metabolites

Special Issue Reprint

Plant Metabolic Genetic Engineering

Edited by
Yanjie Zhang and Yan Li

mdpi.com/journal/metabolites

MDPI

Plant Metabolic Genetic Engineering

Plant Metabolic Genetic Engineering

Guest Editors

Yanjie Zhang
Yan Li

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors
Yanjie Zhang
School of Agricultural Sciences
Zhengzhou University
Zhengzhou
China

Yan Li
The Center of Advanced Analysis and Gene Sequencing
Zhengzhou University
Zhengzhou
China

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Metabolites* (ISSN 2218-1989), freely accessible at: https://www.mdpi.com/journal/metabolites/special_issues/QCSWLS4HOX.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-3935-3 (Hbk)
ISBN 978-3-7258-3936-0 (PDF)
https://doi.org/10.3390/books978-3-7258-3936-0

Contents

metabolites

MDPI

Article

SmMYB4 Is a R2R3-MYB Transcriptional Repressor Regulating the Biosynthesis of Phenolic Acids and Tanshinones in *Salvia miltiorrhiza*

Qian Tian [1], Limin Han [2], Xiaoya Zhu [3], Caijuan Zhang [1], Yunyun Li [1], Xiaoshan Xue [1], Yueyue Wang [1], Donghao Wang [1], Junfeng Niu [1], Wenping Hua [2], Bin Li [1,4,*] and Zhezhi Wang [1,*]

[1] Key Laboratory of the Ministry of Education for Medicinal Resources and Natural Pharmaceutical Chemistry, National Engineering Laboratory for Resource Development of Endangered Crude Drugs in Northwest of China, Shaanxi Normal University, Xi'an 710062, China
[2] College of Life Science and Food Engineering, Shaanxi Xueqian Normal University, Xi'an 710062, China
[3] Daan Gene Co., Ltd., Guangzhou 510665, China
[4] Institute of Botany of Shaanxi Province, Xi'an Botanical Garden of Shaanxi Province, Xi'an 710061, China
* Correspondence: libin1989@snnu.edu.cn (B.L.); zzwang@snnu.edu.cn (Z.W.)

Abstract: *Salvia miltiorrhiza* Bunge is one of the most famous traditional Chinese medicinal plants. The two most important classes of pharmaceutically relevant compounds in *S. miltiorrhiza* are phenolic acids and tanshinones. The MYB family of transcription factors may efficiently regulate the secondary metabolism in plants. In this study, a subgroup 4 R2R3MYB transcription factor gene, *SmMYB4*, was isolated from *S. miltiorrhiza* and functionally characterized using overexpression and a RNAi-mediated silencing. We achieved a total of six overexpressions and eight RNAi transgenic lines from the *Agrobacterium* leaf disc method. The content of the total phenolics, rosmarinic acid, and salvianolic acid B markedly decreased in the *SmMYB4*-overexpressing lines but increased in the *SmMYB4*-RNAi lines. The content of the total tanshinones, cryptotanshinone, and tanshinone IIA decreased in the *SmMYB4*-overexpressing transgenic lines but increased in the *SmMYB4*-RNAi lines. A gene expression analysis demonstrated that SmMYB4 negatively regulated the transcription of the critical enzyme genes involved in the phenolic acid and tanshinone biosynthesis. The genetic control of this transcriptional repressor may be used to improve the content of these bioactive compounds in the cultivated *S. miltiorrhiza*.

Keywords: *Salvia miltiorrhiza* Bunge; SmMYB4; transcription factor; phenolic acids; tanshinones; transgenic plants

Citation: Tian, Q.; Han, L.; Zhu, X.; Zhang, C.; Li, Y.; Xue, X.; Wang, Y.; Wang, D.; Niu, J.; Hua, W.; et al. SmMYB4 Is a R2R3-MYB Transcriptional Repressor Regulating the Biosynthesis of Phenolic Acids and Tanshinones in *Salvia miltiorrhiza*. *Metabolites* **2022**, *12*, 968. https://doi.org/10.3390/metabo12100968

Academic Editors: Yanjie Zhang and Yan Li

Received: 8 September 2022
Accepted: 9 October 2022
Published: 12 October 2022

1. Introduction

Salvia miltiorrhiza Bunge, called 'Dan-Shen' in Chinese, is one of the most widely used traditional herbal medicines for various cardiovascular and cerebrovascular diseases. Water-soluble phenolic acids and lipid-soluble tanshinones are the two major classes of pharmaceutical ingredients in *S. miltiorrhiza* [1]. With an increasing market demand for *S. miltiorrhiza*, the wild resources are rapidly decreasing and becoming extremely scarce. However, the quality of cultivated *S. miltiorrhiza* is deteriorating. Therefore, improving the content of the bioactive components in cultivated *S. miltiorrhiza* using modern plant genetic engineering and metabolic engineering technology is important. The biosynthesis, accumulation, and regulation of these active ingredients have become intensively attractive research areas.

The transcription factors (TFs) can bind the cis-elements of a target gene's promoter region to activate or inhibit its transcription. They offer an effective strategy for enhancing the production of pharmaceutically active metabolites in medicinal plants [2]. Following the publication of the *S. miltiorrhiza* genome [3], many transcription factor gene families,

such as MYB, bHLH, WRKY, ERF, and bZIP have been identified and investigated [4–8]. Their important roles in biosynthesis or bioactive ingredient accumulation have been a research focus in recent years. *SmbHLH10*, *SmWRKY1*, and *SmWRKY2* could upregulate the tanshinone biosynthesis [9–11]. *SmbHLH51* could positively regulate the biosynthesis of phenolic acids [12]. *SmbHLH148*, *SmbHLH3*, *SmERF115*, and *SmGRAS1/2* not only regulated the biosynthesis of the phenolic acid but also the tanshinone biosynthesis [13–16].

The MYB transcription factor is one of the largest members of the TF families in plants. Based on the number of imperfect repeats in the DNA-binding domain, the MYB transcription factors are classified into four subfamilies known as 1RMYB, R2R3-MYB, R1R2R3-MYB, and R1R2R3R1/2-MYB. Most MYB proteins in plants are mainly R2R3-MYBs. Various studies have reported that they play critical roles in plant hormone and stress responses, secondary metabolic regulation, organ morphogenesis, and nutrient absorption [4]. *IbMYB116* from the sweet potato enhances drought tolerance in the transgenic *Arabidopsis thaliana* [17]. *MsMYB4* significantly increased the salinity tolerance of alfalfa [18]. The overexpression of *PbrMYB5* in tobacco confers an enhanced tolerance to chilling stresses, whereas the down-regulation of it in the *P. betulaefolia* by a virus-induced gene silencing, results in an elevated chilling sensitivity [19]. In regulating the secondary metabolism, the R2R3-MYB transcription factors have an essential function in producing flavonoid metabolites. MYB1 and AcMYB10 regulate the biosynthesis of anthocyanin in onion and kiwifruit, respectively [20,21]. VrMybA1 and VrMYBCS1 positively correlate with the anthocyanin accumulation in the berry skin [22]. NnMYB5 is a transcription activator of the anthocyanin synthesis contributing to the flower coloration in the lotus (*Nelumbo Adans*) [23]. CcMYB12 and FeMYBF1 regulate the flavonol biosynthesis in the globe artichoke and buckwheat, respectively [24,25]. NtMYB3 represses the biosynthesis of the flavonols in the narcissus and leads to a proanthocyanin accumulation [26]. The MYB transcription factors also modulate the biosynthesis of lignin, chlorophyll, and carotenoid [27–29]. Additionally, MYB can regulate the root and leaf development, fruit ripening, Fe homeostasis, and resistance against pathogens [30–37]. With the development of the genome sequencing technology, many species 'genome data and information have been reported, which provides powerful tools and resources for the identification and in-depth study of the MYB gene family. Now, more and more MYB gene families have been identified from different species, including *Solanum lycopersicum* [38], *Gossypium hirsutum* [39], *Prunus persica* [40], *Solanum tuberosum* [41], *Medicago sativa* [42], and so on. It has become a hot spot to study and understand the function of each member and their mutual relationship.

There are more than 110 R2R3-MYB TFs classified into 33 subgroups in *S. miltiorrhiza* [4], some of which are supposed to regulate the biosynthesis of either the phenolic acids or tanshinones. For instance, SmPAP1 and SmMYB111 are positive regulators of the phenolic acids, while *SmMYB39* negatively regulates their biosynthesis [43–45]. *SmMYB36* promotes the tanshinone accumulation but inhibits the phenolic acid biosynthesis [46]. Nevertheless, *SmMYB98* simultaneously improves the accumulation of the tanshinones and salvianolic acids [47]. *SmMYB9b* enhances the tanshinone concentration in hairy roots, too [48]. However, the detailed functions of many other members in the SmMYB TF family remain unknown.

We previously isolated a gene of the subgroup 4 R2R3 MYB, named *SmMYB4* (GU586494.1), from *S. miltiorrhiza* and proposed its possible involvement in the phenolic acid biosynthesis. Herein, we used the overexpression and RNAi-mediated gene silencing to investigate its role in regulating the phenolic acid and tanshinone biosynthesis. The manipulation of the SmMYB4 protein levels offers a potential tool to regulate the metabolic flux responsible for the biosynthesis of the phenolic acids and tanshinones in *S. miltiorrhiza*.

2. Materials and Methods

2.1. Plant Materials

Mature seeds of *S. miltiorrhiza*, purchased from Shaanxi Tasly Plants Pharmaceutical Co., Ltd. (Shangluo, China), were sewn in a mixed medium containing an equal

proportion of vermiculite and nutrient soil and germinated at 25 ± 2 °C, using a cycle of light (16 h at 8000 lx) and dark (8 h). The two-month-old seedlings were frozen in liquid nitrogen for the subsequent total RNA extraction. The seeds were sterilized and cultured on a MS (Murashige and Skoog) medium to obtain sterile plantlets, as described by Yan and Wang [49]. The leaves from these plantlets were used for the genetic transformation experiments.

2.2. Subcellular Localization of SmMYB4

The cDNA sequence (with no termination codon) of SmMYB4 was obtained from a public database and then cloned and sequenced from the cDNA library of *S. miltiorrhiza* with the designed primers. The sequence was inserted into pHBT-GFP-NOS vector at KpnI and BamHI restricted sites to form SmMYB4-GFP. The plasmids were transiently transformed into onion epidermal cells using the PDS-1000 Gene Gun (American Bole BIO-RAD, Santa Clara, CA, USA). The onion epidermis cells were isolated and observed under a Leica DM6000B microscope (Leica, Leica Microsystems Co., Ltd., Wetzlar, Germany) at 475 nm.

2.3. Construction of the SmMYB4-OE and SmMYB4-RNAi Expression Vectors

According to the manufacturer's protocol, the total RNA was extracted using E.Z.N.A. OMEGA TM Plant RNA Kit and converted into cDNA using PrimeScript^{TM} RT Master Mix Kit (Takara, Bao Biological Engineering (Dalian) Co., Ltd., Dalian, China).

To construct the overexpression vector, the entire open reading frame (ORF) of *SmMYB4* was amplified with the primers OE*SmMYB4-Bgl*II and OE*SmMYB4-Bst*PI (Supplementary Material Table S1). The PCR (polymerase chain reaction) products were ligated into the pMD19-T vector (Takara) and transformed into *Escherichia coli* DH5α cells. The recombinant plasmids were digested with *Bgl*II/*Bst*PI and then ligated into the pCAMBIA1302 vector to generate the overexpression plasmid OE*SmMYB4*-1302. To construct the RNAi plasmid, a specific 246 bp fragment of *SmMYB4* was selected as the RNAi target fragment. The forward and reverse RNAi target fragments were amplified with the primers *ISmMYB4-Hind*III/ISmMYB4-*Bam*HI and *ISmMYB4-Kpn*I/*ISmMYB4-Xho*I, respectively, and inserted into the pKANNIBAL vector. An interfering box was cloned into pART27 to generate the *SmMYB4*-RNAi vector named *ISmMYB4*-pART27. To verify whether the overexpression and RNAi vectors were successfully constructed, OE*SmMYB4*-1302 and IS*mMYB4*-pART27 were digested with *Bgl*II /*Bst*PI and *Xho*I/*Not*I, respectively.

2.4. Transformation and Detection of S. Miltiorrhiza

The OE*SmMYB4*-1302 and I*SmMYB4*-pART27 vectors were separately transformed into the *Agrobacterium tumefaciens* strain *GV3101* using a freeze-thaw method [50]. The *A. tumefaciens GV3101* harboring the recombinant vectors of OE*SmMYB4*-1302 or I*SmMYB4*-pART27 were used to generate the transgenic plants according to the protocol established in our laboratory [49].

To confirm whether the target sequences had been integrated into the genome of *S. miltiorrhiza*, the genomic DNA was extracted (following OMEGA DNA extraction kit instructions) from the fresh leaves of one-month-old transgenic plantlets and non-transformed (WT) plants. The primers of 35S promoter and Kan resistance gene were used for identifying the overexpression and the RNAi transgenic plantlets, respectively (Supplementary Material Table S2). The PCR products were detected using 1% agarose gel electrophoresis.

The total RNA was isolated from the young leaves and converted into cDNA using the above method. The reverse transcripts were utilized as templates for analyzing the expression of *SmMYB4* in the transgenic lines and the WT lines, using real-time quantitative PCR (qPCR). The primers Q-*SmMYB4*-F and Q-*SmMYB4*-R were synthesized by Beijing Aoke peak Biotechnology Co., Ltd., Beijing, China (Supplementary Material Table S2). *SmUbiquitin* (JF760206.1) served as an internal control to normalize the expression levels. The qPCR reactions were

performed in triplicate under the same conditions, as described previously [44]. The relative gene expression was calculated using the comparative C_T method.

2.5. Determination of the Total Phenolic Acids and Flavonoids

The transgenic and wild *S. miltiorrhiza* were sub-cultured for 90 d. Then, the roots were dried in the shade. The sample preparation followed the methods described by Zhang [45]. The total phenolic acid and flavonoid contents were measured using the Folin–Ciocalteu and $NaNO_2$-$AlCl_3$ methods, respectively [51,52]. The gallic acid and epicatechin (EC) were used as a standard for determining the total phenolic and flavonoid content, respectively.

2.6. Evaluation of the Salvianolic Acid and Tanshinone Contents

The sample preparation followed the previously described method [45]. The contents of the salvianolic acid and tanshinones were determined using a HPLC (high-performance liquid chromatography) system (Agilent Zorbax, Agilent, Santa Clara, CA, USA). The chromatographic separation was performed using a C18 column (Agilent Eclipse XDB-C18-USP L1, 4.6 mm × 250 mm, 5 μm particle size) at 30 °C and a sample injection volume of 20 μL. The detection at 280 nm, using a flow rate of 1.0 mL/min over a gradient of methanol containing 0.4% acetic acid (buffer A), triple distilled water containing 0.4% acetic acid (buffer B), and acetonitrile (Fisher Scientific) containing 0.4% acetic acid (buffer C). The gradient conditions are shown in Supplementary Material Table S3. The standards of rosmarinic acid (RA), salvianolic acid B (SAB), cryptotanshinone (CT), and tanshinone II A (T-II A) were purchased from the National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China).

2.7. Expression Analysis of the Essential Genes in the Phenolic Acid and Tanshinone Pathways

Five essential enzyme genes involved in the biosynthesis of salvianolic acids were selected to elucidate the target gene of SmMYB4, including *SmPAL1* (encoding phenylalanine ammonia-lyase, PAL), *SmC4H* (encoding cinnamate-4-hydroxylase, C4H), *Sm4CL2* (encoding 4-coumaric acid CoA-ligase, 4CL), *SmTAT* (encoding tyrosine aminotransferase, TAT), and *SmHPPR* (encoding hydroxyphenylpyruvate reductase, HPPR). Meanwhile, seven key enzyme genes related to the biosynthesis of tanshinones were investigated, including *SmHMGR1-3* (encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGR), *SmDXS1*, and *SmDXS3* (encoding 1-deoxy-D-xylulose5-phosphate synthase, DXS), and *SmGGPPS1* and *SmGGPPS3* (encoding geranylgeranyl diphosphate synthase, GGPPS). The sequences of these genes were downloaded from the NCBI database and used to design the qPCR primers (Supplementary Material Table S4).The qPCR was performed according to the method mentioned above.

2.8. Analysis of the Target Genes Regulated by SmMYB4

To elucidate the targets of the transcriptional regulation by SmMYB4, the 2.5 kb upstream sequences of the key enzyme genes were extracted from their corresponding scaffolds. The published genome of *S. miltiorrhiza* was obtained from the National Data Center of Traditional Chinese Medicine of China (https://ngdc.cncb.ac.cn/search/?dbId=gwh&q=+PRJCA003150/accessed accessed on 18 August 2020). The amino acid sequence of SmMYB4 was used as the input to the online profile inference tool (http://jaspar.genereg.net/ accessed on 18 August 2020) to search for its binding sequences. The matrix with the smallest E-value was used to scan the promoter sequences of the candidate genes and to ascertain whether they were targets of SmMYB4 (the threshold was set to the default value of 80%).

2.9. Yeast One-Hybrid (Y1H) Assay

The Y1H assay was performed, as described in Li et al., 2020 [53].

3. Results

3.1. Characterization Analysis of SmMYB4

The isolated full-length SmMYB4 contained an ORF comprising 696 bp, encoding a polypeptide of 231 amino acids with a molecular weight of 26 kDa and a pI of 8.865. SmMYB4 is predicted to localize in the nucleus (ProtComp 9.0 server, http://linux1.softberry.com/ accessed on 3 May 2021), and is demonstrated by the further subcellular localization experiment (Supplementary Material Figure S1). The phylogenetic analysis (MEGA6.0) indicated that all R2R3-MYB repressors from various plants are divided into two clades (Figure 1). SmMYB4 is grouped with AtMYB4, AtMYB32, NtMYB2, NtMYB3, and VvMYB4a, belonging to subgroup 4 of the R2R3-MYB transcriptional factors. SmMYB39, which was previously functionally characterized as a repressor of phenolic acid biosynthesis, belongs to the same clade as SmMYB4.

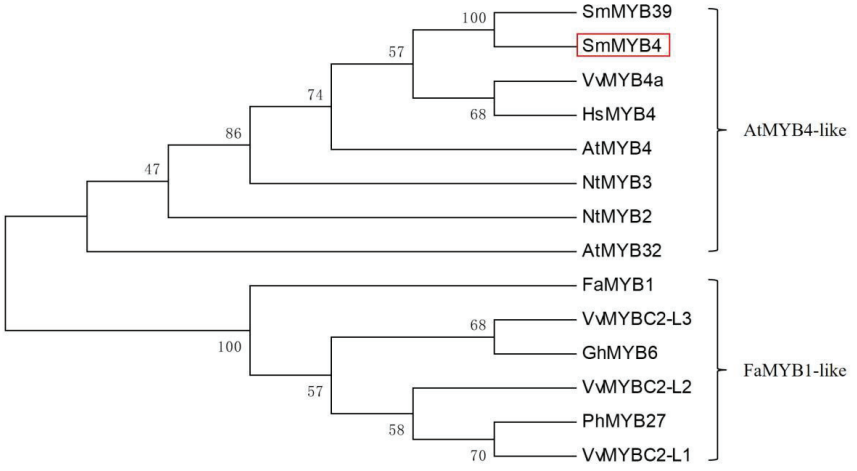


Figure 1. Phylogenetic analysis of the MYB repressors. SmMYB4 is shown by the red rectangle. *S. miltiorrhiza* SmMYB39 (AGS48990.1), SmMYB4 (ADG46002.1), *Arabidopsis thaliana* AtMYB4 (NP_195574.1), AtMYB32 (NP_195225.1), *Fragaria ananassa* FaMYB1 (AAK84064.1), *Narcissus tazetta* NtMYB3 (AGO33166.1), NtMYB2 (ATO58377.1), *Hibiscus syriacus* HsMYB4 (KAE8690675.1), *Petunia hybrid* PhMYB27 (AHX24372.1), *Vitis vinifera* VvMYB4a (NP_001268129.1), VvMYBC2-L1 (ABW34393), VvMYBC2-L2 (ACX50288), VvMYBC2-L3 (AIP98385.1), *Gossypium hirsutum* GhMYB6 (AAN28286.1).

The protein sequence alignment of SmMYB4 and the numerous R2R3-MYB repressors indicated that SmMYB4 consists of conserved R2 and R3 DNA binding domains at the N-terminus. A bHLH binding motif ([D/E]Lx_2[R/K]x_3Lx_6Lx_3R) was found in the R3 DNA binding domain of SmMYB4. Additionally, the amino acid sequence of SmMYB4 has two typically conserved motifs at the C-terminus, C1 (KLIsrGIDPxT/SHRxI/L) and C2 (pdLNLD/ELxiG/S) (Figure 2), which were previously identified in the subgroup 4 R2R3-MYB transcriptional factors [26].

3.2. Molecular Identification of the Transgenic Plantlets

The fragments of about 700 bp were captured from the digestion products of the detected OE*SmMYB4*-1302 vectors, and the expected fragments of approximately 1300 bp and 3400 bp were obtained from the cutting products of the detected IS*mMYB4*-pART27 vectors, which indicated that the overexpression and the RNAi vectors were successfully constructed (Supplementary Material Figure S2). Then, the constructed vectors were sequenced using their respective primers. The transgenic plants were generated to examine the function of SmMYB4. In this research, fourteen overexpression transgenic lines (O-4-1-14) and seventeen interference transgenic lines (I-4-1-16) were screened at the DNA

level (Supplementary Material Figure S3). At the level of transcription, six overexpression lines (O-4-3, O-4-5, O-4-6, O-4-7, O-4-8, O-4-9) and eight silent lines (I-4-1, I-4-2, I-4-3, I-4-5, I-4-7, I-4-8, I-4-9, I-4-10) with a significant alteration, were obtained (Figure 3), in which the highest and the lowest expression level of *SmMYB4* was 201.33-fold and 0.207-fold of that in WT, respectively.

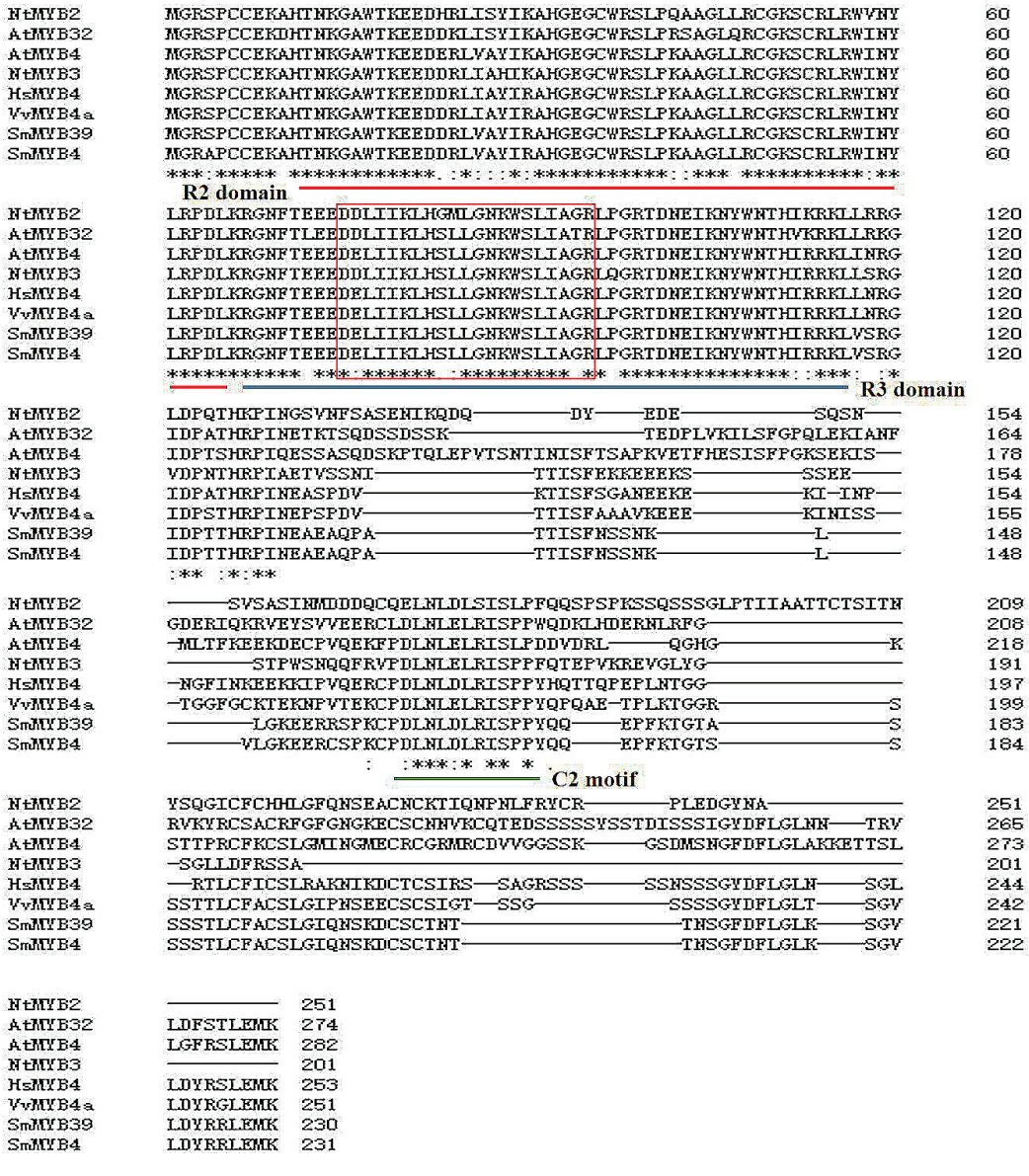


Figure 2. Multiple alignments of the amino acid sequences of SmMYB4 and other R2R3-MYB repressors belonging to subgroup 4. R2 domain, R3 domain, and C2 motif are marked in red, blue, and green lines, respectively. The bHLH binding motif is shown by the red rectangle (*, ** and *** Represents that they have the same conserved structure and function.).

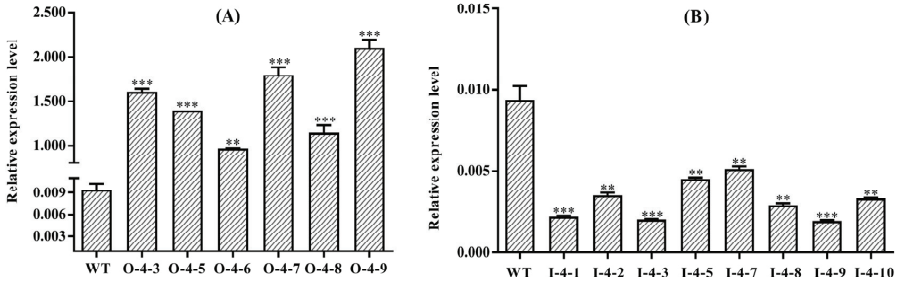


Figure 3. Expression of *SmMYB4* in the transgenic lines. (**A**,B) indicates the expression level of *SmMYB4* in the overexpressing lines and RNAi lines, respectively (** $p < 0.01$, *** $p < 0.001$).

3.3. SmMYB4 Suppresses the Accumulation of Phenolic Acids and Flavonoids

To study the function of SmMYB4 in the secondary metabolites biosynthesis, the total phenolic acid and flavonoid contents of *S. miltiorrhiza* were detected (Figure 4A,B). Compared with the WT, the total phenolic acid content was significantly lower in the *SmMYB4* overexpression lines and markedly higher in the RNAi lines. The total phenolic acid content was minimal (7.38 mg/g, DW) in line O-4-3 and maximal (17.67 mg/g, DW) in line I-4-1. The overexpression of SmMYB4 had no significant effect on the flavonoid content, but its silencing resulted in the accumulation of flavonoids, which reached a peak level (19.10 mg/g, DW) in line I-4-1.

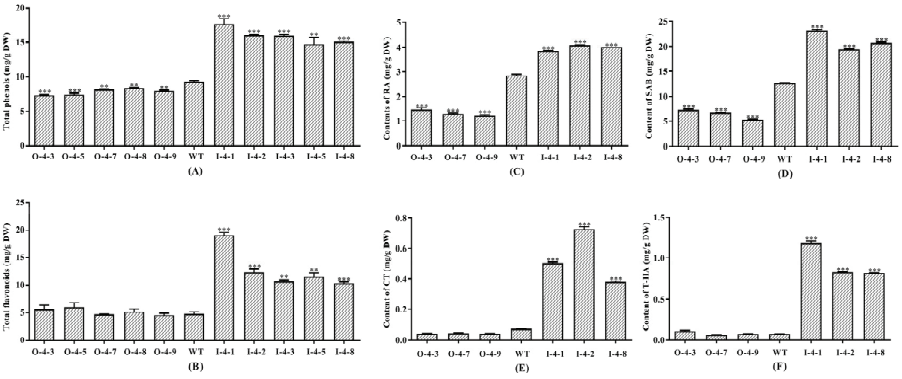


Figure 4. The content of metabolites in the transgenic plants. (**A**–**F**) indicate the content of the total phenolic acids, total flavonoids, RA, SAB, CT, and T-IIA, respectively (** $p < 0.01$, *** $p < 0.001$).

