



Article

Transcriptomic Analysis of Flower Color Changes in *Impatiens uliginosa* in Response to Copper Stress

Yi Tan [†], Xiaoli Zhang [†], Qinmei Li, Xinyi Li, Liang Luo, Haihao He, Guangrong Liang, 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

^{*} Correspondence: haiquanl@163.com (H.H.); xmhhq2001@163.com (M.H.)

[†] These authors contributed equally to this work.

Abstract: *Impatiens uliginosa* is a native and potential water body-restoring ornamental plant. In this study, RNA-Seq technology was used to analyze the transcriptome of its floral organs. Candidate genes related to flower color changes under copper stress were investigated through transcriptome screening, and the intrinsic mechanism of the effects of different concentrations of copper on *I. uliginosa* was revealed at the molecular level. The main findings were as follows: (1) Transcriptome sequencing analysis was performed on the flower organs of *I. uliginosa* treated with different concentrations of copper (0 mg·L⁻¹, 10 mg·L⁻¹, and 20 mg·L⁻¹). A total of 70,319 transcripts and 39,949 unigenes were obtained. An analysis of differentially expressed genes revealed structural genes including *GT*, *ANS*, *CHI*, and *PAL*; transcription factors including *MYB* and *WD40*; and transport factors including *GST* and *ABC*. (2) The gene expression levels of flower color changed in the flowering period of *I. uliginosa* at different copper concentrations. The expression levels of *IuGT* and *IuGST* genes in *I. uliginosa* were significantly different under different concentrations of copper treatments. Their expression levels were the highest at a copper concentration of 0 mg·L⁻¹ and the lowest at 20 mg·L⁻¹. In summary, the low expression of *IuGT* and *IuGST* genes was more conducive to the formation of white flowers of *I. uliginosa*.

Keywords: *Impatiens uliginosa*; copper stress; transcriptome analysis; flower pigmentation changes



Citation: Tan, Y.; Zhang, X.; Li, Q.; Li, X.; Luo, L.; He, H.; Liang, G.; Huang, H.; Huang, M. Transcriptomic Analysis of Flower Color Changes in *Impatiens uliginosa* in Response to Copper Stress. *Horticulturae* **2024**, *10*, 412. <https://doi.org/10.3390/horticulturae10040412>

Academic Editor: Jiafu Jiang

Received: 3 March 2024

Revised: 14 April 2024

Accepted: 14 April 2024

Published: 19 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rapid development of industry and agriculture, as well as the unreasonable discharge of domestic sewage and the burning of fossil fuels, has caused urban and rural water bodies in my country to be polluted by heavy metals, hindering the utilization of water resources. Cu is one of the common heavy metal pollutants in polluted water bodies. It is widely distributed, difficult to degrade, and extremely biologically toxic. It can threaten the health of humans and other aquatic organisms through food chain accumulation and biomagnification. In southwest China, Dianchi Lake, as the largest freshwater body, carries important ecological functions. However, since the 1980s, the water quality of Dianchi Lake has begun to be seriously polluted, especially due to heavy metal pollutants, such as zinc, copper, cadmium, chromium, lead, etc. The copper content in the water is as high as 126.36 µg·L⁻¹. However, there is no normal pollution. The average copper concentration in polluted natural water bodies is 2 µg·L⁻¹, indicating that the current copper pollution level in Dianchi Lake water body is quite serious [1,2]. Compared with physical and chemical methods, phytoremediation technology is more popular due to its advantages of small environmental interference, easy control of secondary pollution, and low cost. It is considered to be one of the current research hotspots in the field of environmental remediation of heavy metal pollution [3]. The selection of tolerant plant species is one

of the basic steps for green plant remediation of Cu-polluted water bodies. Therefore, in-depth study of the molecular mechanism of plant defense against Cu stress can provide theoretical reference for the phytoremediation of Cu water pollution.

Research on plant response to adverse stress can help reveal the response mechanism of plants to adverse environments so that preventive measures can be taken to reduce the damage to plants caused by adverse stress. Transcriptomics is a discipline that studies gene transcription and transcriptional regulation in cells at an overall level and can reveal the molecular mechanisms underlying specific biological processes and diseases. Revealing the expression at the entire genome level under stress from the entire transcription level is of great significance to increase the understanding of complex regulatory networks related to stress adaptation and tolerance and to construct the stress genome transcription regulatory network [4].

Plant glycosyltransferases (GTs) are a multi-member gene superfamily that specifically catalyzes glycosylation reactions and are responsible for transferring sugar moieties to various small molecules and controlling numerous metabolic processes. Glycosylation is one of the key modification steps in plants to produce a broad spectrum of flavonoids with different structures and colors. This process allows anthocyanins to be easily transferred from the production site in the cytoplasm to the vacuole, which shows that glycosylation plays a key role in the formation of flower color [5,6]. It also plays a crucial role in enhancing the stability and solubility of plant anthocyanins. At present, glycosyltransferases have been successfully cloned in model plants, ornamental plants, and horticultural crops such as *Arabidopsis thaliana* [7], *Nelumbo nucifera* [8], and *Prunus persica* [9].

Studies have shown that the glutathione transferase (GST) family is involved in the transport of anthocyanins [10,11]. At present, GST genes have been found to be crucial for anthocyanin transport in a variety of plants, such as *Chrysanthemum morifolium*, lily, and *Pericallis × hybrida*. For example, in *C. morifolium*, members of the GST and MATE families are involved in anthocyanin accumulation, thereby affecting the coloration of floral organs [12]; *LhGST* plays a key role in the transport and accumulation of anthocyanins in the perianth [13]; Jin [14] found that *PhGST3* in *Pericallis × hybrida* may be a candidate gene related to anthocyanin transport. Therefore, the study of plant GSTs in flavonoid accumulation is of great significance.

Impatiens uliginosa is an annual or perennial plant characterized by its unique flower shape, long flowering period, wide adaptability, and high ornamental value [15]. It is widely distributed in Yunnan, Guizhou, Guangxi, and other places. It often grows in forests, moist places in ditches, or by streams, at an altitude of 1500–2600 m. Research on *I. uliginosa* mainly focuses on the cloning of flower color and flower development regulatory genes [16,17], determination of metal elements [18], and the spur development [19,20]. The fact is that discussions on changes in *I. uliginosa* flower color under abiotic stress are rare, especially heavy metal stress. *I. uliginosa* is an important local plant distributed in wetlands around Dianchi Lake, serving as an aquatic flower and water purification plant. Flower color is also a significant factor in determining the ornamental value of *Impatiens*. Based on the previous studies of our research group, it was found that the color of *I. uliginosa* flowers changed significantly under different concentrations of copper. It was observed that the flower color gradually became lighter as the copper concentration increased. However, there are no relevant studies that have reported on the impact of the flower color change so far.

In view of this, this study conducted transcriptome sequencing of *I. uliginosa* petals under Cu stress, and further explored genes related to color changes in petals under copper stress. It lays a theoretical foundation for future in-depth study of the molecular mechanism of *I. uliginosa* in defense against heavy metal toxicity, and also provides a theoretical basis for the practical application of *I. uliginosa* in restoring heavy metal-polluted water bodies.

2. Materials and Methods

2.1. Plant Materials and Treatments

I. uliginosa seeds were collected from Laoyuhe Wetland Park in Kunming, Yunnan Province. They were sown and propagated by cuttings in the Arboretum of Southwest Forestry University. The growth conditions of these plants are an average temperature of about 25 °C, 14 h of sunshine, and a relative humidity of 60% to 75%.

