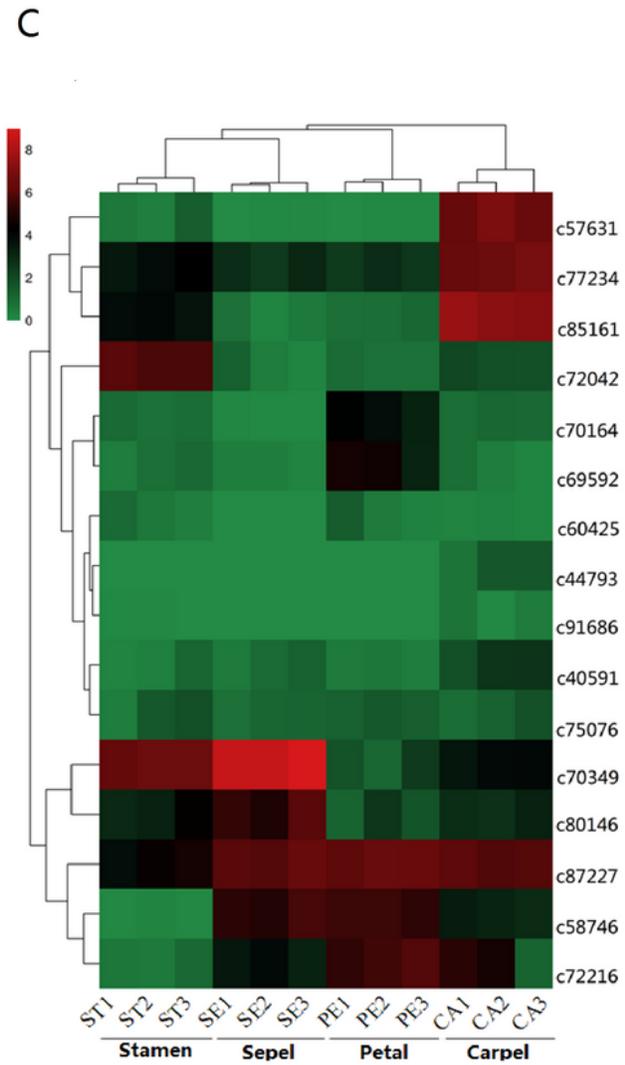
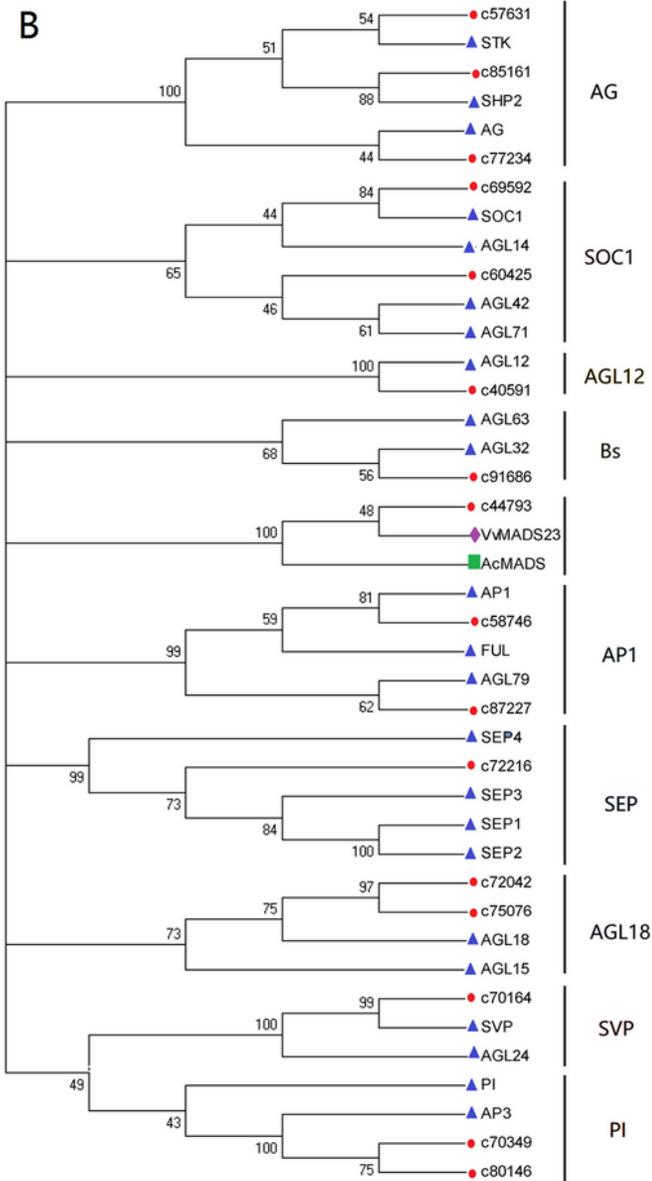
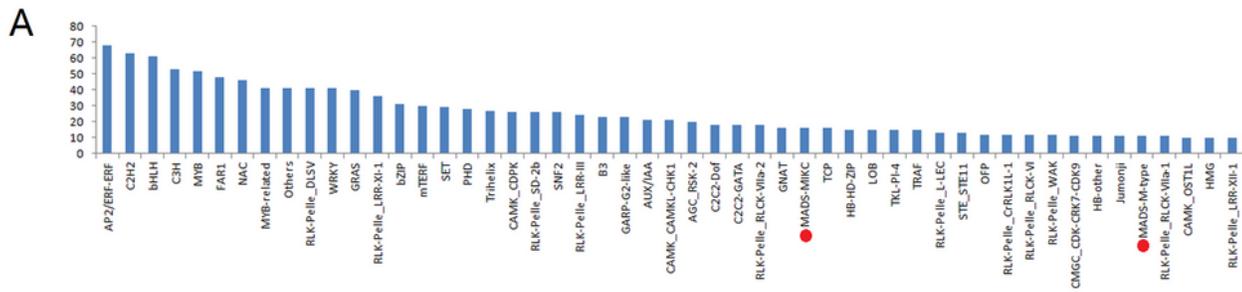