3.4. SmMYB4 Negatively Regulates the Salvianolic Acid and Tanshinone Biosynthesis

RA and SAB are the most abundant water-soluble active constituents in *S. miltiorrhiza*. It was found that their content had a significant change in the transgenic lines compared with the WT, which had 2.88 mg/g RA and 12.77 mg/g SAB (Figure 4). Compared to the WT, the RA and SAB contents were lower in the overexpression lines (O-4-3, O-4-7, O-4-9). Their respective levels in line O-4-9 notably decreased by 0.566-fold and 0.576-fold. Conversely, their content in the interference lines was significantly higher, being 1.41- and 1.80-fold of their respective contents in the WT (Figure 4C,D).

The total tanshinone content in the overexpression lines was lower than that in the WT, especially in line O-4-9, which was 0.32-fold lower. In contrast, their content in all interference lines was significantly higher, up to 2.98-fold that of WT (Supplementary Material Figure S4). Compared with the WT, the CT content in the three overexpression lines was lower, the lowest being measured in line O-4-9 (0.04 mg/g), whereas the T-IIA content was not significantly changed. It is noteworthy that the content of CT and T-IIA

was markedly higher in three interference lines (Figure 4E,F), with line I-4-2 having the most CT (10.14-fold of WT) and line I-4-1 having the most T-IIA (17.2-fold of WT).

3.5. SmMYB4 Inhibits the Expression of Key Enzyme Genes

To uncover the target genes regulated by SmMYB4, the expression of twelve key enzyme genes involved in the salvianolic acid and tanshinone biosynthetic pathways was analyzed. The qPCR analysis demonstrated that silencing *SmMYB4* markedly upregulated *SmPAL1*, *SmC4H*, *Sm4CL2*, and *SmTAT* (Figure 5). On the contrary, the overexpression of *SmMYB4* downregulated these genes except for *Sm4CL2*. The highest levels of *SmPAL1*, *SmC4H*, *Sm4CL2*, and *SmTAT* were 5.46, 2.67, 4.68, and 10.28 times their respective levels observed in the WT. The lowest levels of *SmPAL1*, *SmC4H*, and *SmTAT* were 41.45%, 9.89%, and 17.02% of their respective levels in the WT. In the tanshinone biosynthetic pathway, *SmDXS1*, *SmDXS3*, *SmGGPPS1*, and *SmGGPPS3* were significantly upregulated in all interference lines, while only *SmDXS1* was downregulated in all overexpressing lines. The transcription levels of other genes increased in the overexpressing lines but were not significantly different from their levels in the WT.

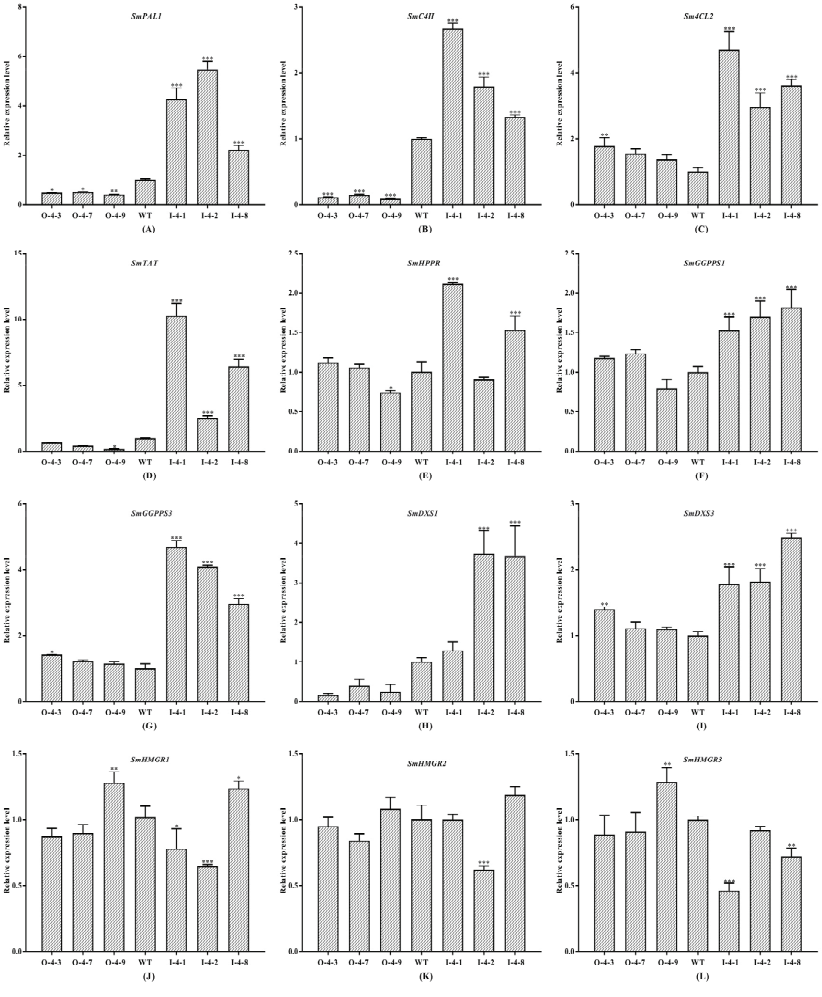


Figure 5. The relative expression level of the key enzyme genes in the salvianolic acid and tanshinone biosynthetic pathways. (**A**–**E**) represent the relative expression level of *SmPAL1*,*SmC4H*, *Sm4CL2*, *SmTAT*,

and *SmHPPR* in the salvianolic biosynthetic pathway, respectively. (**F**–**L**) represent the relative expression level of*SmGGPPS1*, *SmGGPPS3*, *SmDXS1*, *SmDXS3*, *SmHMGR1*, *SmHMGR2*, and *SmHMGR3* in the tanshinone biosynthetic pathway, respectively. (* $p < 0.5$, ** $p < 0.01$, *** $p < 0.001$).

3.6. SmMYB4 Possibly Binds to the Promoter of Key Enzyme Genes

The matrix {(G/A) (G/T) TG (G/T) T (G/A)} was selected to scan the promoters of *SmPAL1*, *SmC4H*, *Sm4CL2*, *SmTAT*, *SmDXS1*, *SmDXS3*, *SmGGPPS1*, and *SmGGPPS3*. Numerous possible SmMYB4 binding sites were found to exist within each promoter. (Supplementary Material Table S5). The Y1H assay demonstrated that SmMYB4 directly bind the promoters of *SmTAT1*, *SmC4H*, and *SmPAL1* (Figure 6). Therefore, we deduced that SmMYB4 might carry out its inhibitory function in the salvianolic acid and tanshinone biosynthetic pathways by binding to the promoter of these genes.

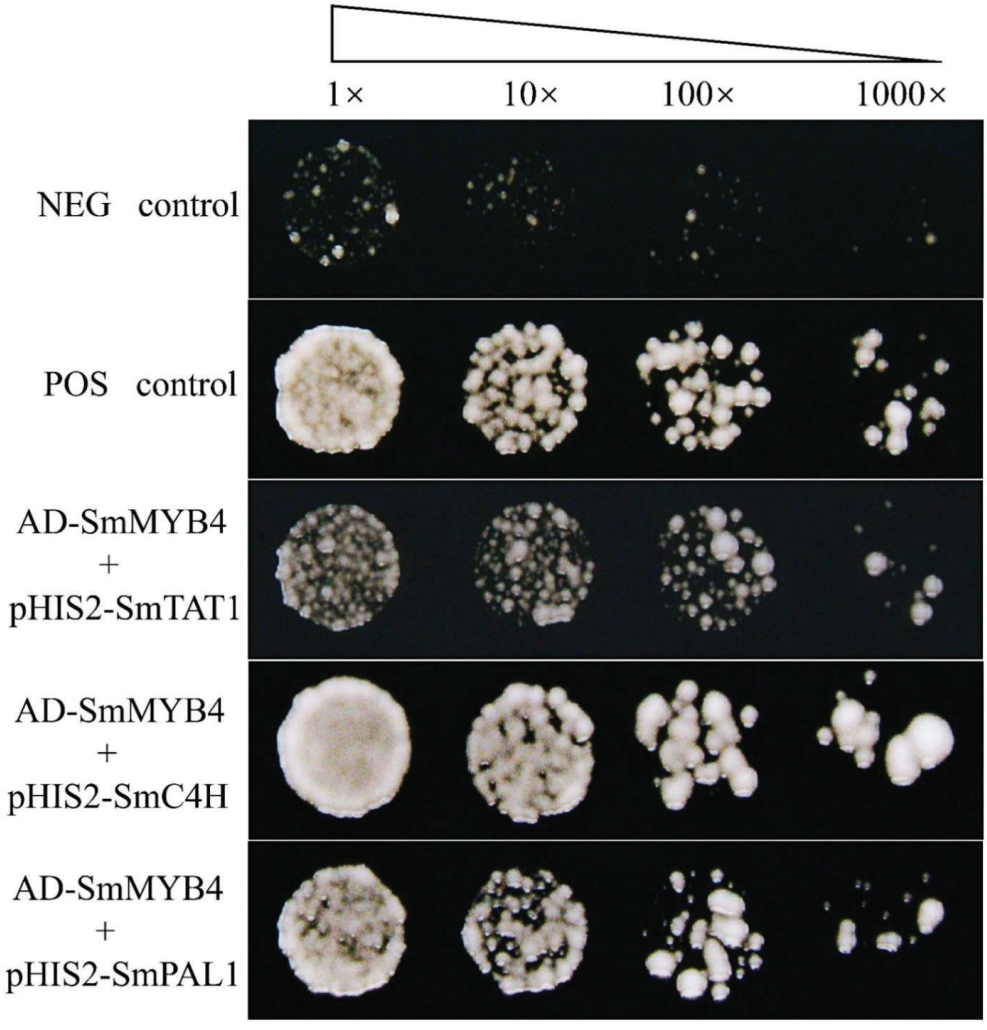


Figure 6. SmMYB4 binds the promoters of SmTAT1, SmC4H, and SmPAL1. The first and second rows represent the negative and positive controls, respectively. NEG control: pGADT7 + p53HIS2; POS control: pGADT7-p53 + p53HIS2.

4. Discussion

S. miltiorrhiza, an important medicinal material for treating cardiovascular and cerebrovascular diseases, has a growing market demand. The content of the bioactive components in the roots is one of the most crucial quality evaluation indexes. The exclusive information demonstrated that the secondary metabolites biosynthesis is regulated by diverse transcription factors and microRNAs [54,55]. The R2R3-MYB transcriptional factors are a large gene family in plants and have been reported to play crucial roles in regulating secondary metabolites. They can positively and negatively regulate the generation of secondary metabolites.

The R2R3 MYB proteins are more conserved in the N-terminal DNA binding domains, while the protein sequences at the C-terminus are divergent and believed to be responsible for several kinds of regulatory functions. Based on the specific conserved regions in the C-terminus, R2R3 (MYBs TFs) have been classified into many subgroups, in which subgroup 4 R2R3-MYB transcriptional factors function as negative regulators in the phenylpropanoid and flavonoid biosynthesis [56]. Many subgroup 4 R2R3-MYB transcription factors have been identified as transcriptional repressors. AtMYB4 suppressed the expression of *AtC4H* and resulted in the downregulated accumulation of the sinapic acid [57]. The ectopic overexpression of PvMYB4 from *Panicum virgatum* in the transgenic switch grass resulted in a reduced lignin content [58]. BrMYB4 negatively regulated the *BrC4H* gene involved in phenylpropanoid biosynthesis [59]. *NtMYB2* and *NtMYB3* repressed the biosynthesis of anthocyanin and flavonols in the Chinese narcissus, respectively [26,60]. *SmMYB39* and *SmMYB36* negatively regulated the phenolic acid and flavonoid biosynthesis in *S. miltiorrhiza* [45,46]. In the current study, the overexpression of *SmMYB4* significantly decreased the salvianolic acid and tanshinone content, while its silencing enhanced their accumulation, suggesting that *SmMYB4* negatively regulates their biosynthesis. Moreover, a C2 motif that is believed to be the key characteristic responsible for the transcriptional inhibition was harbored in the C terminus of SmMYB4.

The variations in the salvianolic acid and tanshinone content are closely related to the expression of genes encoding key enzymes involved in their biosynthesis. The expression levels of *SmPAL1*, *SmC4H*, *Sm4CL2*, and *SmTAT* were significantly downregulated in the *SmMYB4*-overexpressing transgenic lines and markedly upregulated in the RNAi transgenic lines. Combined with the promoter analysis, it was concluded that SmMYB4 affects the accumulation of the salvianolic acids by binding to the promoters of the genes mentioned above. Nevertheless, SmPAP1 activated the promoters of two key genes, *SmPAL1* and *SmC4H* [43]. *SmMYB98* activated the promoter of *SmPAL1* and *SmRAS1* in the salvianolic acid biosynthesis [47]. The overexpression of *SmMYB36* downregulated most genes in the phenylpropanoid pathway (*SmPAL1*, *SmC4H1*, *Sm4CL2*) and tyrosine pathway (*SmTAT1*, *SmHPPR1*). However, the expression level of *RAS1* and *CYP98A14* did not change significantly. A similar phenomenon occurred in the *SmMYB36*-overexpressing lines [46]. The expression of *SmPAL1*, *SmPAL2*, *SmPAL3*, *SmC4H*, *Sm4CL2*, *Sm4CL3*, *SmRAS1*, and *CYP98A14* was significantly induced in the *SmMYB111*-overexpressing lines but markedly reduced in the corresponding silenced lines [44]. The overexpression of *SmMYB9b* in the hairy roots significantly upregulated the transcription of *SmDXS2*, *SmDXR*, *SmGGPPS*, and *SmKSL1*, suggesting it enhanced the tanshinone synthesis through the stimulation of the MEP (methylerythritol phosphate) pathway [48]. The expression of *SmDXS1*, *SmDXS2*, *SmDXR*, *SmMCT*, *SmMDS*, *SmHDS*, *SmCMK*, *SmHDR1*, *SmHMGR2*, *SmGGPPS1*, *SmCPS1*, *SmKSL1*, and *SmCYP76AHI* was enhanced when *SmMYB36* was overexpressed [46]. *SmMYB98* activated the expression of *SmGGPPS1* in the tanshinone biosynthesis [47]. In our study, the expression level of *SmDXS1*, *SmDXS3 SmGGPPS1*, and *SmGGPPS3* increased significantly in the RNAi transgenic lines but was not changed in the overexpressing lines. We conclude that SmMYB4 affects the accumulation of the tanshinones by negatively regulating the expression of these key enzyme genes. The inability to detect any significant changes in the expression of other genes may be due to the competitive binding of the activator-type MYBs and the repressor-type MYBs to those targets (Figure 7).

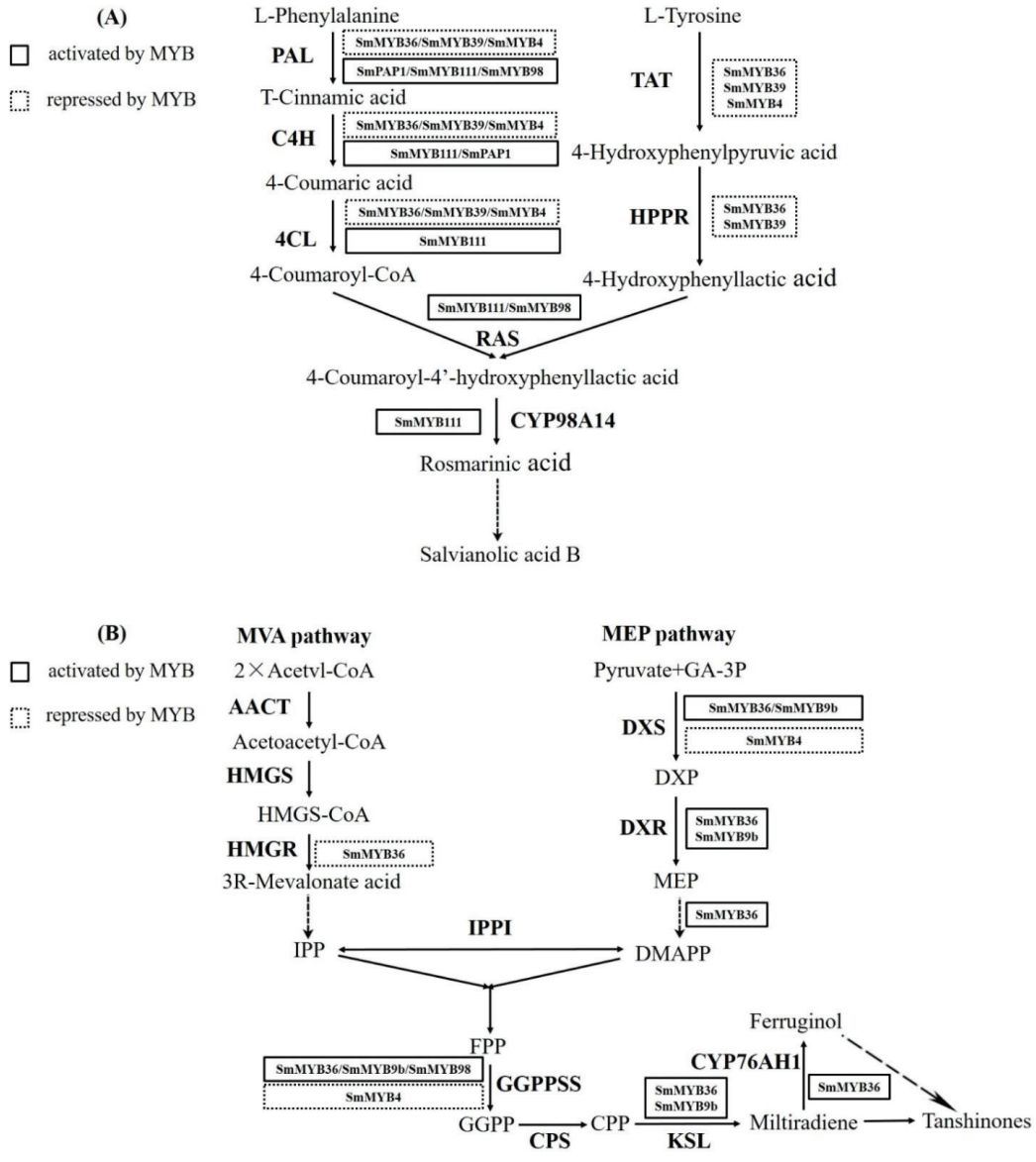


Figure 7. Multiple R2R3-MYBs regulating the salvianolic acid and tanshinone biosynthetic pathways in *S. miltiorrhiza*. (**A**) indicates multiple R2R3-MYBs regulating salvianolic acid synthesis pathway, and (**B**) indicates multiple R2R3-MYBs regulating tanshinone biosynthetic pathway.

The secondary metabolism in plants involves a complex three-dimensional regulatory network. A pathway or a key enzyme gene is usually regulated by many different transcription factors, which can interact and coordinatively regulate the gene expression. For example, PavMYB10.1 interacts with PavbHLH and PavWD40, and binds to the promoter regions of the anthocyanin biosynthesis genes *PavANS* and *PavUFGT* in sweet cherry [61]. The anthocyanin synthesis in the lotus is regulated by a NnMYB5-NnbHLH1-NnTTG1 complex [23]. The MYB activator WHITE PETAL1 (WP1) physically interacts with MtTT8 and MtWD40-1 proteins to form the conserved MYB-basic-helix-loop-helix-WD40 regula-

tory module, and regulates the floral carotenoid pigmentation through the transcriptional activation of the carotenoid biosynthetic genes in *Medicago truncatula* [29]. SmPAP1 regulates the phenolic acid biosynthetic pathway in *S. miltiorrhiza* by interacting with SmMYC2 and activating *SmPAL1* and *SmC4H* [43]. SmMYB111 positively regulates the phenolic acid biosynthesis in *S. miltiorrhiza* by interacting with SmTTG1 and SmbHLH51 [44]. However, the specific mechanism of SmMYB4 needs further study. Our work will entail not only the elucidation of the secondary metabolic networks, but also enhance the understanding of the gene regulatory networks integrated with the metabolic networks.

5. Conclusions

In summary, a subgroup 4 R2R3 MYB transcription factor gene named *SmMYB4* was cloned and functionally identified in the present study. Six overexpressions and eight RNA interference transgenic lines were successfully obtained via the *Agrobacterium*-mediated leaf-disk transformed method. Based on the analysis of the content of the phenolic acids and tanshinones in transgenic plants and the expression of genes in the related pathway, we speculated that SmMYB4 negatively regulates the biosynthesis of the phenolic acids by repressing the expression of *SmPAL1*, *SmC4H*, *Sm4CL2*, and *SmTAT*, and simultaneously negatively regulates the biosynthesis of tanshinones by inhibiting the expression of *SmDXS1*, *SmDXS3 SmGGPPS1*, and *SmGGPPS3*. SmMYB4 may directly bind to the promoter of these genes. Our results will be helpful in better understanding the regulation mechanism of the phenolic acids and tanshinones production in *S. miltiorrhiza* and provide a train of thought to improve the contents of bioactive compounds in this traditional herbal.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12100968/s1, Table S1: Primers used in the *SmMYB4*-OE and *SmMYB4*-RNAi vector construction; Table S2: Primers for the molecular detection of transgenic plants; Table S3: Gradient conditions of the mobile phase of HPLC; Table S4: Primers for the key enzyme genes in the salvianolic acid and tanshinone biosynthetic pathways; Table S5: Information of the SmMYB4 binding sites in the promoters of the key enzyme genes; Figure S1: Subcellular localization of SmMYB4 in onion epidermal cells; Figure S2: Restriction endonuclease digestion of the overexpression and RNAi expression vectors. A represents the Bgl II/BstP I digestion of the overexpression vectors, in which P represents plasmids, M refers to Marker DL2000, and 1–4 represents the digest results of Bgl II/BstP I. B represents the digestion of the RNAi expression vectors, in which P represents plasmids, M refers to Marker DL2000, 1 refers to the digestion of Xho I, and 2 refers to the digestion of Not I. The target fragment is marked with a red line or arrows; Figure S3: Detection of the transgenic lines at the DNA level Detection. Results of the overexpression and interference transgenic lines were individually shown in Figure S3A,B: P indicates positive control (Plasmids), WT indicates negative control (untransformed plants), and M indicates Marker DL2000; Figure S4: Total tanshinone contents in the roots of transgenic plants (** $p < 0.01$, *** $p < 0.001$).

Author Contributions: Conceptualization, Q.T., L.H. and B.L.; methodology, Q.T., L.H., X.Z., C.Z. and Y.L.; formal analysis, Q.T., L.H., X.X. and Y.W.; investigation, Q.T., D.W. and W.H.; data Curation, Q.T., L.H. and X.Z., writing—original draft preparation, Q.T. and L.H.; writing—review and editing, Q.T., D.W., J.N., W.H. and B.L.; project administration, B.L. and Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (31870276, 31900254, and 32170378), the Fundamental Research Funds for the Central Universities (GK202205002, GK202205003, GK202205004), the Project of Shaanxi Science and Technology Department (2020JM-628), Shaanxi Administration of Traditional Chinese Medicine Projects (Grant number 2021-QYZL-02, 2021-QYPT-002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the article and supplementary materials.

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

References

1. Ren, J.; Fu, L.; Nile, S.H.; Zhang, J.; Kai, G. *Salvia miltiorrhiza* in treating cardiovascular diseases: A review on its pharmacological and clinical applications. *Front. Pharmacol.* **2019**, *10*, 753. [CrossRef] [PubMed]
2. Gantet, P.; Memelink, J. Transcription factors: Tools to engineer the production of pharmacologically active plant metabolites. *Trends Pharmacol. Sci.* **2002**, *23*, 563–569. [CrossRef]
3. Xu, H.; Song, J.; Luo, H.; Zhang, Y.; Li, Q.; Zhu, Y.; Xu, J.; Li, Y.; Song, C.; Wang, B.; et al. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. *Mol. Plant* **2016**, *9*, 949–952. [CrossRef] [PubMed]
4. Li, C.; Lu, S. Genome-wide characterization and comparative analysis of R2R3-MYB transcription factors shows the complexity of MYB-associated regulatory networks in *Salvia miltiorrhiza*. *BMC Genom.* **2014**, *15*, 277. [CrossRef]
5. Zhang, X.; Luo, H.; Xu, Z.; Zhu, Y.; Ji, A.; Song, J.; Chen, S. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza*. *Sci. Rep.* **2015**, *5*, 11244. [CrossRef]
6. Yu, H.; Guo, W.; Yang, D.; Hou, Z.; Liang, Z. Transcriptional profiles of *SmWRKY* family genes and their putative roles in the biosynthesis of tanshinone and phenolic acids in *Salvia miltiorrhiza*. *Int. J. Mol. Sci.* **2018**, *19*, 1593. [CrossRef]
7. Ji, A.J.; Luo, H.M.; Xu, Z.C.; Zhang, X.; Zhu, Y.J.; Liao, B.S.; Yao, H.; Song, J.Y.; Chen, S.L. Genome-wide identification of the AP2/ERF gene family involved in active constituent biosynthesis in *Salvia miltiorrhiza*. *Plant Genome* **2016**, *9*, 77. [CrossRef]
8. Zhang, Y.; Xu, Z.; Ji, A.; Luo, H.; Song, J. Genomic survey of bZIP transcription factor genes related to tanshinone biosynthesis in *Salvia miltiorrhiza*. *Acta Pharm. Sin. B* **2018**, *8*, 295–305. [CrossRef]
9. Xing, B.; Yang, D.; Yu, H.; Zhang, B.; Yan, K.; Zhang, X.; Han, R.; Liang, Z. Overexpression of *SmbHLH10* enhances tanshinones biosynthesis in *Salvia miltiorrhiza* hairy roots. *Plant Sci.* **2018**, *276*, 229–238. [CrossRef]
10. Cao, W.; Wang, Y.; Shi, M.; Hao, X.; Zhao, W.; Wang, Y.; Ren, J.; Kai, G. Transcription factor SmWRKY1 positively promotes the biosynthesis of tanshinones in *Salvia miltiorrhiza*. *Front Plant Sci.* **2018**, *9*, 554. [CrossRef]
11. Deng, C.; Hao, X.; Shi, M.; Fu, R.; Wang, Y.; Zhang, Y.; Zhou, W.; Feng, Y.; Makunga, N.P.; Kai, G. Tanshinone production could be increased by the expression of *SmWRKY2* in *Salvia miltiorrhiza* hairy roots. *Plant Sci.* **2019**, *284*, 1–8. [CrossRef] [PubMed]
12. Wu, Y.; Zhang, Y.; Li, L.; Guo, X.; Wang, B.; Cao, X.; Wang, Z. AtPAP1 interacts with and activates SmbHLH51, a positive regulator to phenolic acids biosynthesis in *Salvia miltiorrhiza*. *Front Plant Sci.* **2018**, *9*, 1687. [CrossRef]
13. Xing, B.; Liang, L.; Liu, L.; Hou, Z.; Yang, D.; Yan, K.; Zhang, X.; Liang, Z. Overexpression of *SmbHLH148* induced biosynthesis of tanshinones as well as phenolic acids in *Salvia miltiorrhiza* hairy roots. *Plant Cell Rep.* **2018**, *37*, 1681–1692. [CrossRef] [PubMed]
14. Zhang, C.; Xing, B.; Yang, D.; Ren, M.; Guo, H.; Yang, S.; Liang, Z. SmbHLH3 acts as a transcription repressor for both phenolic acids and tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots. *Phytochemistry* **2020**, *169*, 112183. [CrossRef] [PubMed]
15. Sun, M.; Shi, M.; Wang, Y.; Huang, Q.; Yuan, T.; Wang, Q.; Wang, C.; Zhou, W.; Kai, G. The biosynthesis of phenolic acids is positively regulated by the JA-responsive transcription factor ERF115 in *Salvia miltiorrhiza*. *J. Exp. Bot.* **2019**, *70*, 243–254. [CrossRef] [PubMed]
16. Li, W.; Bai, Z.; Pei, T.; Yang, D.; Mao, R.; Zhang, B.; Liu, C.; Liang, Z. SmGRAS1 and SmGRAS2 regulate the biosynthesis of tanshinones and phenolic acids in *Salvia miltiorrhiza*. *Front Plant Sci.* **2019**, *10*, 1367. [CrossRef]
17. Zhou, Y.; Zhu, H.; He, S.; Zhai, H.; Zhao, N.; Xing, S.; Wei, Z.; Liu, Q. A novel sweetpotato transcription factor gene *IbMYB116* enhances drought tolerance in transgenic *Arabidopsis*. *Front Plant Sci.* **2019**, *10*, 1025. [CrossRef]
18. Dong, W.; Liu, X.; Li, D.; Gao, T.; Song, Y. Transcriptional profiling reveals that a MYB transcription factor *MsMYB4* contributes to the salinity stress response of alfalfa. *PLoS ONE* **2018**, *13*, e0204033. [CrossRef]
19. Xing, C.; Liu, Y.; Zhao, L.; Zhang, S.; Huang, X. A novel MYB transcription factor regulates ascorbic acid synthesis and affects cold tolerance. *Plant Cell Environ.* **2019**, *42*, 832–845. [CrossRef]
20. Schwinn, K.E.; Ngo, H.; Kenel, F.; Brummell, D.A.; Albert, N.W.; McCallum, J.A.; Pither-Joyce, M.; Crowhurst, R.N.; Eady, C.; Davies, K.M. The onion (*Allium cepa* L.) *R2R3-MYB* gene *MYB1* regulates anthocyanin biosynthesis. *Front Plant Sci.* **2016**, *7*, 1865. [CrossRef]
21. Yu, M.; Man, Y.; Wang, Y. Light- and temperature-induced expression of an R2R3-MYB gene regulates anthocyanin biosynthesis in red-fleshed kiwifruit. *Int. J. Mol. Sci.* **2019**, *20*, 5228. [CrossRef]
22. Oglesby, L.; Ananga, A.; Obuya, J.; Ochieng, J.; Cebert, E.; Tsolova, V. Anthocyanin accumulation in muscadine berry skins is influenced by the expression of the MYB transcription factors, *MybA1*, and *MYBCS1*. *Antioxidants* **2016**, *5*, 35. [CrossRef]
23. Sun, S.S.; Gugger, P.F.; Wang, Q.F.; Chen, J.M. Identification of a *R2R3-MYB* gene regulating anthocyanin biosynthesis and relationships between its variation and flower color difference in lotus (*Nelumbo* Adans.). *PeerJ* **2016**, *4*, e2369. [CrossRef]
24. Blanco, E.; Sabetta, W.; Danzi, D.; Negro, D.; Passeri, V.; Lisi, A.; Paolocci, F.; Sonnante, G. Isolation and characterization of the flavonol regulator CcMYB12 from the globe artichoke [*Cynara cardunculus* var. scolymus (L.) Fiori]. *Front Plant Sci.* **2018**, *9*, 941. [CrossRef]
25. Matsui, K.; Oshima, Y.; Mitsuda, N.; Sakamoto, S.; Nishiba, Y.; Walker, A.R.; Ohme-Takagi, M.; Robinson, S.P.; Yasui, Y.; Mori, M.; et al. Buckwheat R2R3 MYB transcription factor FeMYBF1 regulates flavonol biosynthesis. *Plant Sci.* **2018**, *274*, 466–475. [CrossRef] [PubMed]
26. Anwar, M.; Yu, W.; Yao, H.; Zhou, P.; Allan, A.C.; Zeng, L. *NtMYB3*, an R2R3-MYB from Narcissus, regulates flavonoid biosynthesis. *Int. J. Mol. Sci.* **2019**, *20*, 5456. [CrossRef] [PubMed]