I. uliginosa plants with consistent traits and robust growth were selected. The plant height was about 30 cm. The root soil was rinsed with deionized water and then hydroponically cultivated for a week. Then, we used 1/2 Hoagland (without CuSO₄) to acclimate and culture them. Different concentrations of CuSO₄·5H₂O were then applied, and a total of 5 copper concentration gradients were set: 0 mg·L⁻¹, 5 mg·L⁻¹, 10 mg·L⁻¹, 15 mg·L⁻¹, and 20 mg·L⁻¹. The nutrient solution and treatment solution were all prepared with deionized water. In addition, in order to ensure the consistency of the experimental results and the copper stress suffered by the plants during the experiment, all treatments were replaced with newly prepared nutrient solutions containing corresponding copper concentrations every 6 days, with 10 plants in each treatment. After being treated with copper at different concentrations for 30 days, full petals at the peak flowering stage were collected for subsequent related experiments.

2.2. Experimental Methods

2.2.1. Transcriptome Sequencing and Analysis

The petals of *I. uliginosa* were stored in liquid nitrogen, and samples treated with copper concentrations of 0mg·L⁻¹, 10mg·L⁻¹, and 20mg·L⁻¹ were selected, because we found that under these three concentrations, the color change in *I. uliginosa* flowers was the most significant (Figure 1), and sent to the company for sequencing and subsequent analysis.

Total RNA was extracted from the *I. uliginosa* tissue using Plant RNA Purification Reagent for plant tissue according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA) and genomic DNA was removed using DNase I (TaKara TaKaRa, Tokyo, Japan). Then, the integrity and purity of the total RNA quality was determined by 2100 Bioanalyser (Agilent Technologies, Inc., Santa Clara, CA, USA) and quantified using the ND-2000 (NanoDrop Thermo Scientific, Wilmington, DE, USA). Only high-quality RNA sample (OD260/280 = 1.8~2.2, OD260/230 ≥ 2.0, RIN ≥ 8.0, 28S:18S ≥ 1.0, >1 µg) was used to construct sequencing library.

RNA purification, reverse transcription, library construction, and sequencing were performed at Shanghai Majorbio Bio-pharm Biotechnology Co., Ltd. (Shanghai, China) according to the manufacturer's instructions (Illumina, San Diego, CA, USA). The *I. uliginosa* RNA-seq transcriptome libraries were prepared using Illumina TruSeq™ RNA sample preparation kit (San Diego, CA, USA). Poly(A) mRNA was purified from total RNA using oligo-dT-attached magnetic beads and then fragmented by fragmentation buffer. Taking these short fragments as templates, double-stranded cDNA was synthesized using a SuperScript double-stranded cDNA synthesis kit (Invitrogen, CA, USA) with random hexamer primers (Illumina). Then, the synthesized cDNA was subjected to end-repair, phosphorylation, and 'A' base addition according to Illumina's library construction protocol. Libraries were size selected for cDNA target fragments of 200–300 bp on 2% Low Range Ultra Agarose followed by PCR amplified using Phusion DNA polymerase (New England Biolabs, Boston, MA, USA) for 15 PCR cycles. After being quantified by TBS380, two RNAseq libraries were sequenced in a single lane on an Illumina HiSeq xten/NovaSeq 6000 sequencer (Illumina, San Diego, CA) for 2 × 150 bp paired-end reads.

The raw paired-end reads were trimmed and quality controlled by SeqPrep (<https://github.com/jstjohn/SeqPrep>, accessed on 15 January 2022) and Sickle (<https://github.com/najoshi/sickle>, accessed on 17 January 2022) with default parameters. Then, clean data from the samples (*I. uliginosa*) were used to perform de novo assembly with Trinity (<http://trinityrnaseq.sourceforge.net/>, accessed on 18 January 2022). All the assembled transcripts were searched against the NCBI protein nonredundant (NR), COG, and KEGG

databases using BLASTX to identify the proteins that had the highest sequence similarity with the given transcripts to retrieve their function annotations and a typical cut-off E-value less than 1.0×10^{-5} was set. BLAST2GO (<http://www.blast2go.com/b2ghome>, accessed on 22 January 2022) program was used to obtain GO annotations of unique assembled transcripts for describing biological processes, molecular functions, and cellular components. Metabolic pathway analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG, <http://www.genome.jp/kegg/>, accessed on 24 January 2022).

To identify DEGs (differential expression genes) between two different samples, the expression level of each transcript was calculated according to the transcripts per million reads (TPM) method. RSEM (<http://deweylab.biostat.wisc.edu/rsem/>, accessed on 27 January 2022) was used to quantify gene abundances. Essentially, differential expression analysis was performed using the DESeq2/DEGseq/EdgeR with Q value ≤ 0.05 ; DEGs with $|\log_2FC| > 1$ and Q value ≤ 0.05 (DESeq2 or EdgeR)/ Q value ≤ 0.001 (DEGseq) were considered to be significantly different expressed genes. In addition, functional enrichment analysis including GO and KEGG were performed to identify which DEGs were significantly enriched in GO terms and metabolic pathways at Bonferroni-corrected p -value ≤ 0.05 compared with the whole-transcriptome background. GO functional enrichment and KEGG pathway analysis were carried out by Goatools (<https://github.com/tanghaibao/Goatools>, accessed on 29 January 2022) and KOBAS (<http://kobas.cbi.pku.edu.cn/home.do>, accessed on 31 January 2022).



Figure 1. Variation in flower color of *I. uliginosa* treated with different concentrations of copper. Note: Copper concentration: (A). $0 \text{ mg}\cdot\text{L}^{-1}$; (B). $5 \text{ mg}\cdot\text{L}^{-1}$; (C). $10 \text{ mg}\cdot\text{L}^{-1}$; (D). $15 \text{ mg}\cdot\text{L}^{-1}$; (E). $20 \text{ mg}\cdot\text{L}^{-1}$.

2.2.2. *IuGT* and *IuGST* Gene Cloning and Sequence Analysis

The total RNA of *I. uliginosa* was extracted using Plant RNA Kit (Plant RNA kit, Omega, Norcross, GA, USA), and the first strand of *I. uliginosa* cDNA was transcribed and synthesized using a reverse transcription kit (TransGen, Beijing, China) and stored at -80 °C. The design of gene cloning primers is detailed in Table 1. PCR reaction system ($20 \mu\text{L}$): $2\times$ Taq PCR StarMix (Dye) $10 \mu\text{L}$, ddH₂O $7 \mu\text{L}$, forward and reverse primers $1 \mu\text{L}$ each, template cDNA $1 \mu\text{L}$. The specific program is as follows: pre-denaturation at 94 °C for 2 min, denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, extending at 72 °C for 2 min, and final extending at 72 °C for 10 min, 35 cycles. Perform gel recovery on PCR amplification products using the gel recovery kit (Biomed, Beijing, China). Connect the recovered product to the pMDTM19-T Vector carrier (TaKaRa, Tokyo, Japan) according to the instructions, and connect the carrier to the product in DH5 α in receptive cells. Use PCR to screen positive clones and send them to the company for sequencing (Sangon Biotech (Shanghai) Co., Ltd., Shanghai, China). Analyze the physicochemical properties of proteins using online software ExPasy ProtParam (<https://web.expasy.org/protparam/>, accessed on 31 January 2022). Construct a phylogenetic tree of *IuGST* and *IuGT* using MAGA7.0.