Figure 5

Heat map of key genes involving in flowering time pathways.

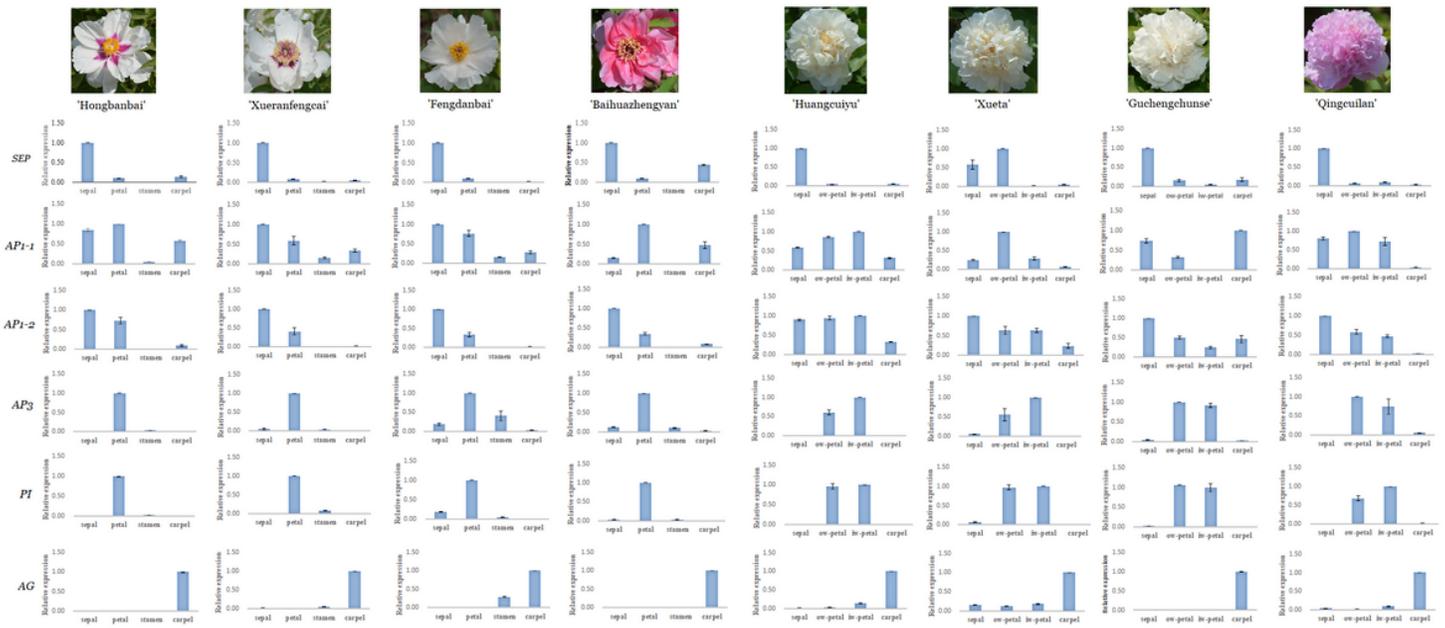


**Figure 6**

Analysis of 16 MADS-box genes A. Identified TFs gene families. Red points indicate MADS-box gene family. B. Phylogenetic tree map of 16 MADS-box genes. C. Expression profiles of 16 MADS-box genes.

*Non-stamen petalody*

*Stamen petalody*



**Figure 7**

The expression profiles of six key genes in stamen petalody and non-stamen petalody groups



**Figure 8**

Figure caption not provided with this version.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalData.docx](#)