horticulturae

MDPI

Article

Analysis of Transcriptome and Expression of *C4H* and *FLS* Genes on Four Flower Colors of *Impatiens uliginosa*

Xiaoli Zhang †, Yi Tan †, Xinyi Li, Zengdong Liu, Fan Li, Haiquan Huang * and Meijuan Huang *

College of Landscape Architecture and Horticulture Sciences, Southwest Research Center for Engineering Technology of Landscape Architecture (State Forestry and Grassland Administration), Yunnan Engineering Research Center for Functional Flower Resources and Industrialization, Research and Development Center of Landscape Plants and Horticulture Flowers, Southwest Forestry University, Kunming 650224, China; m18468178186@163.com (X.Z.); tanyi@swfu.edu.com (Y.T.); xinyili@swfu.edu.cn (X.L.); 18287762119@163.com (Z.L.); lifan@swfu.edu.cn (F.L.)

* Correspondence: haiquan_huang@swfu.edu.cn (H.H.); meijuanhuang@swfu.edu.cn (M.H.)

† These authors contributed equally to this work.

Abstract: Flower color is a major feature of ornamental plants, and the rich flower color of plants is an important factor in determining their ornamental and economic values, so flower color is an important research target for gardening and horticulture breeders at home and abroad. Our research group collected four colors of *Impatiens uliginosa* (white, pink, red, and deep red) during the collection of germplasm resources in the field. In this study, we analyzed the transcriptomes of the four flower colors of *I. uliginosa* by using RNA-Seq technology. The transcriptomes were screened to identify candidate genes related to flower color, and the coloring mechanisms of four flower colors were revealed at the molecular level. The main findings were as follows: (1) The number of the four different transcripts ranged from 64,723 to 93,522 and contained a total of 100,705 unigenes. (2) The analysis of differentially expressed genes revealed structural genes including *C4H*, *FLS*, *PAL*, and *ANS* and transcription factors including *MYB*, *MYB*-related, *AP2-EREBP*, and *bHLH*. (3) Among the four flower colors of *I. uliginosa*, the *C4H1* gene had the highest expression in pink flowers, and the *C4H2* gene had the highest expression in red flowers. This indicated that *C4H* genes positively regulated the red flower color of *I. uliginosa*. However, *FLS* expression was the highest in white flowers, and with deepening flower color, *FLS* gene expression gradually weakened, acting as a negative regulator. The results of this study could lay the theoretical foundation for investigating the mechanism of coloration and flower color variation in *I. uliginosa*.

Keywords: *Impatiens uliginosa*; flower color; transcriptome analysis; *C4H* gene; *FLS* gene

Citation: Zhang, X.; Tan, Y.; Li, X.; Liu, Z.; Li, F.; Huang, H.; Huang, M. Analysis of Transcriptome and Expression of *C4H* and *FLS* Genes on Four Flower Colors of *Impatiens uliginosa*. *Horticulturae* **2024**, *10*, 415. https://doi.org/10.3390/horticulturae10040415

Academic Editor: Antonio Ferrante

Received: 25 March 2024
Revised: 13 April 2024
Accepted: 18 April 2024
Published: 19 April 2024

1. Introduction

Flower color is an important ornamental trait of garden plants, which determines the economic and ornamental value of plants and has important biological significance such as attracting insect pollinators, reducing stress, and perpetuating the race [1]. The main coloring pigments of plant flowers and fruits are carotenoids, flavonoids, and betalains [2]. The formation of flower color is the result of a combination of many factors. pH, pigment type, auxin effect, metal ions, petal epidermal cell structure, and pigment number affect the coloration and variation mechanism of plant flower color [3,4].

There are large differences in the coloration mechanism of anemones originating from different regions and with different flower colors, including pigment types, mineral elements, and developmental stages. Researchers have isolated kaempferol, quercetin, geranium pigments, cornflower pigments, delphinium pigments, and their derivatives from the petals of red *Impatiens balsamina* [5,6]. Clevenger found that cornflower pigments were the most widely distributed, followed by mallow pigments, by determining the anthocyanins of 19 species of *Impatiens* [7]. Hasan identified nine flavonoids, three of which

were new, by determining the leaves of the *Impatiens* bicolor [8]. Tatsuzawa et al. isolated mallow pigments and their derivatives from purple *Impatiens textori* Miqs [9]. Isolation of two new quercetin glycoside compounds from *Impatiens balsamina* was achieved by Lei et al. [10]. Mariana found dihydromyricetin, quercetin, kaempferol, and their derivatives by determining the flavonoids in the petals of *Impatiens glandulifera* [11].

Genes play an important role in the biosynthesis of anthocyanin glycosides, and the expression of these genes is closely related to anthocyanin content, thus affecting the flower color of plants [12]. Cinnamate-4-hydroxylase (*C4H*) alias trans-cinnamate-4-monooxygenase is an important enzyme for step 2 in the synthesis of anthocyanin. It can form a complex with *PAL* and *4CL* and then immobilize it through the N-terminal hydrophobic helix region, which plays an important role in the chemical transfer and subcellular localization of the electron chain [13], catalyzing the conversion of trans-cinnamic acid to 4-coumaric acid, which then efficiently modulates carbon metabolism pathways [14,15]. C4H is a member of the cytochrome P450 oxidase (Cytochrome P450, CYP450) family and belongs to the CYP73 subfamily [16]. The first P450 enzyme in plants to be clonally characterized and functionally determined is *C4H*, which maintains high activity in plant tissues compared to other P450s [17]. It has been shown that the protein activity of *C4H* directly affects the synthesis of flavonoid compounds, lignin, aromatic compounds, etc., in plants [18] and is considered to be an important regulatory point. Baek found that the changes in *C4H* gene expression during the development of blackberries (Korean black raspberry) were the same as the changes in flavonoid accumulation [19]. Studies on the expression pattern of the *C4H* gene in plants inferred that its expression level is also closely related to lignin biosynthesis [20].

In addition, the flavonol synthase gene (*FLS*) has an effect on the accumulation of anthocyanin glycosides and the synthesis of flavonols as well as the color presentation of plants [21], and the *FLS* gene determines the production of plant flavonoids. The *FLS* gene dihydroflavonols was catalyzed by *FLS* to produce flavonols. *FLS* is a member of the 2-ketoglutarate-dependent dioxygenase (2-ODD) family of enzymes, to which anthocyanidinsynthase (ANS) and flavonol synthase (FLS) in the flavonoid metabolic pathway belong [22]. Tanaka isolated and cloned the *FLS* gene from *Petunia hybrida* and also enabled its successful expression in yeast [23]. Subsequently, isolation and cloning of *FLS* genes have been reported in a large number of plants such as *Solanum tuberosum*, *Arabidopsis thaliana*, *Ginkgo biloba*, strawberry (*Fragaria* × *ananassa*), etc. [24–28]. It has also been explored that the reaction substrates of the *ANS* gene and the *FLS* gene are identical, i.e., there was competition and the amino acid sequence similarity between the two was high [29], which further confirms that the *FLS* gene has an effect on the synthesis of both phytoflavonols and anthocyanins.