Table 1. Primers of *IuGT* and *IuGST* genes of *I. uliginosa*.

Primer Name	Primer Sequence (5'-3')
IuGTF	GATCCAGTGCCACTCTGAAACAAAGAAAG
IuGTR	TTAGACATGAGACAGTAAGACACTTATGAGC
IuGSTF	ATGGCAGGAAGAGATTCTTCCGTC

Table 1. Cont.

Primer Name	Primer Sequence (5'-3')
IuGSTR	TTAGTTCTTAGCTTGAGATTCAACCATGGC
qIuGTF	GTCACCTTCGTCAACACCGA
qIuGTR	AGTCCATCGGGAATGGCTTG
qIuGSTF	CACTTAATCCTCCCCTCGG
qIuGSTR	TCAACCACCCCAAGAAGCAT
IuActinF	TGAATGTCCCTGCTGTTG
IuActinR	ACCTCCGCATAACTTTACC

2.2.3. Related Gene Expression and Analysis

Primers were designed based on the sequences of *IuGT* and *IuGST*, and *IuActin* was used as the internal reference gene. qRT-PCR was performed on samples treated with different concentrations of copper during the flowering period of *I. uliginosa* to analyze the expression differences of *IuGT* and *IuGST*. Roche: LightCycler[®] 480 II fluorescence quantitative PCR instrument (Roche, Basel, Switzerland) was used for amplification, and qRT-PCR reaction system (20 μ L): qPCR SYBR Green Master Mix 10 μ L, ddH₂O 7 μ L, forward and reverse primers 1 μ L each, and template cDNA 1 μ L. The specific program is as follows: pre-denaturation at 95 °C for 15 s, denaturation at 60 °C for 30 s, and annealing at 72 °C for 30 s, 40 cycles. Primer information is shown in Table 1. The $2^{-\Delta\Delta C_t}$ calculation method was used to analyze the relative expression of genes.

2.2.4. Data Processing

Microsoft Excel 2019 and SPSS 27.0 software were used for data analysis, and Origin 26.0 software was used for graphing.

3. Results

3.1. Transcriptome Sequencing and Assembly from De Novo

The transcriptome analysis of floral organs was conducted across various copper concentrations designated as Cu0 (0 mg·L⁻¹), Cu10 (10 mg·L⁻¹), and Cu20 (20 mg·L⁻¹). The sequencing process generated a substantial dataset, comprising 169,755,424 raw reads and 25.63 GB of raw data. After rigorous quality control measures, a refined dataset of 24.86 GB of high-quality clean data was obtained, with approximately 56 million clean reads per sample. Notably, the Q30 base quality score ranged from 94.5 to 94.58%, indicating excellent sequence quality, and the GC content exceeded 44.55%. Furthermore, the mapping ratio of all three samples exceeded 83.04%, as detailed in Table 2, demonstrating the reliability and accuracy of our transcriptome analysis.

Table 2. Summary of sequencing data of *I. uliginosa* transcriptome.

Attributes	Cu0	Cu10	Cu20
Raw reads	55,657,964	58,102,838	55,994,622
Clean reads	55,287,682	57,653,418	55,597,734
Clean bases	8,188,113,484	8,477,755,077	8,199,632,407
Q30 (%)	94.57	94.58	94.5
GC content (%)	44.55	44.69	44.82
Total mapped	22,954,371 (83.04%)	24,066,941 (83.49%)	23,599,843 (84.89%)

A de novo assembly approach was applied to the high-quality reads, resulting in the generation of 70,319 transcripts, encompassing a total of 84.6 Mb of sequence data after rigorous optimization and filtering processes. The transcripts exhibited a diverse range of lengths, with the longest transcript measuring 15,701 bp and the shortest 201 bp. The average transcript length was 1203.37 bp, and the N50 value was 1810 bp, indicating a

high-quality assembly. Furthermore, a total of 40,610 unigenes were identified from the assembled transcripts, encompassing 44.2 Mb of sequence data. The length distribution of these unigenes was similar to that of the transcripts, with the maximum and minimum lengths matching those of the transcripts. The average unigene length was 1088.24 bp, and the N50 value was 1812 bp, reflecting the consistency and reliability of the assembly process (Table 3). In terms of unigene size distribution, 83% of the unigenes fell within the range of 200 to 2000 bp, while 15% ranged from 2000 to 4000 bp. Notably, 827 unigenes exceeded 4000 bp in length, representing approximately 2% of the total unigenes identified. This comprehensive length distribution analysis provides valuable insights into the transcriptome complexity and gene expression patterns of the studied floral organs (Figure S1).

Table 3. Statistics of transcriptome assembly.

Attributes	Unigenes	Transcripts
Total number	40,610	70,319
Total base	44,193,278	84,620,119
Largest length (bp)	15,701	15,701
Smallest length (bp)	201	201
Average length (bp)	1088.24	1203.37
N50 length (bp)	1812	1810

To annotate the transcriptome, BLAST searches were conducted to align the assembled sequences with various databases, including NR (NCBI Non-Redundant Protein Sequence Database), Swiss-Prot, Pfam (Protein families), COG (Clusters of Orthologous Groups of proteins), GO (Gene Ontology), and KEGG (Kyoto Encyclopedia of Genes and Genomes). This comprehensive analysis allowed us to obtain valuable annotation information for the transcriptome, enhancing our understanding of the gene expression patterns and functional characteristics of the studied floral organs. Of all the 40,610 unigenes, 98.37% (39,949 unigenes) were successfully annotated in the six databases (Table 4). *Camellia sinensis* (5711, 23.49%), *Actinidia chinensis* (4171, 17.15%), *Nyssa sinensis* (3332, 13.70%), *Vitis vinifera* (631, 2.6%), and *Rhododendron williamsianum* (538, 2.21%) were the top five species that showed similarity with the unigenes of *I. uliginosa* (Figure 2). In total, 19,465 and 19,695 unigenes were assigned to Swiss-Prot and Pfam databases, respectively.

Table 4. Functional annotation of *I. uliginosa* unigenes.

	Unigene Number	Unigene Percent
GO	20,926	52.38%
KEGG	10,945	27.40%
COG	22,343	55.93%
NR	24,238	60.67%
Swiss-Prot	19,456	48.70%
Pfam	19,695	49.30%
Total	39,949	100%

A total of 22,343 unigenes were categorized into 23 distinct COG functional groups. Notably, 12,437 of these unigenes exhibited ambiguous characteristics and were associated with unknown functional roles. Among the effectively annotated unigenes, the largest category was devoted to 'Transcription', closely followed by 'Posttranslational modification, protein turnover, chaperones', and 'Signal transduction mechanisms' (Figure 3).

A total of 20,926 unigenes were assigned to at least one GO term, providing a comprehensive overview of their functional roles. Within the biological process (BP) category, 'Cellular process' and 'Metabolic process' emerged as the most prevalent subcategories. In terms of cellular component (CC), a significant proportion of unigenes were grouped under 'Cell part'. Finally, within the molecular function (MF) category, 'Binding' and 'Catalytic activity' stood out as the most significant subcategories (Figure 4).

