forests

MDPI

*Article*

# Integrated SMRT Technology with UMI RNA-Seq Reveals the Hub Genes in Stamen Petalody in *Camellia oleifera*

**Huie Li [1], Yang Hu [2,3,4], Chao Gao [2,3,4,*], Qiqiang Guo [2,3,4], Quanen Deng [2,3,4], Hong Nan [1], Lan Yang [1], Hongli Wei [2,3,4], Jie Qiu [2,3,4] and Lu Yang [2,3,4]**

1 College of Agriculture, Guizhou University, Guiyang 550025, China; heli@gzu.edu.cn (H.L.); nanwyyy00912@163.com (H.N.); yl18886072834@163.com (L.Y.)
2 Institute for Forest Resources and Environment of Guizhou, Guizhou University, Guiyang 550025, China; xfy_gzhy@163.com (Y.H.); hnguoqiqiang@126.com (Q.G.); dc20170707@163.com (Q.D.); gzuweihongli@163.com (H.W.); QQQIUJ@163.com (J.Q.); yanglu202020@163.com (L.Y.)
3 Key Laboratory of Forest Cultivation in Plateau Mountain of Guizhou Province, Guizhou University, Guiyang 550025, China
4 College of Forestry, Guizhou University, Guiyang 550025, China
* Correspondence: gaochao@gzu.edu.cn

**Abstract:** Male sterility caused by stamen petalody is a key factor for a low fruit set rate and a low yield of *Camellia oleifera* but can serve as a useful genetic tool because it eliminates the need for artificial emasculation. However, its molecular regulation mechanism still remains unclear. In this study, transcriptome was sequenced and analyzed on two types of bud materials, stamen petalody mutants and normal materials, at six stages of stamen development based on integrated single-molecule real-time (SMRT) technology with unique molecular identifiers (UMI) and RNA-seq technology to identify the hub genes responsible for stamen petalody in *C. oleifera*. The results show that a large number of alternative splicing events were identified in the transcriptome. A co-expression network analysis of *MADSs* and all the differentially expressed genes between the mutant stamens and the normal materials showed that four *MADS* transcription factor genes, *CoSEP3.1*, *CoAGL6*, *CoSEP3.2*, and *CoAP3*, were predicted to be the hub genes responsible for stamen petalody. Among these four, the expression patterns of *CoAGL6* and *CoSEP3.2* were consistently high in the mutant samples, but relatively low in the normal samples at six stages, while the patterns of *CoSEP3.1* and *CoAP3* were initially low in mutants and then were upregulated during development but remained relatively high in the normal materials. Furthermore, the genes with high connectivity to the hub genes showed significantly different expression patterns between the mutant stamens and the normal materials at different stages. qRT-PCR results showed a similar expression pattern of the hub genes in the RNA-seq. These results lay a solid foundation for the directive breeding of *C. oleifera* varieties and provide references for the genetic breeding of ornamental *Camellia* varieties.

**Keywords:** petaloid stamen; male sterility; double flower; transcriptome; WGCNA



**Citation:** Li, H.; Hu, Y.; Gao, C.; Guo, Q.; Deng, Q.; Nan, H.; Yang, L.; Wei, H.; Qiu, J.; Yang, L. Integrated SMRT Technology with UMI RNA-Seq Reveals the Hub Genes in Stamen Petalody in *Camellia oleifera*. *Forests* **2021**, *12*, 749. https://doi.org/10.3390/f12060749

Academic Editors: Giovanni Emiliani and Alessio Giovannelli

Received: 8 May 2021
Accepted: 3 June 2021
Published: 6 June 2021



## 1. Introduction

*Camellia oleifera* (also known as oil-seed camellia), is an evergreen shrub or small evergreen tree belonging to the genus *Camellia* of the family Theaceae. It is distributed in the Yangtze River basin and further south of China. In southern China, *C. oleifera* is an important source of edible oil derived from woody plants. The oil of *C. oleifera* is high in unsaturated fatty acids and has shown medicinal effects, making it a high-quality edible oil. When the major cultivar of *C. oleifera*, "Huashuo", is compared with normal *C. oleifera* flowers, the stamen of the male sterile mutant shows remarkable petalody, leading to sterility. In normal *C. oleifera* flowers, the stamen shows no abnormal morphology, allowing for the normal dispersal of pollen. Male sterility caused by stamen petalody is a key factor for a low fruit set rate and a low yield of *C. oleifera* [1,2]. However, male sterility—a

relatively common reproductive trait in plants—can serve as a useful genetic tool because it eliminates the need for artificial emasculation. Thus, this condition affects research and field applications in economic forests due to the outbreeding enhancement. Petaloid stamens are an extremely important ornamental characteristic in plants of the same genus.

*MADS* genes (*MADSs*) have been reported to be widely involved in the regulation of flower organ development in plants [3–6]. The development in *Arabidopsis thaliana* can be interpreted with an ABCE model. In this model, the development of stamens is regulated by B, C, and E class genes, all of which belong to the *MADSs*; however, some of them have a short expression duration and a low abundance [7].

RNA sequencing (RNA-seq) technology has rapidly become a useful tool for the analysis of differential gene expressions among samples at the transcriptome level in the past decade. With the development of this technology, single-molecule real-time (SMRT) sequencing based on Pacific Biosciences System has made it possible to obtain a high-quality full-length transcriptome due to its advantage of ultra-long reads. At the same time, it can provide reference information for the analysis of the alternative splicing (AS) of genes among samples of non-model species without genomes [8–11]. Polymerase chain reaction (PCR)-induced errors in the next-generation sequencing can be corrected by introducing unique molecular identifiers (UMI) into the transcriptome library as a molecular barcode; this process results in a more accurate expression of data, especially for the quantitative expression of low abundance genes [12–15].

Petaloid stamens of *C. oleifera* have a crucial value for artificial hybrid breeding. However, their molecular regulation mechanism still remains unclear. In this study, SMRT technology combined with UMI RNA-seq technology was used to compare the differentially expressed genes (DEGs) between the stamen petalody mutant and the normal stamen at six stamen development stages in order to identify the hub genes responsible for stamen petalody in *C. oleifera*. The gene regulatory network was comprehensively analyzed.

## 2. Materials and Methods

### 2.1. Plant Materials

Developing flower buds of *C. oleifera* were sampled from ten-year-old trees with petaloid stamen mutants and normal stamens in the germplasm in Dongcheng Town, Wangcheng District, Changsha, Hunan Province (113°21′ E, 28°05′ N). This area has an average annual precipitation of 1380 mm, an average annual temperature of 19.3 °C, and an annual accumulated temperature of 5463 °C. In accordance with our previous studies on the development cycle of the bud samples [16], three biological replicates were collected separately at six stages in 2020 as follows: from 20 June to 28 June, stage 1 (S1) was the formation stage of the stamen primordium (when the stamen primordium cells elongate longitudinally and the primordium protrudes from the torus); from 29 June to 6 July, stage 2 (S2) was the inner stamen formation stage (the innermost 1–2 whorls of stamens are the first to form); from 7 July to 13 July, stage 3 (S3) was the formation stage of the outer stamens (the rest of the stamens are then formed); from 14 July to 21 July, stage 4 (S4) was the differentiation stage of anthers and filaments (the anthers and filaments are separated with clear boundaries); from 22 July to 2 August, stage 5 (S5) was the differentiation stage of the pollen sac (four pollen sacs begin to appear with distinct differences from the septum); from 3 August to 11 August, stage 6, (S6) was the formation stage of the stamen (formation of the pollen sac). The sepals and pedicels were quickly removed from all buds before further study.