The group collected *I. uliginosa*, which belongs to the genus *Impatiens* of the family Balsaminaceae of the herbaceous plants, mainly distributed in Yunnan, Guizhou, Guangxi, and other places in China in the early stage of the collection of germplasm resources in the field [30]. There are four different flower colors of *I. uliginosa*, deep red, red, pink, and white, in different habitats or in the same habitat, and the color is bright. The study of the coloring mechanism of this mutant strain is of great significance to the molecular breeding of flower color in *Impatiens*. Currently, the study of flower color in *Impatiens* is mainly focused on the determination of flavonoids in petals and the cloning and expression of structural genes in the synthesis pathway of anthocyanin, while the transcriptional regulation of the molecular mechanism of flower color presentation in the same species with multiple colors has not yet been resolved. Therefore, in this study, we mined the transcriptome data of four different flower colors of *I. uliginosa* to screen the genes related to the regulation of flower color.

2. Materials and Methods

2.1. Plant Materials

The test material used in this experiment was *I. uliginosa*, a wild species of *Impatiens*, which has four different flower colors: white, pink, red, and deep red. They were collected

from Aziying, Laoyuhe Wetland Park, Kunming City, Yunnan Province, and Dadieshui, Shilin Yi Autonomous County, Yunnan Province, respectively. Five plants of each of the four flower colors of *I. uliginosa* were selected to be robust and free of pests and diseases, and the whole flower organ of one flower in the state of full bloom was extracted from each plant, which was separated from the plant and quickly put into liquid nitrogen. The four flower colors were mixed separately, labeled, and then stored at −70 °C (Figure 1).

Figure 1. Blooming period of four different color flowers of *I. uliginosa.* Bars = 1 cm.

2.2. Transcriptome Sequencing and Analysis

2.2.1. RNA Extraction and Detection

Total RNA was extracted from 1 mg of the combined samples by a grinder using the OMEGA Quick RNA Isolation Kit according to the instructions. Contaminating genomic DNA was removed using RNAase-free DNase I. RNA purity was determined using a spectrophotometer, and RNA integrity was verified using an agarose gel. Library and RNA-Seq construction were performed by UW Genetics; the constructed libraries were sequenced using the Illumina sequencing platform after passing quality control.

2.2.2. Second-Generation Sequencing and Assembly

The cDNA library was constructed by Illumina's NEB Next®Ultra™ RNA Library Prep to obtain raw data, followed by QC to obtain clean data to assemble a non-redundant unigene from scratch. In this study, we sequenced four different flower colors and assembled the *I. uliginosa* flower color transcriptomes named IUDR (deep red), IUR (red), IUP (pink), and IUW (white), respectively.

2.2.3. Sequence Analysis and Annotation

After the sequencing data were offloaded, it was necessary to assemble the nucleotide sequences of unigenes of *I. uliginosa* through quality control, filtering, and assembly steps. In order to fully obtain the functional annotation of unigene, this study used the comparison software BLASTX v2.2.23 (BasicLocal Alignment Search Tool) (E-value $\leq 1 \times 10^{-5}$) to compare the unigene sequences with the open databases, so that they can be functionally annotated and categorized. The public alignment databases include GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg, accessed on 3 December 2023), Swiss-Prot (A manually annotated and reviewed protein sequence database, http://www.genome.jp/kegg, accessed on 4 December 2023), and the Unigene Sequence Search Tool (E-value $\leq 1 \times 10^{-5}$). And the reviewed protein sequence database (http://www.ebi.ac.uk/uniprot/, accessed on 6 December 2023), COG/KOG (EuKaryotic Orthologous Group, http://ftp.ncbi.nih.gov/pub/COG/KOG, accessed on 7 December 2023), eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups, http://eggnogdb.embl.de/, accessed on 10 December 2023), Pfam (a database of conserved Protein families or domains), and NR (NCBI non-redundant proteins, https://blast.ncbi.nlm.nih.gov/Blast.cgi, accessed on 11 December 2023) databases were compared (E-value $\leq 1.0 \times 10^{-5}$) [30].

2.2.4. Screening and Enrichment Analysis of Differentially Expressed Genes (DEGs)

The assembled sequences were compared with the protein database for Blastx (http://www.ncbi.nlm.nih.gov/BLAST/, accessed on 15 December 2023), and based on the comparison results, the expression levels of the genes were quantified using the software RSEM v1.2.12, after which the expression levels of the genes were normalized using the FPKM method, and the differential genes were screened on the basis of the expression amount.

2.2.5. Related Gene Expression and Analysis

Primers were designed according to the sequences of *IuC4H1*, *IuC4H2,* and *IuFLS*, and *IuActin* was used as the internal reference gene (Table S1). The expression differences of *IuC4H1*, *IuC4H2,* and *IuFLS* were analyzed by qRT-PCR on the bloom samples of four flower colors of *I. uliginosa*, and three biological replicates were set for each sample. qRT-PCR was carried out on a Roche: LightCycler® 480 II Fluorescent Quantitative PCR Instrument. The qRT-PCR reaction system (20 μL) included the following: 10 μL of qPCR SYBR Green Master Mix, 7 μL of double-distilled water, 1 μL each of forward and reverse primers, and 1 μL of template cDNA. The specific program was as follows: predenaturation at 95 °C for 15 s, denaturation at 60 °C for 30 s, and annealing at 72 °C for 1 min, for 40 cycles. The relative expression of genes was analyzed by using the $2^{-\Delta\Delta Ct}$ calculation method. Data were analyzed with the help of SPSS software 27.0 and plotted by Origin software 64.

3. Results

3.1. RNA Sequencing and Transcriptome De Novo Assembly

In order to understand the molecular mechanism of flower color variation in *I. uliginosa*, the transcripts were assembled from scratch using a reference-free genome approach to construct four high-throughput sequencing libraries of four different flower colors of *I. uliginosa*, namely, IUDR (deep red), IUR (red), IUP (pink), and IUW (white). After removing the articulated subsequence, ambiguous reads, and low-quality reads, the sequencing depths from the transcriptome data of IUDR, IUR, IUP, and IUW, respectively, ranged from 6.65 to 6.77 GB with an average depth of 6.5 GB. The Q30 values ranged from 95.94 to 96.09% with an average of 96.00%, and the net number of reads ranged from 44.35 to 45.11 M, with a net reads ratio of 83.02~87.71% (Table 1).

Table 1. Sequencing quality of transcripts of four different colors.

Sample	Total Raw Reads (Mb)	Total Clean Reads (Mb)	Total Clean Bases (Gb)	Clean Reads Q20 (%)	Clean Reads Q30 (%)	GC (%)	Clean Reads Ratio (%)
IUDR	52.69	44.62	6.69	98.78	96.09	41.46	84.68
IUP	54.33	45.11	6.77	98.75	96	42.77	83.02
IUR	52.69	44.35	6.65	98.78	96.11	42.2	84.17
IUW	51.04	44.77	6.71	98.72	95.94	41.71	87.71

A count of the length and number of transcripts revealed that the number of the four different transcripts ranged from 64,723 to 93,522, with an average length of around 900 bp and an N50 of around 1500 bp (Table 2). The total sample of the four different transcriptomes of unigenes contained 100,705 full-length cDNA sequences with 115,842,379 bases, of which the average length of transcript unigene was 1150 bp (Figure 2), the N50 was 1890 bp, and the highest number of fragments distributed at 300 bp was 22,613, which was 22,613, which accounted for a total unigene percentage of 22.45%. The number of unigenes larger than 3000 bp accounted for 6.54% (6592 fragments) of the total number (Table 3).

Table 2. Quality indicators for transcripts of four different colors.