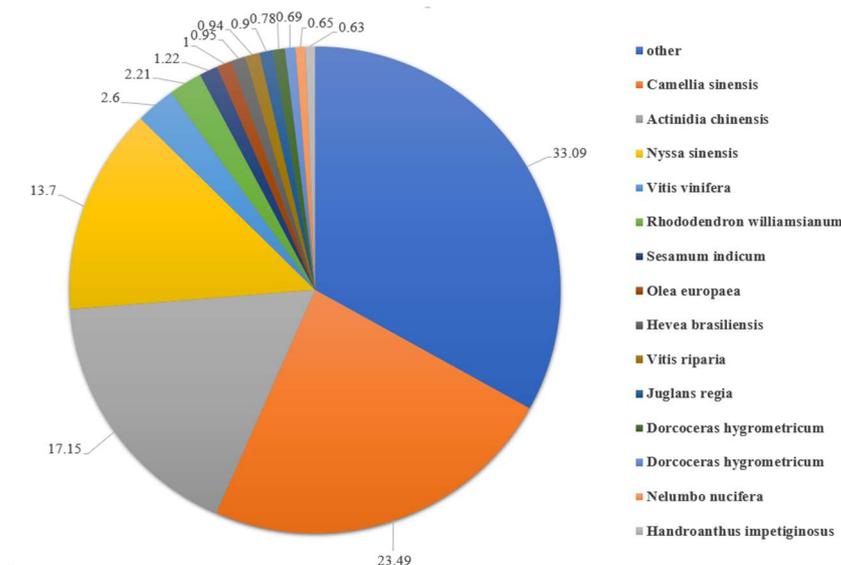


Figure 2. Distribution of NR-annotated species.

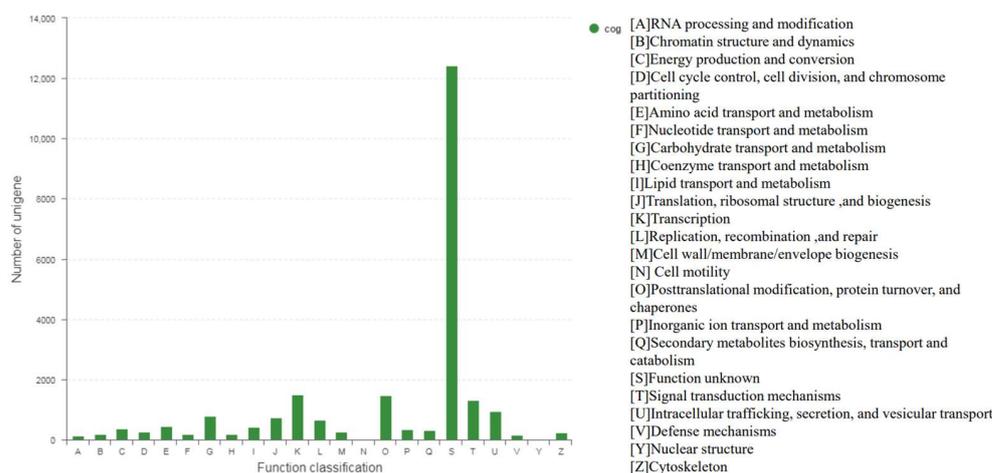


Figure 3. COG classification of unigenes in *I. uliginosa*.

A total of 10,945 unigenes were successfully annotated to KEGG and categorized into 20 distinct pathways across six primary categories. Notably, the majority of these genes were associated with the categories of ‘Metabolism’ and ‘Genetic Information Processing’. Among the various pathways, ‘Translation’, ‘Carbohydrate Metabolism’, and ‘Folding, Sorting, and Degradation’ emerged as the most prominent representatives. This comprehensive analysis, as depicted in Figure 5, offers valuable insights into the function of the transcriptome and highlights the key biological processes and metabolic pathways involved.

In order to explore the genes that may be involved in regulating the flower color changes in *I. uliginosa* under copper stress, the differential expression at different copper concentrations was analyzed. The results showed that a total of 15,610 genes were differentially expressed among three groups. There were 4974 differentially expressed genes between the Cu0 group and the Cu10 group, including 3290 up-regulated genes and 1684 down-regulated genes; there were 4017 differentially expressed genes between the Cu0 group and the Cu20 group, including 1699 up-regulated genes and 2318 down-regulated genes. There were 6619 differentially expressed genes between the Cu10 group and the Cu20 group, including 1846 up-regulated genes and 4773 down-regulated genes (Table 5). There were 760 DEGs shared among these three groups (Figure 6).

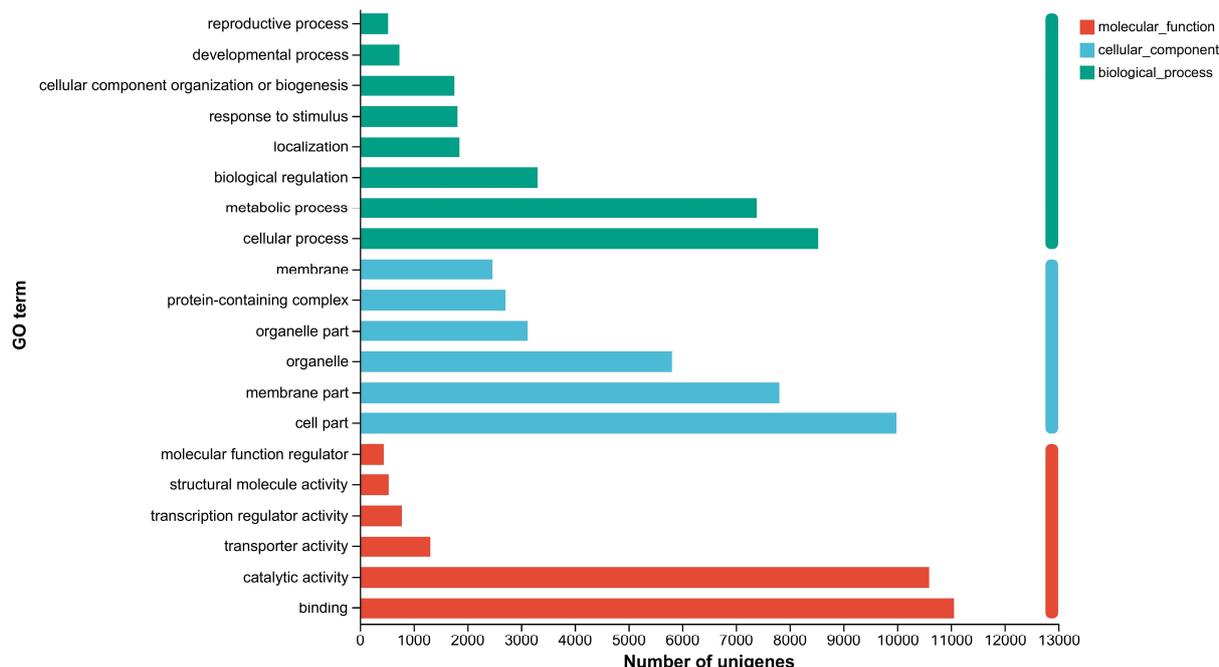


Figure 4. Main GO categories of unigenes in *I. uliginosa* transcriptome.