27. An, C.; Sheng, L.; Du, X.; Wang, Y.; Zhang, Y.; Song, A.; Jiang, J.; Guan, Z.; Fang, W.; Chen, F.; et al. Overexpression of *CmMYB15* provides chrysanthemum resistance to aphids by regulating the biosynthesis of lignin. *Hortic. Res.* **2019**, *6*, 84. [CrossRef]
28. Ampomah-Dwamena, C.; Thrimawithana, A.H.; Dejnoprat, S.; Lewis, D.; Espley, R.V.; Allan, A.C. A kiwifruit (*Actinidia deliciosa*) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytol.* **2019**, *221*, 309–325. [CrossRef] [PubMed]
29. Meng, Y.; Wang, Z.; Wang, Y.; Wang, C.; Zhu, B.; Liu, H.; Ji, W.; Wen, J.; Chu, C.; Tadege, M.; et al. The MYB activator WHITE PETAL1 associates with MtTT8 and MtWD40-1 to regulate carotenoid-derived flower pigmentation in *Medicago truncatula*. *Plant Cell* **2019**, *31*, 2751–2767. [CrossRef] [PubMed]
30. Ge, L.; Dou, Y.; Li, M.; Qu, P.; He, Z.; Liu, Y.; Xu, Z.; Chen, J.; Chen, M.; Ma, Y. SiMYB3 in foxtail millet (*Setaria italica*) confers tolerance to low-nitrogen stress by regulating root growth in transgenic plants. *Int. J. Mol. Sci.* **2019**, *20*, 5741. [CrossRef]
31. Yang, Y.; Zhang, L.; Chen, P.; Liang, T.; Li, X.; Liu, H. UV-B photoreceptor UVR8 interacts with MYB73/MYB77 to regulate auxin responses and lateral root development. *EMBO J.* **2020**, *39*, e101928. [CrossRef] [PubMed]
32. Zhang, L.; Dong, C.; Zhang, Q.; Zhao, G.; Li, F.; Xia, C.; Zhang, L.; Han, L.; Wu, J.; Jia, J.; et al. The wheat MYB transcription factor TaMYB18 regulates leaf rolling in rice. *Biochem. Biophys. Res. Commun.* **2016**, *481*, 77–83. [CrossRef] [PubMed]
33. Fan, Z.Q.; Ba, L.J.; Shan, W.; Xiao, Y.Y.; Lu, W.J.; Kuang, J.F.; Chen, J.Y. A banana R2R3-MYB transcription factor MaMYB3 is involved in fruit ripening through modulation of starch degradation by repressing starch degradation-related genes and *MabHLH6*. *Plant J.* **2018**, *96*, 1191–1205. [CrossRef]
34. Wang, F.P.; Wang, X.F.; Zhang, J.; Ma, F.; Hao, Y.J. MdMYB58 modulates Fe homeostasis by directly binding to the *MdMATE43* promoter in plants. *Plant Cell Physiol.* **2018**, *59*, 2476–2489. [CrossRef]
35. Yu, Y.; Guo, D.; Li, G.; Yang, Y.; Zhang, G.; Li, S.; Liang, Z. The grapevine R2R3-type MYB transcription factor VdMYB1 positively regulates defense responses by activating the *stilbene synthase gene 2* (*VdSTS2*). *BMC Plant Biol.* **2019**, *19*, 478. [CrossRef]
36. Zhang, Y.L.; Zhang, C.L.; Wang, G.L.; Wang, Y.X.; Qi, C.H.; Zhao, Q.; You, C.X.; Li, Y.Y.; Hao, Y.J. The R2R3 MYB transcription factor MdMYB30 modulates plant resistance against pathogens by regulating cuticular wax biosynthesis. *BMC Plant Biol.* **2019**, *19*, 362. [CrossRef]
37. Qiu, Z.; Yan, S.; Xia, B.; Jiang, J.; Yu, B.; Lei, J.; Chen, C.; Chen, L.; Yang, Y.; Wang, Y.; et al. The eggplant transcription factor MYB44 enhances resistance to bacterial wilt by activating the expression of *spermidine synthase*. *J. Exp. Bot.* **2019**, *70*, 5343–5354. [CrossRef] [PubMed]
38. Li, Z.; Peng, R.; Tian, Y.; Han, H.; Xu, J.; Yao, Q. Genome-wide identification and analysis of the MYB transcription factor superfamily in *Solanum lycopersicum*. *Plant Cell Physiol.* **2016**, *57*, 1657–1677. [CrossRef] [PubMed]
39. Salih, H.; Gong, W.; He, S.; Sun, G.; Sun, J.; Du, X. Genome-wide characterization and expression analysis of MYB transcription factors in *Gossypium hirsutum*. *BMC Genet* **2016**, *17*, 129. [CrossRef]
40. Zhang, C.; Ma, R.; Xu, J.; Yan, J.; Guo, L.; Song, J.; Feng, R.; Yu, M. Genome-wide identification and classification of *MYB* superfamily genes in peach. *PLoS ONE* **2018**, *13*, e0199192. [CrossRef] [PubMed]
41. Sun, W.; Ma, Z.; Chen, H.; Liu, M. MYB gene family in potato (*Solanum tuberosum* L.): Genome-wide identification of hormone-responsive reveals their potential functions in growth and development. *Int. J. Mol. Sci.* **2019**, *20*, 4847. [CrossRef] [PubMed]
42. Zhou, Q.; Jia, C.; Ma, W.; Cui, Y.; Jin, X.; Luo, D.; Min, X.; Liu, Z. MYB transcription factors in alfalfa (*Medicago sativa*): Genome-wide identification and expression analysis under abiotic stresses. *PeerJ* **2019**, *7*, e7714. [CrossRef] [PubMed]
43. Hao, G.; Jiang, X.; Feng, L.; Tao, R.; Li, Y.; Huang, L. Cloning, molecular characterization and functional analysis of a putative R2R3-MYB transcription factor of the phenolic acid biosynthetic pathway in *S. miltiorrhiza* Bge. f. alba. *Plant Cell Tissue Organ. Cult. (PCTOC)* **2015**, *124*, 151–168. [CrossRef]
44. Li, S.; Wu, Y.; Kuang, J.; Wang, H.; Du, T.; Huang, Y.; Zhang, Y.; Cao, X.; Wang, Z. SmMYB111 is a key factor to phenolic acid biosynthesis and interacts with both SmTTG1 and SmbHLH51 in *Salvia miltiorrhiza*. *J. Agric. Food Chem.* **2018**, *66*, 8069–8078. [CrossRef]
45. Zhang, S.; Ma, P.; Yang, D.; Li, W.; Liang, Z.; Liu, Y.; Liu, F. Cloning and characterization of a putative R2R3 MYB transcriptional repressor of the rosmarinic acid biosynthetic pathway from *Salvia miltiorrhiza*. *PLoS ONE* **2013**, *8*, e73259. [CrossRef] [PubMed]
46. Ding, K.; Pei, T.; Bai, Z.; Jia, Y.; Ma, P.; Liang, Z. SmMYB36, a novel R2R3-MYB transcription factor, enhances tanshinone accumulation and decreases phenolic acid content in *Salvia miltiorrhiza* hairy roots. *Sci. Rep.* **2017**, *7*, 5104. [CrossRef]
47. Hao, X.; Pu, Z.; Cao, G.; You, D.; Zhou, Y.; Deng, C.; Shi, M.; Nile, S.H.; Wang, Y.; Zhou, W.; et al. Tanshinone and salvianolic acid biosynthesis are regulated by *SmMYB98* in *Salvia miltiorrhiza* hairy roots. *J. Adv. Res.* **2020**, *23*, 1–12. [CrossRef]
48. Zhang, J.; Zhou, L.; Zheng, X.; Zhang, J.; Yang, L.; Tan, R.; Zhao, S. Overexpression of *SmMYB9b* enhances tanshinone concentration in *Salvia miltiorrhiza* hairy roots. *Plant Cell Rep.* **2017**, *36*, 1297–1309. [CrossRef]
49. Yan, Y.-P.; Wang, Z.-Z. Genetic transformation of the medicinal plant *Salvia miltiorrhiza* by *Agrobacterium tumefaciens*-mediated method. *Plant Cell Tissue Organ Cult.* **2007**, *88*, 175–184. [CrossRef]
50. McCormac, A.C.; Elliott, M.C.; Chen, D.F. A simple method for the production of highly competent cells of *Agrobacterium* for transformation via electroporation. *Mol. Biotechnol.* **1998**, *9*, 155–159. [CrossRef]
51. Yan, Q.; Shi, M.; Ng, J.; Wu, J.Y. Elicitor-induced rosmarinic acid accumulation and secondary metabolism enzyme activities in *Salvia miltiorrhiza* hairy roots. *Plant Sci.* **2006**, *170*, 853–858. [CrossRef]
52. Dewanto, V.; Wu, X.; Adom, K.K.; Liu, R.H. Thermal processing enhances the nutritional value of tomatoes by increasing total antioxidant activity. *J. Agric. Food Chem.* **2002**, *50*, 3010–3014. [CrossRef]

53. Li, L.; Wang, D.; Zhou, L.; Yu, X.; Yan, X.; Zhang, Q.; Li, B.; Liu, Y.; Zhou, W.; Cao, X.; et al. JA-responsive transcription factor SmMYB97 promotes phenolic acid and tanshinone accumulation in *Salvia miltiorrhiza*. *J. Agric. Food Chem.* **2020**, *68*, 14850–14862. [CrossRef] [PubMed]
54. Bai, Z.; Wu, J.; Huang, W.; Jiao, J.; Zhang, C.; Hou, Z.; Yan, K.; Zhang, X.; Han, R.; Liang, Z.; et al. The ethylene response factor *SmERF8* regulates the expression of *SmKSL1* and is involved in tanshinone biosynthesis in *Saliva miltiorrhiza* hairy roots. *J. Plant Physiol.* **2020**, *244*, 153006. [CrossRef] [PubMed]
55. Xu, X.; Jiang, Q.; Ma, X.; Ying, Q.; Shen, B.; Qian, Y.; Song, H.; Wang, H. Deep sequencing identifies tissue-specific microRNAs and their target genes involving in the biosynthesis of tanshinones in *Salvia miltiorrhiza*. *PLoS ONE* **2014**, *9*, e111679. [CrossRef]
56. Zhou, M.; Zhang, K.; Sun, Z.; Yan, M.; Chen, C.; Zhang, X.; Tang, Y.; Wu, Y. LNK1 and LNK2 corepressors interact with the MYB3 transcription factor in phenylpropanoid biosynthesis. *Plant Physiol.* **2017**, *174*, 1348–1358. [CrossRef]
57. Jin, H.; Cominelli, E.; Bailey, P.; Parr, A.; Mehrtens, F.; Jones, J.; Tonelli, C.; Weisshaar, B.; Martin, C. Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *Embo J.* **2000**, *19*, 6150–6161. [CrossRef]
58. Shen, H.; He, X.; Poovaiah, C.R.; Wuddineh, W.A.; Ma, J.; Mann, D.G.J.; Wang, H.; Jackson, L.; Tang, Y.; Neal Stewart, C., Jr.; et al. Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor *PvMYB4* for improvement of lignocellulosic feedstocks. *New Phytol.* **2012**, *193*, 121–136. [CrossRef]
59. Zhang, L.; Wang, Y.; Sun, M.; Wang, J.; Kawabata, S.; Li, Y. *BrMYB4*, a suppressor of genes for phenylpropanoid and anthocyanin biosynthesis, is down-regulated by UV-B but not by pigment-inducing sunlight in turnip cv. Tsuda. *Plant Cell Physiol.* **2014**, *55*, 2092–2101. [CrossRef]
60. Anwar, M.; Wang, G.; Wu, J.; Waheed, S.; Allan, A.C.; Zeng, L. Ectopic overexpression of a novel R2R3-MYB, *NtMYB2* from Chinese narcissus represses anthocyanin biosynthesis in tobacco. *Molecules* **2018**, *23*, 781. [CrossRef]
61. Jin, W.; Wang, H.; Li, M.; Wang, J.; Yang, Y.; Zhang, X.; Yan, G.; Zhang, H.; Liu, J.; Zhang, K. The R2R3 MYB transcription factor *PavMYB10.1* involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (*Prunus avium* L.). *Plant Biotechnol. J.* **2016**, *14*, 2120–2133. [CrossRef] [PubMed]

metabolites

MDPI

Article

GC-TOF-MS-Based Non-Targeted Metabolomic Analysis of Differential Metabolites in Chinese Ultra-Long-Term Industrially Fermented Kohlrabi and Their Associated Metabolic Pathways

Xin Nie [1,†], Hongfan Chen [1,2,3,†], Lu Xiang [1], Yulin Zhang [2,3], Dayu Liu [2,3] and Zhiping Zhao [2,3,*]

[1] College of Food Science and Technology, Sichuan Tourism University, Chengdu 610100, China
[2] College of Food and Biological Engineering, Chengdu University, Chengdu 610106, China
[3] Sichuan Key Laboratory of Solid State Fermentation Resource Utilization, Yibin Univerisity, Yibin 644000, China
* Correspondence: zhaozhiping@cdu.edu.cn
† These authors contributed equally to this work.

Abstract: Fermented kohlrabi is a very popular side dish in China. Chinese kohlrabies industrially fermented for 0 years (0Y), 5 years (5Y), and 10 years (10Y) were employed and analyzed by non-targeted metabolomics based on GC-TOF-MS, and the differential metabolites were screened using multivariate statistical analysis techniques, including principal component analysis (PCA) and orthogonal partial least squares discrimination analysis (OPLS-DA). The results showed that 47, 38, and 33 differential metabolites were identified in the three treatment groups of 0Y and 5Y (A1), 0Y and 10Y (A2), and 5Y and 10Y (A3), respectively (VIP > 1, $p < 0.05$). The metabolites were mainly carbohydrates, amino acids, and organic acids. Furthermore, 13 differential metabolites were screened from the three groups, including L-glutamic acid, L-aspartic acid, γ-aminobutyric acid, and other compounds. Four metabolic pathways termed alanine, aspartate, and glutamate metabolism, arginine biosynthesis, arginine and proline metabolism, and glycolysis/gluconeogenesis were the most significant pathways correlated with the differential metabolites, as analyzed according to the Kyoto Encyclopedia of Genes and Genomes (KEGG). The odors for the three ultra-long-term industrially fermented kohlrabies were significantly different, as detected by E-nose. The present work describes the changes in metabolites between different ultra-long-term industrially fermented kohlrabies and the associated metabolic pathways, providing a theoretical basis for the targeted regulation of characteristic metabolite biosynthesis in Chinese fermented kohlrabi.

Keywords: kohlrabi; metabolomics; GC-TOF-MS; differential metabolites; metabolic pathway

Citation: Nie, X.; Chen, H.; Xiang, L.; Zhang, Y.; Liu, D.; Zhao, Z. GC-TOF-MS-Based Non-Targeted Metabolomic Analysis of Differential Metabolites in Chinese Ultra-Long-Term Industrially Fermented Kohlrabi and Their Associated Metabolic Pathways. *Metabolites* **2022**, *12*, 991. https://doi.org/10.3390/metabo12100991

Academic Editors: Yanjie Zhang and Yan Li

Received: 27 September 2022
Accepted: 16 October 2022
Published: 19 October 2022

1. Introduction

Kohlrabi (*Brassica juncea* var. *megarrhiza* Tsen et Lee) is a root-type variant of mustard, belonging to the cruciferous *Brassica* annual herb, which is rich in vitamins, fiber, mineral elements, and other nutrients [1]. Kohlrabi originated in China as a cross-pollinating crop with small flower organs and is mainly used as a raw material to make Chinese pickles. Kohlrabi is widely distributed in China and mainly grown in Sichuan, Yunnan, Guizhou, Hunan, Hubei, Jiangsu, and Zhejiang Provinces. The crop is usually sown from mid-to-late August to early September and harvested from December to January of the next year. The suitable growth temperature for kohlrabi is 15–20 °C. In addition to its edible value, kohlrabi also has certain medicinal values. It has been suggested that regular consumption of *Brassica* plants can effectively reduce the incidence of colon cancer and rectal cancer, as well as cardiovascular and degenerative diseases, immune dysfunction, and age-related macular degeneration [2]. Raw kohlrabi has a strong mustard spicy taste and thus it is not suitable as a side dish before fermentation. The mustard spicy taste is generally removed

by fermentation and fermented kohlrabi becomes crisp and tender, with a rich aroma and a delicious taste, making it a very popular side dish in Chinese cuisine. Yang et al. [3] analyzed the effects of high-hydrostatic-pressure treatment and heat treatment on the texture of kohlrabi, and suggested that high-hydrostatic-pressure treatment could better maintain the textural properties of kohlrabi. The flavor of fermented kohlrabi is closely correlated with production processes and raw materials. Jiang et al. [4] found that raw materials could produce significantly different flavors of kimchi. By using headspace solid-phase microextraction combined with gas chromatography–mass spectrometry, Zhang and co-workers [5] determined and analyzed the types and contents of flavor substances in mustard cabbage using traditional and modern fermentation processes. A total of 78 flavor substances were detected, including 16 esters, 11 aldehydes, 18 alcohols, 8 ketones, 6 acids, 5 olefins, 7 phenols, and 7 others. The main different flavor components between the two different fermented mustard cabbages were allyl isothiocyanate, ethyl acetate, 3-butyronitrile, phenol, ethanol, and acrolein.

Metabolomics is a qualitative and quantitative technique used to analyze changes in small-molecule metabolites with a relative molecular mass less than 1000 in organisms and their changes after stimulation and interference [6]. According to experimental purposes, metabolomics is normally divided into targeted metabolomics and non-targeted metabolomics. In recent years, metabolomics has been widely used in drug evaluation, disease diagnosis and treatment, plant stress-resistance research, and functional food development. Nuclear magnetic resonance spectroscopy, gas chromatography–mass spectrometry, gas chromatography–time-of-flight mass spectrometry (GC-TOF-MS), and liquid chromatography–mass spectrometry hyphenated techniques are the commonly used techniques in metabolomics. GC-TOF-MS exhibits more advantages, such as high mass resolution and accuracy, high scanning speed, and good sensitivity [7], and has been widely used to quantify and identify metabolites in various fields of research, such as carbohydrates, sugar alcohols, organic acids, and amino acids [8]. Metabolomics can also be employed to evaluate the flavor and nutritional value of food [9]. The development of dishes based on fermented kohlrabi depends on the metabolite contents.

To date, research on the metabolomics of Chinese ultra-long-term industrially fermented kohlrabi has not been reported in the literature, and the metabolic pathways closely related to the metabolites in Chinese ultra-long-term industrially fermented kohlrabi are poorly understood. In this study, non-targeted metabolomics of GC-TOF-MS combined with multivariate statistical analysis methods were used to measure and analyze the related metabolites of Chinese ultra-long-term (0-year, 5-year, and 10-year) industrially fermented kohlrabi, for the first time, and to elucidate their associated metabolic pathways. The 10-year fermented kohlrabi is normally used for soup, while, the 5-year fermented kohlrabi is suitable for Chinese braised dishes. The three different kohlrabies possess significantly different sensory properties. The experimental results will provide some theoretical support for differential metabolite synthesis in industrially fermented kohlrabi and the product development of functional kohlrabi.

2. Materials and Methods

2.1. Production of Ultra-Long-Term Industrially Fermented Kohlrabi

The kohlrabies were cultivated by local farmers in Chengjia Town (latitude 29 degrees 39′ north and longitude 104 degrees 21′ east) in the same area of Zigong City, China, and were of the same variety (Queyecai). After maturity, the kohlrabies were harvested and washed with water. Then, the washed kohlrabies were air-dried outside for about 30 days. After air-drying, the kohlrabies were transferred into fermentation tanks. Eight percent (*w*/*w*) salt (upper layer 60%, middle layer 30%, and lower layer 10%) was added to the kohlrabies and they were pickled for 5 days. Then, 5% (*w*/*w*) salt (upper layer 10%, middle layer 30%, and lower layer 60%) was added to the once-pickled kohlrabies and they were pickled for a further 4 days. Finally, all the fermentation tanks were completely sealed, and the pickled kohlrabies were fermented for 5 years (5Y) and 10 years (10Y). The

kohlrabies were fermented in 200 kg tanks. There were 1000–1600 kohlrabies per 200 kg tank, depending on the size of the kohlrabies. All tested samples were vacuum-packaged in aluminum foil bags and stored at −80 °C for further analysis.

2.2. Physicochemical Analysis

pH values were measured with a pH meter (PHS-3C, Shanghai Sanxin Instrument, Shanghai, China). Color measurements for redness (*a**), yellowness (*b**), and brightness (*L**) were performed using a colorimeter (RC-10, Konica Minolta, Tokyo, Japan). The total acid contents, reducing sugar contents, and soluble protein contents were respectively determined by the acid–base titration method [10], the direct titration of the alkaline copper tartrate solution method [11], and the spectrophotometry method [12]. The salt contents were determined by Dought's method [13].

2.3. Extraction of Metabolites from Ultra-Long-Term Industrially Fermented Kohlrabi

Extraction of metabolites from ultra-long-term industrially fermented kohlrabies was performed according to the method of Oliver and Tobias [14]. Fifty milligrams of kohlrabi was transferred into a 2 mL Eppendorf tube and 0.5 mL of solution (acetonitrile:isopropanol:water (3:3:2, $v/v/v$)) was added to each tube. Then, four zirconium beads with sizes of 2 mm were placed in the tube and the kohlrabi was ground using a high-throughput tissue grinder (Scientz Biotechnology, SCIENTZ-48, Ningbo, China). The sample was centrifuged at 12,000 rpm for 2 min after grinding, and 2 mL of the supernatant was transferred into a clean Eppendorf tube, which was subsequently dried in a vacuum concentrator. The dried powder was resuspended in 80 µL of 20 mg/mL methoxamine pyridine and incubated at 60 °C for 60 min. Finally, 100 µL of BSTFA-TMCS (v/v, 99:1) was added and incubated at 70 °C for 90 min after 30 s of vortexing. The resuspension was centrifuged at 14,000 rpm (Cence, H1850-R, Changsha, China) for 3 min, and 100 µL of the supernatant was transferred into a detection bottle.

2.4. Detection of Metabolites from Ultra-Long-Term Industrially Fermented Kohlrabi by GC-TOF-MS

GC-TOF-MS was performed as described in a previous study [15]. A DB-5MS capillary column (30 m × 250 µm i.d., 0.25 µm film thickness; Agilent J & W Scientific, Folsom, CA, USA) was used for gas chromatography. Derivatized species were separated at a constant flow of helium at 1 mL/min, and 1 µL of the sample was injected through the autosampler in a split ratio of 1:10. The inlet temperature was 280 °C, and the transfer line and ion source temperatures were 320 °C and 230 °C, respectively. The heating program was: initial temperature 50 °C, lasting 0.5 min, increased to 320 °C at a rate of 15 °C/min and held at 320 °C for 9 min. A full-scan method was used, with a scan rate of 10 spec/s, an electron energy of −70 V, and a solvent delay of 3 min.

2.5. GC-TOF-MS Data Preprocessing and Analysis

The obtained GC-TOF-MS raw data were converted into a common format by Analysis Base File Converter software and subsequently pre-processed by MSDIAL software. Differential metabolites were characterized according to molecular weight (molecular weight error < 30 ppm). Fragmentation information obtained in MS/MS mode was further matched and annotated in the HMDB, Metlin, Massbank, LipMaps, mzclound, and BioNovoGene (Suzhou, China) databases. The matrix containing metabolite names and peak intensities was processed according to missing-value imputation and normalization and subsequently imported into SIMCA (version 14.1) for PCA and OPLS-DA analysis.

2.6. Metabolic Pathway Analysis

Various public databases were searched, including the KEGG and PubChem databases, to further validate the details of the identified metabolites. The screened differential

metabolites were mapped to MetaboAnalyst 5.0 (www.metaboanalyst.ca, accessed on 10 March 2022) for related metabolic pathway analysis.

2.7. E-Nose Analysis

Electronic-nose analysis was performed using a Fox 4000 Sensory Array Fingerprint Analyzer (Alpha M.O.S., Toulouse, France), which contains 18 metal oxide semiconductors. A description for all 18 sensors was given in a previous study [16]. For E-nose analysis, 0.5 g of fermented kohlrabies was ground and transferred into a 10 mL headspace bottle. The balance time was 5 min at 70 °C. The temperature of the injection module was 70 °C. The injection speed was 500 μL/s, and the injection period was 1 s. The data-acquisition time was 120 s. The detection time was 180 s. A PCA plot of the E-nose was employed to further analyze all the samples. Five parallel experiments were performed for every sample.

2.8. Statistical Analysis

Student's *t*-tests were performed on metabolites with VIPs > 1 in the OPLS-DA model using IBM SPSS Statistics (version 24), and the significance level was set at $p < 0.05$. The fold change for each metabolite was calculated by EXCEL (version 2016). After log2 processing, the differential metabolites were defined with $|\log2\ FC| > 1.1$, VIP > 1, and $p < 0.05$ as the standard.

3. Results

3.1. Physicochemical Analysis

The physicochemical analyses of the three ultra-long-term industrially fermented kohlrabies are listed in Table 1. The pH and reducing sugar contents gradually decreased and the total acid contents thus constantly increased. The protein contents became lower and lower, which was possibly the result of proteolysis catalyzed by proteases. The colors of the industrially fermented kohlrabies changed significantly, as shown in Figure S1. The color of the 5Y fermented kohlrabi was nearly red, and its *L** value was smaller than that of the 0Y fermented kohlrabi but bigger than that of the 10Y fermented kohlrabi.

Table 1. Changes in physicochemical properties of ultra-long-term industrially fermented kohlrabies.