Sample	Total Number	Total Length	Mean Length	N50	N70	N90	GC (%)
IUDR	79,277	69,699,393	879	1545	946	336	41.46
IUP	91,109	76,919,883	844	1521	868	318	42.77
IUR	93,522	80,820,572	864	1530	903	329	42.2
IUW	64,723	59,663,736	921	1600	982	361	41.71

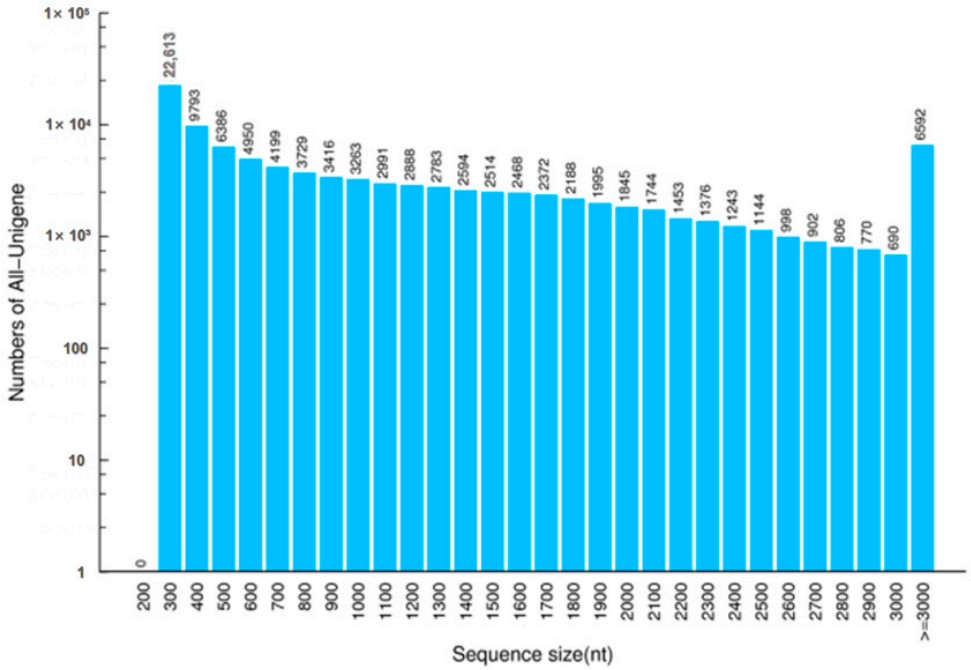

Figure 2. Length distribution of All-Unigene.

Table 3. Unigene data summary of four different colors.

Sample	Total Number	Total Length	Mean Length	N50	N70	N90	GC (%)
IUDR	55,003	58,568,930	1064	1689	1135	462	41.44
IUP	67,735	66,343,559	979	1646	1016	390	42.76
IUR	70,685	70,639,298	999	1640	1046	405	42.2
IUW	48,345	51,364,234	1062	1702	1121	453	41.68
All-Unigene	100,705	115,842,379	1150	1890	1267	495	41.98

3.2. Functional Annotation of Transcriptome Genes of I. uliginosa

The results of sequence match analysis with other species showed that *Vitis vinifera* had 10,213 homologous genes, accounting for 14.32% of the total, followed by *Sesamum indicum* with 4318 homologous genes, accounting for 6.06% of the total, *Coffea canephora* with 3458 homologous genes, accounting for 4.85% of the total, and *Solanum tuberosum* with 2983 homologous genes, accounting for 4.18% of the total (Figure 3). A total of 100,705 unigenes were successfully annotated by searching five public databases based on BLASTx (E-value $< 1 \times 10^{-5}$), among which the NCBI-NR database contained 71,307 unigenes, Swiss-Prot contained 50,638 unigenes, the KEGG database contained 55,276 unigenes, the InterPro database contained 59,058 unigenes, and the KOG database contained 32,641 unigenes (Figure 4, Table 4).

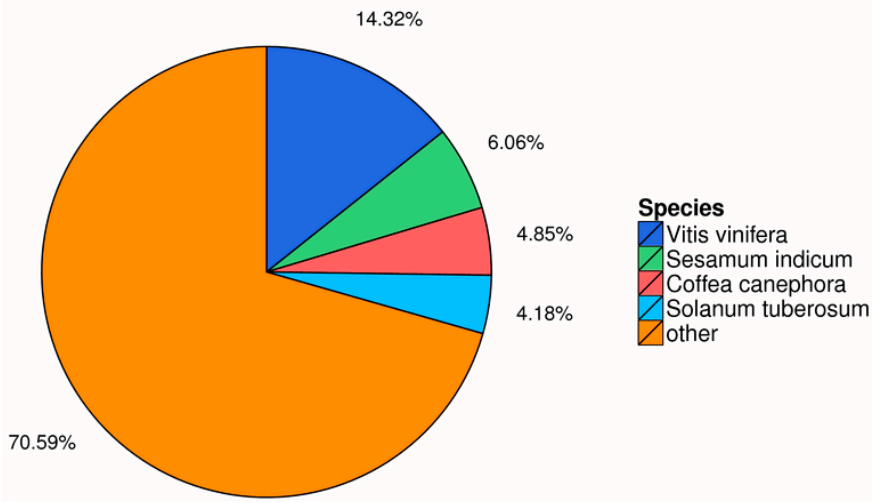

Figure 3. Percentages of unigenes that align with sequences from top five species in the NR database.

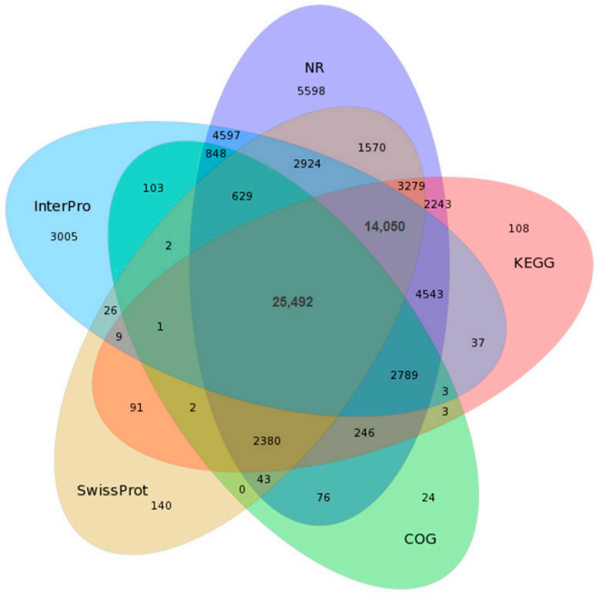

Figure 4. Annotation from five databases. Note: KEGG: Kyoto Encyclopedia of Genes and Genomes; COG: Clusters of Orthologous Groups of proteins; NR: NCBI non-redundant protein sequences.

Table 4. Annotation of unigenes searched against NR, Nt, Swiss-Prot, KEGG, KOG, InterPro, and GO databases.

Values	Number	Percentage
NR-Annotated	71,307	70.81%
Nt-Annotated	59,001	58.59%
Swissprot-Annotated	50,638	50.28%
KEGG-Annotated	55,276	54.89%
COG-Annotated	32,641	32.41%
Interpro-Annotated	59,058	58.64%
GO-Annotated	29,473	29.27%
Overall	77,877	77.33%
Total	100,705	100%

3.3. GO (Gene Ontology) and COG (Clusters of Orthologous Group of Proteins) Annotation Analysis of I. uliginosa Transcriptome Data

The GO gene pool contained three gene attribute sections: molecular function, biological process, and cellular component. The transcriptome of *I. uliginosa* contains 71,307 GO-annotated genes, accounting for 70.81% of the whole annotation, and consists of 58 different annotated gene functional regions. Among them, the biological process section contained 14,490 genes, the metabolic process contained 14,948 genes, and the single-organism process contained 9082 genes; the molecular function section contained 13,079 genes and the catalytic activity process contained 14,757 genes; the cellular component section contained 9905 genes, the cell cellular component contained 9905 genes, the cell part contained 11,934 genes, and the cell contained 12,094 genes. In the three categories of "translation regulator activity", "locomotion", and "cell killing", the unigenes were annotated to 1, 2, and 4 in each category (Figure 5).