### 2.2. Paraffin Section Microscopy

Paraffin section microscopy was performed in accordance with our previous study [17] with a modification. The buds were placed in Carnoy's solution for 10 h to fully eliminate the air from the material. They were then transferred to a 70% alcohol solution and stored in a refrigerator at 4 °C for later use. The fixed material was properly trimmed and then dyed with a 70% haematoxylin solution, washed with water, and dehydrated using alcohol of various

concentrations. The sample was made transparent with xylene, embedded in paraffin, and then sliced with a paraffin microtome (Leica RM2235, Heidelberg, Germany) with a thickness of 10 μm. A permanent cover sheet was made with Canada balsam, and photographs were acquired using an optical microscope (Leica DM2500, Heidelberg, Germany).

### 2.3. Scanning Electron Microscopy

The buds were fixed with a 2.5% glutaraldehyde fixative solution (prepared with 0.1 mol·$L^{-1}$ phosphate buffer) for 2 h, washed with a phosphate buffer (0.1 mol·$L^{-1}$), and then fixed with a 1% osmium fixative solution (prepared with 0.1 mol·-$L^{-1}$ phosphate buffer) for another 2 h. After washing with a phosphate buffer (0.1 mol·$L^{-1}$), the materials were dehydrated with 30%, 50%, 70%, 80%, 90%, 95%, and 100% ethanol, and then transitioned to 100% tert-butanol to be freeze-dried. The samples were placed on a sample table in an ion sputterer for 30 s and then photographed using a scanning electron microscope (HiTACHI TM4000, Tokyo, Japan).

### 2.4. RNA Isolation, Library Construction and Sequencing

The RNAs were isolated from a total of 36 bud samples (two types of materials at six stages with three biological replicates) by using a RNAprep Pure Plant Kit (Polysaccharides and Polyphenolics-rich) (Qiangen, Beijing, China) in accordance with the kit instructions. The integrity of all the RNA samples was checked through agarose gel electrophoresis and the Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). The integrity of all the samples was greater than 8.00, and these samples were used for further study. All RNA samples were mixed in equal mole amounts into an RNA pool. mRNAs were enriched by Oligo(dT) magnetic beads for a full-length cDNA library. These samples were prepared in accordance with the Isoform Sequencing protocol (Iso-Seq) using the Clontech SMARTer PCR cDNA Synthesis Kit and the BluePippin Size Selection System protocol as described by Pacific Biosciences (PN 100-092-800-03), and then sequenced using a SMRT approach based on the PacBio sequencing platform. The 36 libraries labeled with UMI were constructed. In contrast to the traditional sequencing library construction, the mRNAs were enriched by the rRNA deletion method, and each cDNA molecule was ligated with a unique barcode sequence adapter before amplification. The constructed libraries were sequenced using the pair-end approach based on the Illumina NovaSeq system. All sequencing processes were completed in Novogene Co., LTD. (Beijing, China).

### 2.5. Sequence Processing and Analysis

The PacBio official software package SMRTLink v10.1 was used to process the raw SMRT data. Subread sequences were first obtained, and circular consensus sequences were obtained by correction among subreads. The sequences were then divided into full-length sequences and non-full-length sequences in accordance with whether the sequences contained 5′ and 3′ end primers and a polyA tail. Full-length sequences were clustered using the hierarchical n*log (n) algorithm to obtain the cluster consensus sequence. These sequences were then polished to obtain high-quality consensus sequences, which were corrected by Illumina data, and redundancy was eliminated using software CD-HIT v4.8.1 [18] for subsequent analysis.

The processed SMRT sequences were used as a reference for the assembly of UMI Illumina reads. The clean reads of each sample were compared with the reference with the parameters of the comparison software Bowtie2 v2.4.3 [19] of the RSEM software package [20] with end-to-end and sensitive modes. Other parameters were set to default. The results were counted to obtain the read count value of each gene, and the fragments per kilobase of transcript sequence per million base pairs (FPKM) transformation was performed to obtain the gene expression level. DEGs were screened by comparing the genes expressed in two types of materials at the same developmental stage.

The gene function annotation was conducted on the sequences that were not redundant with CD-HIT software [18] versus databases including NR, NT, PFAM, KOG/COG,

Swiss-Prot, KEGG, and GO. The GO enrichment analysis of the DEGs was implemented using the GOseq R package (http://www.bioconductor.org/packages/release/bioc/html/goseq.html (accessed on 27 July 2020)). The DEGs were then screened by comparing the genes expressed in two types of materials at the same developmental stage by using the DEGSeq R package (http://www.bioconductor.org/packages/release/bioc/html/DEGseq.html (accessed on 29 July 2020)). The *p* values were adjusted using the Benjamini Hochberg method. A corrected *p*-value of 0.05 and a log2 (Fold change) of 1.5 were set as the thresholds for significantly differential expressions.

### 2.6. Alternative Splicing Analysis

SUPPA v2.3 (https://github.com/comprna/SUPPA (accessed on 1 August 2020)) was used to calculate the expression weight (psi) of AS on the basis of the TPM values. The differential AS of the two types was performed using the significance test of psi. The dpsi value was adjusted using the Mann–Whitney U test method. The absolute dpsi value of 0.1 and a *p* value of 0.05 were set as the thresholds for significantly differential alternative splices.

### 2.7. Identification of MADSs

The MADS transcription factors (TFs) were predicted on ITAK software [21]. The MADS TFs of the model plants *A. thaliana* were downloaded from TAIR (https://www.arabidopsis.org/ (accessed on 3 February 2021)). The phylogenetic tree of the MADSs from both *C. oleifera* and *A. thaliana* was inferred on the basis of protein sequences under maximum likelihood on MEGA v6 [22] and then visualized on software iTOL v5 [23].

### 2.8. Co-Expression Network Analysis

The expression data of all the DEGs and the *MADSs* were selected for weighted co-expression network analysis (WGCNA) on the R package WGCNA v 1.47 [24,25] with a merge cut height value of 0.25 and a minimum module size of 50. The Kendall model was used to analyze the correlation between the gene expression and the phenotype in order to reduce the sensitivity to outliers [26,27]. The module with the highest correlation with the phenotype was selected, and all the genes in the module were annotated. They were enriched and related to the flower organ development. The genes related were then extracted and visualized on Cytospace software, and the hub genes were screened in accordance with the degree of node connectivity [28,29].

### 2.9. Expression Analysis

The expression patterns of the hub gene and the relatively high connectives were analyzed on the basis of log2 (FPKM). A heatmap of the expression patterns was constructed on TBtools v1.092 software [30] for the mean expression in three biological replicates of each gene at each stage. The clustering of similar gene expression patterns was also completed by TBtools v1.092. Quantitative real-time PCR (qRT-PCR) reactions were conducted by using a Talent qPCR kit in accordance with the operating manual (TIAGEN, Beijing, China) with three technical replicates. The FPKM of the housekeeping gene *CoGAPDH* was relatively stable in this study. Thus, it was used as an internal control to normalize the relative expression of all the verified genes. The specific primers are listed in the Supplementary Materials section in Table S1.