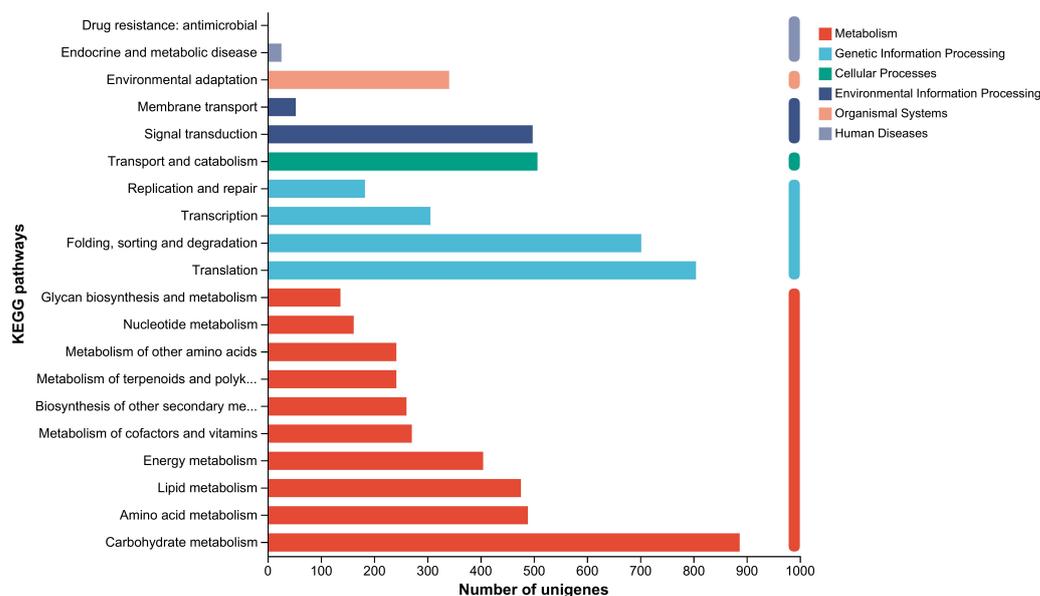


Figure 5. KEGG metabolic pathway of unigenes in *I. uliginosa*.

GO classification was performed on the differentially expressed genes of *I. uliginosa* under different concentrations of copper treatment. The GO classification results showed that there was a total of 20 secondary function entries. In BP, the number of genes annotated to secondary functional items such as cellular process, metabolic process, and biological regulation was relatively large, 197, 162, and 68, respectively. In MF, the number of genes classified into the catalytic activity and binding secondary function entries was relatively large, 303 and 269, respectively. The most significantly enriched Cu0 vs. Cu10 vs. Cu20 are clathrin coat assembly (GO:0048268), 1-phosphatidylinositol binding (GO:0005545), and clathrin-coated vesicle (GO:0030136) (Figure S1).

KEGG classification was performed on the differentially expressed genes of *I. uliginosa* under different concentrations of copper treatment. The annotated DEGs were distributed

in 62 metabolic pathways and belong to five categories. Among them, the metabolism category contains the largest number of genes and the largest number of secondary pathways. This category includes Carbohydrate metabolism, Lipid metabolism, Energy metabolism and biosynthesis of other secondary metabolites, with the number of genes being 38, 12, 10, and 10, respectively. KEGG enrichment pathway analysis was performed on the differentially expressed genes of *I. uliginosa* under different concentrations of copper treatment. A total of six significant metabolic pathways were enriched, among which the most significantly enriched were Pentose and glucuronate interconversions (map00040), nitrogen metabolism (map00910) and Phenylpropanoid biosynthesis (map00940), with 22, 6, and 10, respectively (Figure S1).

Table 5. Statistical table of the number of differential genes.

Diff_Group	Total DEG	Up	Down
Cu0 vs. Cu10	4974	3290	1684
Cu0 vs. Cu20	4017	1699	2318
Cu10 vs. Cu20	6619	1846	4773

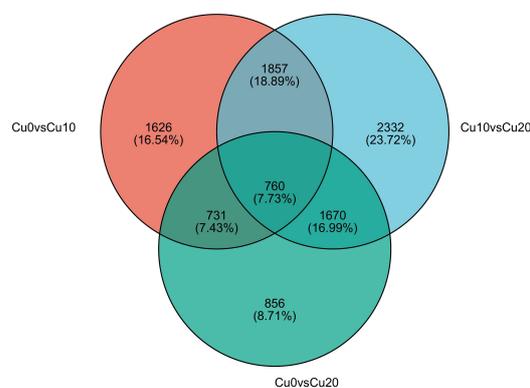


Figure 6. Venn diagram of DEGs.

In order to further study the differentially expressed genes related to the flower color of *I. uliginosa* under different concentrations of copper treatment, the flavonoid biosynthetic pathway related to flower color regulation was analyzed, the DEGs were compared with the metabolic pathways, and the metabolic pathway diagram of DEGs was obtained (Figure S2).

In order to further analyze the role of structural genes in anthocyanin accumulation under different concentration treatments, we can see from Figure S3 that a total of 359 *ANS* genes, 138 *CHI* genes, 4 *PAL* genes, and 3 *GT* genes were screened. From Figure S3, we can see that only the expression pattern of *GT* genes in Cu0 vs. Cu10 vs. Cu20 is down-regulated; that is, as the copper concentration increases, the expression level gradually decreases. The expression level was the highest when the copper concentration was 0 mg·L⁻¹, and the expression level was low at high concentrations (20 mg·L⁻¹), which is consistent with the flower color phenotype and anthocyanin content under different concentrations of copper treatment.

Based on GO and KEGG functional classification and enrichment analysis of differentially expressed genes, factors related to flower color transcription were further analyzed. From Figure S3, we can see that a total of nine MYB genes and three WD40 protein genes were screened from DEGs. From Figure S3, we can see that only the expression pattern of the MYB (TRINITY_DN7677_c0_g2) gene in Cu0 vs. Cu10 vs. Cu20 is down-regulated; that is, as the copper concentration increases, the expression level gradually decreases. The expression level was the highest when the copper concentration was 0 mg·L⁻¹, and the expression level was low at high concentrations (20 mg·L⁻¹), which is consistent with the flower color phenotype and anthocyanin content under different concentrations of copper treatment.

In order to further analyze the role of transport genes in anthocyanin accumulation under different concentration treatments, we can see from Figure S3 that a total of four *ABC* genes and three *GST* genes were screened. As can be seen from Figure S3, only the *ABC* (TRINITY_DN14115_c0_g1) gene and the *GST* (TRINITY_DN5047_c0_g1) gene have a down-regulated expression pattern in Cu0 vs. Cu10 vs. Cu20; that is, as the copper concentration increases, the expression level gradually decreases. The expression level was the highest when the copper concentration was $0 \text{ mg}\cdot\text{L}^{-1}$, and the expression level was low at high concentrations ($20 \text{ mg}\cdot\text{L}^{-1}$), which is consistent with the flower color phenotype and anthocyanin content under different concentrations of copper treatment.

3.2. Physical and Chemical Properties and Physiological Analysis of Candidate Genes

The *IuGT* gene sequence of *I. uliginosa* has a length of 1482 bp, encoding 493 amino acids with a molecular formula of $\text{C}_{2469}\text{H}_{3808}\text{N}_{654}\text{O}_{718}\text{S}_{21}$. The theoretical isoelectric point PI value is 5.41; the instability index is 44.99, indicating that the protein is an unstable protein. The total average hydrophilicity index is -0.151 , indicating that the protein belongs to hydrophilic protein. The total length of the cDNA of the *IuGST* gene in *I. uliginosa* is 696 bp, encoding 231 amino acids with the molecular formula $\text{C}_{1202}\text{H}_{1859}\text{N}_{293}\text{O}_{335}\text{S}_6$. The theoretical isoelectric point PI value is 5.42, and the instability index is 40.86. Therefore, it is speculated that the protein is an unstable protein with a total average hydrophilicity index of 0.015, indicating that it belongs to hydrophobic proteins (Table 6). To explore the evolutionary relationships of candidate genes, a phylogenetic tree was constructed using the amino acid sequences of *IuGT* and *IuGST*, respectively. As shown in the figure, in the evolutionary tree, *IuGT* and *Vitis riparia* (XP:034674887.1) are clustered together with the closest genetic relationship; *IuGST* is closely related to *Jatropha curcas* (KDP21251.1) (Figure 7).