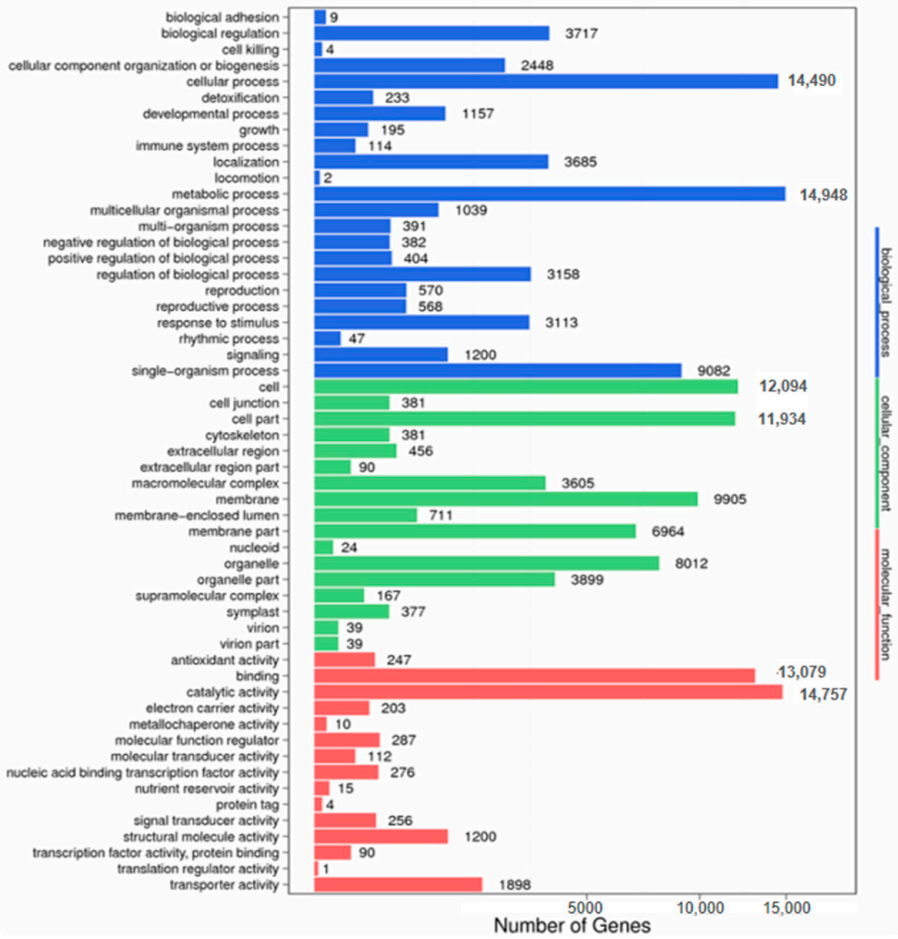

Figure 5. Annotation of unigenes searched against GO databases.

The COG database had the ability to infer unknown gene functions and compare homologous protein sequences, which accounted for 32.41% of the total annotated genes in the *I. uliginosa* transcriptome COG database, of which a total of 32,641 genes were annotated. The annotated genes were categorized into 25 classes according to their functions. General function prediction only contained a total of 9925 genes; replication, recombination, and repair contained 4816 genes; transcription contained 5406 genes. Post-translational modification, protein turnover, and chaperones contained 4029 genes (Figure 6).

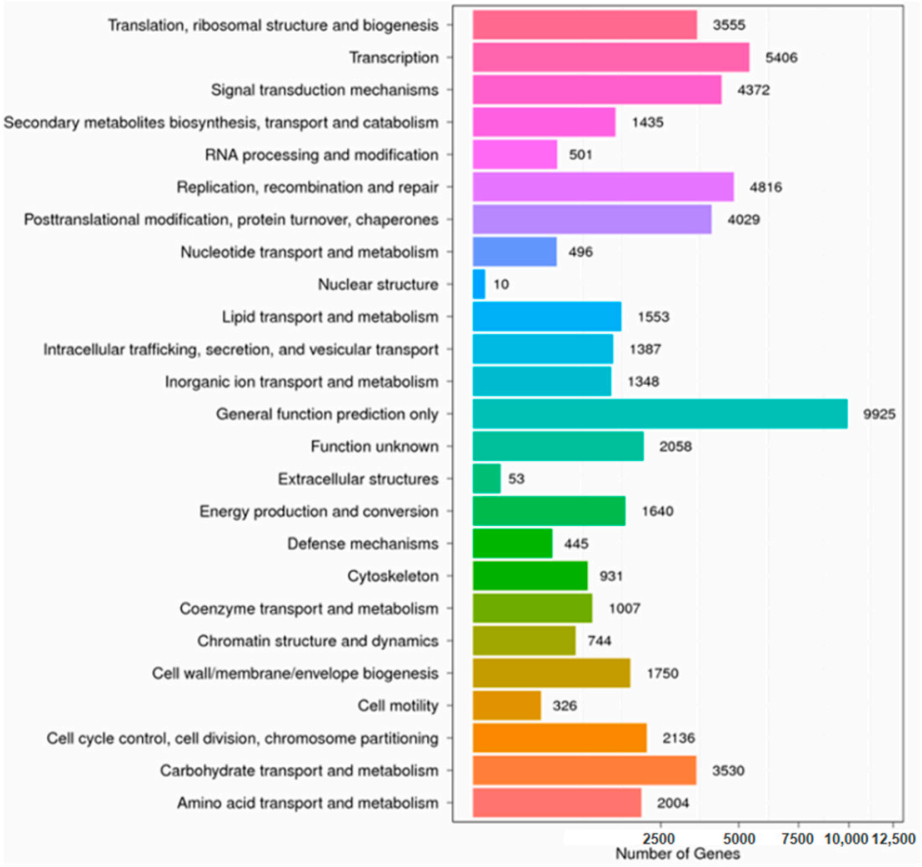

Figure 6. Annotation of unigenes searched against KOG databases.

3.4. Identification and Functional Annotation of Differential Genes (DEGs) in I. uliginosa

Based on the results of the gene expression levels of each sample, the number of genes obtained for each different flower color was analyzed using a Venn diagram, and the results showed that there was a large gap in the expression of genes in the four samples of different flower colors, with a total of 1402 genes between IUDR and IUP; a total of 1639 genes between IUDR and IUR; and a total of 1842 genes between IUDR and IUW; a total of 992 genes between IUP and IUR. The number of genes between IUP and IUP totaled 9929; the number of genes between IUP and IUW totaled 1157; and the number of genes between IUR and IUW totaled 1087. There were a total of 38,191 genes between the four samples, namely, IUDR, IUP, IUW, and IUR. Among them, the *C4H* and *FLS* genes were better expressed (Figure 7A).