## 3. Results

### 3.1. Observation of Stamen Development

At the first two stages, the difference between the mutant and the normal stamens was unremarkable (Figure 1A2–F2). At (**S3**), a slight difference was observed on the top of the stamens in the mutant, which showed subtle directional tendencies of a finger-shape (Figure 1A3–C3), whereas the normal stamens were only in a finger-shape (Figure 1D3–F3). At (**S4**), the top of the finger-shaped stamens in the mutant group showed petaloid direction

(Figure 1A4–C4), whereas the normal stamens began to form anthers (Figure 1D4–F4). At (**S5**), the upper pollen sacs in the mutant stamens differed significantly from those of the normal stamens (Figure 1A5–C5). At (**S6**), the petaloid stamens in the mutants spread downward (Figure 1A6–C6), whereas the apical part of the anther in the normal stamens retained meristematic tissue (Figure 1D6–F6). Our previous studies showed that the development of microspores occurs in a normal bud [31]. In the fully opened flower, the anthers of the mutant stamens did not form complete or normal pollen sac structures, resembling petals (Figure 1A7,B7). By contrast, the anthers of the normal stamens formed full pollen sacs (Figure 1C7,D7).

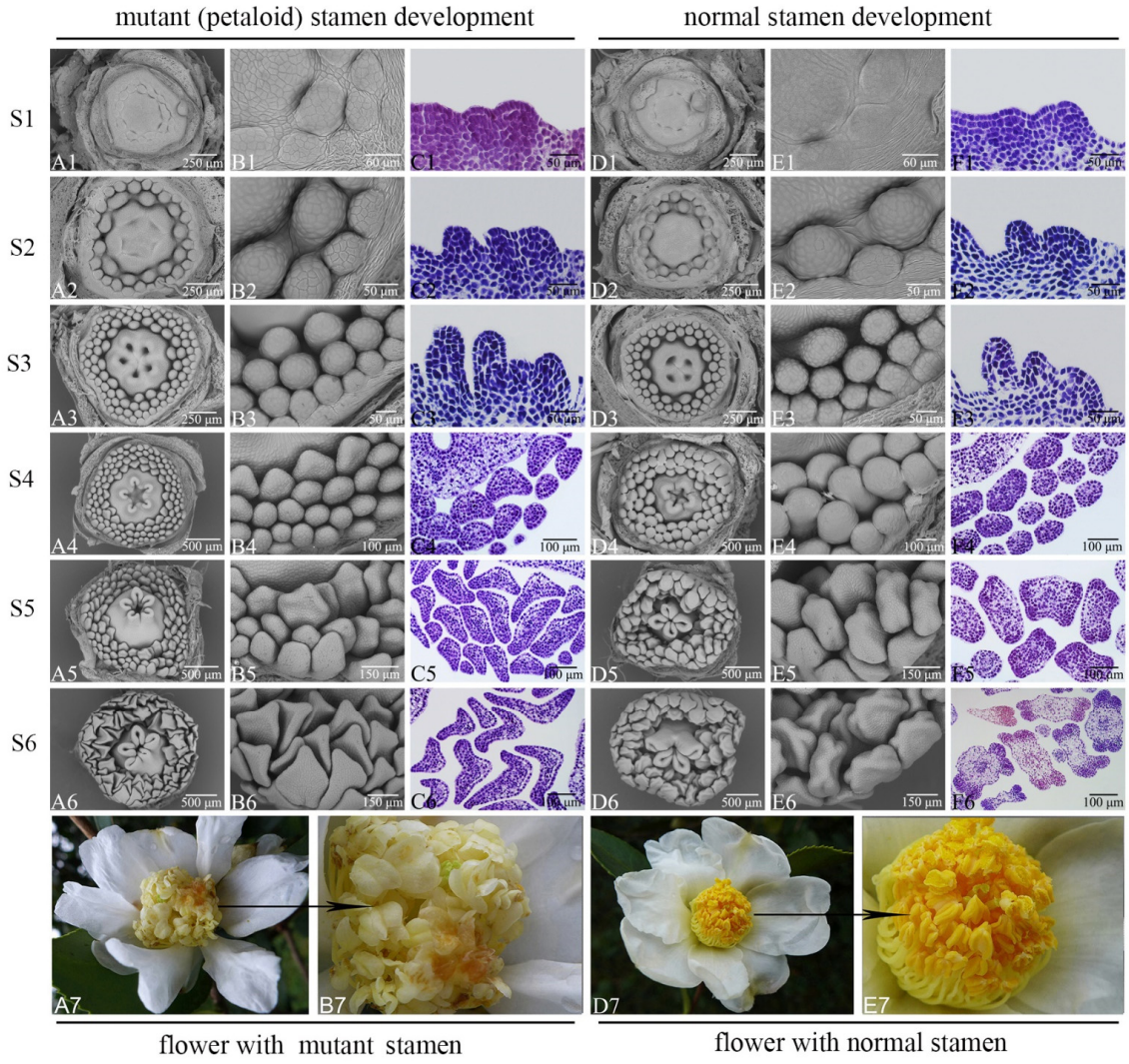


**Figure 1.** Development of the stamen at different stages in the mutant and the normal *C. oleifera*. Gray images show the scanning electron microscopy of the stamen development, blue images show the paraffin sections microscopy of stamen development, and color images show the stamens in the fully opened flowers. At (**S1**), the stamen primordia of the two types of materials were distinct cell clusters with one or two concentric circles of anther clusters centered on the torus (**A1**–**F1**); At (**S2**), the first two whorls of stamen primordia of the two types of materials are further developed, slightly protruding from the torus (**A2**–**F2**); At (**S3**), the top stamens with subtle directional tendencies in the mutant stamen (**A3**–**C3**), finger-shaped stamens in the normal stamens (**D3**–**F3**). At (**S4**), the innermost petaloid and abortion stamens with a triangle cross section at the apical region in the mutant stamens (**A4**–**C4**), and anthers with a slightly butterfly-shaped cross section begin to form at the apical region of normal stamens (**D4**–**F4**); At (**S5**), no pollen sacs forming in the outer region of the upper pollen sacs of stamen in mutant stamens (**A5**–**C5**), and anthers forming with a symmetrically butterfly-shaped cross-section in further developed normal stamens (**D5**–**F5**); At (**S6**), obvious petaloid stamens without pollen sacs in mutant stamens (**A6**–**C6**), and complete anthers form in normal stamen (**D6**–**F6**). Petaloid stamens without anther in the fully opened flower (**A7**,**B7**), and normal stamens containing anthers in fully open flowers (**C7**,**D7**).