	0Y	5Y	10Y
pH	6.26 ± 0.01^{a}	3.59 ± 0.04^{b}	3.47 ± 0.02^{c}
Total acid (%)	0.26 ± 0.02^{a}	0.88 ± 0.06^{b}	0.91 ± 0.01^{b}
Reducing sugar (g/100 g)	5.47 ± 0.1^{a}	2.47 ± 0.01^{b}	2.05 ± 0.04^{c}
Protein (g/100 g)	2.68 ± 0.15^{a}	1.87 ± 0.13^{b}	1.24 ± 0.13^{c}
Salt content (%)	12.06 ± 0.2^{a}	12.64 ± 0.14^{b}	12.66 ± 0.16^{b}
Brightness (*L**)	56.31 ± 1.2^{a}	31.54 ± 1.08^{b}	22.03 ± 0.52^{b}
Redness (*a**)	3.47 ± 0.17^{a}	10.4 ± 0.19^{b}	5.75 ± 0.29^{c}
Yellowness (*b**)	20.78 ± 0.65^{a}	14.34 ± 0.4^{b}	9.64 ± 0.16^{c}

Different superscript letters in the same row indicate significant differences ($p < 0.05$).

3.2. PCA Analysis for the Metabolomics of Ultra-Long-Term Industrially Fermented Kohlrabies

Based on GC-TOF-MS combined with multivariate statistical methods, the differential metabolites of the three kohlrabies were explored. Twenty microliters of each extract from the three different kohlrabies were mixed together and the obtained mixture was used as a quality control (QC). QC is often required to obtain high-quality metabolomic data, as revealed in Figure 1. The aggregation degree of the QC was better than that of the 0Y, 5Y, and 10Y ultra-long-term fermented kohlrabies, indicating that the experimental data were reliable. All data for the 0Y, 5Y, and 10Y kohlrabies were distributed in different regions, indicating significant differences in the metabolites among the three different groups of fermented kohlrabi. R^2X and Q^2 were the main parameters for judging the PCA model,

$R^2X = 0.72 > 0.5$ and $Q^2 = 0.536$, suggesting that the PCA model could well interpret the significant differences among the differential metabolites in the three kohlrabies.

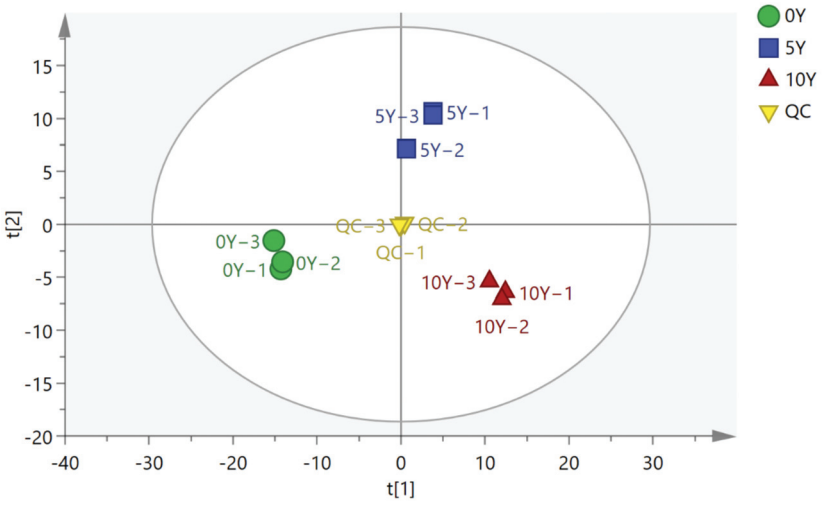


Figure 1. The scatter plots of the PCA model for the 0Y, 5Y, and 10Y fermented kohlrabi samples representing 0-year, 5-year and 10-year ultra-long-term industrially fermented kohlrabies, respectively.

3.3. OPLA-DA Analysis of Metabolites from Three Different Kohlrabies

For further analysis of the three treatment groups (A1, A2, and A3), OPLS-DA analysis was performed and obtained good separation, as shown in Figure 2. Figure 2A,C,E indicated the OPLS-DA score maps for the three kohlrabi groups, respectively. It was obvious that all groups were significantly separated, implying that the significant differences in the types and contents of metabolites from the three kohlrabies were caused by the different ultra-long-term fermentation periods. The OPLS-DA model parameters listed in Table 2 could explain and predict the differences between every two kohlrabi metabolites. In order to further verify the model was reliable, 200-loop-iteration permutation tests were performed, as can be seen in Figure 2B,D,F. All the intersections for the Q^2 regression line and Y-axis were on the negative half-axis, indicating that the OPLS-DA model was stable and reliable. Moreover, no overfitting phenomenon was observed.

Table 2. Parameters of the OPLS-DA models for the industrially fermented kohlrabi samples.

Group	R^2X	R^2Y	Q^2
A1 (0Y–5Y)	0.719	0.999	0.980
A2 (0Y–10Y)	0.775	1.000	0.994
A3 (5Y–10Y)	0.688	1.000	0.979

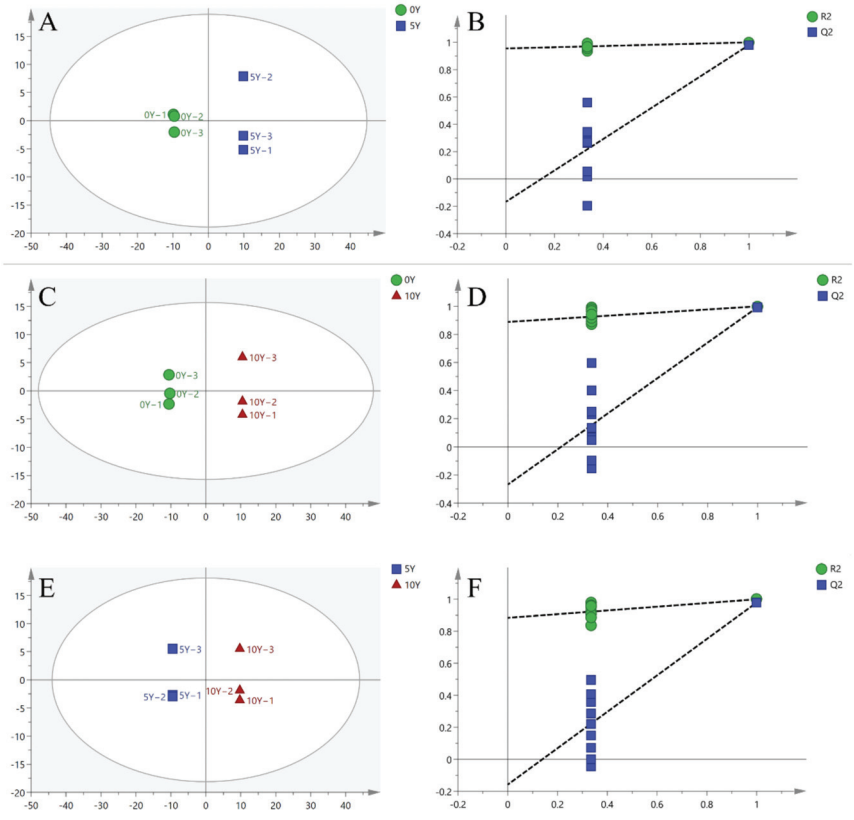


Figure 2. Scatter plots and permutation tests of the OPLS-DA models for fermented kohlrabi samples. (**A**,**C**,**E**) represent the OPLS-DA analyses for A1 (0Y–5Y), A2 (0Y–10Y), and A3 (5Y–10Y), while (**B**,**D**,**F**) represent permutation tests for A1 (0Y–5Y), A2 (0Y–10Y), and A3 (5Y–10Y).

3.4. Screening and Identification of Differential Metabolites from the Three Industrially Fermented Kohlrabies

In this study, VIP > 1 and $p < 0.05$ were used as the screening standards for differential metabolites from ultra-long-term industrially fermented kohlrabi. Based on the set standards, 47, 38, and 33 differential metabolites were respectively screened from group A1, A2, and A3, as shown in Figure 3A. Carbohydrates and carbohydrate conjugates, amino acids and their derivatives, and organic acids and their derivatives were the main differential metabolites in the three industrially fermented kohlrabies. Additionally, more organic acids were produced with the extension of fermentation time. The differential metabolites of the three industrially fermented kohlrabies were analyzed using a Venn diagram, as shown in Figure 3B. Thirteen major differential metabolites were identified from the three industrially fermented kohlrabies, including six carbohydrates and carbohydrate conjugates (2-amino-3,4,5,6-tetrahydroxyhexanal, 3,6-anhydrogalactose, *α*-D-quinolulose, *β*-lactose, D-(+)-cellobiose, methylpyran glycosides), one amino acid and one derivative (DL-aspartic acid, 2-amino-3-methylsuccinic acid), one nucleic acid derivative (uracil), one fatty acid derivative (lactitol), one alkaloid (leucine), and two other metabolites (2,3-butanediol, galactosylglycerol).

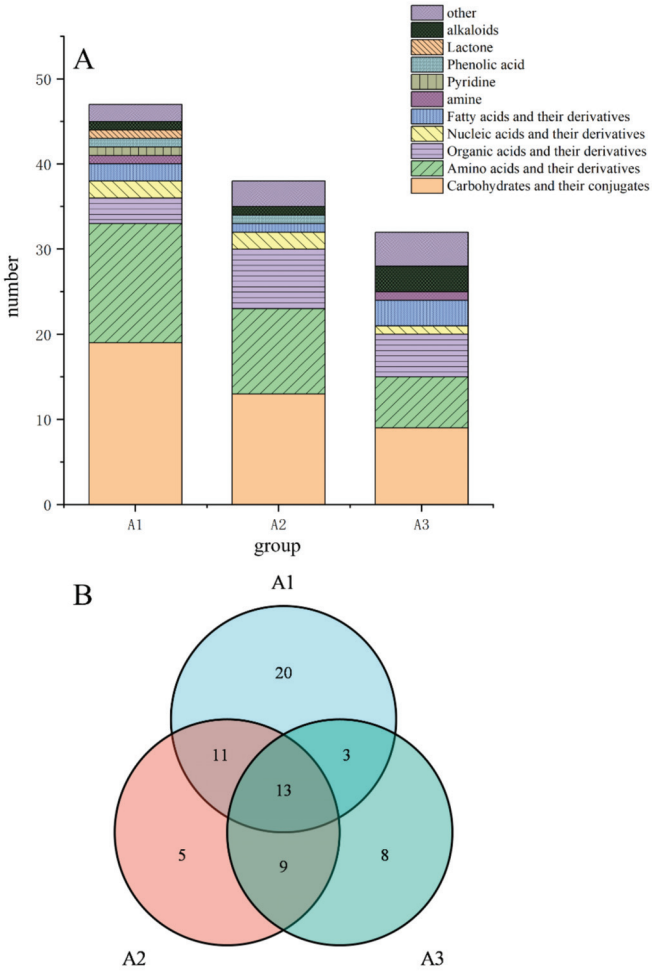


Figure 3. Classification information (**A**) and Venn diagram (**B**) for the differential metabolites in industrially fermented kohlrabi. (A1: Chinese kohlrabies industrially fermented for 0 years and 5 years, A2: Chinese kohlrabies industrially fermented for 0 years and 10 years, A3: Chinese kohlrabies industrially fermented for 5 years and 10 years.)

3.5. Heat Map Analysis for Differential Metabolites

In order to better evaluate the metabolic patterns of the main differential metabolites in industrially fermented kohlrabi, the relative contents of the 13 differential metabolites in the industrially fermented kohlrabies were analyzed by hierarchical clustering analysis and displayed in the form of a heat map, as shown in Figure 4. It was obvious that all the main differential metabolites increased with the extension of fermentation, except 2,3-butanediol. The main differential metabolites with higher contents in 10Y kohlrabi were β-lactose, lactitol, DL-aspartic acid, methylhexoside, 2-amino-3-methylsuccinic acid, 3,6-anhydrogalactose, D-(+)-cellobiose, α-D-quinovopirose, uracil, dictamnine, 2-amino-3,4,5,6-tetrahydroxyhexanal, and galactosylglycerol. However, the main differential metabolite with a higher content in 5Y kohlrabi was 2,3-butanediol.

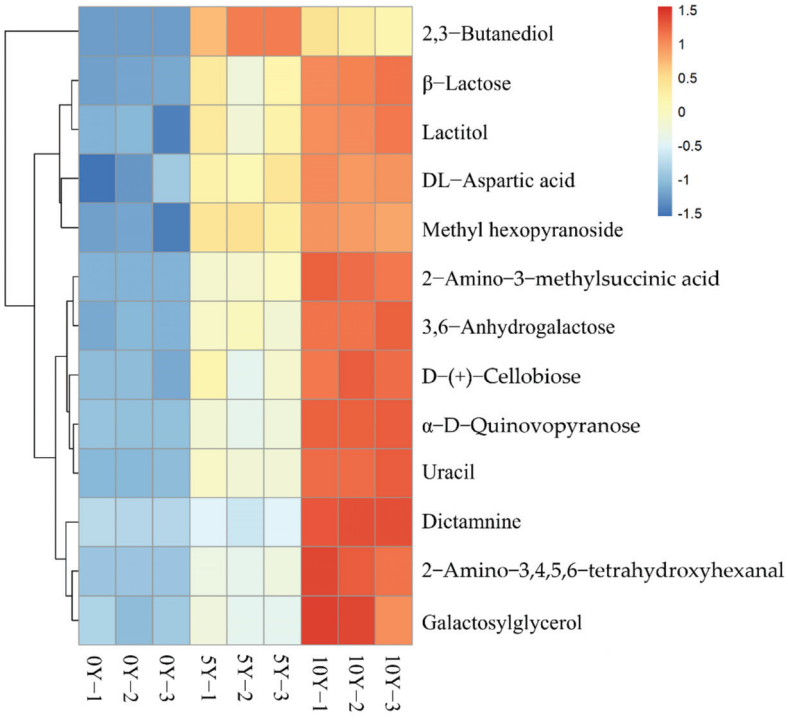


Figure 4. Hierarchical clustering analysis heat map of the main differential metabolites in ultra-long-term industrially fermented kohlrabi samples.

3.6. Metabolic Pathway Analysis for Differential Metabolites in Ultra-Long-Term Industrially Fermented Kohlrabi

The identified differential metabolites were imported into the KEGG database to obtain perturbed metabolic pathway information. Consequently, four main pathways were enriched, termed alanine, aspartate and glutamate metabolism (impact = 0.82), arginine biosynthesis (impact = 0.26), arginine and proline metabolism (impact = 0.24), and glycolysis/gluconeogenesis (impact = 0.13), as shown in Figure 5. Alanine and aspartic acid and glutamic acid were metabolized with L-aspartic acid as a raw material, which was converted into oxaloacetate under the catalysis of α-amino acid oxidase. 2-oxoglutaric acid was generated from oxaloacetate through the citric acid cycle, which was subsequently converted into L-glutamic acid, catalyzed by L-glutamic acid synthase. L-glutamic acid acted as a raw material in the arginine biosynthetic pathway and was catalyzed by glutamate dehydrogenase (NADP+) to generate NH_3, which was subsequently converted into L-citrulline under the catalysis of carbamoyl phosphate synthase, acetylornithine transcarbamylase, and acetylornithinease [17]. The metabolism of arginine and proline employed L-glutamic acid as raw material, which produced L-ornithine under the catalysis of glutamate kinase, glutamate semialdehyde dehydrogenase, and ornithine aminotransferase. The L-ornithine was catalyzed by ornithine decarboxylase to produce putrescine, which was subsequently converted into γ-aminobutyric acid through putrescine-pyruvate aminotransferase and the catalysis of other enzymes. Oxaloacetate generated by the citric acid cycle was used as the precursor for the glycolysis/gluconeogenesis pathway, which was subsequently converted into phosphoenolpyruvate under the catalysis of phosphoenolpyruvate carboxykinase. Phosphoenolpyruvate further generated pyruvate by phosphoenolpyruvate carboxyl synthase, which was finally converted into lactic acid, catalyzed by lactate dehydrogenase.

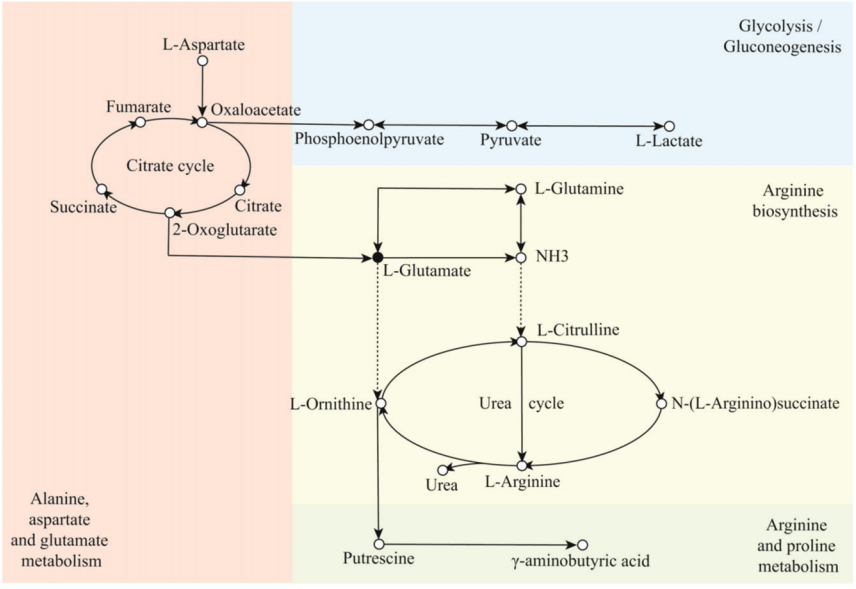


Figure 5. Integrated analysis diagram of the main metabolic pathways in ultra-long-term industrially fermented kohlrabi.

3.7. E-Nose Analysis

As analyzed by E-nose, the differences in nitrogen oxides, amines, sulfides, and acetone among the three different fermented kohlrabies were not significant, as revealed in Figure 6A. However, significant differences in polar substances, non-polar substances (hydrocarbons, ammonia, chlorine), aromatics (toluene, xylene), amines, and chlorines were observed in the three different industrially fermented kohlrabies. With the extension of fermentation, the sensor signals became stronger, suggesting that the contents of the related compounds increased, which results agree well with the changes in the metabolites.

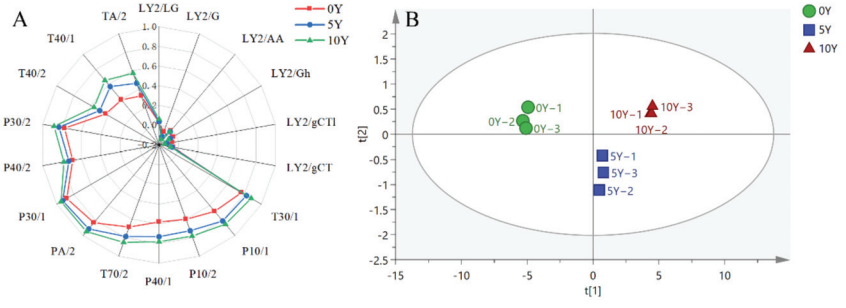


Figure 6. E-nose analysis. (**A**) Radar graph of the three different industrially fermented kohlrabies. Blue, red, and black indicate 0Y, 5Y, and 10Y industrially fermented kohlrabies, respectively. (**B**) PCA plot of the three different industrially fermented kohlrabies.

PCA is the method normally used to generate principal component variables to eliminate correlations between original feature variables. As shown in Figure 6B, the three different fermented kohlrabies were presented in a PCA spatial distribution map. The first two components (PC1 and PC2 were 96.70% and 2.08%, respectively) explained nearly 99% of the total variance, indicating that the PCA model can better reflect the overall

information for the samples [18]. The contribution of PC1 was much higher than that of PC2, implying that the farther the samples were on the horizontal axis, the greater the difference in smell. The spatial regions of these samples showed that the odors for the 0Y, 5Y, and 10Y industrially fermented kohlrabies were significantly different.

4. Discussion

During fermentation, the carbohydrates were metabolized by microorganisms to generate acids; consequently, the total acid contents increased and the pH and reducing sugar contents decreased [19]. Generally, the growth of most spoilage microorganisms will be inhibited when the pH is lower than 3.8 and the total acid is more than 0.5%, resulting in the safety of fermented vegetables. The flavor and metabolites will be significantly changed with the extension of fermentation. It has been well demonstrated that microorganisms continuously utilize the substrates in kohlrabi through enzymatic reactions to maintain their normal physiological functions [20], leading to the hydrolysis of proteins and fats to generate amino acids and fatty acids. As shown in Figure S2, during the fermentation process, the contents of organic acids and their derivatives, amino acids and their derivatives, and carbohydrates and carbohydrate conjugates increased with the extension of fermentation. It has been suggested that the contents of organic acids, amino acids, aldehydes, and ketones would significantly increase in fermented wild mustard greens with the extension of fermentation, which would promote the formation of unique flavors [21]. Organic acids are important factors in the flavors and aromas of fermented foods and play an irreplaceable role in the formation of flavor and overall sourness during the fermentation process. Lactic acid is the most important acid-tasting substance in Chinese paocai, which is produced by the anaerobic fermentation and metabolism of lactic acid bacteria. Lactic acid is an important flavor substance in fermented vegetables, which provides the fresh, sour taste to products. With the extension of fermentation time, the contents of lactic acid in the fermented kohlrabies increased. During the fermentation process of Chinese suancai, the pH kept decreasing and the total titratable acidity thus kept increasing [22]. Organic acids are considered the intermediate products of many life activities in plants and play important roles in regulating plant life activities. For example, salicylic acid and benzene-ring-containing carboxylic acids and their derivatives could systemically induce and produce resistance in a variety of plants [23]. Similarly, malic acid could increase tolerance to water stress in different plants [24]. On the other hand, organic acids could prevent the accumulation of reactive oxygen species in cells and give plants the ability to avoid oxidative stress [25] through maintaining intracellular pH, potential, and redox balance [26].

Amino acids are a class of important biologically active organic compounds. As shown in Figure S2, among the 69 differential metabolites detected, 10 free amino acids can be divided into 4 groups: sweet amino acids (L-ornithine), bitter amino acids (L-valine, L-isoleucine, L-leucine, L-methionine), tasty amino acids (DL-aspartic acid, L-glutamate), and tasteless amino acids (L-glutamine, DL-homoserine, γ-aminobutyric acid). During fermentation, sweet amino acid contents first increased and then decreased. Except for L-methionine, all the other bitter amino acids showed increasing trends in the A1 (0Y–5Y) group. However, their levels stabilized in the A3 (5Y–10Y) group. The fresh amino acids showed an upward trend during fermentation, which possibly explained why the 10Y fermented kohlrabi was suitable for soup. Besides their important contributions to the taste of fermented vegetables [27], free amino acids are also important flavor precursors. Some free amino acids produced by protein hydrolysis can be converted to small molecular compounds, such as aldehydes, ketones, and hydrocarbons, through a series of reactions, such as deamination and decarboxylation, which have important impacts on the taste of fermented food. The contents of free amino acids in kohlrabi changed significantly during the ultra-long-term fermentation, indicating that the fermentation period was closely related to product taste. Generally, the free amino acids generally showed a trend of first rising and then slightly declining during fermentation, which was consistent with

a previous study [28], possibly due to the conversion of free amino acids into volatile flavor substances by microorganisms. The contents of amino acids were generally affected by a variety of factors. It has been reported that wheat-grain free amino acids were affected by soil type and irrigation water salinity levels [29]. However, free amino acids in tomato plants were highly dependent on nitrogen formation and concentration in the rooting medium. Li and co-workers demonstrated that the contents of free amino acids in mustard were significantly affected by different light treatments. Furthermore, the contents of free amino acids in sprouted mustard greens were significantly higher than in mustard seeds. Meanwhile, light treatment promoted the synthesis of free amino acids in mustard [30]. The community structure of microorganisms and their physiological metabolism have a significant impact on amino acid contents in fermentation systems. Lactic acid bacteria decompose proteins into various free amino acids under the catalysis of secreted polypeptide enzymes [31]. Zhao et al. suggested that *Lactobacillus plantarum* and *Lactobacillus brucei* could significantly increase levels of glutamic acid, glycine, and gamma-aminobutyric acid in Chinese paocai, as measured by GC-MS [32]. Amino acids are used as the main biological macromolecular materials for building organisms and play important roles in plants, including building cells and repairing tissue damage, enhancing plant metabolism, and improving plant stress resistance. Amino acids could affect the growth of plants by influencing the coordinated distribution of nitrogen sources and nitrogen in plants. Increasing leaf nitrogen allocation could significantly enhance leaf nitrogen fixation and nitrogen-use efficiency by photosynthesis [33].

Amino acids also play a positive role in the stress response of many plants as a kind of regulator. For example, proline, tryptophan, leucine, isoleucine, and valine were considered to be related to drought stress in plants [34]. However, L-tryptophan, L-glutamic acid, and L-phenylalanine played important roles in salt stress in Tibetan highland barley [35]. It is well known that glutamate and homocysteine participate in cadmium stress in mustard leaves [36]. On the other hand, amino acids have effects on plant flowering under stress conditions. For example, aspartic acid, glutamic acid, alanine, glycine, and serine could promote plant flowering, while cysteine, threonine, and phenylalanine inhibit plant flowering [37].

Carbohydrates participate in the physiological reaction processes of cells in plants and can also be used as structural components and metabolic components in living organisms [38]. For example, carbohydrates act as important components of signaling pathways that connect and construct networks of pathways to control metabolic responses in plants. It has been suggested that carbohydrates induce stress responses and increase plant resistance to abiotic stress [39].

Secondary metabolites are a class of organic compounds with complex structures which are synthesized through a series of metabolic pathways using a large number of primary metabolites as precursors, including alkaloids, pigments, antibiotics, phenolic acids, and other substances [40]. Benzoate was the only differential metabolite among phenolic acids detected in ultra-long-term industrially fermented kohlrabi. Under enzymatic reactions, phosphoenolpyruvate and erythrose-4-phosphate generate phenylalanine through the shikimate pathway, which is subsequently converted into trans-cinnamic acid, benzoic acid, salicylic acid, and 4-hydroxychalcone, and thus enters the flavonoid metabolic pathway [41]. It has been suggested that alkaloids play important anti-tumor and analgesic roles, among which are a class of nitrogen-containing basic organic compounds, mainly including isoquinolines, pyridines, etc. [42]. In this study, two differential metabolites of alkaloids, dictamnine and 2-hydroxypyridine, were detected in ultra-long-term industrially fermented kohlrabi. Furthermore, the contents of the two alkaloids were dramatically increased during the fermentation process. Dictamnine is a small-molecule alkaloid with various biological activities, which is normally synthesized through the shikimic acid pathway and can be extracted from various plants. On the other hand, dictamnine has many beneficial biological activities, such as anti-platelet aggregation, anti-hypertension, antibacterial, anti-mitotic, and anti-fungal activities [43].

The change in 2,3-butanediol observed in this work was consistent with a previous study in which, with the extension of fermentation, the content of 2,3-butanediol in Chinese suancai first increased and then decreased [44]. Carbohydrates could generate acetoin through a series of enzymatic reactions, which is subsequently converted into by-products, such as 2,3-butanediol and lactic acid, under the action of acetoin reductase in a reversible way [45]. Xiong et al. indicated that during the fermentation process of Chinese paocai, lactic acid bacteria consumed a large amount of sucrose and glucose to produce lactic acid. The accumulation of lactic acid resulted in a decrease in the content of 2,3-butanediol [46], which finding possibly supports our result that the content of 2,3-butanediol in 10Y ultra-long-term industrially fermented kohlrabi was lower than that in 5Y ultra-long-term industrially fermented kohlrabi. 2,3-Butanediol was the common differential metabolite among the three ultra-long-term fermented kohlrabies, and its contents were significantly different among the three kinds of fermented kohlrabies, as revealed in Figure 4, such that it might be used as an important biomarker to distinguish the different ultra-long-term fermented kohlrabies.

The 5Y and 10Y kohlrabies are respectively used for making soup and Chinese braised dishes, which possibly correlates with the metabolites in the two different kohlrabies. The contents of dictamnine, galactosylglycerol, linoleate, L-Rhamnose, β-Lactose, (9Z,12Z,15Z)-Octadecatrienoic acid, galacturonan, lactitol, L-Isoleucine, and threonate continuously increased, with the highest level in 10Y kohlrabi (Figure S2), among which the dictamnine, galactosylglycerol, L-rhamnose, and (9Z,12Z,15Z)-octadecatrienoic acid possess the function of inhibiting cancer cell proliferation [47]. Moreover, linoleate, galacturonan, and hexadecanoic acid exhibit antibacterial and anti-inflammatory activities [48]. Lactitol and L-isoleucine are used to treat liver disease and hyperinsulinism [49,50]. Threonate has a certain therapeutic effect on Alzheimer's disease [51]. L-rhamnose, lactitol, lactic acid, linoleate, hexadecanoic acid, and pentose are flavor substances, with sweet, sour, fatty, and smoky-sweet-salty-coffee tastes, respectively, playing very important roles in flavor formation in 10Y kohlrabi. However, the levels of D-tagatose, L-leucine, tranexamic acid, D-(−)-fructose and 2,3-butanediol were higher in 5Y kohlrabi than in the 0Y and 10Y kohlrabies (Figure S2), among which D-tagatose, L-leucine, and tranexamic acid have the function of promoting weight loss, neuron regeneration, and hemostasis, respectively [52–54], while D-(−)-fructose brings a clean sweet taste and 2,3-butanediol gives the fruity and creamy aroma which contributes to the flavor of 5Y kohlrabi.