Differentially expressed genes (DEGs) were detected using the DEseq2 algorithm, and the results of DEG detection are shown in Figure 7B. Subsequently, gene expression in the IUDR, IUP, IUW, and IUR samples was compared two-by-two, and the results showed that downregulated expression of differentially expressed genes (DEGs) predominated in the comparison of the four different flower colors (Figure 7B). A total of 6197 DEGs were expressed in IUDR vs. S IUW, of which 6098 were upregulated and 14,650 were downregulated. A total of 6197 DEGs were expressed in IUDR vs. IUP, of which 15,519 were upregulated and 13,894 were downregulated. A total of 6197 DEGs were expressed in IUW vs. IUP, of which 6234 were upregulated and 13,790 were downregulated. There were a total of 6197 DEGs in IUDR vs. IUR, with 14,353 upgrades and 15,166 downgrades; a total of 6197 DEGs in IUR vs. IUP, of which 5844 were upgraded and 7216 were downgraded; and a total of 6197 DEGs in IUR vs. IUW, with 6496 upgrades and 16,618 downgrades. In the IUDR vs. IUP, IUDR vs. IUR, and IUR vs. IUP two-by-two comparisons, there was little difference between the number of upward and downward adjustments in DEGs, while the

number of downward adjustments was greater than the number of upward adjustments in the other three comparisons.

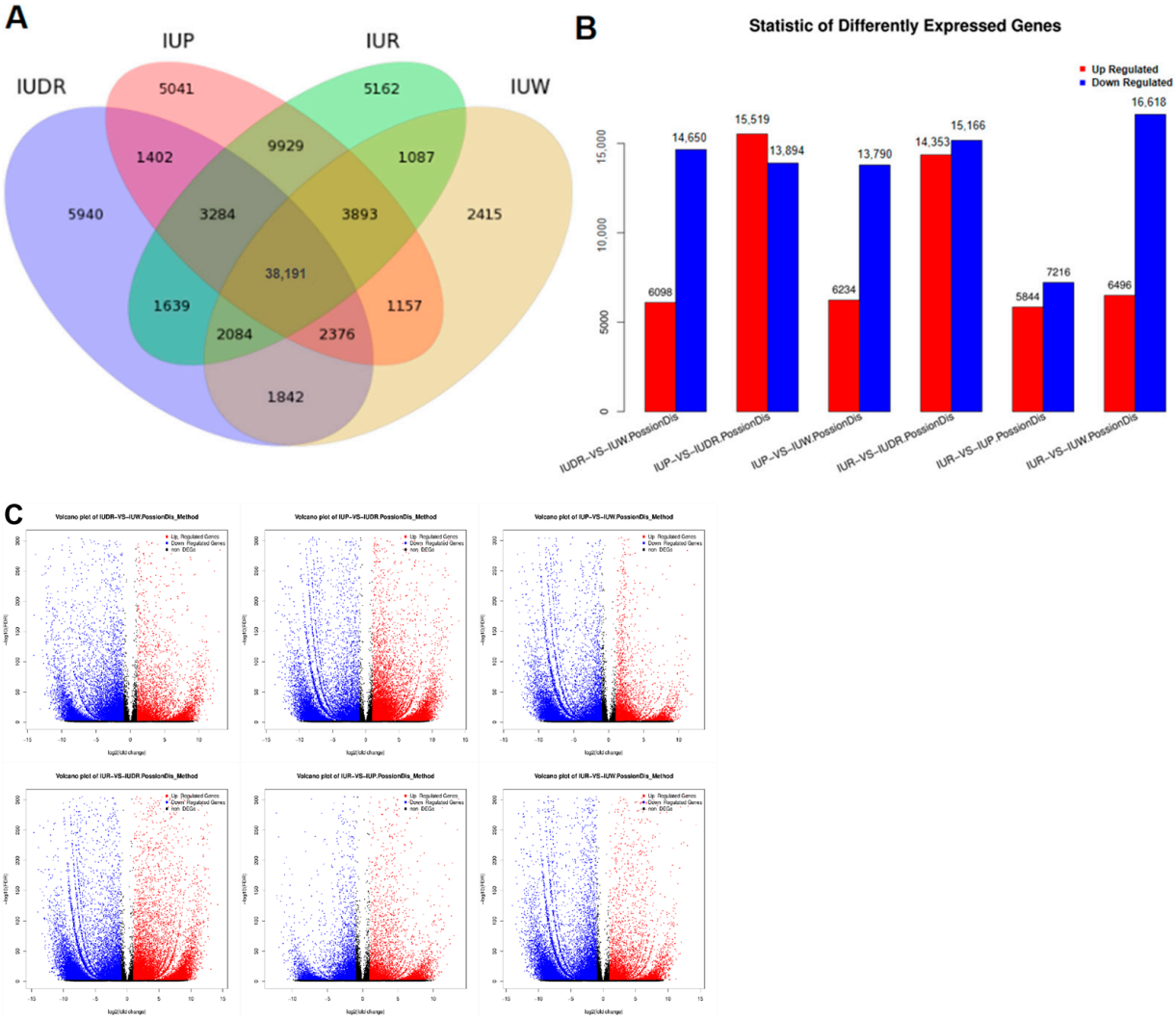

Figure 7. The expressed genes analysis between petals at four different stages: (**A**) Venn diagram shows the expression of genes in IUDR, IUP, and IUR samples. (**B**) Venn diagram showing the DEGs between IUDR vs. IUP, IUDR vs. IUR, and IUP vs. IUR. (**C**) The numbers of up- and downregulated genes between IUDR vs. IUP, IUDR vs. IUR, and IUP vs. IUR. Volcano plots of transcriptome between IUDR vs. IUP, IUDR vs. IUR, and IUP vs. IUR. Red dots represent upregulated DEGs. Blue dots represent downregulated DEGs. Black points represent non-DEGs.

The distribution of DEGs was also demonstrated using a volcano plot. In the IUDR vs. IUR comparison, the number of DEGs had a statistically significant change in the expression level in the volcano plot over the other comparisons (Figure 7C).

3.5. Identification and Expression Analysis of Transcription Factors in I. uliginosa

Transcription factors in plants are involved in the regulation of various metabolic pathways in plants, which enable plants to have a wider range of adaptations, and transcription factors were thought to have a wider regulation in anthocyanin biosynthesis than structural genes. Therefore, we predicted the unigenes encoding transcription factors from the transcriptome data of four different flower colors of *I. uliginosa*, and a total of 3493 unigenes were identified, which could be classified into 59 families (Figure 8), most of which were from the "*MYB*" family (429 unigenes), which accounted for 12.47% of the

total number of unigenes, and the "MYB-unigenes" family (12.47% of the total number of unigenes). The "*MYB*-related" family (355 unigenes) accounted for 10.32% of the total, the "*AP2*-EREBP" family (227 unigenes) accounted for 6.60% of the total, and the "*bHLH*" family (227 unigenes) accounted for 6.60% of the total. The "*bHLH*" family (221 unigenes) accounted for 6.42% of the total. *MYB* and *MYB*-related transcription factors are essential for the plant flavonoid metabolic pathway.