### 3.2. Sequence Processing

Full-length transcriptome sequencing was performed on 36 bud materials with the same amount of mixed RNA samples to complete gene sequences. All raw data were deposited in the National Genomics Data Center, China, under BioProject acc. PRJCA004583. A total of 23,603,965 clean reads were generated from SMRT sequencing, resulting in 41,717 genes with an N50 of 3131 after processing, where 15,704 genes were more than 3000 bp long. The length distribution of subreads is shown in the Supplementary Materials section, Figure S1, and the length distribution of the unigenes is shown in the Supplementary Materials section, Figure S2. The annotation results are shown in the Supplementary Materials section, Figure S3, after annotation with different databases. The numbers of the reads generated from 36 samples by UMI sequencing and mapped reads against the reference sequence are listed in the Supplementary Materials section, Table S2. The DEGs at each stage are shown in the Supplementary Materials section, Figure S4.

### 3.3. Alternative Splicing

AS can be accurately identified and analyzed with the advantage of full-length transcriptome sequencing reading length. The results show that a large number of AS events are found in the expressed sequences (Figure 2). A total of 4088 genes containing two AS isoforms were found, accounting for 31% of the total. Three of the seven AS types were dominant, where 165 genes belonged to the retained intron type, accounting for 4.17% of the total AS genes, followed by the alternative-3′ splice-sites type with 317 genes, and the alternative-5′ splice-sites type with 135 genes. A total of 341 AS isoforms were differentially expressed between the mutant and the normal samples.

### 3.4. MADS TFs

MADS TFs are vital regulators of floral organ development, especially in *A. thaliana*, a well-studied model plant. In this study, a total of 42 *MADSs* were identified. Phylogenetic analysis showed that the *MADSs* expressed in the development of flower buds in *C. oleifera* were clustered with multiple *A. thaliana MADS* subfamilies including *SEP, AGL*, *AP3*, *FLC*, *AGL17*, *SVP*, *SOC1*, *AG*, *AP1*, and *AGL6* (Figure 3), indicating the diverse regulatory functions of the *MADSs* in *C. oleifera*.

### 3.5. Identification of Hub Genes

WGCNA could cluster genes into modules with similar expression patterns, analyze the correlation between the clustering modules and specific traits, and could further identify the hub genes in the network. The results show that the highest correlation coefficient between the module and the stamen trait was 0.74. The gene network of the module indicated that four genes show a high connectivity. These four genes were the hub genes in the stamen petaloid in *C. oleracea* (Figure 4). Further annotation analysis indicated that the four hub genes belonged to the MADS TF family, including two *SEP3*, a *AGL6*, and a *AP3*, which were designated as *CoSEP3.1*, *CoSEP3.2*, *CoAGL6*, and *CoAP3*. Genes with a relatively high connectivity to the hub genes were involved in flower organ development, including floral organ morphogenesis (*HUA2*, *SEU*, *TSO1*), hormone response (*ARF2, GID1C*), AS (*CLO*, *PAPS1*), and signal transduction (*MKK5*, *SERK*).

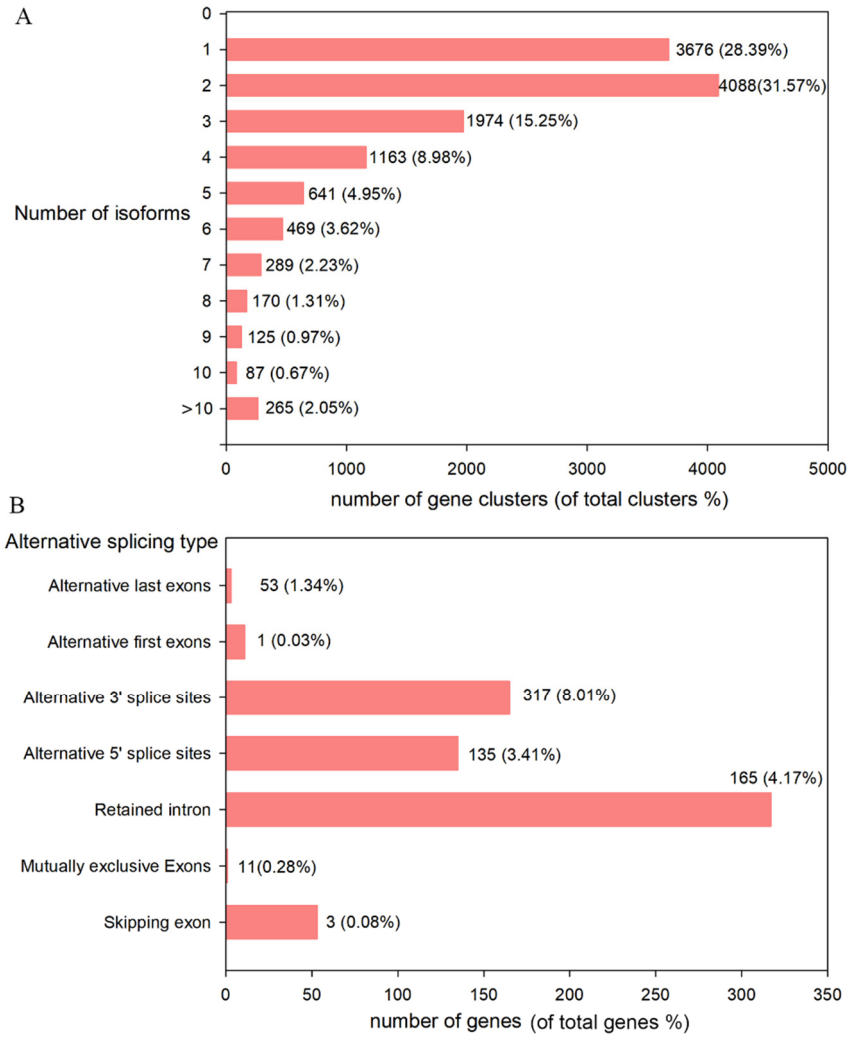


**Figure 2.** AS events in the transcriptome of mutant and normal *C. oleifera*. Numbers of AS events (**A**) and AS types (**B**) in the transcriptome.