Table 6. The basic physical and chemical properties of *IuGT* and *IuGST*.

Basic Physicochemical Properties	<i>IuGT</i>	<i>IuGST</i>
Molecular formula	C2469H3808N654O718S21	C1202H1859N293O335S6
Theoretical isoelectric point	5.41	5.42
Molecular weight	54,814.61 Da	25,967.10 Da
Total number of atoms	7670	3695
Negatively charged residue (Asp + Glu)	60	31
Positively charged residue (Arg + Lys)	44	25
Instability index	44.99	40.86
Total average hydrophilicity index	-0.151	0.015

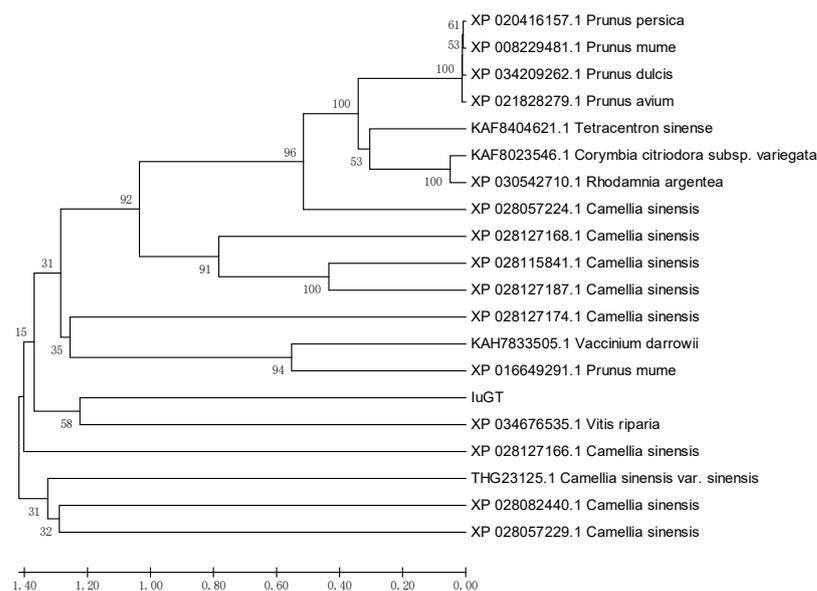


Figure 7. Cont.

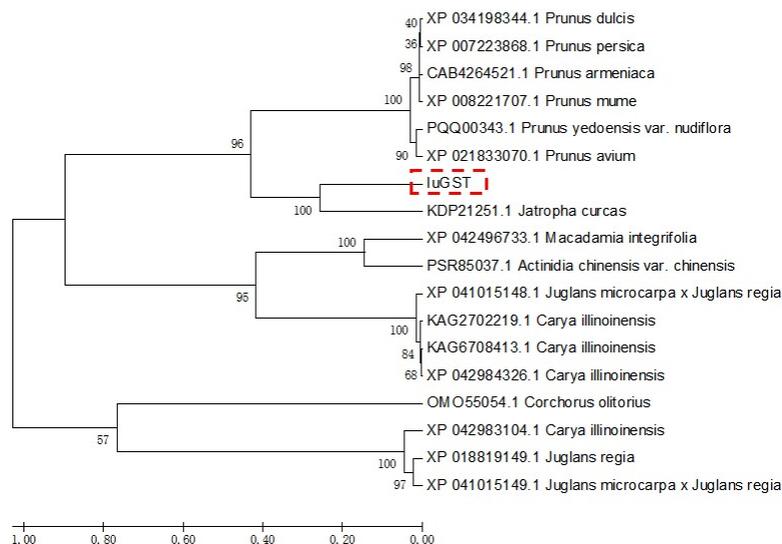


Figure 7. Phylogenetic tree of the amino acid sequence of IuGT and IuGST.

3.3. qRT-PCR Validation of the Candidate Genes

Two candidate genes were selected for qRT-PCR to verify the accuracy of transcriptome data (Figure 8). The expression levels of *IuGT* and *IuGST* genes in *I. uliginosa* were significantly different under different concentrations of copper. As the copper concentration increased, the expression levels of *IuGT* and *IuGST* genes showed a gradually decreasing trend. When the copper concentration was $0 \text{ mg}\cdot\text{L}^{-1}$, the expression level was the highest, and when the copper concentration was $20 \text{ mg}\cdot\text{L}^{-1}$, the expression level was the lowest. The expression levels of *IuGT* and *IuGST* genes had decreased by more than 10 times.

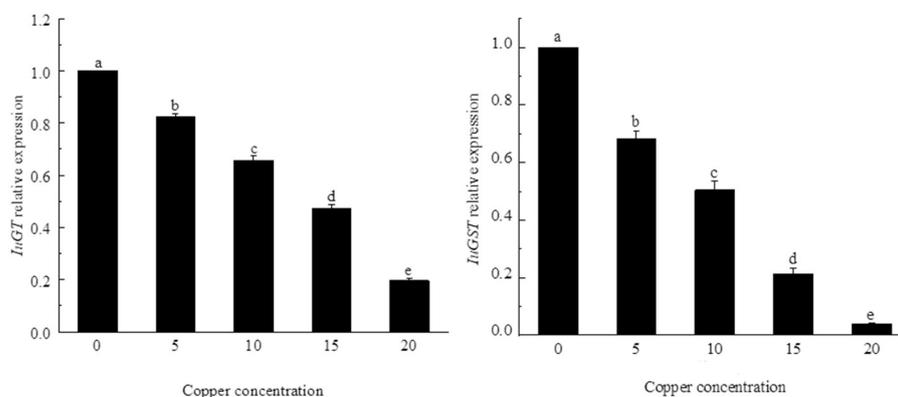


Figure 8. The relative expression levels of *IuGT* and *IuGST* genes under different copper concentrations. Different letters represent statistically significant differences ($\alpha = 0.05$).

4. Discussion

Different concentrations of different metal ions have different effects on plants. For example, Li [18] found that the *Impatiens hawkeri* color of red was affected by the concentration of Cu^{2+} and Al^{3+} . When the Cu^{2+} concentration was $3.2 \times 10^{-7} \text{ mol}\cdot\text{L}^{-1}$, the flowers were the darkest in color; as the Al^{3+} concentration increased, the flower color first became lighter and then gradually deepened [21]. In previous studies, we measured the chromaticity value, anthocyanins, total flavonoids, carotenoids, copper ion content in petals, and related physiological and biochemical indicators of *I. uliginosa* under copper stress. The color of the *I. uliginosa* organs changed under different concentrations of copper treatment. The flower color gradually became lighter as the copper concentration increased. It was also found that as the copper concentration increased, the color of *I. uliginosa* petals

became lighter, and the total flavonoid content and anthocyanin content showed a downward trend [22,23]. When the copper concentration was $20 \text{ mg}\cdot\text{L}^{-1}$, the copper ion content of *I. uliginosa* was $2363.90 \text{ mg}\cdot\text{kg}^{-1}$, which was 39 times that of CK.