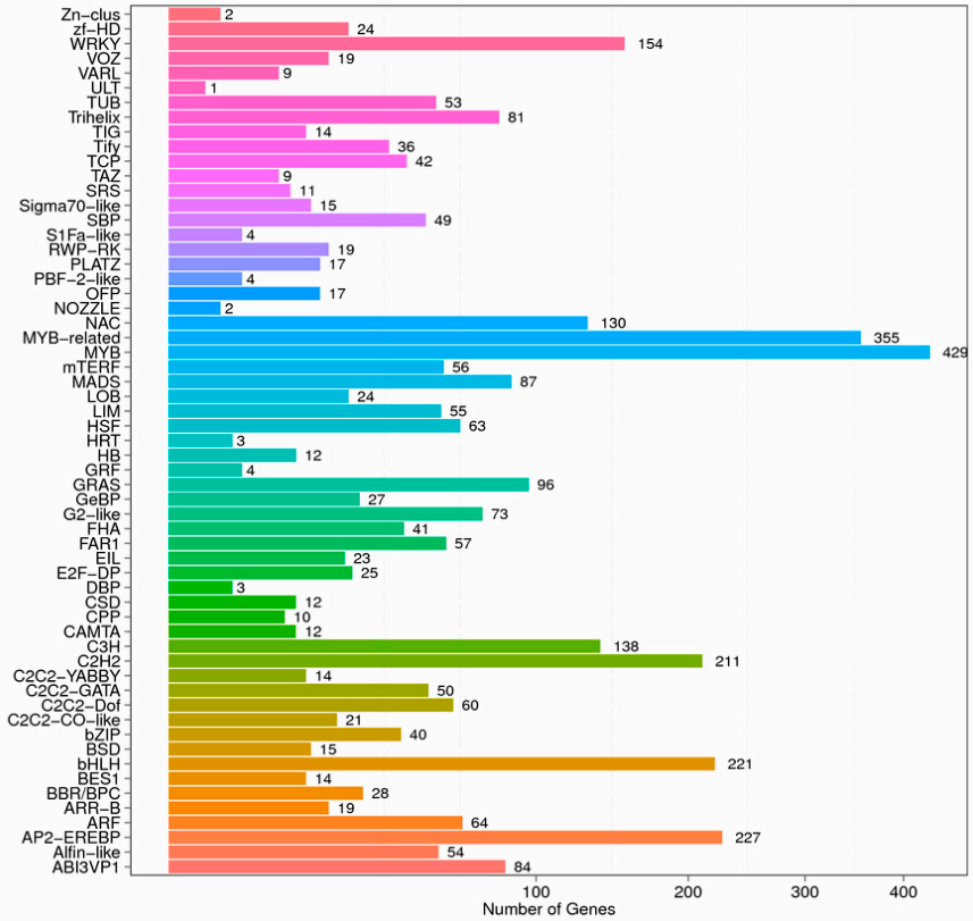

Figure 8. Transcription factor family classification of unigenes.

3.6. qRT-PCR Validation of the Candidate Genes

Two candidate genes were selected for qRT-PCR to verify the accuracy of the transcriptome data (Figure 9). There were significant differences in the expression of *IuC4H1*, *IuC4H2*, and *IuFLS* genes in different flower colors of *I. uliginosa*. The results showed that the *C4H1* gene was most expressed in pink flowers, slightly more in white flowers than in deep red flowers, and least in red flowers, with a significant difference in expression between pink and red. Expression of the *C4H2* gene was highest in red flowers, slightly higher in pink flowers than in deep red flowers, and weaker in white flowers, significantly different from the other three flower colors. *FLS* expression was highest in white flowers and lowest in deep red flowers and was significantly different from the other two flower colors. As the flower color deepened, the expression of *FLS* gradually weakened. Thus, among the four flower colors of *I. uliginosa*, the *C4H1* gene was expressed most frequently in pink flowers, the *C4H2* gene was expressed most frequently in red flowers, and the *FLS* was expressed most frequently in white flowers, suggesting that both the *C4H* and *FLS* genes are involved in the formation of *I. uliginosa* flower color.

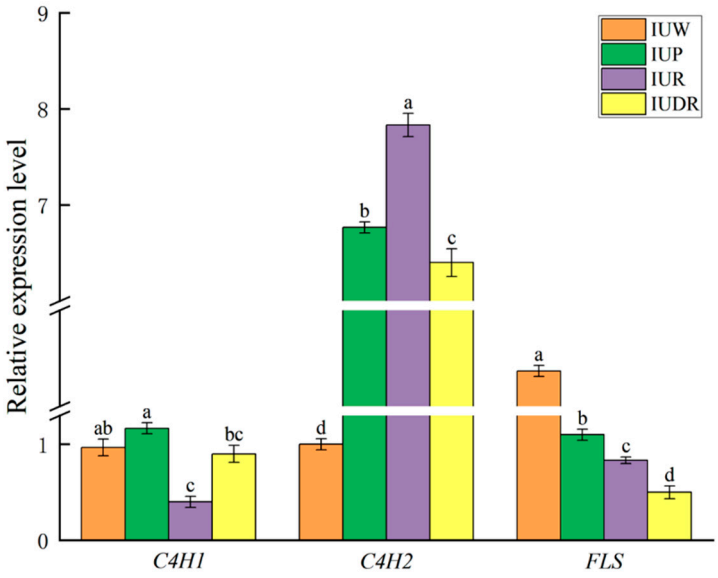

Figure 9. The relative expression of *C4H1*, *C4H2*, and *FLS* genes in four flower colors of *I. uliginosa*. Data are shown as the mean ± standard error of the mean, based on three replicates. Different letters indicate significant differences at $p < 0.05$ according to Duncan's multiple range test.

4. Discussion

The origin and evolution of flowers is a key issue in exploring the origin and evolution of angiosperm species. Meanwhile, flower color is an epigenetic and evolutionarily adapted phenotypic trait in ornamental plants, and the quality of flower color determines the ornamental and economic value of plants. Flowering plants will show a wide range of flower colors, with multiple colors of the same species and multiple species of the same color [2,31]. From previous studies, the genus *Rhododendron* (*Ericaceae*) has more than 1000 flower colors with unique ornamental value and a model system for flower color studies. Nie et al. investigated the differences in color-presenting mechanisms of two rhododendron species, yellow-flowered *R. molle* and red-flowered *R. simsii*, through a multilocus study [32]. For plants of the same genus, differences in key enzymes or gene regulatory mechanisms for pigment synthesis can also result in differences in flower color [33]. Researchers also explored genetically related genes for flower color and pigmentation by studying small-flowered chrysanthemums with different flower colors [34]. Studies on the flower color of homozygous ornamental plants are still scarce.

In this study, four flower colors of *I. uliginosa* were used as research materials, and transcriptome sequencing was performed to lay the foundation for investigating the molecular mechanism of flower color synthesis in *I. uliginosa*. The length and number of transcripts were counted, and it was found that the number of different transcripts of the four flower colors of *I. uliginosa* ranged from 64,723 to 93,522, with an average length of about 900 bp and an N50 of about 1500 bp. The total sample of the four different transcriptomes of unigenes contained 100,705 full-length cDNA sequences with 115,842,379 bases, of which the average length of transcript unigene was 1150 bp, the N50 was 1890 bp, and the highest number of fragments distributed in the 300 bp was 22,613 fragments, which accounted for 22.45% of the total number of unigenes. The number of unigenes larger than 3000 bp was 6592, accounting for 6.54% of the total number of unigenes.

The *C4H* gene is a key enzyme gene in the second step of the anthocyanin biosynthesis pathway [35], and also, the activity of this enzyme can affect the synthesis of compounds such as lignin and aromatic species, which has an important role in plant secondary metabolism [36]. In the four flower colors of *I. uliginosa* at full bloom, *C4H1* was expressed highest in pink and lowest in red, whereas *C4H2* was expressed highest in red and lowest in white, and thus it was hypothesized that the *C4H2* gene had a stronger positive effect than *C4H1* on flower color formation in *I. uliginosa*, in the same way that only one of the

two copies of osmanthus is highly expressed in petals [37] and may separately fulfill their respective functions.