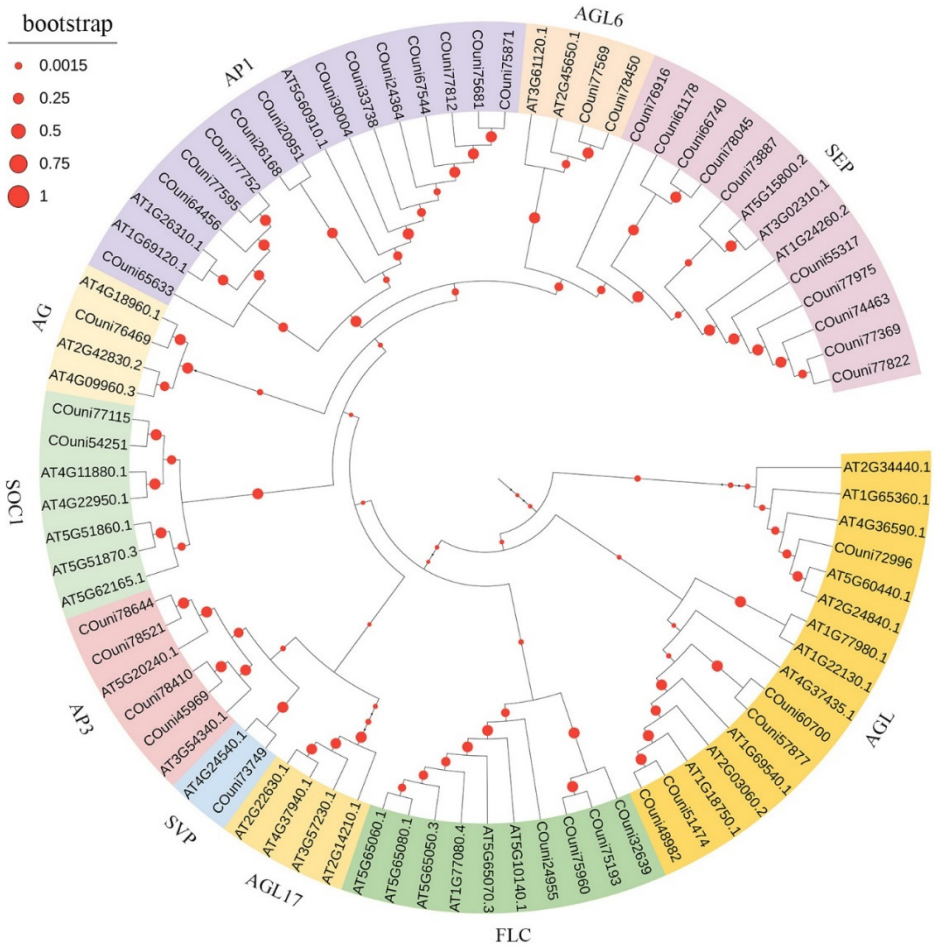


**Figure 3.** Phylogenetic relationship of the *MADSs* in the transcriptome of *C. oleifera*. Genes with the same background color belong to the same subfamily.

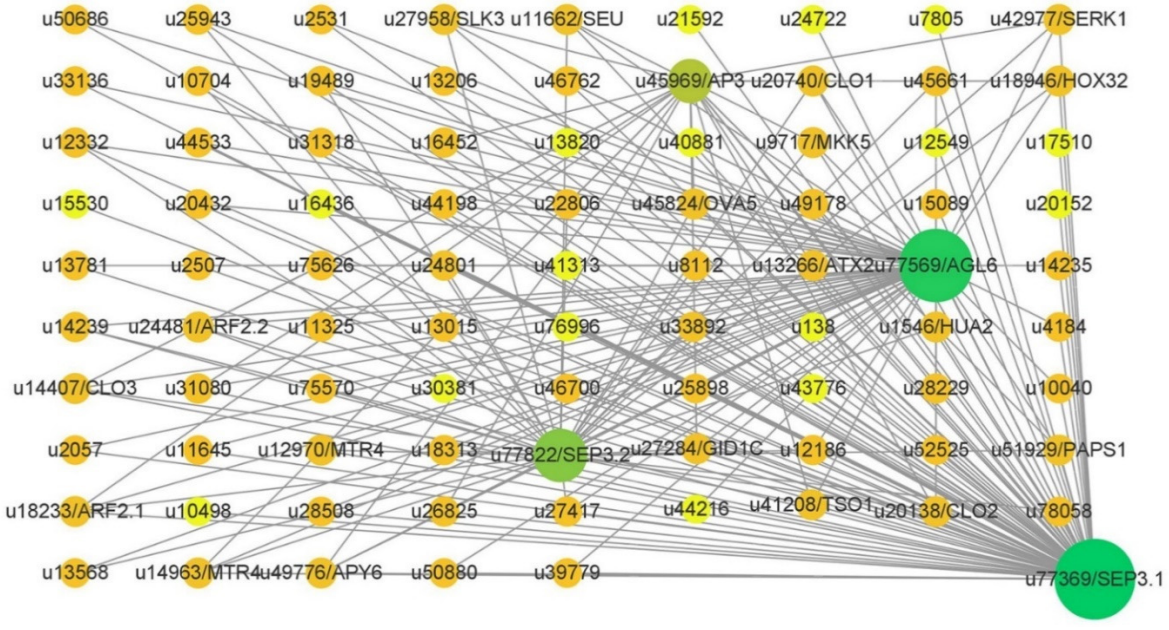


**Figure 4.** Co-expressed network of genes involved in stamen petalody of *C. oleifera*. Only the genes with a relatively high connectivity with the hub genes are shown, and the large green circles represent the hub genes with high connectivity to other genes.

### *3.6. Expression Analysis*

The expression patterns were obviously different in the mutant and normal samples of the hub genes and their connectivities were detected by RNA-seq (Figure 5A). These expression profiles were clustered into three main groups, and the four hub genes were all clustered in the third main group. The expressions of two hub genes, *CoSEP3.2* and *CoAGL6*, were consistently high in the mutant samples, but were relatively low in the normal samples, while expressions of the other two hub genes, *CoSEP3.1* and *CoAP3*, were initially low in the mutant buds and then upregulated with development, but remained relatively high in the normal buds. Furthermore, the genes with a high connectivity to the hub genes showed significantly different expression patterns at different stages between the mutant and the normal samples. qRT-PCR results showed a similar expression pattern of the hub genes in the RNA-seq (Figure 5B).