The colors of the three different kohlrabies were significantly different, as can be seen in Figure S1. Tyrosine was detected in all the three different kohlrabies and its contents gradually decreased, with the lowest level in 10Y kohlrabi (Figure S2). Tyrosine can be converted into dopa under the catalysis of tyrosinase, which can generate melanin through oxidation [55] and probably has a potential impact on the color of kohlrabi.

5. Conclusions

In the present study, GC-TOF-MS combined with multivariate statistical analysis was used to determine the differential metabolites in Chinese ultra-long-term industrially fermented kohlrabies that underwent fermentation for different periods and to clarify their associated metabolic pathways. There were 47, 38, and 33 differential metabolites in the three groups of kohlrabi industrially fermented for 0 year and 5 years, 0 year and 10 years, and 5 years and 10 years, respectively (VIP > 1, $p < 0.05$). The differential metabolites were mainly carbohydrates, amino acids, and organic acids, and there were 13 common differential metabolites, including L-glutamic acid, L-aspartic acid, γ-aminobutyric acid, etc. Through KEGG metabolic pathway analysis, four metabolic pathways termed alanine, aspartate and glutamate metabolism, arginine biosynthesis, arginine and proline metabolism, and glycolysis/gluconeogenesis were finally identified. The correlation with differential metabolites was the most significant, which may provide some theoretical support for subsequent research on fermented kohlrabi. Traditionally fermented kohlrabi undergoes complex biochemical reactions during the fermentation process. The metabolic

pathways are normally complex, and the metabolites are numerous. Moreover, the odors of the three ultra-long-term industrially fermented kohlrabies were significantly different. Further in-depth analysis of the changes in the metabolites and metabolome of kohlrabi during fermentation is required.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12100991/s1, Figure S1: The appearance of the Chinese fermented kohlrabi. A, B, and C, respectively, indicated 0Y, 5Y, and 10Y fermented kohlrabi; Figure S2: Clustering heat map of differential metabolites of industrially fermented kohlrabi with different ultra-long-term fermentation periods.

Author Contributions: X.N., conceptualization, formal analysis, methodology, writing—original draft; H.C., data curation, formal analysis, methodology; L.X., writing—review and editing; Y.Z., formal analysis; D.L., formal analysis; Z.Z., conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, validation, visualization, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the Science and Technology Department of Sichuan Province (2022NSFSC1702, 2021YFQ0072, 2021YJ0275), the Solid State Fermentation Resource Utilization Key Laboratory of Sichuan Province (2019GTY004), and the Guangdong Province Key Laboratory of Marine Biotechnology (GPKLMB202104).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the results of this study are included in the present article.

Acknowledgments: We thank Huailiang Luo, Chief Manager of the Zigong Taifu Agricultural and Sideline Products Processing Plant, for supplying all the kohlrabi samples.

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

References

1. Pasko, P.; Galanty, A.; Żmudzki, P.; Gdula-Argasińska, J.; Zagrodzki, P. Influence of Different Light Conditions and Time of Sprouting on Harmful and Beneficial Aspects of Rutabaga Sprouts in Comparison to Their Roots and Seeds. *J. Sci. Food Agric* **2018**, *99*, 302–308. [CrossRef] [PubMed]
2. Kapusta-Duch, J.; Florkiewicz, A.; Leszczyńska, T.; Borczak, B. Directions of Changes in the Content of Selected Macro- and Micronutrients of Kale, Rutabaga, Green and Purple Cauliflower Due to Hydrothermal Treatment. *Appl. Sci.* **2021**, *11*, 3452. [CrossRef]
3. Yang, Z.; Duan, X.; Yang, J.; Wang, H.; Liu, F.; Xu, X.; Pan, S. Effects of high hydrostatic pressure and thermal treatment on texture properties of pickled kohlrabi. *LWT* **2022**, *157*, 113078. [CrossRef]
4. Jiang, L.; Xian, S.; Liu, X.; Shen, G.; Zhang, Z.; Hou, X.; Chen, A. Metagenomic Study on Chinese Homemade Paocai: The Effects of Raw Materials and Fermentation Periods on the Microbial Ecology and Volatile Components. *Foods* **2021**, *11*, 62. [CrossRef] [PubMed]
5. Zhang, J.; Zhang, C.; Xin, X.; Liu, D.; Zhang, W. Comparative Analysis of Traditional and Modern Fermentation for Xuecai and Correlations Between Volatile Flavor Compounds and Bacterial Community. *Front. Microbiol.* **2021**, *12*, 631054. [CrossRef] [PubMed]
6. Liu, S.; Wu, Y.; Chan, L. Application of Metabonomics Approach in Food Safety Research-A Review. *Food Rev. Int.* **2020**, *36*, 547–558. [CrossRef]
7. Ding, M.; Cheng, J.; Xiao, W.; Qiao, B.; Yuan, Y. Comparative metabolomic analysis on industrial continuous and batch ethanol fermentation processes by GC-TOF-MS. *Metabolomics* **2008**, *5*, 229–238. [CrossRef]
8. Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A.R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* **2006**, *1*, 387–396. [CrossRef] [PubMed]
9. Zhu, G.; Wang, S.; Huang, Z.; Zhang, S.; Liao, Q.; Zhang, C.; Lin, T.; Qin, M.; Peng, M.; Yang, C.; et al. Rewiring of the Fruit Metabolome in Tomato Breeding. *Cell* **2018**, *172*, 249–261.e212. [CrossRef]
10. *GB/T 12456—2008*; Determination of Total Acid in Foods. State Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, China National Standardization Administration Committee. China Standards Press: Beijing, China, 2008.

11. *GB/T 5009.7—2016*; Determination of Reducing Sugar in Foods. State Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, China National Standardization Administration Committee. China Standards Press: Beijing, China, 2016.
12. *GB/T 5009.5—2016*; Determination of Protein in Foods. State Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, China National Standardization Administration Committee. China Standards Press: Beijing, China, 2016.
13. Doughty, H.W. Mohr's method for the determination of silver and halogens in other than neutral solutions. *J. Am. Chem. Soc.* **1924**, *46*, 2707–2709. [CrossRef]
14. Fiehn, O.; Kind, T. Metabolite profiling in blood plasma. *Methods Mol. Biol.* **2007**, *358*, 3–17. [CrossRef]
15. Fiehn, O.; Wohlgemuth, G.; Scholz, M.; Kind, T.; Lee, D.Y.; Lu, Y.; Moon, S.; Nikolau, B. Quality control for plant metabolomics: Reporting MSI-compliant studies. *Plant J.* **2008**, *53*, 691–704. [CrossRef] [PubMed]
16. Chen, J.; Tao, L.; Zhang, T.; Zhang, J.; Wu, T.; Luan, D.; Ni, L.; Wang, X.; Zhong, J. Effect of four types of thermal processing methods on the aroma profiles of acidity regulator-treated tilapia muscles using E-nose, HS-SPME-GC-MS, and HS-GC-IMS. *LWT* **2021**, *147*, 111585. [CrossRef]
17. Minocha, R.; Majumdar, R.; Minocha, S.C. Polyamines and abiotic stress in plants: A complex relationship. *Front. Plant Sci.* **2014**, *5*, 175. [CrossRef]
18. Xiao, Y.; Zhang, Y.; Guo, B.; Chen, H.; Sun, B.; Zhang, Y. Comparison of Volatile Flavor Compounds in Yu-Shiang Shredded Pork Processed by Two Different Methods by SDE-GC-MS. *Food Sci.* **2015**, *36*, 70–75. [CrossRef]
19. Wu, R.; Yu, M.; Liu, X.; Meng, L.; Wang, Q.; Xue, Y.; Wu, J.; Yue, X. Changes in flavour and microbial diversity during natural fermentation of suan-cai, a traditional food made in Northeast China. *Int. J. Food Microbiol.* **2015**, *211*, 23–31. [CrossRef] [PubMed]
20. Dong, Y.; Jiang, H. Microbial production of metabolites and associated enzymatic reactions under high pressure. *World J. Microbiol. Biotechnol.* **2016**, *32*, 178. [CrossRef]
21. Liu, D.; Zhang, C.; Zhang, J.; Xin, X.; Liao, X. Metagenomics reveals the formation mechanism of flavor metabolites during the spontaneous fermentation of potherb mustard (*Brassica juncea* var. multiceps). *Food Res. Int.* **2021**, *148*, 110622. [CrossRef] [PubMed]
22. Zhang, A.; Zhang, Z.; Zhang, K.; Liu, X.; Lin, X.; Zhang, Z.; Bao, T.; Feng, Z. Nutrient consumption patterns of *Lactobacillus plantarum* and their application in suancai. *Int. J. Food Microbiol.* **2021**, *354*, 109317. [CrossRef]
23. Zhang, Y.; Xu, S.; Ding, P.; Wang, D.; Cheng, Y.; He, J.; Gao, M.; Xu, F.; Li, Y.; Zhu, Z.; et al. Control of salicylic acid synthesis and systemic acquired resistance by two members of a plant-specific family of transcription factors. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18220–18225. [CrossRef] [PubMed]
24. Gao, Q.; Wu, C.; Wang, M.; Xu, B.; Du, L. Effect of drying of jujubes (*Ziziphus jujuba* Mill.) on the contents of sugars, organic acids, α-tocopherol, β-carotene, and phenolic compounds. *J. Agric. Food Chem.* **2012**, *60*, 9642–9648. [CrossRef] [PubMed]
25. Airaki, M.; Leterrier, M.; Mateos, R.M.; Valderrama, R.; Chaki, M.; Barroso, J.B.; Del Río, L.A.; Palma, J.M.; Corpas, F.J. Metabolism of reactive oxygen species and reactive nitrogen species in pepper (*Capsicum annuum* L.) plants under low temperature stress. *Plant Cell. Environ.* **2012**, *35*, 281–295. [CrossRef] [PubMed]
26. Igamberdiev, A.U.; Eprintsev, A.T. Organic Acids: The Pools of Fixed Carbon Involved in Redox Regulation and Energy Balance in Higher Plants. *Front. Plant Sci.* **2016**, *7*, 1042. [CrossRef]
27. Shang, Z.; Ye, Z.; Li, M.; Ren, H.; Cai, S.; Hu, X.; Yi, J. Dynamics of microbial communities, flavor, and physicochemical properties of pickled chayote during an industrial-scale natural fermentation: Correlation between microorganisms and metabolites. *Food Chem.* **2022**, *377*, 132004. [CrossRef] [PubMed]
28. Jeong, S.H.; Jung, J.Y.; Lee, S.H.; Jin, H.M.; Jeon, C.O. Microbial succession and metabolite changes during fermentation of dongchimi, traditional Korean watery kimchi. *Int. J. Food Microbiol.* **2013**, *164*, 46–53. [CrossRef]
29. Labanauskas, C.K.; Stolzy, L.H.; Handy, M.F. Protein and free amino acids in wheat grain as affected by soil types and salinity levels in irrigation water. *Plant Soil.* **1981**, *59*, 299–316. [CrossRef]
30. Li, X.; Kim, Y.B.; Uddin, M.R.; Lee, S.; Kim, S.J.; Park, S.U. Influence of light on the free amino acid content and γ-aminobutyric acid synthesis in *Brassica juncea* seedlings. *J. Agric. Food Chem.* **2013**, *61*, 8624–8631. [CrossRef] [PubMed]
31. Lu, W.; Wen, Z.; Li, H.; Yuan, D.; Li, J.; Zhang, H.; Huang, Z.; Cui, S.; Du, W. Identification of the quantitative trait loci (QTL) underlying water soluble protein content in soybean. *Theor. Appl. Genet.* **2013**, *126*, 425–433. [CrossRef]
32. Zhao, N.; Zhang, C.; Yang, Q.; Guo, Z.; Yang, B.; Lu, W.; Li, D.; Tian, F.; Liu, X.; Zhang, H.; et al. Selection of Taste Markers Related to Lactic Acid Bacteria Microflora Metabolism for Chinese Traditional Paocai: A Gas Chromatography-Mass Spectrometry-Based Metabolomics Approach. *J. Agric. Food Chem.* **2016**, *64*, 2415–2422. [CrossRef]
33. Perchlik, M.; Tegeder, M. Leaf Amino Acid Supply Affects Photosynthetic and Plant Nitrogen Use Efficiency under Nitrogen Stress. *Plant Physiol.* **2018**, *178*, 174–188. [CrossRef] [PubMed]
34. Khan, N.; Ali, S.; Zandi, P.; Mehmood, A.; Ullah, S.; Ikram, M.; Ismail, I.; Shahid, M.; Babar, M. Role of sugars, amino acids and organic acids in improving plant abiotic stress tolerance. *Pak J. Bot.* **2020**, *52*, 355–363. [CrossRef]
35. Wang, Y.; Zeng, X.; Xu, Q.; Mei, X.; Yuan, H.; Jiabu, D.; Sang, Z.; Nyima, T. Metabolite profiling in two contrasting Tibetan hulless barley cultivars revealed the core salt-responsive metabolome and key salt-tolerance biomarkers. *AoB Plants* **2019**, *11*, plz021. [CrossRef] [PubMed]

36. D'Alessandro, A.; Taamalli, M.; Gevi, F.; Timperio, A.M.; Zolla, L.; Ghnaya, T. Cadmium stress responses in *Brassica juncea*: Hints from proteomics and metabolomics. *J. Proteome Res.* **2013**, *12*, 4979–4997. [CrossRef] [PubMed]
37. Tanaka, O.; Nakayama, Y.; Emori, K.; Takeba, G.; Sugino, M. Flower-Inducing Activity of Lysine in *Lemna paucicostata* 6746. *Plant Cell Physiol* **1997**, *38*, 124–128. [CrossRef]
38. Ahmad, P.; Abdel Latef, A.A.; Hashem, A.; Abd Allah, E.F.; Gucel, S.; Tran, L.-S.P. Nitric Oxide Mitigates Salt Stress by Regulating Levels of Osmolytes and Antioxidant Enzymes in Chickpea. *Front. Plant Sci.* **2016**, *7*, 347. [CrossRef]
39. Sami, F.; Yusuf, M.; Faizan, M.; Faraz, A.; Hayat, S. Role of sugars under abiotic stress. *Plant Physiol. Biochem.* **2016**, *109*, 54–61. [CrossRef]
40. Shitan, N. Secondary metabolites in plants: Transport and self-tolerance mechanisms. *Biosci. Biotechnol. Biochem.* **2016**, *80*, 1283–1293. [CrossRef]
41. Abu-Reidah, I.M.; Arráez-Román, D.; Al-Nuri, M.; Warad, I.; Segura-Carretero, A. Untargeted metabolite profiling and phytochemical analysis of *Micromeria fruticosa* L. (Lamiaceae) leaves. *Food Chem.* **2019**, *279*, 128–143. [CrossRef]
42. Zhang, J.; Morris-Natschke, S.L.; Ma, D.; Shang, X.; Yang, C.; Liu, Y.; Lee, K.H. Biologically active indolizidine alkaloids. *Med. Res. Rev.* **2021**, *41*, 928–960. [CrossRef]
43. Wang, P.; Zhao, Y.; Zhu, Y.; Sun, J.; Yerke, A.; Sang, S.; Yu, Z. Metabolism of dictamnine in liver microsomes from mouse, rat, dog, monkey, and human. *J. Pharm. Biomed.* **2016**, *119*, 166–174. [CrossRef]
44. Zhao, D.; Du, R.; Song, G.; Sun, J.; Ping, W.; Ge, J. Chemical Composition and Bacterial Diversity of *Lactobacillus casei 11MZ-5-1* Fermented Chinese Cabbage Pickle. *Food Sci.* **2018**, *39*, 116–122. [CrossRef]
45. Tong, Y.; Wu, W.; Peng, H.; Liu, L.; Huang, H.; Ji, X. Metabolic engineering for efficient microbial production of 2,3-butanediol. *CIESC J.* **2016**, *67*, 2656–2671. [CrossRef]
46. Xiong, T.; Li, J.; Liang, F.; Wang, Y.; Guan, Q.; Xie, M. Effects of salt concentration on Chinese sauerkraut fermentation. *LWT* **2016**, *69*, 169–174. [CrossRef]
47. Dátilo, M.N.; Sant'Ana, M.R.; Formigari, G.P.; Rodrigues, P.B.; de Moura, L.P.; da Silva, A.S.R.; Ropelle, E.R.; Pauli, J.R.; Cintra, D.E. Omega-3 from Flaxseed Oil Protects Obese Mice Against Diabetic Retinopathy Through GPR120 Receptor. *Sci. Rep.* **2018**, *8*, 14318. [CrossRef]
48. Liu, Q.; Yin, G.; Zhang, J. The suitable condition for *Lactobacillus parafarraginis ZH1* producing hexadecanoic acid and inhibiting pathogenic and spoilage yeasts. *Biol. Control* **2020**, *149*, 104318. [CrossRef]
49. Martínez-Monteagudo, S.I.; Enteshari, M.; Metzger, L. Lactitol: Production, properties, and applications. *Trends Food Sci. Technol.* **2019**, *83*, 181–191. [CrossRef]
50. Gart, E.; van Duyvenvoorde, W.; Caspers, M.P.M.; van Trigt, N.; Snabel, J.; Menke, A.; Keijer, J.; Salic, K.; Morrison, M.C.; Kleemann, R. Intervention with isoleucine or valine corrects hyperinsulinemia and reduces intrahepatic diacylglycerols, liver steatosis, and inflammation in Ldlr-/-.Leiden mice with manifest obesity-associated NASH. *FASEB J.* **2022**, *36*, e22435. [CrossRef]
51. Sun, Q.; Weinger, J.G.; Mao, F.; Liu, G. Regulation of structural and functional synapse density by L-threonate through modulation of intraneuronal magnesium concentration. *Neuropharmacology* **2016**, *108*, 426–439. [CrossRef]
52. Jayamuthunagai, J.; Gautam, P.; Srisowmeya, G.; Chakravarthy, M. Biocatalytic production of D-tagatose: A potential rare sugar with versatile applications. *Crit. Rev. Food Sci. Nutr.* **2017**, *57*, 3430–3437. [CrossRef]
53. Ma, C.; Teng, L.; Lin, G.; Guo, B.; Zhuo, R.; Qian, X.; Guan, T.; Wu, R.; Liu, Y.; Liu, M. L-leucine promotes axonal outgrowth and regeneration via mTOR activation. *FASEB J.* **2021**, *35*, e21526. [CrossRef]
54. Hunt, B.J. The current place of tranexamic acid in the management of bleeding. *Anaesthesia* **2015**, *70*, 50–53.e18. [CrossRef] [PubMed]
55. Park, J.; Jung, H.; Kim, K.; Lim, K.M.; Kim, J.Y.; Jho, E.H.; Oh, E.S. D-tyrosine negatively regulates melanin synthesis by competitively inhibiting tyrosinase activity. *Pigment. Cell. Melanoma Res.* **2018**, *31*, 374–383. [CrossRef] [PubMed]

metabolites

MDPI

Article

Integrated Metabolomic and Transcriptomic Analyses Reveal the Basis for Carotenoid Biosynthesis in Sweet Potato (*Ipomoea batatas* (L.) Lam.) Storage Roots

Qingming Ren [1,†], Xiaoxi Zhen [1,2,†], Huiyu Gao [1], Yinpei Liang [1,2], Hongying Li [1,2], Juan Zhao [1], Meiqiang Yin [1], Yuanhuai Han [1,2,*] and Bin Zhang [1,2,3,*]

[1] College of Agriculture, Shanxi Agricultural University, Jinzhong 030801, China
[2] Shanxi Key Laboratory of Minor Crops Germplasm Innovation and Molecular Breeding, Shanxi Agricultural University, Jinzhong 030801, China
[3] Ministerial and Provincial Co-Innovation Centre for Endemic Crops Production with High-Quality and Efficiency in Loess Plateau, Shanxi Agricultural University, Jinzhong 030801, China
* Correspondence: binzhang@sxau.edu.cn (B.Z.); hanyuanhuai@sxau.edu.cn (Y.H.)
† These authors contributed equally to this work.

Abstract: Carotenoids are important compounds of quality and coloration within sweet potato storage roots, but the mechanisms that govern the accumulation of these carotenoids remain poorly understood. In this study, metabolomic and transcriptomic analyses of carotenoids were performed using young storage roots (S2) and old storage roots (S4) from white-fleshed (variety S19) and yellow-fleshed (variety BS) sweet potato types. S19 storage roots exhibited significantly lower total carotenoid levels relative to BS storage roots, and different numbers of carotenoid types were detected in the BS-S2, BS-S4, S19-S2, and S19-S4 samples. β-cryptoxanthin was identified as a potential key driver of differences in root coloration between the S19 and BS types. Combined transcriptomic and metabolomic analyses revealed significant co-annotation of the carotenoid and abscisic acid (ABA) metabolic pathways, *PSY* (phytoene synthase), *CHYB* (β-carotene 3-hydroxylase), *ZEP* (zeaxanthin epoxidase), *NCED3* (9-cis-epoxycarotenoid dioxygenase 3), *ABA2* (xanthoxin dehydrogenase), and *CYP707A* (abscisic acid 8′-hydroxylase) genes were found to be closely associated with carotenoid and ABA content in these sweet potato storage roots. The expression patterns of the transcription factors OFP and FAR1 were associated with the ABA content in these two sweet potato types. Together, these results provide a valuable foundation for understanding the mechanisms governing carotenoid biosynthesis in storage roots, and offer a theoretical basis for sweet potato breeding and management.

Keywords: sweet potato; carotenoid; abscisic acid; metabolome; transcriptome

Citation: Ren, Q.; Zhen, X.; Gao, H.; Liang, Y.; Li, H.; Zhao, J.; Yin, M.; Han, Y.; Zhang, B. Integrated Metabolomic and Transcriptomic Analyses Reveal the Basis for Carotenoid Biosynthesis in Sweet Potato (*Ipomoea batatas* (L.) Lam.) Storage Roots. *Metabolites* **2022**, *12*, 1010. https://doi.org/10.3390/metabo12111010

Academic Editors: Yanjie Zhang and Yan Li

Received: 29 September 2022
Accepted: 21 October 2022
Published: 23 October 2022

1. Introduction

Sweet potato (*Ipomoea batatas* (L.) Lam.) is an annual worldwide member of the Convolvulaceae family, representing an important food source and cash crop throughout the world [1]. Sweet potato storage roots contain high levels of carbohydrates, minerals, dietary fiber, and bioactive ingredients [2]. Yellow-fleshed sweet potato types are rich in β-carotene and other provitamin A carotenoids [3]. The color of sweet potato storage roots is considered highly important, in part because it can inform consumer choices [4]. There is thus a clear need to fully explore the interplay between carotenoid biosynthesis and storage root color development in order to better guide the selective breeding of sweet potatoes with a high level of commercial value.

Carotenoids are isoprenoid-derived pigments that are vital for the growth of the plants [5]. In addition to determining the coloration of the flesh of fruits and roots in which they accumulate, these carotenoids can exert a range of antioxidant and antitumor activities in humans, with some studies suggesting that these bioactive compounds can prevent a

variety of ocular diseases [6,7]. The pigmentation of sweet potato storage roots is primarily determined by the apparent carotenoid and anthocyanin discrepancy [8,9]. While orange- and yellow-fleshed storage roots contain similar carotenoids, the actual differing levels of particular carotenoids within the roots of different sweet potato types have been extensively studied in many plant species. [10–12].

Carotenoid biosynthesis has been studied at length in plants including *Arabidopsis* [13], tomatoes [14], peppers [15], rice [16], and maize [17]. This process is governed by a series of controlled reactions including condensation [14], dehydrogenation [18], cyclization [19], hydroxylation [20], and epoxidation [17]. The phytoene synthase (PSY)-mediated condensation of 20-carbon geranylgeranyl diphosphate (GGPP) molecules to generate colorless phytoene (C40) represents a critical step in this biosynthetic process [21]. Lycopene is then generated through successive phytoene desaturation and isomerization. Phytoene desaturase (PDS), z-carotene isomerase (Z-ISO), z-carotene desaturase (ZDS), and carotenoid isomerase (CRTISO) catalyze four different dehydrogenation reactions [22]. During the process of carotenoid cyclization reactions, competition between lycopene beta cyclase (LCY-β) and lycopene epsilon cyclase (LCY-ε) governs the relative production of β-carotene and α-carotene [12], with the double hydroxylation of these two respective products yielding lutein and zeaxanthin, after which further modification of zeaxanthin can yield neoxanthin and violaxanthin [17]. These latter two carotenoids serve as precursors for the biosynthesis of the phytohormone abscisic acid (ABA), which relies on the enzymatic cleavage of 9-cis-violaxanthin or 9′-cis-neoxanthin (9-cis-epoxy-xanthophylls), mediated by NCED, to produce C15-xanthoxin and C25-apocarotenoid [23]. After transfer from plastids into the cytoplasm, xanthoxin then undergoes two further processing steps to yield ABA, whereas apocarotenoid can be converted by ABA2 (short-chain dehydrogenase/reductase-like enzyme ABA-deficient 2), after which it can be oxidated by AAO (ABA by abscisic aldehyde oxidase) to yield ABA [24]. Further metabolic processing of ABA can then be performed by enzymes, including ABA-8′-hydroxylases and CYP enzymes of the 707A clade, producing major catabolites, including phaseic acid (PA) and dihydrophaseic acid (DPA) [25].

Carotenoid production is primarily shaped by the expression of genes that regulate the different steps in these biosynthetic pathways [26], with differential gene expression ultimately accounting for the observed differences in carotenoid content among species and types. In yellow celery, lutein and β-carotene content is closely associated with AgLCYB and AgPSY2 expression levels [20]. Analyses of mutant Oranzheva kaoia types, exhibiting high concentrations of β-carotene, have revealed the presence of a biosynthetic pathway breakdown owing to CrtZchr03 gene deletion, ultimately contributing to increased β-carotene content and affirming the regulatory role of particular genes in this metabolic context [27]. Studies of mutant Cara cara navel oranges have further demonstrated that the DXS1, DXR, GGPPS2, PSY1, and LCYB genes are the primary determinants of carotenoid biosynthesis [9].

Comprehensive multi-omics analyses have been widely employed to study the mechanisms whereby plant pigmentation is established. In *Cyclocarya paliurus*, for example, the MYB transcription factors (TFs) and two bHLH TFs were identified via a multi-omics approach as important regulators of flavonoid biosynthesis [28]. Similarly, integrative metabolomic and transcriptomic analyses have enabled the determination of the mechanisms that regulate anthocyanin and flavonoid accumulation within sweet potato root skin and leaf vein base tissues [29]. By leveraging these multi-omics techniques, researchers have also established a putative transcriptional regulatory network that dictates flavonoid and carotenoid biosynthesis in navel oranges, enabling the development of a hierarchical model for proposed pathway-related genes and TFs [9]. Similar strategies have also supported studies concerning the relationship between carotenoid biosynthesis and petal color in *Brassica napus* [26].

To date, most studies of sweet potato carotenoids have primarily focused on phenotypic characteristics or derivation rather than on underlying metabolic and molecular processes. In the present study, the storage roots of the white-fleshed S19 and the yellow-

fleshed BS sweet potato types were selected for systematic metabolomic and transcriptomic analyses aimed at exploring the mechanisms underpinning carotenoid accumulation in these sweet potato storage roots. Ultimately, this approach highlighted key metabolites and genes associated with the coloration of these sweet potato storage roots while offering novel mechanistic insight into the basis for sweet potato carotenoid biosynthesis.

2. Materials and Methods

2.1. Plant Materials and Treatments

The white-fleshed, red-skinned S19 variety (*Ipomoea batatas* (L.) Lam. cv 'Shangshu 19') and the yellow-fleshed, light yellow-skinned BS variety (*Ipomoea batatas* (L.) Lam. cv 'Baishu') were selected for experimental use.

On 30 May 2020, these two sweet potato types were planted on the experimental agricultural farm of Shanxi Agricultural University (coordinates: 112.58′ E, 37.42′ N). They were planted in a random block arrangement, with two rows of 50 plants per variety using a single ridge planting approach with 35 cm between plants, and a row spacing of 90 cm. Sample collection was performed on day 90 (S1), day 100 (S2), day 110 (S3), and day 120 (S4) after planting. In total, five plants per variety were harvested at random for each time point, yielding 20 storage roots. Similarly sized storage roots with smooth skin were selected for subsequent characterization.