Flavonols are one of the major co-pigments that enhance the stability of anthocyanin glycosides and influence flower color formation. In the metabolic pathway of flavonoid compounds, flavonols share a presynthesis pathway with anthocyanin glycoside biosynthesis. Flavonol synthase (FLS) catalyzes the conversion of dihydroflavonols to flavonols [38]. The relative expression of the *FLS* gene in the four flower colors differed more significantly, and the highest expression was found in white flowers of *I. uliginosa*. This inference was consistent with the results of Ping in his functional validation of the *FLS* gene in roses, where they found that transgenic tobacco had increased flavonol content and flower differentiation to white [39]. Suppression of the *FLS* gene in *Chinese bellflower* was followed by substantial red coloration at the bud stage of the transgenic strain and pigment accumulation in the pistil, a phenotype that was evident in the progeny after crosses [40]. It can be hypothesized that the *FLS* gene may play an important role in the formation of white flowers of *I. uliginosa*.

5. Conclusions

In this study, four flower colors (white, pink, red, and deep red) of *I. uliginosa* were used as materials, and transcriptome sequencing was performed on the petals of *I. uliginosa* at the full bloom stage to investigate the intrinsic mechanism of four flower colors in *I. uliginosa* from the point of view of molecular biology. At present, there are still few studies on the specific color formation mechanism of the four flower colors of *I. uliginosa*, and the previous studies mainly focused on physiological and biochemical experiments, such as the determination of anthocyanin and pH. This study is the first to perform transcriptome sequencing of the four flower colors of *I. uliginosa*, and it can lay a theoretical foundation for further study on the molecular mechanism of different flower colors of *I. uliginosa* and other ornamental plants.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae10040415/s1, Table S1: Sequences of primers related to qRT-PCR amplification.

Author Contributions: Software, F.L.; Investigation, Z.L.; Data curation, X.L.; Writing—original draft, X.Z.; Writing—review & editing, Y.T.; Supervision, H.H. and M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out with the support of the National Natural Science Foundation of China [32060364, 32060366, 31860230], Joint Special Key Project for Agricultural Basic Research in Yunnan Province(202101BD070001-018), Major scientific and technological projects in Yunnan Province [202102AE090052], and Program for Doctoral Supervisors Team in Genetic Improvement and High-efficient Propagation of Landscape Plants in Yunnan Province.

Data Availability Statement: Data are contained within the article and Supplementary Materials.

Acknowledgments: Thanks to Southwest Forestry University for providing experimental instruments and equipment for this experiment, and also thanks to teachers Haiquan Huang and Meijun Huang for their patient guidance during the experiment.

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

References

1. Kellenberger, R.T.; Glover, G.B. The evolution of flower colour. *Curr. Biol.* **2023**, *33*, R484–R488. [CrossRef] [PubMed]
2. Li, H. The of Anthocyanidin, Metal Elementsand pH on the Colour Formation of Blue Flowers from Monocots. Master's Thesis, North West Agriculture and Forestry University, Xianyang, China, 2013; pp. 52–56.
3. Han, Y.; Huang, K.; Liu, Y.; Jiao, T.; Ma, G.; Qian, Y.; Wang, P.; Dai, X.; Gao, L.; Xia, T. Functional Analysis of Two Flavanone-3-Hydroxylase Genes from *Camellia sinensis*: A Critical Role in Flavonoid Accumulation. *Genes* **2017**, *8*, 300. [CrossRef] [PubMed]
4. Lou, Q.; Liu, Y.; Qi, Y.; Jiao, S.; Tian, F.; Jiang, L.; Wang, Y. Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* **2014**, *65*, 3157–3164. [CrossRef] [PubMed]