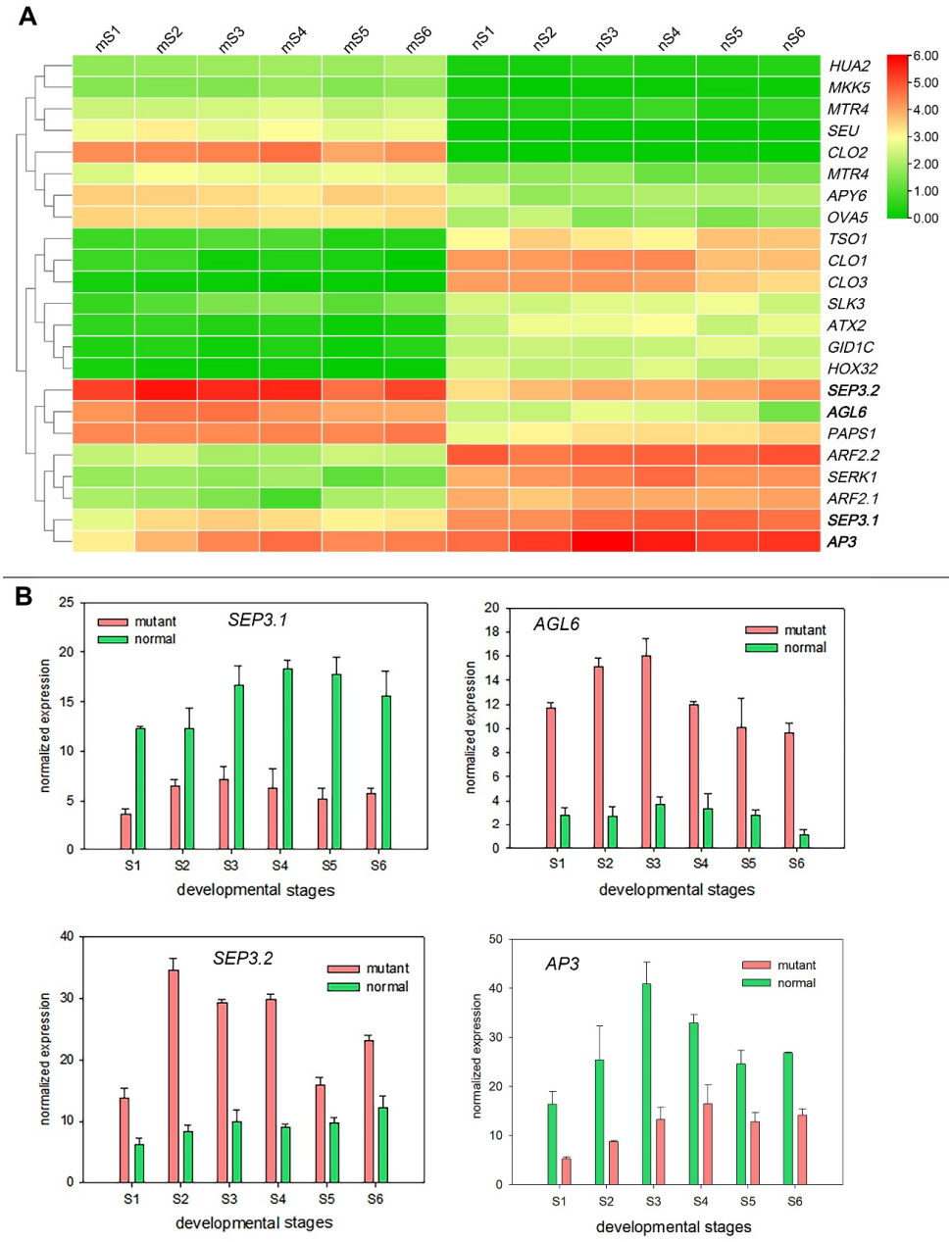


**Figure 5.** Expression patterns of the selected genes. (**A**) Expression patterns of the hub genes and their relatively high connectivities on the basis of log2 (FPKM) values which are means of three biological replicates. (**B**) Verification of the hub genes expression by qRT-PCR, values are means of three replicates, and bars represent SD.

## 4. Discussion

The stamen petalody phenomenon is common in plants, some of which are stable, and some of which change with the variation in environmental conditions [32,33]. The mutant of *C. oleifera* has a stable phenotype of stamen petrification according to our previous observation, enabling this study to be conducted.

AS is a common post-transcriptional process to enhance the proteome diversity in eukaryotes and is not exempted in the development of plant flower organs. Up to 60% of intron-containing genes in *A. thaliana* undergo AS. More than 1700 isoforms were differentially expressed between the inflorescent meristem stages and the flower development stages, suggesting that AS is involved in the transition during the floral development in this model plant [34]. In *Magnolia stellate*, one of the three AS isoforms of *AG* orthologous, expressed in developing stamens and carpels, can mimic the flower phenotype of the *ag* mutant and can produce double flowers with the homeotic transformation of stamens into petals in ectopic expressed *A. thaliana*, showing different functions from the two other isoforms [35]. In rice, *OsMADS3*, an AG ortholog, has two AS isoforms that differ in only one serine residue. Only the isoform, lacking serine residue, can specify stamens and carpels in *ag* mutant flowers [36]. Consistent with the previous reports, the numbers of AS isoforms were identified in the transcriptomes of all samples in this study by using the full-length transcriptome as a reference. These AS isoforms are probably involved in the development of certain flower organs.

Gene expression studies based on RNA-seq technology usually generate a large amount of expression data, and WGCNA enables the identification of the hub genes that have a high connectivity with other genes, which are usually the TFs that play a crucial role [24]. On the basis of RNA-Seq and WGCNA, four *MADSs* as hub genes were identified in the regulation of stamen petaloid in double flowers of *Lagerstroemia speciosa* [37]. Eighteen genes, including 7 *MADSs* and 11 other TF genes, were found to be involved in petaloid stamens in *Paeonia lactiflora* [38]. This finding indicated that *MADSs* are widely involved in the regulation of stamen petalody in non-model plant species. Similar to previous reports, in this study, four *MADSs*, including two *SEP*s, a *AGL6*, and a AP3, were identified to be involved in petaloid stamen determination in *C. oleracea*.

The molecular mechanism of flower organ development is best understood in the model plant *A. thaliana*. This mechanism was proposed as a ABCE model, where B-class MADSs specified either petals or stamens, and C-class MADSs specified stamens and carpels; they act in a combinatorial manner to determine the floral meristem identity [39–41]. The specification of the stamen is mainly regulated by the expression of a C-class gene *AGAMOUS* [42,43]. However, plants have intensely expanded the functional diversity of the MADSs during evolution by the formation of heterotetrameric complexes. SEPALLATA3 (SEP3), an E-class gene, has been regarded as a redundant mediator in flower organ specification [44], but has been later reported to function as a central protein–protein interaction hub, driving tetramerization with other MADSs, such as AG, PISTILLATA (PI), AP3, and even other SEP, to specify the determination of floral organs [45–48]. In petunia, the phenotype of *AG*-silenced plants has a similar flower phenotype to that of *SEP3*-silenced plants. Third-whorl stamens in the plants were transformed into petaloid organs, and the two genes were verified to interact [49]. In Zingiberales plants, *SEP-like* genes have undergone several duplication events giving rise to multiple copies, and *AGL6*, sister to the *SEP-like* genes, play a crucial role in the alteration of stamen into petaloid organs [50]. In line with previous studies, in this study, four *MADSs*, including two *SEP3*, a *AP3*, and a *AGL6*, were identified as hub genes in petaloid stamen determination in *C. oleracea.* Although the homologous of *AG* of *A. thaliana* was found in this study, it was not identified as the hub gene. This finding suggests that this determination may not be made by a single gene but by a combination of MADSs in *C. oleracea*. However, the expression patterns of the two *SEP3* genes were almost opposite in the two types of materials, which might be because their expressions needed to reach a certain balance and interact with other MADSs to determine the petaloid stamen in

mutant *C. oleracea* or other unknown reasons. Therefore, the aim of future research should be to verify the function of the hub genes through genetic transformation.

## 5. Conclusions

Transcriptome sequencing was performed and analyzed on two types of materials, stamen petalody mutants and normal materials, at six stamen development stages on the basis of integrated SMRT technology with UMI RNA-seq technology. This process was performed to identify the hub genes responsible for the stamen petalody of *C. oleifera*. The results showed that numbers of AS isoforms were identified in the transcriptomes. Four *MADSs* as hub genes responsible for stamen petalody were identified. Among them, *CoSEP3.1* was the most important, followed by *CoAGL6*, *CoSEP3.2*, and *CoAP3*. These results lay a solid foundation for the directive breeding of *C. oleifera* varieties and provide references for the genetic breeding of ornamental *Camellia* varieties.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/f12060749/s1, Figure S1: length distribution of subreads from SMRT sequencing; Figure S2: length distribution of unigenes. The X-axis represents length (bp), the Y-axis represents the number of genes; Figure S3: Venn diagram of unigene annotation. The sum of the numbers in each large circle represents the number of unigenes for the database annotation, and the overlapping parts of the circles represent unigenes annotation results that are common among databases; Figure S4: DEGs at each stage. DEGs were then screened by comparing the genes expressed in two types of materials at the same developmental stage. Corrected p-value of 0.05 and log2 (Fold change) of 1.5 were set as the threshold for significantly differential expression, values are means of three biological replicates; Table S1: specific primers used for qRT-PCR verification; Table S2: total reads of each sample and mapped reads against the reference sequence. Total reads represent clean data of sequencing reads after quality control; Total mapped represent the number of clean reads that can be matched to the reference sequence.