This study used copper-treated ($0 \text{ mg}\cdot\text{L}^{-1}$, $10 \text{ mg}\cdot\text{L}^{-1}$, and $20 \text{ mg}\cdot\text{L}^{-1}$) flower organs of *I. uliginosa* during the blooming period as research materials and conducted transcriptome sequencing so as to explore the molecular mechanism of flower color fading as copper concentration increases. After sequencing, a total of 40,610 unigenes were obtained. In order to obtain the functional information of *I. uliginosa* unigene, these genes were annotated on NR, NT, Swissprot, KEGG, KOG, Pfam, and GO databases, and a total of 39,949 unigenes were assembled. Analysis of differentially expressed genes found that 1626, 856, and 2332 genes only existed in Cu0 vs. Cu10, Cu0 vs. Cu20, and Cu10 vs. Cu20, respectively. Meanwhile, only 760 differentially expressed genes appeared in all comparison groups. This study analyzed differentially expressed genes and screened out structural genes that play a key role in the accumulation of flavonoids in *I. uliginosa*, including *ANS*, *CHI*, *PAL*, and *GT*. Among them, the expression pattern of *IuGT* (TRINITY_DN3511_c0_g1) is consistent with the flower color phenotype and anthocyanin content under different concentrations of copper, namely, the transcription level of the *IuGT* gradually decreases as the copper concentration increases. *GTs* have been identified in many plants, such as *Fagopyrum esculentum* [24], *Rosa rugosa* [25], and *Arabidopsis thaliana* [26]. The above confirmed that glycosylation is very important for flower color formation. In addition, phytochrome transport factors have also become important factors affecting plant color, such as the GST protein family. Studies have shown that the glutathione transferase (*GST*) family is the key factor of the accumulation of anthocyanins in the vacuole, and it can directly bind to anthocyanins and thus become a carrier for anthocyanin transport [8]. A total of three *GST* genes were screened out in this study. The expression pattern of *IuGST* (TRINITY_DN5047_c0_g1) in Cu0 vs. Cu10 vs. Cu20 is down-regulated, namely, as the copper concentration increases, the expression level gradually decreases. The expression level is the highest when the copper concentration is $0 \text{ mg}\cdot\text{L}^{-1}$, and the expression level is low at high concentrations ($20 \text{ mg}\cdot\text{L}^{-1}$), which is consistent with the flower color phenotype and anthocyanin content under different concentrations of copper treatment.

Different genes directly affect the direction and metabolic process of flavonoids according to their different expression levels, producing different metabolites, which in turn affect the color of plants. Hassani et al. [27] studied the regulation of peach flower color and discovered that the *GT* gene expression level was the highest in red peach flowers and almost not expressed in white flowers. This study demonstrated that as the copper concentration increased, the expression level of *IuGT* in the organs of *I. uliginosa* showed a gradually decreasing trend. When the copper concentration was $20 \text{ mg}\cdot\text{L}^{-1}$, the expression level was the lowest, and the flower color showed a change from pink to white. Studies have shown that *GST* family genes are involved in responses to stress in many plants. Tasaki et al. [28] found that *GST1* is not only a gene responsible for anthocyanin transport in Japanese gentian under sugar-induced stress conditions but also essential for the accumulation of gentsione in flowers. This study found that as the copper concentration increased, the expression of *IuGST* in the organs of *I. uliginosa* showed a gradually decreasing trend. When the copper concentration was $20 \text{ mg}\cdot\text{L}^{-1}$, the expression level was the lowest. Meanwhile, as the copper concentration increased, the anthocyanin content continued to decrease, which was consistent with the expression pattern of *IuGST*.

5. Conclusions

This study used *I. uliginosa* as the material and treated it with different concentrations of copper ($0 \text{ mg}\cdot\text{L}^{-1}$, $5 \text{ mg}\cdot\text{L}^{-1}$, $10 \text{ mg}\cdot\text{L}^{-1}$, $15 \text{ mg}\cdot\text{L}^{-1}$, and $20 \text{ mg}\cdot\text{L}^{-1}$). After 30 days, the petals of *I. uliginosa* were taken to explore the effect of different concentrations of copper on *I. uliginosa* flower color from a molecular biology perspective, revealing the mechanism of the effect of different concentrations of copper on Dianshui Jinfeng flower color. It lays a theoretical foundation for future in-depth study of the molecular mechanism of *I. uliginosa*

in defense against heavy metal toxicity, and also provides a theoretical basis for the practical application of *I. uliginosa* in restoring heavy metal-polluted water bodies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/horticulturae10040412/s1>, Figure S1. (a) Length distribution of unigenes; (b) Bubble chart of GO enrichment of differential genes; (c) KEGG Pathway enrichment of differentially expressed genes. Figure S2. Key metabolic pathways involved in color changes of *I. uliginosa* under copper stress. Figure S3. (a) Venn diagram of structural genes; (b) Differential expression of GT Unigenes; (c) Venn diagram of transcription factors; (d) Differential expression of MYB Unigenes; (e) Venn diagram of transport factors; (f) Differential expression of ABC and GST Unigenes.

Author Contributions: Y.T. and H.H. (Haiquan Huang) were responsible for the experimental design. Y.T. and X.Z. carried out sample collection, experiments, data analysis, and article writing. Q.L., X.L., L.L., H.H. (Haihao He) and G.L. participated in the experiment. M.H. supervised the research and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Joint Special Key Project for Agriculture Basic Research in Yunnan Province (202101BD070001-018), the National Natural Science Foundation of China (32060364, 32060366), and the Doctoral Supervisor Team Project of Landscape Plants Genetic Improvement and Efficient Breeding in Yunnan Province.

Data Availability Statement: All data generated in this article are included within the article and its additional files.

Acknowledgments: We thank Shanghai Majorbio Bio-pharm Technology Co., Ltd. (Shanghai, China) for its help in sequencing. The data were analyzed through the free online platform of Majorbio Cloud Platform (www.majorbio.com, accessed on 15 February 2022).