5. Hagen, C.W. The differentiation of pigmentation in flower parts I. The flavonoid pigments of *Impatiens balsamina*, genotype *llHHPrPr*, and their distribution within the plant. *Am. J. Bot.* **1966**, *53*, 46–54. [CrossRef]
6. Miles, C.D.; Hagen, C.W. The differentiation of pigmentation in flower parts, IV. Flavonoid elaborating enzymes from petals of Impatiens balsamina. *Plant Physiol.* **1968**, *43*, 1347–1354. [PubMed]
7. Clevenger, S. Anthocyanidins of some *Impatiens* species. *Evolution* **1971**, *25*, 669–677. [PubMed]
8. Hasan, A.; Tahir, M.N. Flavonoids from the leaves of *Impatiens* bicolor. *Turk. J. Chem.* **2005**, *29*, 65–70.
9. Tatsuzawa, F.; Saito, N.; Mikanagi, Y.; Shinoda, K.; Toki, K.; Shigihara, A.; Honda, T. An unusual acylated malvidin 3-glucoside from flowers of *Impatiens textori* Miq. (*Balsaminaceae*). *Phytochemistry* **2009**, *70*, 672–674. [CrossRef]
10. Lei, J.; Qian, S.H.; Jiang, J.Q. Two new flavone glycosides from the seeds of *Impatiens balsamina* L. *J. Asian Nat. Prod. Res.* **2010**, *12*, 1033–1037. [CrossRef]
11. Mariana, N.V.; Peter, W.; Gerold, J. Flavonoids from the flowers of *Impatiens glandulifera* Royle isolated by high performance countercurrent chromatography. *Phytochem. Anal. PCA* **2016**, *27*, 116–125.
12. Zhao, L.Q.; Liu, Y.; Huang, Q.; Gao, S.; Huang, M.J.; Huang, H.Q. Effects of cell morphology, physiology, biochemistry and *CHS* genes on four flower colors of Impatiens uliginosa. *Front. Plant Sci.* **2024**, *15*, 151343830. [CrossRef] [PubMed]
13. Liu, K.; Zhao, Z.; Gao, A.; Chen, Y.; Huang, J.; Dang, Z.; Luo, R. Cloning of mango *C4H* gene and its expression analysis. *Jiangsu Agric. Sci.* **2017**, *45*, 8–12.
14. Ma, J.; Su, L.; Yuan, M.; Ji, H.; Li, A.; Wang, X. Cloning and expression studies of peanut *C4H* and *ANR* genes. *J. Nucl. Agric. Sci.* **2012**, *26*, 43–48.
15. Yao, S.; Wang, W.; Li, M.; Xu, Y.; Wang, Y.; Liu, Y.; Gao, L.; Xia, T. The Gene Cloning and expression analysis of tea tree cinnamic acid 4-hydroxylase gene. *J. Tea Sci.* **2015**, *35*, 35–44.
16. Zhao, L.; Ma, L.; Yang, Z.; Feng, W.; Zheng, X. Cloning and expression analysis of the *C4H* gene of *Solanum vulgare*. *Acta Pharm. Sin.* **2017**, *52*, 821–831.
17. Li, L.; Zhao, Y.; Ma, L. New advances in the study of key enzymes of the phenylalanine metabolic pathway: PAL, C4H, 4CL. *Chin. J. Bioinform.* **2007**, *20*, 187–189.
18. Liang, L.; Han, X.; Zhang, Z.; Guo, Q.; Xu, Y.; Liu, J.; Lian, Y. Cloning and expression analysis of the cinnamic acid-4-hydroxylase (*C4H*) gene from *Cinnamomum album*. *Chin. J. Tradit. Chin. Med.* **2014**, *39*, 1767–1771.
19. Baek, M.H.; Chung, B.Y.; Kim, J.H.; Kim, J.S.; Lee, S.S.; An, B.C.; Lee, I.J.; Kim, T.H. cDNA cloning and expression pattern of cinnamate-4-hydroxylase in the Korean black raspberry. *BMB Rep.* **2008**, *41*, 529–536. [CrossRef]
20. Tao, S.; Wang, D.; Jin, C.; Sun, W.; Liu, X.; Zhang, S.; Gao, F.; Khanizadeh, S. Cinnamate-4-hydroxylase gene is involved in the step of lignin biosynthesis in Chinese white pear. *J. Am. Soc. Hortic. Sci.* **2015**, *140*, 573–579. [CrossRef]
21. Li, Y.; Ran, L.; Mo, T.; Liu, N.; Zeng, J.; Liang, A.; Wang, C.; Suo, Q.; Chen, Z.; Wang, Y.; et al. Yellow Petal locus GaYP promotes flavonol biosynthesis and yellow coloration in petals of Asiatic cotton (*Gossypium arboreum*). *TAG Theor. Appl. Genet.* **2023**, *136*, 98. [CrossRef]
22. Schwinn, K.E.; Boase, M.R.; Bradley, J.M.; Lewis, D.H.; Deroles, S.C.; Martin, C.R.; Davies, K.M. *MYB* and *bHLH* transcription factor transgenes increase anthocyanin pigmentation in petunia and lisianthus plants, and the petunia phenotypes are strongly enhanced under field conditions. *Front. Plant Sci.* **2014**, *5*, 603. [CrossRef]
23. Tanaka, Y. Cloning and expression of flavonol synthase from *Petunia hybrida*. *Plant J.* **2010**, *4*, 1003–1010.
24. Pelletier, M.K.; Shirley, B.W. Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in Arabidopsis. Further evidence for differential regulation of "early" and "late" genes. *Plant Physiol.* **1997**, *113*, 1437. [CrossRef]
25. Nielsen, K.; Deroles, S.C.; Markham, K.R.; Bradley, M.J.; Podivinsky, E.; Manson, D. Antisense flavonol synthase alters copigmentation and flower color in lisianthus. *Mol. Breed.* **2002**, *9*, 217–229. [CrossRef]
26. Almeida, J.R.; D'Amico, E.; Preuss, A.; Carbone, F.; de Vos, C.H.; Deiml, B.; Mourgues, F.; Perrotta, G.; Fischer, T.C.; Bovy, A.G.; et al. Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria xananassa*). *Arch. Biochem. Biophys.* **2007**, *465*, 61–71. [CrossRef] [PubMed]
27. Takahashi, R.; Githiri, S.M.; Hatayama, K.; Dubouzet, E.G.; Shimada, N.; Aoki, T.; Ayabe, S.; Iwashina, T.; Toda, K.; Matsumura, H. A single-base deletion in soybean flavonol synthase gene is associated with magenta flower color. *Plant Mol. Biol.* **2007**, *63*, 125–135. [CrossRef]
28. Luo, P. Molecular Cloning and Functional Analysis of Fiavonoids Biosynthesis Related Genes from *Rosa rugosa*. Ph.D. Thesis, Huazhong Agricultural University, Wuhan, China, 2016; pp. 2–4.
29. Lukačin, R.; Wellmann, F.; Britsch, L.; Martens, S.; Matern, U. Flavonol synthase from Citrus unshiu is a bifunctional dioxygenase. *Phytochemistry* **2003**, *62*, 287–292. [CrossRef] [PubMed]
30. Li, Y.; Wei, C.M.; Li, X.Y.; Meng, D.C.; Gu, Z.J.; Qu, S.P.; Huang, M.J.; Huang, H.Q. De novo transcriptome sequencing of *Impatiens uliginosa* and the analysis of candidate genes related to spur development. *BMC Plant Biol.* **2022**, *22*, 553. [CrossRef]
31. Allan, A.C.; Hellens, R.P.; Laing, W.A. *MYB* transcription factors that colour our fruit. *Trends Plant Sci.* **2008**, *13*, 99–102. [CrossRef]
32. Nie, S.; Zhao, S.W.; Shi, T.L.; Zhao, W.; Zhang, R.G.; Tian, X.C.; Guo, J.F.; Yan, X.M.; Bao, Y.T.; Li, Z.C.; et al. Gapless genome assembly of azalea and multi-omics investigation into divergence between two species with distinct flower color. *Hortic. Res.* **2023**, *10*, 241. [CrossRef]
33. Nie, S.; Ma, H.Y.; Shi, T.L.; Tian, X.C.; El-Kassaby, Y.A.; Porth, I.; Yang, F.S.; Mao, J.F. Progress in phylogenetics, multi-omics and flower coloration studies in *Rhododendron*. *Ornam. Plant Res.* **2024**, *4*, e003. [CrossRef]

34. Song, X.; Tian, Y.; Gao, K.; Li, J.; Li, Y.; Wang, J.; Deng, C.; Zhang, F.; Kong, D.; Fan, G.; et al. Genetic and QTL analysis of flower color and pigments in small-flowered chrysanthemum based on high-density genetic map. *Ornam. Plant Res.* **2023**, *3*, 17. [CrossRef]
35. Zhang, Q.; Yang, W.; Liu, J.; Liu, H.; Lv, Z.; Zhang, C.; Chen, D.; Jiao, Z. Postharvest UV-C irradiation increased the flavonoids and anthocyanins accumulation, phenylpropanoid pathway gene expression, and antioxidant activity in sweet cherries (*Prunus avium* L.). *Postharvest Biol. Technol.* **2021**, *175*, 111490. [CrossRef]
36. Peng, H.; Yang, T.; Whitaker, B.D.; Trouth, F.; Shangguan, L.; Dong, W.; Jurick, W.M. Characterization of spermidine hydroxycinnamoyl transferases from eggplant (*Solanum melongena* L.) and its wild relative *Solanum richardii* Dunal. *Hortic. Res.* **2016**, *3*, 16062. [CrossRef] [PubMed]
37. Zeng, X.; Zheng, R.; Luo, J.; Wang, C. Cloning and Characterization of Cinnamate 4-hydroxylase (*C4H*) Genes from *Osmanthus fragrans*. *J. Hortic.* **2016**, *43*, 525–537.
38. Liu, H.; Su, B.; Zhang, H.; Gong, J.; Zhang, B.; Liu, Y.; Du, L. Identification and Functional Analysis of a Flavonol Synthase Gene from *Grape Hyacinth*. *Molecules* **2019**, *24*, 1579. [CrossRef] [PubMed]
39. Luo, P.; Ning, G.; Wang, Z.; Shen, Y.; Jin, H.; Li, P.; Huang, S.; Zhao, J.; Bao, M. Disequilibrium of Flavonol Synthase and Dihydroflavonol-4-Reductase Expression Associated Tightly to White vs. Red Color Flower Formation in Plants. *Front. Plant Sci.* **2016**, *6*, 1257. [CrossRef]
40. Davies, K.M.; Schwinn, K.E.; Deroles, S.C.; Manson, D.G.; Lewis, D.H.; Bloor, S.J.; Bradley, J.M. Enhancing anthocyanin production by altering competition for substrate between flavonol synthase and dihydroflavonol 4-reductase. *Euphytica* **2003**, *131*, 259–268. [CrossRef]