**Author Contributions:** H.L. and C.G. designed the experiments. Y.H., C.G., Q.D., H.W., J.Q. and L.Y. (Lu Yang). performed sampling and microscopy. H.L., C.G., Q.G., H.N. and L.Y. (Lan Yang) performed RNA isolations and RNA-seq data analysis. The authors who completed that part of the experiment prepared the draft of their part. All authors have read, revised, and agreed to the published version of the manuscript.

**Funding:** This research was supported by National Natural Science Foundation of China (31800516, 32060331), Science and Technology Planning Projects of Guizhou (Qian Ke He [2019] 2310, Qian Ke He [2020]1Y057, [2018]5781, Qian Ke He Fu Qi [2020] 4011), Science and Technology Project of Guizhou Education Department (Qian jiao He KY [2018]097, Qian jiao He KY [2019] 022), Cultivation Project of Guizhou University ([2019]35), and Research Project of Introducing Talents in Guizhou University ([2017]41), China.

**Data Availability Statement:** The datasets analyzed in the current study are available in Genome Sequence Archive (GSA) of National Genomics Data Center, China, under BioProject PRJCA004583 (https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA004583 (accessed on 1 March 2021)) with BioSamples accession numbers of SAMC343851-SAMC343887.

**Acknowledgments:** We would like to thank Lu Shikui for his help in sampling.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Xiong, H.; Chen, P.; Zhu, Z.; Chen, Y.; Zou, F.; Yuan, D. Morphological and cytological characterization of petaloid type cytoplasmic male sterility in *Camellia oleifera*. *HortScience* **2019**, *54*, 1149–1155. [CrossRef]
2. Tan, X.; Yuan, D.; Yuan, J.; Zou, F.; Xie, P.; Su, Y.; Yang, D.; Peng, J. An elite variety: *Camellia oleifera* 'Huashuo'. *Sci. Silvae Sinicae.* **2011**, *47*, 184, (in Chinese with English Abstract).
3. Wang, D.; Hao, Z.; Long, X.; Wang, Z.; Zheng, X.; Ye, D.; Peng, Y.; Wu, W.; Hu, X.; Wang, G.; et al. The Transcriptome of *Cunninghamia lanceolata* male/female cone reveal the association between MIKC MADS-box genes and reproductive organs development. *BMC Plant Biol.* **2020**, *20*. [CrossRef]

4. Wang, P.; Liao, H.; Zhang, W.; Yu, X.; Zhang, R.; Shan, H.; Duan, X.; Yao, X.; Kong, H. Flexibility in the structure of spiral flowers and its underlying mechanisms. *Nat. Plants* **2016**, *2*. [CrossRef]
5. Hugouvieux, V.; Zubieta, C. MADS transcription factors cooperate: Complexities of complex formation. *J. Exp. Bot.* **2018**, *69*, 1821–1823. [CrossRef] [PubMed]
6. Lai, X.; Daher, H.; Galien, A.; Hugouvieux, V.; Zubieta, C. Structural basis for plant MADS transcription factor oligomerization. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 946–953. [CrossRef] [PubMed]
7. Theissen, G.; Melzer, R.; Ruempler, F. MADS-domain transcription factors and the floral quartet model of flower development: Linking plant development and evolution. *Development* **2016**, *143*, 3259–3271. [CrossRef] [PubMed]
8. Jiang, X.; Hall, A.B.; Biedler, J.K.; Tu, Z. Single molecule RNA sequencing uncovers trans-splicing and improves annotations in Anopheles stephensi. *Insect Mol. Biol.* **2017**, *26*, 298–307. [CrossRef]
9. Li, Y.; Dai, C.; Hu, C.; Liu, Z.; Kang, C. Global identification of alternative splicing via comparative analysis of SMRT- and Illumina-based RNA-seq in strawberry. *Plant J.* **2017**, *90*, 164–176. [CrossRef]
10. Teng, K.; Teng, W.; Wen, H.; Yue, Y.; Guo, W.; Wu, J.; Fan, X. PacBio single-molecule long-read sequencing shed new light on the complexity of the Carex breviculmis transcriptome. *BMC Genom.* **2019**, *20*. [CrossRef]
11. Wu, Y.; Xu, J.; Han, X.; Qiao, G.; Yang, K.; Wen, Z.; Wen, X. Comparative Transcriptome Analysis Combining SMRT- and Illumina-Based RNA-Seq Identifies Potential Candidate Genes Involved in Betalain Biosynthesis in Pitaya Fruit. *Int. J. Mol. Sci.* **2020**, *21*, 3288. [CrossRef] [PubMed]
12. Chen, W.; Li, Y.; Easton, J.; Finkelstein, D.; Wu, G.; Chen, X. UMI-count modeling and differential expression analysis for single-cell RNA sequencing. *Genome Biol.* **2018**, *19*. [CrossRef] [PubMed]
13. Karst, S.M.; Ziels, R.M.; Kirkegaard, R.H.; Sorensen, E.A.; McDonald, D.; Zhu, Q.; Knight, R.; Albertsen, M. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nat. Methods* **2021**, *18*, 165–169. [CrossRef]
14. Posfai, D.; Krishnan, K.; Song, C.; Liu, P.; Naishadham, G.; Langhorst, B.W.; Dimalanta, E.T.; Davis, T.B. Improving transcriptome analysis by incorporating unique molecular identifiers into RNA-sequencing. *Eur. J. Hum. Genet.* **2020**, *28*, 611–612.
15. Smith, T.; Heger, A.; Sudbery, I. UMI-tools: Modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **2017**, *27*, 491–499. [CrossRef]
16. Liao, T.; Yuan, D.-Y.; Zou, F.; Gao, C.; Yang, Y.; Zhang, L.; Tan, X.-F. Self-Sterility in *Camellia oleifera* may be due to the prezygotic late-acting self-incompatibility. *PLoS ONE* **2014**, *9*. [CrossRef]
17. Gao, C.; Yuan, D.; Yang, Y.; Wang, B.; Liu, D.; Zou, F. Pollen Tube Growth and Double Fertilization in *Camellia oleifera*. *J. Am. Soc. Hort. Sci.* **2015**, *140*, 12–18. [CrossRef]
18. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152. [CrossRef]
19. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]
20. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* **2011**, *12*. [CrossRef] [PubMed]
21. Zheng, Y.; Jiao, C.; Sun, H.; Rosli, H.G.; Pombo, M.A.; Zhang, P.; Banf, M.; Dai, X.; Martin, G.B.; Giovannoni, J.J.; et al. iTAK: A program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant.* **2016**, *9*, 1667–1670. [CrossRef] [PubMed]
22. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]
23. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**. [CrossRef] [PubMed]
24. Zhang, B.; Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **2005**, *4*. [CrossRef] [PubMed]
25. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf.* **2008**, *9*. [CrossRef]
26. Moreno-Moral, A.; Mancini, M.; D'Amati, G.; Camici, P.; Petretto, E. Transcriptional network analysis for the regulation of left ventricular hypertrophy and microvascular remodeling. *J. Cardiovasc. Trans. Res.* **2013**, *6*, 931–944. [CrossRef]
27. Hall, T.; Steiner, R.; Wright, H.; Wilmot, B.; Roullet, J.; Peters, M.; Harris, M. Lipid and sterol gene sequence variation in autism and correlates with neurodevelopmental status: A pilot study. *Arch. Clin. Neuropsychol.* **2015**, *2*, 137–146. [CrossRef]
28. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef]
29. Otasek, D.; Morris, J.H.; Boucas, J.; Pico, A.R.; Demchak, B. Cytoscape automation: Empowering workflow-based network analysis. *Genome Biol.* **2019**, *20*. [CrossRef]
30. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* **2020**, *13*, 1194–1202. [CrossRef]
31. Gao, C.; Yuan, D.Y.; Wang, B.F.; Yang, Y.; Liu, D.M.; Han, Z.Q. A cytological study of anther and pollen development in *Camellia oleifera*. *Gen. Mol. Res.* **2015**, *14*, 8755–8765. [CrossRef]
32. Mizunoe, Y.; Ozaki, Y. Effects of growth temperature and culture season on morphogenesis of petaloid-stamen in double-flowered *Cyclamen*. *Horticult. J.* **2015**, *84*, 269–276. [CrossRef]