Harvested sweet potato roots were rinsed, dried with absorbent paper, and cut into 0.5 cm × 0.5 cm × 0.5 cm cubes that were snap-frozen for 30 min in liquid nitrogen and stored at −80 °C.

2.2. Analyses of Total Carotenoid Content and Color

A colorimetric approach was used to assess the average carotenoid content in BS and S19 sweet potato storage root samples. The color was assessed using a X Rite VS450 colorimeter (Xrite, Grand Rapids, MI, USA). Carotenoid content following crude organic solvent-mediated extraction was assessed as per Lambert–Beer's law as follows:

$$A = \alpha CL \quad (1)$$

where A corresponds to the absorbance of a given solution, C corresponds to the concentration of the reactant, L corresponds to liquid layer thickness, and α corresponds to the absorption coefficient.

Each sample was analyzed in triplicate using biological replicates.

2.3. Quantification of Carotenoid Content and ABA Levels

After freeze-drying, samples of detached sweet potato storage roots from the BS and S19 types at the S2 and S4 time points were ground using a mixer mill. An UPLC-APCI-MS/MS system (UPLC, ExionLC^{TM} AD, https://sciex.com.cn/ (accessed on 4 November 2020); MS Applied Biosystems 6500 Triple Quadrupole, https://sciex.com.cn/ (accessed on 4 November 2020)) and the MetWare (http://www.metware.cn/ (accessed on 4 November 2020)) application were then used to assess the carotenoid and phytohormone content in these samples, with all analyses having been performed by the MetWare company. Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) approaches were used to evaluate differences in metabolite profiles among these samples. Differentially abundant metabolites (DAMs) were identified using defined significance criteria: variable importance in projection ((VIP) ≥ 1, $|\log2$ FC (fold change)$| \geq 1$, and $p < 0.05$ (Student's *t*-test)). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was then used to map DAMs and to assess significant enrichment thereof, and to define key enriched pathways.

2.4. qPCR

Prior to the experimental operation, the bench was treated with a Solid RNase scavenger (Coolaber, Beijing, China). Frozen sweet potato storage root samples were ground in

liquid nitrogen using a mortar and pestle to produce a powder from which total RNA was extracted using RNaiso Plus (Takara Biotechnology, Beijing, China). An ultra-low-volume spectrometer (BioDrop, Cambridge, UK) was utilized to measure the concentration and A260/A280 ratios for these samples, with cDNA then being prepared with a PrimeScript™ RT reagent Kit with gDNA Eraser (Takara Biotechnology, Beijing, China), based on provided directions for RNA samples with an A260/A280 of 1.8–2.1. Prior to qPCR analyses, these cDNA samples were subject to 5-fold dilution. A Bio-Rad CFX96 Real-Time PCR instrument and TB Green® Premix Ex Taq™ II (Takara Biotechnology, Dalian, China) were used for qPCR with the following settings: 95 °C for 30 s; 40 cycles of 95 °C for 5 s; 60 °C for 30 s. A melting curve from 60 °C to 95 °C (0.5 °C increments for 5 s) was used. The ΔΔCq approach was used to assess relative gene expression [30] in the Bio-Rad Manager 3.1 software, with Actin serving as a control to normalize gene expression levels. Primer Premier 5.0 was used to design all primers for this study (Table S8).

2.5. RNA-Sequencing

For RNA-Seq analyses, Oligo(dT) magnetic beads were used to isolate polyadenylated mRNA, while the removal of rRNA from total RNA samples was additionally used to isolate mRNA. A splitting buffer was used to break RNA molecules into shorter strands that then served as templates for first-strand cDNA synthesis using random hexamer primers. Then, dNTPs (dUTP, dATP, dGTP, dCTP), DNA polymerase I, and first-strand cDNA were used for second-strand cDNA synthesis, followed by the use of AMPure XP beads to isolate prepared cDNA. Following the addition of a poly-A tail, samples were then concatenated using a sequencing adapter, and a final cDNA library was generated through PCR-based enrichment. A Qubit 2.0 instrument was used to quantify cDNA library size, with insert size being assessed using an Agilent 2100 instrument, and qPCR being used to measure the effective concentration. After cDNA library quality had been confirmed, an Illumina Hi-Seq instrument was used for sequencing.

2.6. Quality Control and Bioinformatics Analyses

After sequencing, cleaned reads were obtained through filtering, error rate analyses, and assessments of GC content in the obtained raw reads. The Trinity software was used to derive reference sequences, and clustering was performed with the Corset (https://code.google.com/p/corset-project/ (accessed on 10 November 2020)) tool, which enables the establishment of gene-level counts based on de novo transcriptomic assemblies in which the most extended cluster sequences are designated as a unigene. MetWare (http://www.metware.cn/ (accessed on 4 November 2020)) performed all analyses. Unigenes were compared with the KEGG and the Gene Ontology (GO) databases using the BLAST software, and predicted amino acid sequences for protein-coding unigenes were established, after which they were compared with the protein family (Pfam) database using the HMMER program. Differentially expressed genes (DEGs) were compared between samples using DESeq2 in R, with Benjamini–Hochberg correction being used to control for the *p*-value false discovery rate (FDR) when performing multiple comparisons. DEGs were ultimately identified based on the following criteria: $|\log2FC| \geq 1$ and FDR < 0.05. The prediction of gene function was performed using cluster analyses, and distribution frequencies were assessed in all functional categories.

2.7. Statistical Analysis

SPSS 22.0 was used for statistical analyses of included samples. Data were compared using the Tukey test for multiple comparisons. Differences in physicochemical indices were compared using a minimum of three biological replicates for all experiments. Microsoft Office Excel 2016 was used for data analyses.

3. Results

3.1. Carotenoid Accumulation in Storage Roots Varies in Different Sweet Potato Types

During different stages of development, the coloration of S19 and BS storage roots were compared, revealing that these two types exhibited white and yellow storage root flesh, respectively (Figure 1a). In line with these phenotypes, BS samples contained 7-fold higher total carotenoid levels relative to S19 samples over the analyzed stages of development (S1–S4) (Figure 1c,d). Changes in color represented by the b * (yellow/white) value were significantly positively correlated with total carotenoid content ($r^2 > 0.95$), and this trend remained consistent from stage S1 to stage S4 (Figure 1b–d). Carotenoid content in these storage root samples was then examined in further detail via LC-MS/MS, leading to the identification of 68 carotenoid types, including 14 types of lutein, 7 types of carotenes, and 47 types of carotenoid esters. The BS-S2, BS-S4, S19-S2, and S19-S4 samples contained 12, 12, 9, and 8 carotenoid types, respectively. Different trends in carotenoid content were observed in these two sweet potato types over the course of development. The most abundant carotenoid in the yellow-fleshed BS sweet potato storage roots was β-cryptoxanthin (81.41% in S2; 82.02% in S4), and these levels were 86- and 111-fold higher than the corresponding levels in white-fleshed S19 storage roots at the S2 and S4 stages, respectively (Figures 2f and S1a,b). Whereas the most abundant carotenoid in S19 samples was β-carotene (69.18% in S2; 77.92% in S4), its total abundance remained significantly lower than that in BS storage roots during all stages of development (Figures 1 and S1c,d). These results suggest that the contents of β-cryptoxanthin, β-carotene, and zeaxanthin palmitate may be the primary determinants of the overall differences in total carotenoid content observed in storage roots from these two types of sweet potato. Significant increases in violaxanthin, lutein palmitate, violaxanthin palmitate, violaxanthin dipalmitate, and echinenone content were also evident in BS storage roots relative to those from S19 sweet potatoes at the S2 and S4 stages of development, whereas no differences in antheraxanthin, apocarotenal, or lutein content were observed when comparing these two types. Principal component analysis (PCA) confirmed that there were significant differences between the metabolite profiles of these two sweet potato types, with PC1 and PC2 accounting for 61.9% and 15.7% of the variability among these samples, respectively (Figure S1e).

3.2. Changes in ABA Metabolism over the Course of Storage Root Development

ABA metabolism and its relationship with carotenoids in sweet potatoes were examined by using LC-MS/MS. Opposing trends in ABA and ABA-glucosyl ester (ABA-GE) content were observed in these storage roots, with ABA-GE levels trending upwards over the course of development in both sweet potato types, while ABA levels trended downwards. Total ABA content was significantly higher in BS storage roots than S19 storage roots (Figure 3a,b). As shown in Figure 3c, changes in the ABA content of these two sweet potato types were positively correlated with the observed trends in most carotenoid compositions. However, the change trend of ABA-GE content was only related to that of antheraxanthin content (Figure 3c).

3.3. Transcriptomic Analyses of BS and S19 Sweet Potato Storage Roots

To fully understand the molecular mechanisms that shape sweet potato carotenoid biosynthetic processes, a transcriptomic analysis of the S19 and BS sweet potato storage roots was conducted at the young (S2) and old (S4) stages of development. In total, analyses of these 12 samples yielded 79.17 GB of clean data (6 GB/sample), with Q30 base percentages of >90% and a GC content of 45% (Table S1). Over 84% of reads were mapped, including 70% uniquely mapped reads and 10% multiply mapped reads (Table S2).

DESeq2 was next used to identify differentially expressed genes (DEGs) in BS and S19 samples at these different stages of development ($|\log2FC| \geq 1$, FDR < 0.05). PCA and cluster analyses revealed close clustering of biological replicate samples, with clear distinctions between samples from different groups (Figure 4a–c). In total, these analyses identified 8643 (4470 upregulated and 4173 downregulated), 12,212 (6154 upregulated

and 6058 downregulated), 3298 (983 upregulated and 2315 downregulated), and 4409 (1391 upregulated and 3018 downregulated) DEGs for the respective S19-S2_vs_BS-S2, S19-S4_vs_BS-S4, BS-S2_vs_BS-S4, and S19-S2_vs_S19-S4 comparisons (Figure 4d). In total, 10 subset classes were clustered based on DEG expression patterns (Figure 4e).

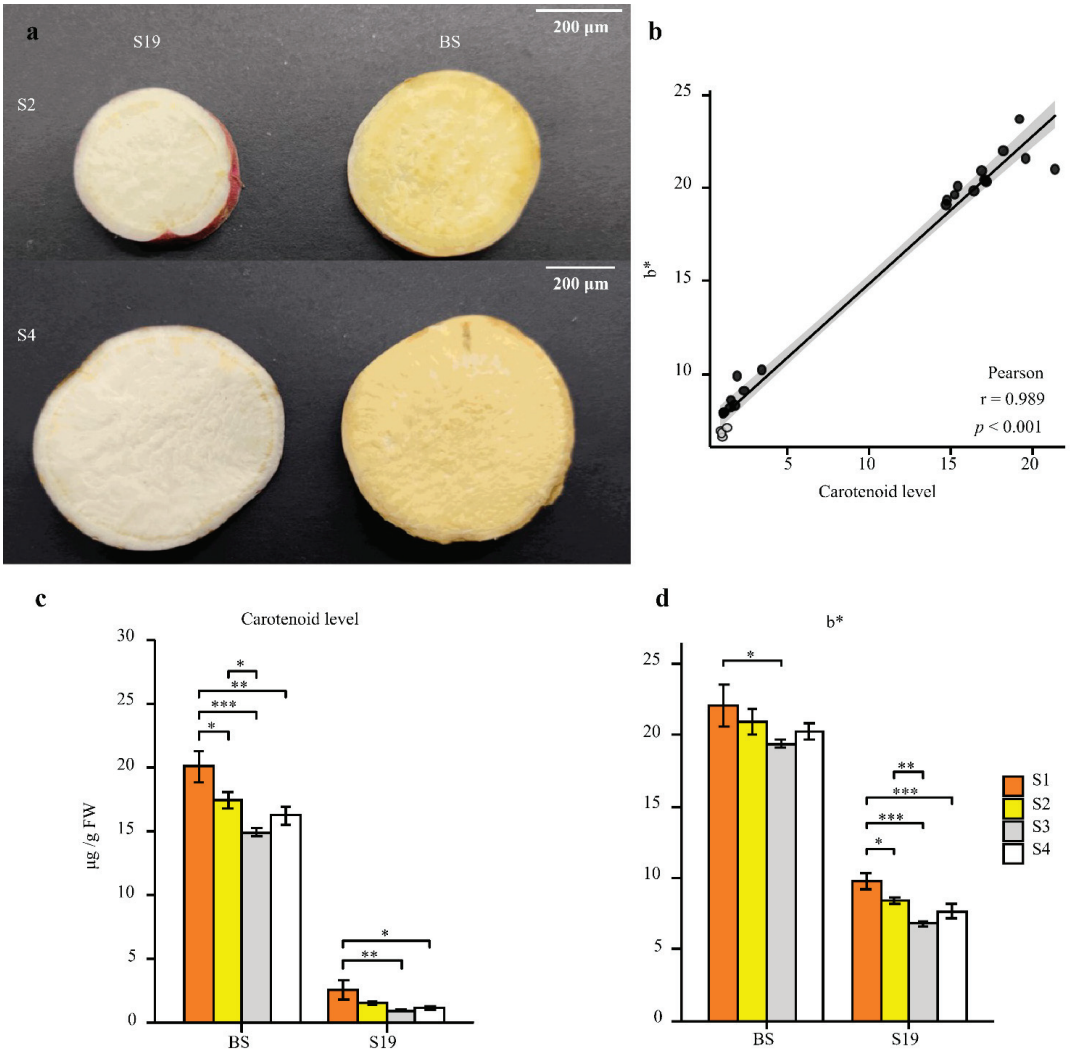


Figure 1. Analyses of sweet potato storage root phenotypic characteristics and carotenoid content. (**a**) Storage root phenotypes were assessed for BS-S2, BS-S4, S19-S2, and S19-S4 samples. Scale bar = 200 µm. (**b**) Correlations between carotenoid content and color change. (**c**,**d**) Total carotenoid content (**c**) and color change (**d**) were analyzed in both sweet potato types at the analyzed time points. Data are means ± SEM ($n = 3$). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

When the annotated DEGs in these groups were subject to GO classification based on the molecular function (MF), biological process (BP), and cellular component (CC) annotation categories (Table S3), 3800 DEGs (885 MF, 2491 BP, and 424 CC), 3883 DEGs (898 MF, 2547 BP, and 438 CC), 3064 DEGs (728 MF, 2048 BP, and 288 CC), and 3329 DEGs

(728 MF, 2222 BP, and 325 CC) were annotated for these four respective comparisons (S19-S2_vs_BS-S2, S19-S4_vs_BS-S4, BS-S2_vs_BS-S4, and S19-S2_vs_S19-S4).

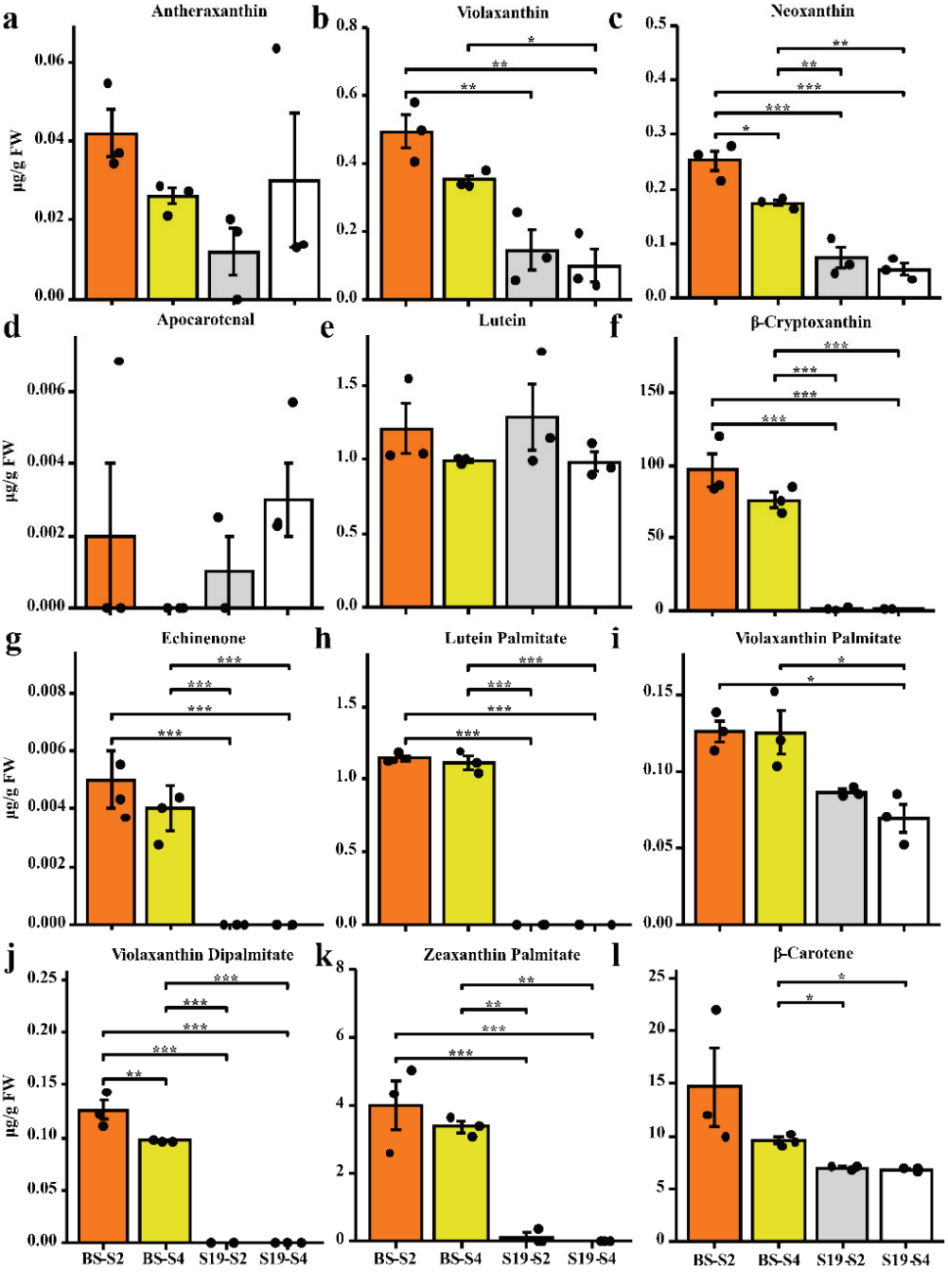


Figure 2. Analysis of carotenoid levels in sweet potato samples. The content of lutein (**a**–**g**), carotenoid esters (**h**–**k**), and carotene (**l**). Data are means ± SEM ($n = 3$). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Specific metabolic pathways involved in the regulation of sweet potato storage root carotenoid biosynthesis in the BS and S19 were further explored through KEGG enrichment analyses for these four comparisons, revealing these DEGs to be significantly enriched in cellular processes, environmental information processing, genetic information processing, metabolism, and organismal systems. In total, 19, 11, 30, and 20 DEGs from the four respective comparisons (BS-S2_vs_BS-S4, S19-S2_vs_S19-S4, S19-S2_vs_BS-S2, and S19-S4_vs_BS-S4) were found to be significantly enriched in the carotenoid biosynthesis pathway (ko00906) (Figures S2–S4).

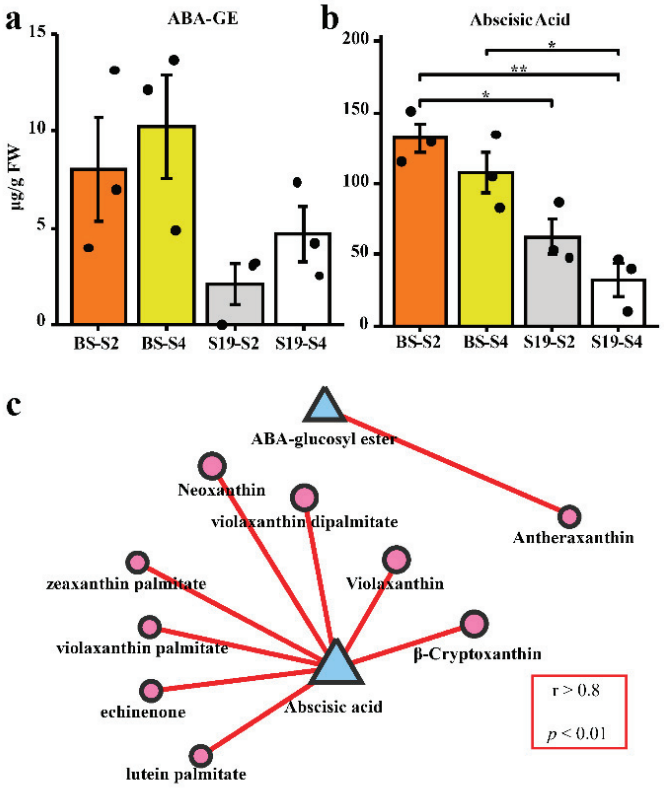


Figure 3. Correlations between ABA levels and carotenoid content. (**a**) Total ABA-GE levels in the indicated samples (μg/g FW). (**b**) Total ABA levels in the indicated samples (μg/g FW). (**c**) The contentsof ABA and ABA-GE were positively correlated with the contents of carotenoid-related components (graphic area is positively correlated with content.). Data are means ± SEM (n = 3). ** $p < 0.01$, * $p < 0.05$.

3.4. Combined Analyses of Metabolites and Genes Associated with Carotenoid Biosynthesis in Sweet Potato Storage Roots

To explore the metabolites and mechanisms associated with sweet potato storage root carotenoid accumulation, a combined transcriptomic and metabolomic analysis of BS and S19 samples during the different stages of development was conducted. In total, 43 DEGs and 7 differentially accumulated metabolites (DAMs) were associated with the carotenoid biosynthesis pathway (Tables S4 and S5), enabling the establishment of a metabolic profiling diagram for these carotenoid pathways (Figure 5). In total, 24 differentially expressed structural genes associated with the biosynthesis and catabolism of carotenoids were identified, including genes encoding PSY (15-cis-phytoene synthase), ZEBRA2 (prolycopene isomerase), ZEP (zeaxanthin epoxidase), VDE (violaxanthin de-epoxidase), NCED (9-cis-

epoxycarotenoid dioxygenase), LCYE (lycopene epsilon-cyclase), and CHYB (β-carotene 3-hydroxylase).

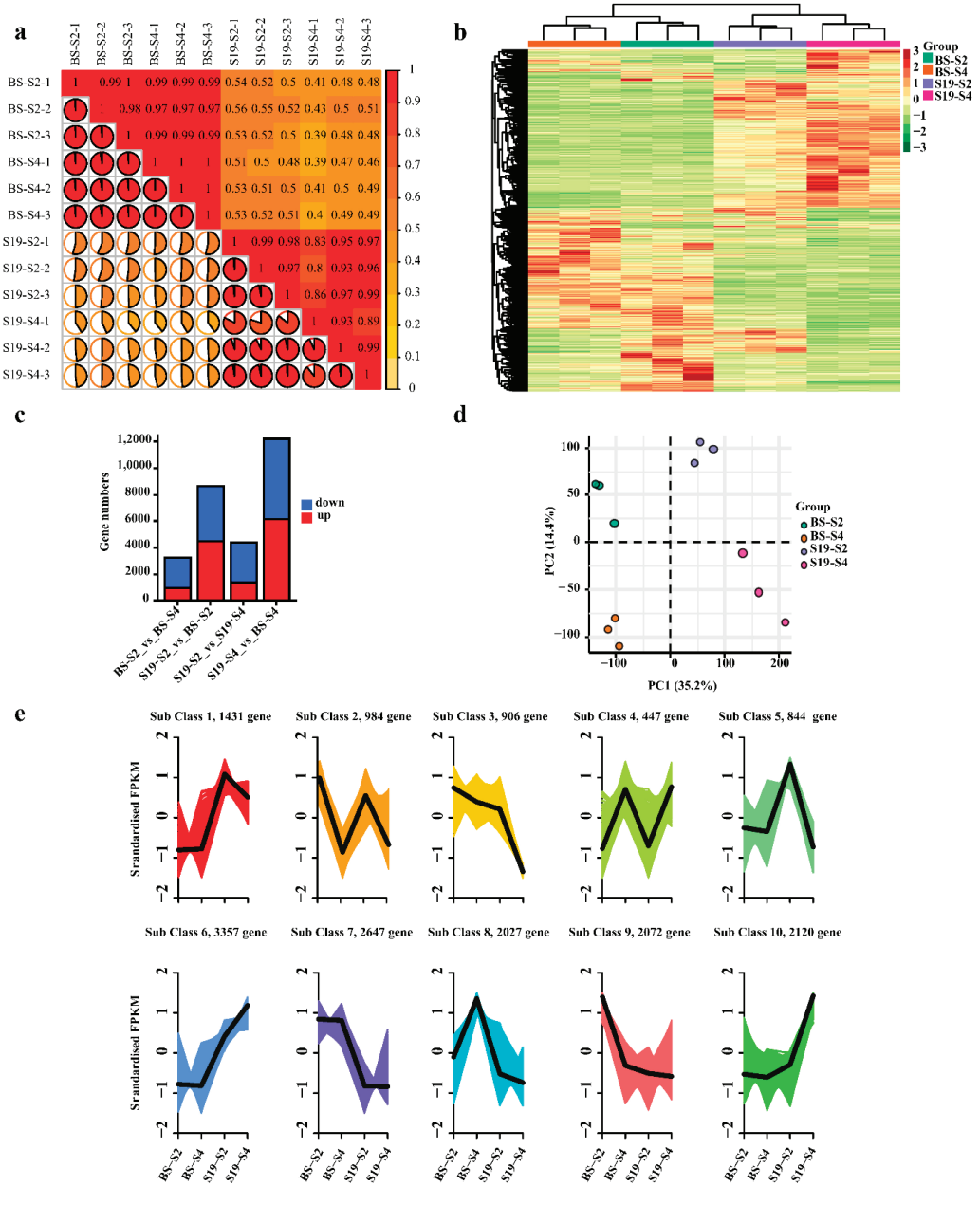


Figure 4. Analyses of DEGs identified in transcriptomic of sweet potato storage root samples and principal component analyses. (**a**) Sample repeatability and correlation analyses. (**b**) Heatmaps representing hierarchically clustered DEGs. (**c**) Statistical analyses of DEGs in the indicated groups. (**d**) PCA for samples included in transcriptomic analyses. (**e**) DEG cluster analyses.

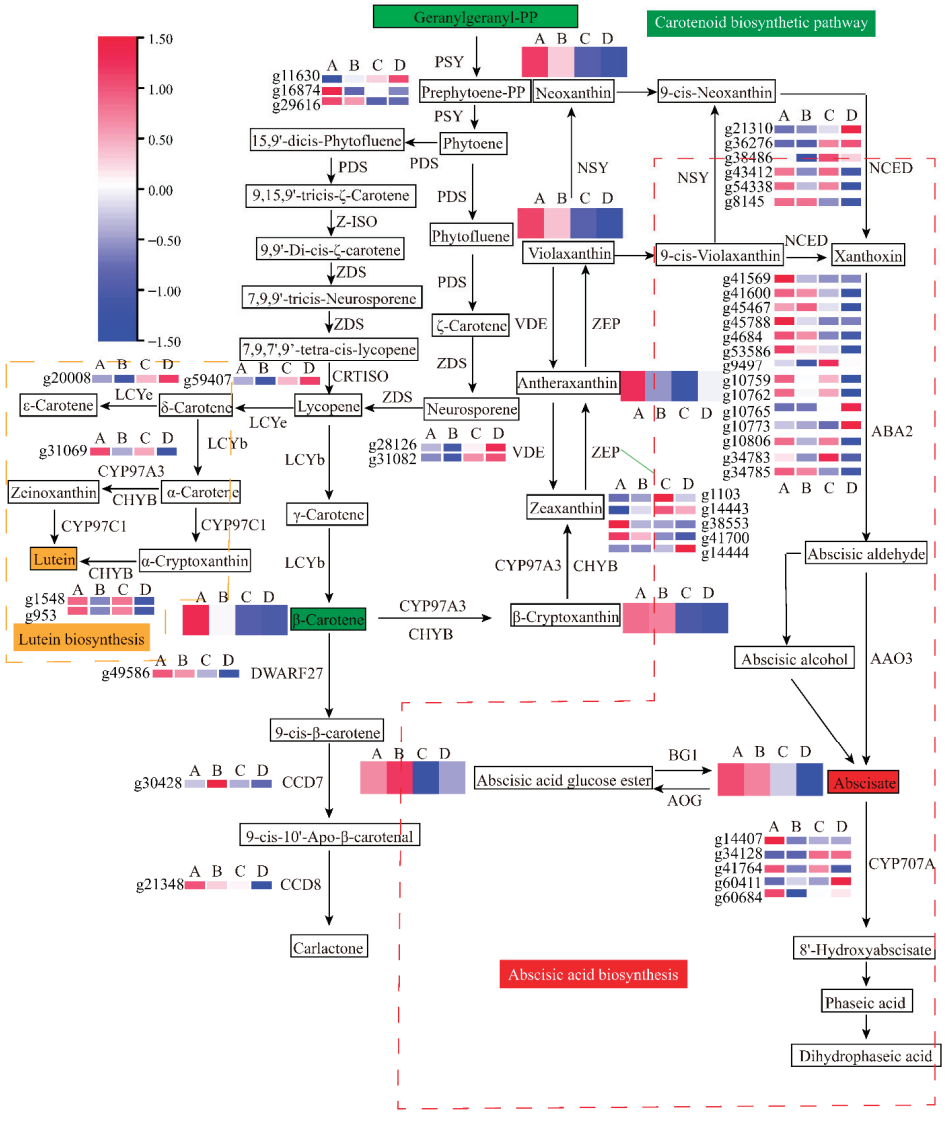


Figure 5. The carotenoid and ABA metabolic pathways exhibiting DEG and DAM enrichment. (A: BS-S2; B: BS-S4; C: S19-S2; D: S19-S4.).