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

References

1. Li, B.; Dao, J.R.; Zhu, R.Y.; He, H.; Meng, X.Q.; Han, F.X. Distribution, accumulation and risk assessment of heavy metal pollution in Dianchi Lake. *Environ. Chem.* **2021**, *40*, 1808–1818.
2. Zhang, G.H.; Xie, Q.; Yan, K.; Li, Z.Y.; Su, T.; Yang, Y.L. Pollution Characteristics and Ecological Risk Assessment of Heavy Metals in Sediments of Dianchi Outer Lake During the ‘13th Five-Year Plan Period’. *J. Soil Water Conserv.* **2023**, *37*, 240–247.
3. Peng, G.Y.; Hu, L.; Huang, C.; Yang, K.; Wan, W.; Huang, C.G. Transcriptome Analysis of Response to Heavy Metal Copper Stress in *Setcreasea purpurea* Root Tissue. *Biotechnol. Bull.* **2022**, *38*, 83–94.
4. Zhang, C.; Tang, C.C.; Wang, J.Y.; Guo, L.; Wang, L.; Li, W. Advances on transcriptome of plants under stresses. *J. Biol.* **2017**, *34*, 86–90.
5. García-Vico, L.; Sánchez, R.; Fernández, G.; Sanz, C.; Pérez, A.G. Study of the olive β -glucosidase gene family putatively involved in the synthesis of phenolic compounds of virgin olive oil. *J. Sci. Food Agric.* **2021**, *101*, 5409–5418. [[CrossRef](#)]
6. Gachon, C.M.; Langlois-Meurinne, M.; Saindrenan, P. Plant secondary metabolism glycosyl transferases: The emerging functional analysis. *Trends Plant Sci.* **2005**, *10*, 542–549. [[CrossRef](#)]
7. Li, P.; Li, Y.J.; Zhang, F.J.; Zhang, G.Z.; Jiang, X.Y.; Yu, H.M.; Hou, B.K. The Arabidopsis UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance via modulating anthocyanin accumulation. *Plant J.* **2017**, *89*, 85–103. [[CrossRef](#)]
8. Hu, Z.; He, J.; Chen, K.; Wang, Z.; Liu, J.; Qiao, X.; Ye, M. Molecular cloning and biochemical characterization of a new flavonoid glycosyltransferase from the aquatic plant lotus. *Biochem. Biophys. Res. Commun.* **2019**, *510*, 315–321. [[CrossRef](#)]
9. Cheng, J.; Wei, G.; Zhou, H.; Gu, C.; Vimolmangkang, S.; Liao, L.; Han, Y. Unraveling the mechanism underlying the glycosylation and methylation of anthocyanins in peach. *Plant Physiol.* **2014**, *166*, 1044–1058. [[CrossRef](#)]
10. Young, L.; Akhoy, L.; Kulkarni, M.; You, F.; Booker, H. Fine-mapping of a putative glutathione S-transferase (GST) gene responsible for yellow seed colour in flax (*Linum usitatissimum*). *BMC Res. Notes* **2022**, *15*, 72. [[CrossRef](#)]
11. Lu, S.W.; Zheng, X.A.; Wang, J.Y.; Fang, J.G. Research Progress on the Metabolism of Flavonoids in Grape. *Acta Hort. Sin.* **2021**, *48*, 2506–2524.
12. Kim, S.H.; Sung, S.Y.; Kim, Y.S.; Jo, Y.D.; Kang, S.Y.; Kim, J.B.; Ahn, J.W.; Ha, B.K.; Kim, D.S. Isolation and characterization of differentially expressed genes in petals of Chrysanthemum mutant cultivars developed by irradiation. *Sci. Hortic.* **2015**, *189*, 132–138. [[CrossRef](#)]
13. Cao, Y.; Xu, L.; Xu, H.; Yang, P.; He, G.; Tang, Y.; Qi, X.; Song, M.; Ming, J. LhGST is an anthocyanin-related glutathione S-transferase gene in Asiatic hybrid lilies (*Lilium* spp.). *Plant Cell Rep.* **2021**, *40*, 85–95. [[CrossRef](#)]

14. Jin, X.H.; Hong, Y.; Huang, H.; Dai, S.L.; Zhu, Y. Isolation and Expression Analysis of GST Gene Encoding Glutathione S-transferase from *Senecio cruentus*. *Acta Hortic. Sin.* **2013**, *40*, 1129–1138.
15. Yu, S.X. *Balsaminaceae of China*; Peking University Press: Beijing, China, 2012; p. 147.
16. Li, Y.; Huang, W.L.; Lin, Y.F.; Luo, C.; Li, X.Y.; Liu, Y.L.; Zhu, J.P.; Feng, Z.X.; Huang, M.J.; Huang, H.Q. Cloning and Expression Analysis of CHI Genes in *Impatiens uliginosa*. *Acta Agric. Univ. Jiangxiensis* **2020**, *42*, 468–474.
17. Liu, Y.L.; Feng, Z.X.; Zhu, J.P.; Luo, C.; Liang, Y.; Li, Y.; Tong, Z.K.; Huang, H.Q.; Huang, M.J. Cloning and Expression Analysis of AP3/DEF Homologous Genes Associated with Floral Development in *Impatiens uliginosa*. *Mol. Plant Breed.* **2020**, *18*, 6626–6632.
18. Guo, J.W.; Huang, Q.; Ma, M.L.; Sun, Z.H.; Wen, Y.H.; Huang, H.Q.; Huang, M.J.; Wang, Q. Determination of Metal Elements in the Petals of Yunnan Wild *Impatiens* in Different Colors. *J. Anhui Agric. Sci.* **2019**, *47*, 179–181.
19. 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](#)]
20. Li, Y.; Huang, W.L.; Li, X.Y.; Zhang, Y.D.; Meng, D.C.; Wei, C.M.; Huang, M.J.; Huang, H.Q. The cellular and molecular basis of the spur development in *Impatiens uliginosa*. *Hortic. Res.* **2024**, *11*, uhae015. [[CrossRef](#)]
21. Li, R.H. Effects of Copper and Aluminum on Growth and Flower Color of *Impatiens hawkeri*. Master's Thesis, Hebei Agricultural University, Baoding, China, 2005.
22. Li, Q.M.; Li, W.X.; Cao, M.H.; Liu, S.; Zhang, T.Y.; Wang, Q.; Huang, M.J.; Huang, H.Q. Physiological and Biochemical Correlations between Color of *Impatiens uliginosa* Flower and Nutrient Supply on Copper. *Fujian J. Agric. Sci.* **2021**, *36*, 1323–1329.
23. Li, Q.M.; Li, W.X.; Li, X.Y.; Li, Y.; Qu, S.P.; Huang, M.J.; Huang, H.Q. Effects of Copper Stress on Flower Development and Physiological and Biochemical Characteristics of *Impatiens uliginosa*. *Shandong Agric. Sci.* **2022**, *54*, 74–79.
24. Matsui, K.; Tomatsu, T.; Kinouchi, S.; Suzuki, T.; Sato, T. Identification of a gene encoding glutathione S-transferase that is related to anthocyanin accumulation in buckwheat (*Fagopyrum esculentum*). *J. Plant Physiol.* **2018**, *231*, 291–296. [[CrossRef](#)] [[PubMed](#)]
25. Sui, X.; Zhao, M.; Han, X.; Zhao, L.; Xu, Z. RrGT1, a key gene associated with anthocyanin biosynthesis, was isolated from *Rosa rugosa* and identified via overexpression and VIGS. *Plant Physiol. Biochem.* **2019**, *135*, 19–29. [[CrossRef](#)]
26. Yonekura-Sakakibara, K.; Fukushima, A.; Nakabayashi, R.; Hanada, K.; Matsuda, F.; Sugawara, S.; Inoue, E.; Kuromori, T.; Ito, T.; Shinozaki, K.; et al. Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. *Plant J.* **2012**, *69*, 154–167. [[CrossRef](#)] [[PubMed](#)]
27. Hassani, D.; Liu, H.L.; Chen, Y.N.; Wan, Z.B.; Zhuge, Q.; Li, S.X. Analysis of biochemical compounds and differentially expressed genes of the anthocyanin biosynthetic pathway in variegated peach flowers. *Genet. Mol. Res.* **2015**, *14*, 13425–13436. [[CrossRef](#)]
28. Tasaki, K.; Yoshida, M.; Nakajima, M.; Higuchi, A.; Watanabe, A.; Nishihara, M. Molecular characterization of an anthocyanin-related glutathione S-transferase gene in Japanese gentian with the CRISPR/Cas9 system. *BMC Plant Biol.* **2020**, *20*, 370. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.