33. Han, Y.; Tang, A.; Wan, H.; Zhang, T.; Cheng, T.; Wang, J.; Yang, W.; Pan, H.; Zhang, Q. An APETALA2 homolog, RcAP2, regulates the number of rose petals derived from stamens and response to temperature fluctuations. *Front. Plant Sci.* **2018**, *9*. [CrossRef]
34. Wang, H.; You, C.; Chang, F.; Wang, Y.; Wang, L.; Qi, J.; Ma, H. Alternative splicing during *Arabidopsis* flower development results in constitutive and stage-regulated isoforms. *Front. Genet.* **2014**, *5*. [CrossRef] [PubMed]
35. Zhang, B.; Liu, Z.X.; Ma, J.; Song, Y.; Chen, F.J. Alternative splicing of the *AGAMOUS* orthologous gene in double flower of *Magnolia stellata* (Magnoliaceae). *Plant. Sci.* **2015**, *241*, 277–285. [CrossRef]
36. Dreni, L.; Ravasio, A.; Gonzalez-Schain, N.; Jacchia, S.; da Silva, G.J.; Ricagno, S.; Russo, R.; Caselli, F.; Gregis, V.; Kater, M.M. Functionally divergent splicing variants of the rice AGAMOUS ortholog *OsMADS3* are evolutionary conserved in grasses. *Front. Plant. Sci.* **2020**, *11*. [CrossRef]
37. Hu, L.; Zheng, T.; Cai, M.; Pan, H.; Wang, J.; Zhang, Q. Transcriptome analysis during floral organ development provides insights into stamen petaloidy in *Lagerstroemia speciosa*. *Plant Physiol. Biochem.* **2019**, *142*, 510–518. [CrossRef]
38. Fan, Y.; Zheng, Y.; Teixeira da Silva, J.A.; Yu, X. Comparative transcriptomics and WGCNA reveal candidate genes involved in petaloid stamens in *Paeonia lactiflora*. *J. Horticult. Sci. Biotechnol.* **2021**. [CrossRef]
39. Chen, D.; Yan, W.; Fu, L.-Y.; Kaufmann, K. Architecture of gene regulatory networks controlling flower development in *Arabidopsis thaliana*. *Nat. Commun.* **2018**, *9*. [CrossRef]
40. O'Maoileidigh, D.S.; Graciet, E.; Wellmer, F. Gene networks controlling *Arabidopsis thaliana* flower development. *N. Phytol.* **2014**, *201*, 16–30. [CrossRef] [PubMed]
41. Wils, C.R.; Kaufmann, K. Gene-regulatory networks controlling inflorescence and flower development in *Arabidopsis thaliana*. *Biochim. Biophys. Acta Gene Regul. Mech.* **2017**, *1860*, 95–105. [CrossRef]
42. Ito, T.; Ng, K.-H.; Lim, T.-S.; Yu, H.; Meyerowitz, E.M. The homeotic protein AGAMOUS controls late stamen development by regulating a jasmonate biosynthetic gene in *Arabidopsis*. *Plant Cell* **2007**, *19*, 3516–3529. [CrossRef]
43. Jack, T. New members of the floral organ identity AGAMOUS pathway. *Trends Plant Sci.* **2002**, *7*, 286–287. [CrossRef]
44. Pelaz, S.; Ditta, G.S.; Baumann, E.; Wisman, E.; Yanofsky, M.F. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* **2000**, *405*, 200–203. [CrossRef]
45. Hugouvieux, V.; Silva, C.S.; Jourdain, A.; Stigliani, A.; Charras, Q.; Conn, V.; Conn, S.J.; Carles, C.C.; Parcy, F.; Zubieta, C. Tetramerization of MADS family transcription factors SEPALLATA3 and AGAMOUS is required for floral meristem determinacy in *Arabidopsis*. *Nucleic Acids Res.* **2018**, *46*, 4966–4977. [CrossRef]
46. Melzer, R.; Theissen, G. Reconstitution of floral quartets in vitro involving class B and class E floral homeotic proteins. *Nucleic Acids Res.* **2009**, *37*, 2723–2736. [CrossRef] [PubMed]
47. Smaczniak, C.; Immink, R.G.H.; Muino, J.M.; Blanvillain, R.; Busscher, M.; Busscher-Lange, J.; Dinh, Q.D.; Liu, S.; Westphal, A.H.; Boeren, S.; et al. Characterization of MADS-domain transcription factor complexes in *Arabidopsis* flower development. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1560–1565. [CrossRef] [PubMed]
48. Soza, V.L.; Snelson, C.D.; Hazelton, K.D.H.; Di Stilio, V.S. Partial redundancy and functional specialization of E-class SEPALLATA genes in an early-diverging eudicot. *Dev. Biol.* **2016**, *419*, 143–155. [CrossRef] [PubMed]
49. Kapoor, M.; Tsuda, S.; Tanaka, Y.; Mayama, T.; Okuyama, Y.; Tsuchimoto, S.; Takatsuji, H. Role of petunia pMADS3 in determination of floral organ and meristem identity, as revealed by its loss of function. *Plant J.* **2002**, *32*, 115–127. [CrossRef]
50. Yockteng, R.; Almeida, A.M.R.; Morioka, K.; Alvarez-Buylla, E.R.; Specht, C.D. Molecular evolution and patterns of duplication in the SEP/AGL6-like lineage of the Zingiberales: A proposed mechanism for floral diversification. *Mol. Biol. Evol.* **2013**, *30*, 2401–2422. [CrossRef]