PSY (*g29616*) expression in BS has been shown to be significantly upregulated relative to levels in S19 storage roots at the S2 and S4 stages. *CHYB* expression in the storage roots of these sweet potato types trended downward from the S2 to the S4 stage, with *CHYB* gene (*g1548* and *g953*) expression levels in BS storage roots that were 3.56-, 3.85-, and 2.52-, 2.59-fold higher than those in the S19 storage root at the analyzed stages of development. Relative to BS storage roots, lower levels of *ZEP* (*g41700*) expression were evident in S19 storage roots, particularly during the S2 stage. However, *ZEP* genes (*g14444* and *g1103*) were expressed at very high levels in S19 samples, with opposing expression trends during different stages of development. *CCD* gene (*g49586*) expression was significantly increased in BS storage roots at the analyzed stages of development, as compared to those from S19 sweet potatoes. Relative *NCED3* (*g36276*) expression was increased by 34.09- and 39.74-fold

in S19 storage roots during the respective S2 and S4 stages relative to levels in BS storage roots. Moreover, 8 ABA metabolism-associated DEGs were identified, including 14 ABA2 (xanthoxin dehydrogenase) and 5 CYP707A (abscisic acid 8′-hydroxylase). Downregulated expression of the *ABA2* gene was observed over the course of development, with significantly lower *ABA2* (*g34785*) expression in S19 storage roots relative to those from BS sweet potatoes. The ABA degradation-associated gene *CYP707A* (*g34128*) was expressed at 302.20- and 59.31-fold higher levels in S19 storage roots relative to BS storage roots at the S2 and S4 stages, respectively. Three key *ABA2* genes (*g4684*, *g34785*, and *g41600*) were also significantly positively correlated with ABA content (Figure 6). These DEGs may be associated with the observed differences in carotenoid metabolic activity in the BS and S19 types, and with differences in the levels of detected DAMS. β-carotene (C02094), β-cryptoxanthin (C08591), antheraxanthin (C08579), violaxanthin (C08614), neoxanthin (C13431), ABA (C06082), and ABA-GE (C15970) were detected (Figure 5, Table S5).

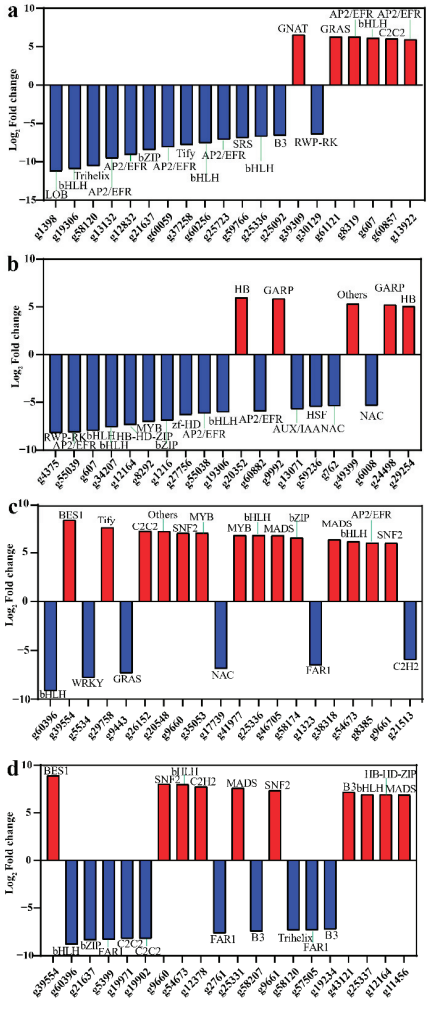


Figure 6. Differentially expressed genes encoding transcription factors in sweet potato samples. (**a**–**d**) The top 20 TF-encoding genes exhibiting the greatest fold-change in expression levels when comparing the BS-S2 and BS-S4 (**a**), S19-S2 and BS-S2 (**b**), S19-S2 and S19-S4 (**c**), and S19-S4 and BS-S4 (**d**) samples.

3.5. Identification of Transcription Factors Related to Carotenoid Accumulation

Transcription factors (TFs) and other transcriptional regulators (TRs) are key regulators of the expression of all genes, including those associated with the biosynthesis of carotenoids. In this study, 284 (88 upregulated, 196 downregulated), 255 (58 upregulated, 197 downregulated), 365 (231 upregulated, 134 downregulated), and 577 (335 upregulated, 242 downregulated) differentially expressed TF-coding genes were identified from the comparisons of BS-S2_vs_BS-S4, S19-S2_vs_S19-S4, S19-S2_vs_BS-S2, and S19-S4_vs_BS-S4, respectively. Moreover, these four respective comparisons yielded 44 (10 upregulated, 34 downregulated), 57 (18 upregulated, 39 downregulated), 89 (41 upregulated, 48 downregulated), and 126 (54 upregulated, 72 downregulated) differentially expressed TR-encoding genes (Table S6). These TFs included members of the AP2/ERF, MYB, bHLH, bZIP, NAC, FAR1, C2H2, PLATZ, WRKY, OFP, and TCP families, while identified TRs included members of the PHD, SET, SNF2, TAZ, and AUX/IAA, and SNF2 families (Table S7). Of these differentially expressed TFs and TRs, the most highly expressed included AP2/ERF, MYB, bHLH, AUX/IAA, SNF2, and PHD. When screening the top 20 DEGs encoding TFs, the genes associated with carotenoid biosynthesis were found to be upregulated when comparing S19 and BS storage roots, particularly during the S2 stage of development (Figure 7). The gene encoding NAC was expressed at higher levels in BS storage roots relative to S19 storage roots during both the S2 and S4 stages, although when comparing S2 vs. S4 samples, *NAC* (*g103*) was downregulated in BS storage roots yet upregulated in S19 storage roots. Strikingly, of the analyzed differentially expressed TF genes, *OFP* (*g10779*) and *FAR1* (*g1323*) expression levels differed significantly in the storage roots from the two sweet potato types, with *OFP* (*g10779*) being expressed at significantly higher levels in BS samples, whereas *FAR1* (*g1323*) was expressed at significantly higher levels in S19 samples at the S2 and S4 stages of development.

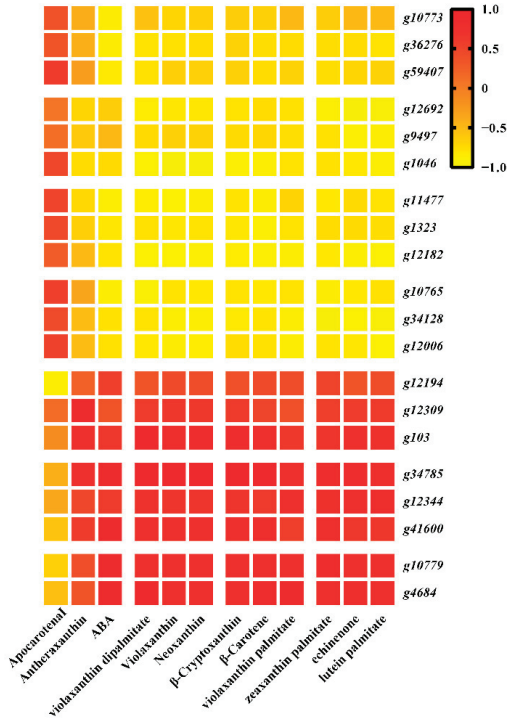


Figure 7. Correlation between DEGs and DAMs trends, with stronger correlations corresponding to a more consistent trend.

3.6. RNA-Seq Result Validation

To confirm the validity of the RNA-Seq results established above, 13 DEGs associated with carotenoid biosynthesis, 8 DEGs associated with ABA biosynthesis, and 9 TFs closely associated with carotenoid biosynthesis were selected for qPCR-based verification analysis (Figure 8). Overall, the observed expression trends for these 30 genes were highly consistent with the RNA-Seq results presented above, confirming the accuracy and reliability of these transcriptomic sequencing analyses.

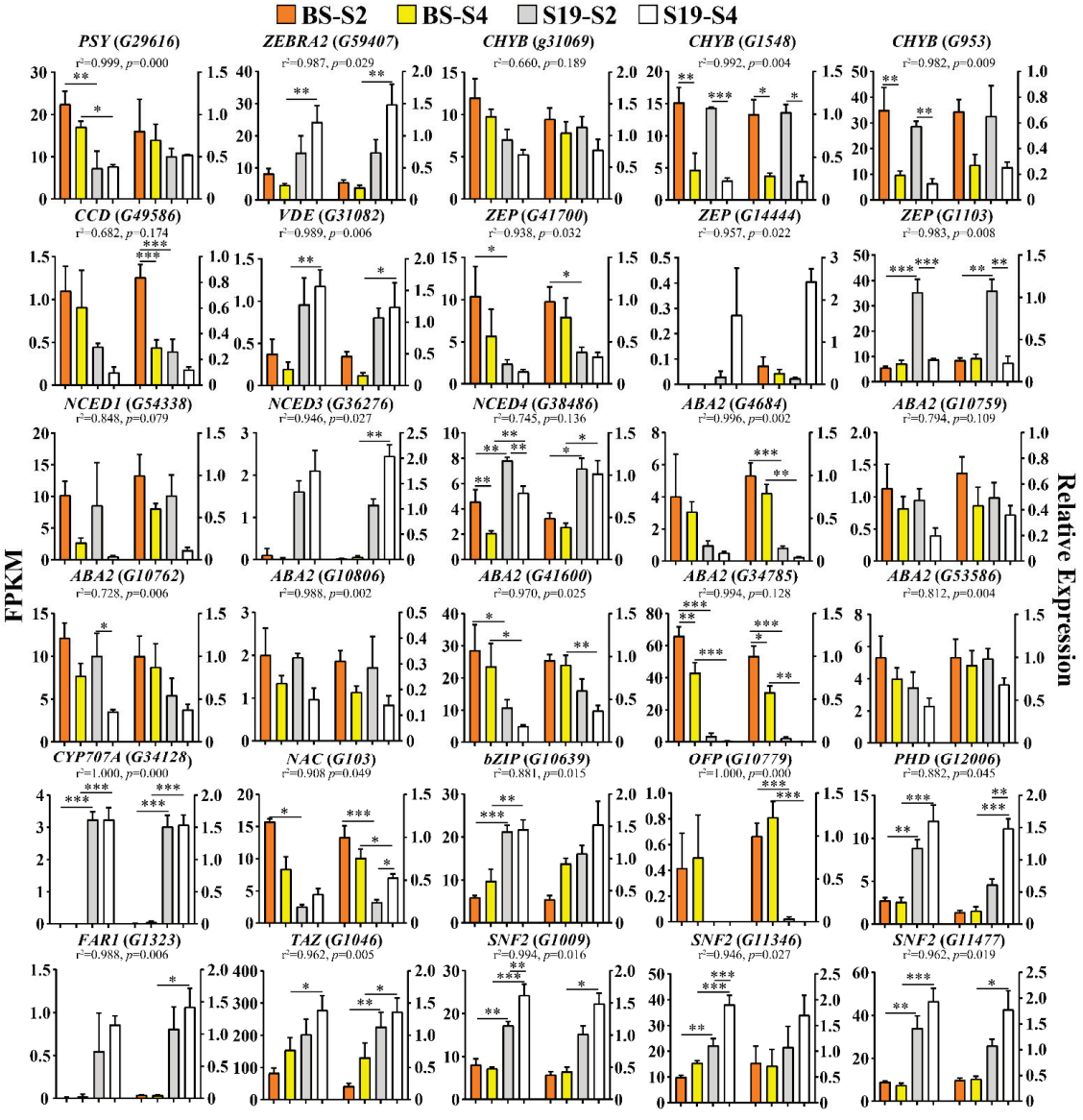


Figure 8. Expression analyses of structural genes and transcription factors associated with the carotenoid and ABA biosynthesis pathways. Data are means ± SEM (n = 3). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4. Discussion

In sweet potato storage roots, carotenoid content is a key determinant of coloration, appearance, and health-related benefits, thereby shaping consumer preference [31]. Indeed, the increase in total carotenoid content is strongly correlated with the changes in the color of these storage roots over the course of development. Sweet potato storage roots contain carotenoids such as β-carotene, β-cryptoxanthin, zeaxanthin, and violaxanthin [32]. Here, characterization of carotenoid content in storage root samples prepared from two sweet potato types at different stages of development revealed that BS storage roots contained significantly higher carotenoid levels relative to S19 storage roots, with marked differences in the levels of the two primary carotenoids in these storage roots (β-carotene and β-cryptoxanthin). The most abundant carotenoids in yellow-fleshed sweet potatoes included β-carotene and β-cryptoxanthin, with the latter being present at higher concentrations [10–12,32]. Consistently, β-cryptoxanthin was the most abundant carotenoid in yellow-fleshed BS storage roots, suggesting that it may be the primary determinant of the observed differences in coloration between these two sweet potato types.

Efforts to characterize carotenoid metabolism-related genes have been made in a range of species including tomatoes [14], carrots [33], oranges [34], and peppers [15]. The mechanisms shaping carotenoid biosynthesis in sweet potatoes, however, have yet to be firmly established. In an effort to identify key carotenoid accumulation-associated genes and metabolites, metabolomic and transcriptomic analyses were thus performed. PSY is a key rate-limiting carotenoid biosynthesis-associated enzyme in plants that can ultimately shape total carotenoid content [21,35]. Both RNA-Seq and qPCR analyses confirmed significantly increased *PSY* (*g29616*) expression in BS storage roots relative to those from S19 sweet potatoes during the tested developmental stages, potentially partially accounting for observed differences in carotenoid content in these two types. The CHYB enzyme is responsible for catalyzing the addition of a hydroxyl residue necessary for esterification, and is vital for chromoplast carotenoid accumulation in a range of plants. CHYB is a rate-limiting enzyme in the zeaxanthin biosynthesis pathway that has been shown to play a key role in chromoplast carotenoid accumulation in *Ipomoea petals* [36]. Consistently, BS storage roots exhibited higher *CHYB* (*g1548* and *g953*) expression levels relative to S19 storage roots, potentially contributing to differences in β-carotene and zeaxanthin content in BS and S19 samples. The ZEP enzyme catalyzes zeaxanthin and antheraxanthin β-rings epoxidation, and BS storage roots were herein found to express significantly higher levels of *ZEP* (*g41700*). In line with the present report, Suematsu et al. suggested that *ZEP* is an important mediator of carotenoid accumulation in yellow-fleshed sweet potatoes [37]. The NCED enzyme also plays an important role in downstream aspects of plant carotenoid biosynthesis pathways [23,38]. In apricots, *NCED* expression is responsible for flesh coloration, causing the low levels of β-carotenoid observed in white apricots [39,40]. Here, white-fleshed S19 storage roots were found to exhibit particularly high levels of *NCED3* (*g36276*) expression, particularly during the S4 stage of development, in line with prior research, and suggesting that this gene may be a central driver of the differences in coloration and carotenoid content between the BS and the S19 sweet potato types.

ABA is a key carotenoid-derived phytohormone that is essential in the context of carotenoid metabolism and biosynthesis [23]. By cleaving the (C_{11}–C_{12}) double bond in 9-cis-violaxanthin and 9′-cis-neoxanthin, NCED generates xanthoxin, which is an ABA precursor, such that NCED catalyzes the initial step necessary for ABA biosynthesis [24]. In the present analysis, the ABA biosynthesis-related genes *ABA2* (*g34785*) and *CYP707A* (*g34128*) were found to exhibit opposing expression patterns in the two analyzed sweet potato types. Specifically, *ABA2* (*g34785*) was significantly upregulated in BS storage roots relative to those from S19 sweet potatoes at the S2 and S4 stages of development, whereas *CYP707A* (*g34128*) exhibited the opposite expression pattern. *ABA2* has previously been reported by González-Guzmán et al. to encode a key enzyme involved in ABA biosynthesis that catalyzes xanothoxin conversion into abscisic aldehyde [41]. In *Capsicum annuum*, Kim et al. determined that ABA hydroxylation mediated by CYP707A was able to promote the

degradation of ABA, and thereby reduce the levels of this phytohormone [42]. Consistently, *ABA2* (*g34785*) and *CYP707A* (*g34128*) expression levels were found to be positively and negatively correlated with the ABA levels in BS and S19 samples. Together, these data support roles for *PSY*, *CHYB*, *ZEP*, *NCED3*, *ABA2*, and *CYP707A* as essential regulators of ABA and carotenoid biosynthesis and metabolism in sweet potato storage roots.

Our study showed that the top 20 DEGs encoding TFs that were identified when comparing the S19 and BS storage roots samples were associated with ABA and carotenoid biosynthesis. MYB, HD-ZIP, bZIP, bHLH, and WRKY have previously been identified as critical TFs involved in coordinating carotenoid biosynthesis [43,44]. Specifically, bHLH has been shown to bind the *PSY* promoter and to thereby suppress the expression of this gene, thus reducing carotenoid accumulation [43]. Moreover, bZIP has been reported to influence total carotenoid content in tomatoes [45], while MYB plays a similar role in wolfberries [46]. In line with these prior results, marked bHLH, bZIP, and MYB upregulation was observed in S19 storage root samples relative to those from BS sweet potatoes.

Strikingly, the two sweet potato types exhibited opposite expression patterns for the TF-encoding *FAR1* (*g1323*) and *OFP* (*g10779*) genes. In a prior report, Tang et al. highlighted the importance of FAR1 as a positive regulator of Arabidopsis ABA signaling [47]. Global transcriptional profiling of Arabidopsis specimens further revealed that OFP is a transcriptional repressor capable of controlling ABA signal transduction [48]. We showed that significant increases in *FAR1* (*g1323*) expression were evident in S19 storage roots relative to those from BS sweet potatoes, whereas *OFP* (*g10779*) expression was markedly reduced in S19 samples, as compared to BS samples. This may at least partially account for the observed differences in ABA content in these two sweet potato types [49]. ABA is a key phytohormone that is derived from carotenoid metabolite precursors [23]. These two TFs may thus serve as important regulators of the interplay between carotenoid metabolism and ABA biosynthesis in sweet potato storage roots.

5. Conclusions

The mechanisms underlying carotenoid biosynthesis were assessed through comprehensive metabolomic and transcriptomic analyses of the storage roots from two sweet potato types exhibiting different levels of carotenoid content. In this study, *PSY*, *CHYB*, *ZEP*, *NCED3*, *ABA2*, and *CYP707A* were identified as critical genes associated with the observed differences in carotenoid and ABA content in these two sweet potato types. The OFP and FAR1 transcription factors were also identified as important regulators of ABA biosynthesis in sweet potatoes. Carotenoid metabolism is closely related to the biosynthesis of the ABA, and carotenoids are precursors of the ABA. These results enabled the establishment of a proposed integrated network controlling the metabolism and biosynthesis of carotenoids and ABA in sweet potato storage roots. However, further work will be critical to validate these results and to clarify the role that other genes and regulatory mechanisms play in shaping these processes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12111010/s1, Figure S1: Proportions of carotenoid levels and PCA analyses for the indicated samples; Figure S2: KEGG classifications between different samples; Figure S3: KEGG classifications for the comparison of S19-S2 and BS-S2 samples; Figure S4: KEGG classifications for the comparison of the S19-S4 and BS-S4 samples; Table S1: Summary of RNA-seq data and mapped reads for mRNA; Table S2: Statistical results of transcriptome data comparison with reference genomes; Table S3: GO enrichment; Table S4: Statistical table of DEGs in four groups of BS and S19; Table S5: Statistical table of DAMs in four groups of samples; Table S6: TFs classification; Table S7: The gene-specific primers; Table S8: The gene-specific primers.

Author Contributions: Conceptualization, Q.R. and X.Z.; methodology, B.Z.; software, Y.H.; validation, Q.R., H.G. and Y.L.; formal analysis, J.Z.; investigation, M.Y.; resources, H.L.; data curation, Q.R.; writing—original draft preparation, Q.R.; writing—review and editing, X.Z.; visualization, Q.R.; supervision, Q.R.; project administration, Q.R.; funding acquisition, Q.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministerial and Provincial Co-Innovation Centre for Endemic Crops Production with High-quality and Efficiency in Loess Plateau (grant number SBGJXTZX-13), High-level Foreign Experts Introduction Project of China (grant number G2022004009L), Shanxi Key Laboratory of Minor Crops Germplasm Innovation and Molecular Breeding, Shanxi Agricultural University (grant number 202204010910001), Grand Science and Technology Special Project in Shanxi Province (grant numbers 202101140601027), and the Innovation Programs of Graduate Education in Shanxi (2021Y332).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in the study are deposited in the NCBI SRA database, accession number PRJNA873221.

Acknowledgments: The authors would like to thank all the reviewers who participated in the review, as well as MJEditor (www.mjeditor.com (accessed on 13 August 2022)) for providing English editing services during the preparation of this manuscript.

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

References

1. Li, C.; Kou, M.; Arisha, M.; Tang, W.; Ma, M.; Yan, H.; Wang, X.; Wang, X.; Zhang, Y.; Liu, Y.; et al. Transcriptomic and metabolic profiling of high-temperature treated storage roots reveals the mechanism of saccharification in sweetpotato (*Ipomoea batatas* (L.) Lam.). *Int. J. Mol. Sci.* **2021**, *22*, 6641. [CrossRef]
2. Banda, L.; Kyallo, M.; Entfellner, J.B.D.; Moyo, M.; Swanckaert, J.; Mwanga, R.O.M.; Onyango, A.; Magiri, E.; Gemenet, D.C.; Yao, N.; et al. Analysis of β-amylase gene (Amyβ) variation reveals allele association with low enzyme activity and increased firmness in cooked sweetpotato (*Ipomoea batatas*) from East Africa. *J. Agric. Food Res.* **2021**, *4*, 100121. [CrossRef] [PubMed]
3. Lewandowski, K.; Zhang, X.; Hayes, M.; Ferruzzi, M.G.; Paton, C.M. Design and nutrient analysis of a carotenoid-rich food product to address vitamin A and protein deficiency. *Foods* **2021**, *10*, 1019. [CrossRef] [PubMed]
4. Agarwal, D.; Mui, L.; Aldridge, E.; McKinney, J.; Hewson, L.; Fisk, I.D. The progression of lipid oxidation, β-carotenes degradation and sensory perception of batch-fried sliced sweet potato crisps during storage. *Food Funct.* **2021**, *12*, 4535–4543. [CrossRef] [PubMed]
5. Quian-Ulloa, R.; Stange, C. Carotenoid biosynthesis and plastid development in plants: The role of light. *Int. J. Mol. Sci.* **2021**, *22*, 1184. [CrossRef]
6. Giannaccare, G.; Pellegrini, M.; Senni, C.; Bernabei, F.; Scorcia, V.; Cicero, A.F.G. Clinical applications of astaxanthin in the treatment of ocular diseases: Emerging insights. *Mar. Drugs* **2020**, *18*, 239. [CrossRef]
7. Ribeiro, D.; Sousa, A.; Nicola, P.; de Oliveira, J.M.P.F.; Rufino, A.T.; Silva, M.; Freitas, M.; Carvalho, F.; Fernandes, E. β-Carotene and its physiological metabolites: Effects on oxidative status regulation and genotoxicity in in vitro models. *Food Chem. Toxicol.* **2020**, *141*, 111392. [CrossRef] [PubMed]
8. He, L.; Liu, X.; Liu, S.; Zhang, J.; Zhang, Y.; Sun, Y.; Tang, R.; Wang, W.; Cui, H.; Li, R.; et al. Transcriptomic and targeted metabolomic analysis identifies genes and metabolites involved in anthocyanin accumulation in tuberous roots of sweetpotato (*Ipomoea batatas* L.). *Plant Physiol. Biochem.* **2020**, *156*, 323–332. [CrossRef] [PubMed]
9. Zhang, H.; Chen, J.; Peng, Z.; Shi, M.; Liu, X.; Wen, H.; Jiang, Y.; Cheng, Y.; Xu, J.; Zhang, H. Integrated Transcriptomic and Metabolomic analysis reveals a transcriptional regulation network for the biosynthesis of carotenoids and flavonoids in '*Cara cara*' navel Orange. *BMC Plant Biol.* **2021**, *21*, 29. [CrossRef]
10. Ishiguro, K.; Yoshinaga, M.; Kai, Y.; Maoka, T.; Yoshimoto, M. Composition, content and antioxidative activity of the carotenoids in yellow-fleshed sweetpotato (*Ipomoea batatas* L.). *Breed. Sci.* **2010**, *60*, 324–329. [CrossRef]
11. Drapal, M.; Fraser, P.D. Determination of carotenoids in sweet potato (*Ipomoea batatas L.* Lam) tubers: Implications for accurate provitamin A determination in staple sturdy tuber crops. *Phytochemistry* **2019**, *167*, 112102. [CrossRef] [PubMed]
12. Yang, Y.; Shi, D.; Wang, Y.; Zhang, L.; Chen, X.; Yang, X.; Xiong, H.; Bhattarai, G.; Ravelombola, W.; Olaoye, D.; et al. Transcript profiling for regulation of sweet potato skin color in Sushu8 and its mutant *Zhengshu20*. *Plant Physiol. Biochem.* **2020**, *148*, 1–9. [CrossRef] [PubMed]
13. Yuan, H.; Pawlowski, E.G.; Yang, Y.; Sun, T.; Thannhauser, T.W.; Mazourek, M.; Schnell, D.; Li, L. Arabidopsis ORANGE protein regulates plastid pre-protein import through interacting with Tic proteins. *J. Exp. Bot.* **2021**, *72*, 1059–1072. [CrossRef] [PubMed]
14. Barja, M.V.; Ezquerro, M.; Beretta, S.; Diretto, G.; Florez-Sarasa, I.; Feixes, E.; Fiore, A.; Karlova, R.; Fernie, A.R.; Beekwilder, J.; et al. Several geranylgeranyl diphosphate synthase isoforms supply metabolic substrates for carotenoid biosynthesis in tomato. *New Phytol.* **2021**, *231*, 255–272. [CrossRef]
15. Wei, X.; Meng, C.; Yuan, Y.; Nath, U.K.; Zhao, Y.; Wang, Z.; Yang, S.; Li, L.; Niu, L.; Yao, Q.; et al. CaPSY1 gene plays likely the key role in carotenoid metabolism of pepper (*Capsicum annuum*) at ripening. *Funct. Plant. Biol.* **2021**, *48*, 141–155. [CrossRef]

16. Mallikarjuna Swamy, B.P.; Jr Marundan, S.; Samia, M.; Ordonio, R.L.; Rebong, D.B.; Miranda, R.; Alibuyog, A.; Rebong, A.T.; Tabil, M.A.; Suralta, R.R.; et al. Development and characterization of GR2E Golden rice introgression lines. *Sci. Rep.* **2021**, *11*, 2496. [CrossRef] [PubMed]
17. Diepenbrock, C.H.; Ilut, D.C.; Magallanes-Lundback, M.; Kandianis, C.B.; Lipka, A.E.; Bradbury, P.J.; Holland, J.B.; Hamilton, J.P.; Wooldridge, E.; Vaillancourt, B.; et al. Eleven biosynthetic genes explain the majority of natural variation in carotenoid levels in maize grain. *Plant Cell* **2021**, *33*, 882–900. [CrossRef] [PubMed]
18. Naing, A.H.; Kyu, S.Y.; Win Pe, P.P.; Il Park, K.; Lee, J.M.; Lim, K.B.; Kim, C.K. Silencing of the phytoene desaturase (PDS) gene affects the expression of fruit-ripening genes in tomatoes. *Plant Methods* **2019**, *15*, 110. [CrossRef]
19. Kössler, S.; Armarego-Marriott, T.; Tarkowská, D.; Turečková, V.; Agrawal, S.; Mi, J.; de Souza, L.P.; Schöttler, M.A.; Schadach, A.; Fröhlich, A.; et al. Lycopene β-cyclase expression influences plant physiology, development, and metabolism in tobacco plants. *J. Exp. Bot.* **2021**, *72*, 2544–2569. [CrossRef]
20. Ding, X.; Jia, L.L.; Xing, G.M.; Tao, J.P.; Sun, S.; Tan, G.F.; Li, S.; Liu, J.X.; Duan, A.Q.; Wang, H.; et al. The accumulation of lutein and β-carotene and transcript profiling of genes related to carotenoids biosynthesis in yellow celery. *Mol. Biotechnol.* **2021**, *63*, 638–649. [CrossRef]
21. Lana, G.; Zacarias-Garcia, J.; Distefano, G.; Gentile, A.; Rodrigo, M.J.; Zacarias, L. Transcriptional analysis of carotenoids accumulation and metabolism in a pink-fleshed lemon mutant. *Genes* **2020**, *11*, 1294. [CrossRef]
22. Chayut, N.; Yuan, H.; Saar, Y.; Zheng, Y.; Sun, T.; Zhou, X.; Hermanns, A.; Elad, O.; Faigenboim, A.; Hui, M.; et al. Comparative transcriptome analyses shed light on carotenoid production and plastid development in melon fruit. *Hortic. Res.* **2021**, *8*, 112. [CrossRef]
23. Felemban, A.; Braguy, J.; Zurbriggen, M.D.; Al-Babili, S. Apocarotenoids involved in plant development and stress response. *Front. Plant Sci.* **2019**, *10*, 1168. [CrossRef]
24. Moreno, J.C.; Mi, J.; Alagoz, Y.; Al-Babili, S. Plant apocarotenoids: From retrograde signaling to interspecific communication. *Plant J.* **2021**, *105*, 351–375. [CrossRef]
25. Chen, K.; Li, G.J.; Bressan, R.A.; Song, C.P.; Zhu, J.K.; Zhao, Y. Abscisic acid dynamics, signaling, and functions in plants. *J. Integr. Plant Biol.* **2020**, *62*, 25–54. [CrossRef] [PubMed]
26. Jia, L.; Wang, J.; Wang, R.; Duan, M.; Qiao, C.; Chen, X.; Ma, G.; Zhou, X.; Zhu, M.; Jing, F.; et al. Comparative transcriptomic and metabolomic analyses of carotenoid biosynthesis reveal the basis of white petal color in *Brassica napus*. *Planta* **2021**, *253*, 8. [CrossRef] [PubMed]
27. Tomlekova, N.; Spasova-Apostolova, V.; Pantchev, I.; Sarsu, F. Mutation associated with orange fruit color increases concentrations of β-carotene in a sweet pepper variety (*Capsicum annuum* L.). *Foods* **2021**, *10*, 1225. [CrossRef] [PubMed]
28. Sheng, X.; Chen, H.; Wang, J.; Zheng, Y.; Li, Y.; Jin, Z.; Li, J. Joint transcriptomic and metabolic analysis of flavonoids in *cyclocarya paliurus* leaves. *ACS Omega* **2021**, *6*, 9028–9038. [CrossRef]
29. Zhao, D.; Zhao, L.; Liu, Y.; Zhang, A.; Xiao, S.; Dai, X.; Yuan, R.; Zhou, Z.; Cao, Q. Metabolomic and transcriptomic analyses of the flavonoid biosynthetic pathway for the accumulation of anthocyanins and other flavonoids in sweetpotato root skin and leaf vein base. *J. Agric. Food Chem.* **2022**, *70*, 2574–2588. [CrossRef] [PubMed]
30. Noguchi, A.; Nakamura, K.; Sakata, K.; Sato-Fukuda, N.; Ishigaki, T.; Mano, J.; Takabatake, R.; Kitta, K.; Teshima, R.; Kondo, K.; et al. Development and interlaboratory validation of a simple screening method for genetically modified maize using a ΔΔC(q)-based multiplex real-time PCR assay. *Anal. Chem.* **2016**, *88*, 4285–4293. [CrossRef] [PubMed]
31. Islam, S.N.; Nusrat, T.; Begum, P.; Ahsan, M. Carotenoids and β-carotene in orange fleshed sweet potato: A possible solution to vitamin A deficiency. *Food Chem.* **2016**, *199*, 628–631. [CrossRef] [PubMed]
32. Kang, L.; Park, S.C.; Ji, C.Y.; Kim, H.S.; Lee, H.S.; Kwak, S.S. Metabolic engineering of carotenoids in transgenic sweetpotato. *Breed. Sci.* **2017**, *67*, 27–34. [CrossRef] [PubMed]
33. Arias, D.; Arenas, M.A.; Flores-Ortiz, C.; Peirano, C.; Handford, M.; Stange, C. Daucus carota *DcPSY2* and *DcLCYB1* as tools for carotenoid metabolic engineering to improve the nutritional value of fruits. *Front. Plant Sci.* **2021**, *12*, 677553. [CrossRef] [PubMed]
34. Zhang, J.; Tao, N.; Xu, Q.; Zhou, W.; Cao, H.; Xu, J.; Deng, X. Functional characterization of Citrus *PSY* gene in Hongkong kumquat (*Fortunella hindsii* Swingle). *Plant Cell Rep.* **2009**, *28*, 1737–1746. [CrossRef] [PubMed]
35. Rodríguez-Villalón, A.; Gas, E.; Rodríguez-Concepción, M. Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown Arabidopsis seedlings. *Plant J.* **2009**, *60*, 424–435. [CrossRef]
36. Yamamizo, C.; Kishimoto, S.; Ohmiya, A. Carotenoid composition and carotenogenic gene expression during *Ipomoea* petal development. *J. Exp. Bot.* **2009**, *61*, 709–719. [CrossRef] [PubMed]
37. Suematsu, K.; Tanaka, M.; Kurata, R.; Kai, Y. Comparative transcriptome analysis implied a *ZEP* paralog was a key gene involved in carotenoid accumulation in yellow-fleshed sweetpotato. *Sci. Rep.* **2020**, *10*, 20607. [CrossRef] [PubMed]
38. Chauffour, F.; Bailly, M.; Perreau, F.; Cueff, G.; Suzuki, H.; Collet, B.; Frey, A.; Clément, G.; Soubigou-Taconnat, L.; Balliau, T.; et al. Multi-omics analysis reveals sequential roles for ABA during seed maturation. *Plant Physiol.* **2019**, *180*, 1198–1218. [CrossRef] [PubMed]
39. Jiang, F.; Zhang, J.; Wang, S.; Yang, L.; Luo, Y.; Gao, S.; Zhang, M.; Wu, S.; Hu, S.; Sun, H.; et al. The apricot (*Prunus armeniaca* L.) genome elucidates Rosaceae evolution and beta-carotenoid synthesis. *Hortic. Res.* **2019**, *6*, 128. [CrossRef] [PubMed]

40. Zhou, W.; Zhao, S.; Xu, M.; Niu, Y.; Nasier, M.; Fan, G.; Quan, S.; Zhang, S.; Wang, Y.; Liao, K. Identification of key genes controlling carotenoid metabolism during apricot fruit development by integrating metabolic phenotypes and gene expression profiles. *J. Agric. Food Chem.* **2021**, *69*, 9472–9483. [CrossRef]
41. González-Guzmán, M.; Apostolova, N.; Bellés, J.M.; Piqueras, P.; Ponce, M.R.; Micol, J.L.; Serrano, R.; Rodríguez, P.L. The short-chain alcohol dehydrogenase ABA2 catalyzes the conversion of xanthoxin to abscisic aldehyde. *Plant Cell* **2002**, *14*, 1833–1846. [CrossRef] [PubMed]
42. Kim, H.M.; Park, S.H.; Ma, S.H.; Park, S.Y.; Yun, C.H.; Jang, G.; Joung, Y.H. Promoted ABA hydroxylation by *capsicum annuum CYP707As* overexpression suppresses pollen maturation in nicotiana tabacum. *Front. Plant Sci.* **2020**, *11*, 583767. [CrossRef]
43. Dhar, M.K.; Sharma, R.; Koul, A.; Kaul, S. Development of fruit color in Solanaceae: A story of two biosynthetic pathways. *Brief. Funct. Genom.* **2014**, *14*, 199–212. [CrossRef] [PubMed]
44. Luo, F.; Cai, J.H.; Kong, X.M.; Zhou, Q.; Zhou, X.; Zhao, Y.B.; Ji, S.J. Transcriptome profiling reveals the roles of pigment mechanisms in postharvest broccoli yellowing. *Hortic. Res.* **2019**, *6*, 74. [CrossRef] [PubMed]
45. Liu, Y.; Roof, S.; Ye, Z.; Barry, C.; van Tuinen, A.; Vrebalov, J.; Bowler, C.; Giovannoni, J. Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. *Proc. Nati. Acad. Sci. USA* **2004**, *101*, 9897–9902. [CrossRef]
46. Yin, Y.; Guo, C.; Shi, H.; Zhao, J.; Ma, F.; An, W.; He, X.; Luo, Q.; Cao, Y.; Zhan, X. Genome-wide comparative analysis of the R2R3-MYB gene family in five solanaceae species and identification of members regulating carotenoid biosynthesis in wolfberry. *Int. J. Mol. Sci.* **2022**, *23*, 2259. [CrossRef] [PubMed]
47. Tang, W.; Ji, Q.; Huang, Y.; Jiang, Z.; Bao, M.; Wang, H.; Lin, R. FAR-RED ELONGATED HYPOCOTYL3 and FAR-RED IMPAIRED RESPONSE1 transcription factors integrate light and abscisic acid signaling in Arabidopsis. *Plant Physiol.* **2013**, *163*, 857–866. [CrossRef]
48. Wang, H.; Zhang, D.; Zhou, X.; Zhou, G.; Zong, W.; Chen, L.; Chang, Y.; Wu, X. Transcription factor AtOFP1 involved in ABA-mediated seed germination and root growth through modulation of ROS homeostasis in arabidopsis. *Int. J. Mol. Sci.* **2022**, *23*, 7427. [CrossRef] [PubMed]
49. Wang, S.; Chang, Y.; Guo, J.; Zeng, Q.; Ellis, B.E.; Chen, J.G. Arabidopsis ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. *PLoS ONE* **2011**, *6*, e23896. [CrossRef]

metabolites

MDPI

Article

Comparative Transcriptome Analysis of MeJA Responsive Enzymes Involved in Phillyrin Biosynthesis of *Forsythia suspensa*

Xiaoran Liu [1,2,†], Jiaqi Zhang [1,2,†], Hao Liu [1,2], Huixiang Shang [3], Xingli Zhao [1,2], Huawei Xu [1,2], Hongxiao Zhang [1,2] and Dianyun Hou [1,2,*]

[1] Agricultural of College, Henan University of Science and Technology, Luoyang 471023, China
[2] The Luoyang Engineering Research Center of Breeding and Utilization of Dao-di Herbs, Luoyang 471023, China
[3] Sanmenxia Academy of Agricultural Sciences, Sanmenxia 472000, China
* Correspondence: dianyun518@163.com
† These authors contributed equally to this work.

Abstract: *Forsythia suspensa* (Thunb.) has been widely used in traditional medicines in Asia. According to the 2020 edition of Chinese Pharmacopoeia, phillyrin is the main active ingredient in *F. suspensa*, which is effective in clearing heat, reducing swelling, and dispersing nodules. *F. suspensa* leaf is a non-toxic substance and it can be used to make a health tea. Here, we combine elicitors and transcriptomics to investigate the inducible biosynthesis of the phillyrin from the *F. suspensa*. After the fruits and leaves of *F. suspensa* were treated with different concentrations of methyl jasmonate (MeJA), the content of phillyrin in the fruits reached a peak at 200 μM MeJA for 12 h, but which was decreased in leaves. To analyze the differences in key enzyme genes involved in the phillyrin biosynthesis, we sequenced the transcriptome of *F. suspensa* leaves and fruits treated with 200 μM MeJA for 12 h. We hypothesized that nine genes related to coniferin synthesis including: *F. suspensa UDP-glycosyltransferase* (*FsUGT*); *F. suspensa 4-coumarate coenzyme CoA ligase* (*Fs4CL*); and *F. suspensa Caffeoyl-CoA O-methyltransferase (FsCCoAOMT)* etc. The qRT-PCR analysis of genes related to phillyrin biosynthesis was consistent with RNA-seq analysis. We also investigated the dynamic changes of genes in *F. suspensa* leaves and fruits at different time points after 200 μM MeJA treatment, which laid the foundation for further study of the molecular mechanisms regulating the biosynthesis of phillyrin.

Keywords: *Forsythia suspensa*; comparative transcriptome; MeJA treatment; phillyrin; phillyrin biosynthesis

Citation: Liu, X.; Zhang, J.; Liu, H.; Shang, H.; Zhao, X.; Xu, H.; Zhang, H.; Hou, D. Comparative Transcriptome Analysis of MeJA Responsive Enzymes Involved in Phillyrin Biosynthesis of *Forsythia suspensa*. *Metabolites* **2022**, *12*, 1143. https://doi.org/10.3390/metabo12111143

Academic Editors: Yanjie Zhang and Yan Li

Received: 22 October 2022
Accepted: 18 November 2022
Published: 20 November 2022

1. Introduction

Forsythia suspensa (Thunb.) is utilized as a common traditional medicine in China, Japan, Korea, and many European countries. It is called 'Lianqiao' in China [1]. Based on the different harvest time, *F. suspensa* fruits can be classified into 'Qingqiao' and 'Laoqiao' forms. In folk medicine, the extract of the dried fruit has long been used to treat a variety of diseases, such as inflammation, pyrexia, gonorrhea, tonsillitis, and ulcers [2]. *F. suspensa* fruit is also the main active ingredient in many widely used classic Chinese patent medicine prescriptions, such as Shuanghuanglian injections [3] and Lianhua Qingwen granules [4]. Lianhua Qingwen is also recommended for the treatment of COVID-19 [5]. The dried ripe fruit of *F. suspensa* has also been prescribed for the treatment of diabetes in China [6]. In addition to the medicinal value of *F. suspensa* fruit, it has also been reported in recent years that the extract of *F. suspensa* leaves has antibacterial and other pharmacological effects [7]. In folk medicine, *F. suspensa* leaves are also used as tea [8]. Therefore, it is of great significance to study the secondary metabolites and their synthetic pathways in *F.*

suspensa. In recent years, many active ingredients have been identified in *F. suspensa*, such as phenylethanoid glycosides, lignans, flavonoids, terpenes, and volatile oils [9]. Lignan and phenylethanol glycoside are the two main representative characteristic components in *F. suspensa*; more than 30 lignans and lignan glycosides, including (+)-forsythin, have been isolated from *F. suspensa* alone [10]. Phillyrin (also named forsythin), a kind of lignan substance, mitigates apoptosis and oxidative stress [11], has anti-viral and anti-inflammatory activity [12], and improves insulin resistance [13]. The biosynthesis of lignan (including phillyrin) consists of three stages: coniferyl alcohol, the precursor of lignan, is synthesized through the phenylpropanoid pathway; then, structurally diverse lignans are synthesized from coniferyl alcohol; and finally, lignans are modified by glycosylation to form lignan glycosides, which are accumulated and stored in plant cells and tissues [14].

The phenylpropane biosynthesis pathway shares some intermediates with the lignan biosynthesis pathways [15]. Based on previous study, phenylalanine ammonia-lyase enzyme (PAL), coumarate 3-hydroxylase (C3H), 4-coumarate CoA ligase (4CL), caffeoyl o-methyltransferase (COMT), caffeoyl-CoA o-methyltransferase (CCoAOMT), cinnamoyl-CoA reductase (CCR), cinnamyl alcohol dehydrogenase (CAD), dirigent protein (DIR), o-methyltransferase (OMT), and udp-glycosyltransferase (UGT), are key enzymes of the phillyrin-related secondary metabolism in plants. PAL is the enzyme that catalyzes the first step reaction of phenylpropanoid metabolism, connects the biological primary metabolism, and is the key and rate-limiting enzyme of phenylpropanoid metabolism [16]. C3H catalyzes the formation of caffeic acid from coumaric acid, which is methylated to form ferulic acid [17]. The enzyme 4CL activates cinnamic acid and its hydroxylated derivatives by forming the corresponding CoA thioesters, and it is one of the key enzymes in lignan biosynthesis [18]. Both COMT and CCoAOMT are substrate methylation enzymes, and lignan monomer synthesis requires a two-step methylation reaction at the 3′ and 5′ positions of the two-step methylation reaction; COMT and CCoAOMT are two methylation enzymes at different substrate levels that are associated with the specific synthesis of lignan monomers [19,20]. As an enzyme that catalyzes the first reaction of the lignan synthesis pathway, CCR is considered to be a potential control point for regulating the flow of carbon to lignans [21]. CAD catalyzes the formation of coniferyl dehyde to coniferyl alcohol, which is one of the key enzymes in lignan synthesis [14]. Although a coniferyl alcohol has yet to be identified, a dirigent protein (DIR) was shown to participate in the stereo-specific dimerization of E-coniferyl alcohol [22]. In the biological pathway for the formation of phillyrin from coniferyl alcohol, after the formation of (+)-epipinoresinol, phillygenin can be produced by oxygen methylation, possibly mediated by a specifc O-methylase [23]. Most lignans exist in plants in the form of glycosylation, which is catalyzed by UGT to generate lignan glycosides [24]. Methyl jasmonate (MeJA), a well-known exogenous inducing factor, participates in many plant processes, ranging from plant defense to growth and development [25]. As an abiotic inducer, MeJA can rapidly and selectively induce the expression of specific gene associated with specific biological process during plant secondary metabolism, thereby regulating the synthesis of secondary metabolites in plant cells [26,27]. MeJA-induced transcriptome changes have been analyzed in many different plant species including wheat and sweet basil [28,29]. In addition, genomic and transcriptomic information of *F. suspensa* is scarce in public databases. To identify the genes involved in the biosynthesis of phillyrin, the fruits and leaves of *F. suspensa* plants after 12 h of MeJA treatment were explored by transcriptional sequencing, based on the Illumina Novaseq 6000 sequencing platform, and the relevant genes in the phillyrin synthesis pathway were analyzed using Quantitative Real-Time PCR (qRT-PCR). This study provides a theoretical basis for further exploring the molecular regulation mechanism of secondary metabolites in *F. suspensa* and reveals the mechanism of the formation of phillyrin in *F. suspensa*.

2. Materials and Methods

2.1. Plant Materials and Treatment

F. suspensa used in this study was grown in 2021 at the *F. suspensa* medicine source base in Yiyang, Luoyang, China. In this study, fruit, and leaf tissues of *F. suspensa* plants, were treated with 50, 200, and 400 μM of MeJA. Preparation of MeJA (Sigma-Aldrich, 95%, Sigma, St. Louis, MO, USA) stock solution: 218 μL MeJA stock solution was dissolved in 2 mL absolute ethanol, then diluted with ddH_2O to make the 10 mM MeJA solution. The high (400 μM), medium (200 μM), and low (50 μM) concentrations of MeJA treatment solution were obtained by diluting the 10 mM MeJA solution, respectively. During the experiment, final concentrations of MeJA solutions (50, 200, and 400 μM) were sprayed on the fruit and leaf surfaces of *F. suspensa* plants until the solution covered the leaf and fruit surfaces just dripping down. Control plants were sprayed with an equal volume of ethanol in ddH_2O. For each treatment, each replicate was considered and three plants were collected. Leaves and fruits from treatments 0 h (as control), 7 h, 12 h, 24 h, 48 h were collected, frozen in liquid nitrogen, and stored at −80 °C. The *F. suspensa* fruit treated with 50 μM MeJA, 200 μM MeJA, and 400 μM MeJA were represented as M1F, M2F and M3F, respectively. The *F. suspensa* leaf treated with 50 μM MeJA, 200 μM MeJA, and 400 μM MeJA, were represented as M1L, M2L, and M3L, respectively.

2.2. Determination of Phillyrin Content in F. suspensa Fruits and Leaves

High performance liquid chromatography (HPLC, Agilent 1260 Infinity, Agilent Technologies Co. Ltd., Palo Alto, CA, USA) was used to determine the content of phillyrin in *F. suspensa* fruit and leaf. The MeJA-treated *F. suspensa* fruits and leaves were lyophilized and homogenized in a sterilized mortar. The extract was sonicated with 15 mL of methanol (70% *v*/*v*) for 30 min for a total of 0.5 g. The extract was cooled to room temperature and centrifuged at 4000 rpm for 10 min. The filtrate was then filtered through a 0.22 μm organic microporous membrane and stored under seal.

HPLC conditions: the mobile phase consisted of acetonitrile (A)-water (B) with gradient elution (0–40 min, 25% A) was at flow rate 1.0 mL min^{-1}, column temperature 30 °C, the column was manufactured by Thermo Fisher Scientific (Waltham, MA, USA) Co., Ltd., with a diameter of 4.6 mm, a length of 250 mm, and a particle size of 5 μm. Detection wavelength 277 nm. The injection volume of the HPLC method was 10 μL. For the quantification of phillyrin, standards were used. The results were described in grams per kilogram of fresh weight. All experiments were repeated in triplicate.

2.3. RNA Extraction, cDNA Library Construction and Illumina Sequencing

We analyzed the content of phillyrin at different time periods after different concentrations of MeJA treatment by HPLC, and selected *F. suspensa* fruits (M2F)/leaves (M2L) treated with 200 μM MeJA for 12 h and *F. suspensa* fruits (CKF)/leaves (CKL) treated with control solution for 0 h for transcriptome sequencing including 12 samples with three replicates for each material. Total RNA was extracted from the tissue using TRIzol® Reagent (Invitrogen, Carlsbad, CA, USA) (Plant RNA Purification Reagent for plant tissue) according the manufacturer's instructions (Invitrogen) and genomic DNA was removed using DNase I (TaKara, Japan). RNA degradation and contamination was monitored on 1% agarose gels. Then, the integrity and purity of the total RNA quality was determined by 2100 Bioanalyser (Agilent Technologies Co. Ltd., Palo Alto, CA, USA) and quantified using the ND-2000 (NanoDrop Technologies, Wilmington, DE, USA). Only high-quality RNA sample (OD260/280 = 1.8~2.2, OD260/230 = 1.8~2.2, RIN = 8.8~10.0, 28S:18S = 1.4~1.8) 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). Paired-end RNA-seq sequencing library was sequenced with the Illumina NovaSeq 6000 sequencer (Illumina, San Diego, CA, USA) (2 × 150 bp read length). Each sample contained three biological replicates.

2.4. Transcriptome Assembly and Annotation

The high-quality clean reads were obtained after removing reads containing adapters and low-quality reads. Transcriptome de novo assembly was performed using the Trinity program [30]. Functional annotation of the unigenes was carried out using BLASTx program with an E-value threshold of 1×10^{-5} against the NCBI non-redundant protein (Nr) datbase [31], Swiss-Prot protein database, COG database [32], and Kyoto Encyclopedia of Genes and Genomes (KEGG) database [33]. GO annotation was analyzed by Blast2GO software (Version 2.2.31+) [34]. Transcription factors (TFs) were identified using PlantTFDB database with default parameters [35].

2.5. Differentially Expressed Genes (DEGs) Analysis

The high-quality RNA reads were counted against bowtie's paired results [36] by RSEM (http://deweylab.biostat.wisc.edu/rsem/). Differential expression analyses among the two treatments (CKL-vs.-M2L and CKF-vs.-M2F with three biological replicates per treatment) were conducted using edgeR software (Version 3.24.3) [37]. Genes with a fold change ≥ 2 and a false discovery rate (FDR) < 0.05 in a comparison were defined as significant DEGs. DEGs were then subjected to enrichment analysis of GO functions and KEGG pathways. Heatmaps of DEGs were analyzed using HemI [38].

2.6. Quantitative Real-Time PCR (qRT-PCR)

The RNA samples (1 μg) with A260/A230 ratios between 1.8 and 2.2 were used to synthesize the first strand of cDNA with the First Strand cDNA Synthesis Kit (CISTRO, Guangzhou, China) in a 20 μL reaction mixture according to the manufacturer's protocol. The cDNA samples were diluted 10-fold with distilled water prior to the qRT-PCR analysis. The qRT-PCR analysis was performed using ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) and Roche Light Cycle96 Real-Time PCR system with the following reaction conditions: initial denaturation at 95 °C for 30 s, followed by a two-step program of 95 °C for 10 s and 60 °C for 30 s for 40 cycles, with a melting curve analysis at 95 °C for 15 s, then from 60 °C to 95 °C at a rate of 0.2 °C/s. Candidate genes related to phillyrin were identified from *F. suspensa* transcriptome data, the longest CDS sequence was selected, and the specific primers for the target genes were designed using Oligo 7.0, and are listed in Table S1. The *UKN1* + *SDH* + *G6PD* genes were used as reference gene group and the relative expression levels were determined according the $2^{-\Delta\Delta Ct}$ method [39,40]. The specificity of the primer pair was verified by the presence of a single peak in the melt curve analysis during the qRT-PCR process (Figure S1). All biological replicates were performed in triplicate.

2.7. Statistical Analyses

Data were analyzed with one-way ANOVA by using SPSS 17.0, and means were contrasted with Duncan's multiple scope test at $p < 0.05$. All the experiments were carried out in triplicate.

3. Results

3.1. Effect of MeJA Induction on the Content of Phillyrin

The effect of MeJA treatment on the content of phillyrin in the leaves and fruits of *F. suspensa* was shown in Figure 1. After 50 μM MeJA treatment, the phillyrin content in *F. suspensa* fruits was slightly increased compared with the control. Meanwhile, the content of phillyrin was gradually increased and then decreased after 200 μM MeJA treatment, peaking at 12 h with 1.63-fold compared with control. After 400 μM MeJA treatment, the content of phillyrin was obviously increased at 7 h and 24 h with 1.60 and 1.51-fold compared with control, respectively (Figure 1a). After 50 μM MeJA treatment of *F. suspensa* leaves, the content of phillyrin decreased compared to the control. After 200 μM MeJA treatment, the content of phillyrin decreased, then increased, and finally decreased with time. 400 μM MeJA treatment of *F. suspensa* leaves showed a decrease in phillyrin content at

12 h compared with the control (Figure 1b). This indicates that different parts of *F. suspensa* responded differently to the induction of different MeJA concentrations. The exogenous MeJA treatment of *F. suspensa* led to the induction of phillyrin in response to MeJA, but the induction effect varied among time and tissue parts. Due to *F. suspensa* leaves having been used to make health-care tea in folk tradition [8], we also treated them with 200 μM MeJA for 12 h to investigate the different mechanisms of phillyrin synthesis in *F. suspensa* leaves and fruits, under the same conditions of treatment.

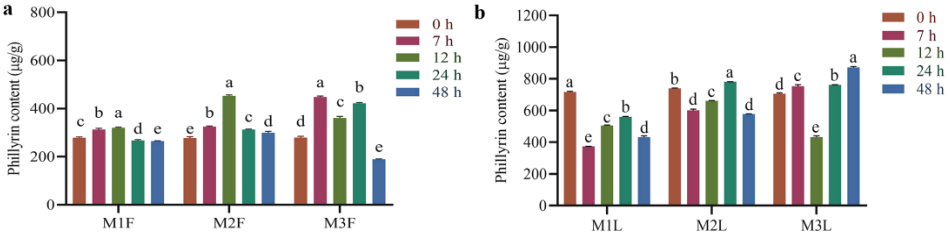


Figure 1. Phillyrin content of *F. suspensa* fruits and leaves treated with different MeJA. (**a**) Phillyrin content of *F. suspensa* fruits treated with different MeJA. (**b**) Phillyrin content of *F. suspensa* leaves treated with different MeJA. M1F: *F. suspensa* fruit treated with 50 μM MeJA; M2F: *F. suspensa* fruit treated with 200 μM MeJA; M3F: *F. suspensa* fruit treated with 400 μM MeJA. M1L: *F. suspensa* leaf treated with 50 μM MeJA; M2L: *F. suspensa* leaf treated with 200 μM MeJA; M3L: *F. suspensa* leaf treated with 400 μM MeJA. Data were analyzed by SPSS, followed by Duncan's honestly significant difference test at $p < 0.05$. All Statistical analyses of data had three biological repeats. All data are displayed as means ± SD.

3.2. Illumina Sequencing and De Novo Assembly

To further understand the molecular mechanism underlying the response of phillyrin to MeJA, an RNA-seq analysis was performed. We established 12 cDNA libraries for transcriptome sequencing. Illumina Hiseq 6000 platform generated 98.21 Gb clean reads with an average Q30 value of 94.55% and an average GC content of 44%, Q20 > 98%, and Q30 > 94% (Table 1). The max length, min length, and N50 length, for the assembled unigenes were 15,856, 201, and 1851 bp, respectively. With regards to the length distribution of the unigenes, the total number of sequences of all unigene lengths of 200–500 bp in the transcriptome data of *F. suspensa* is much higher than the total number of sequences of other respective lengths (Figure S2).

Table 1. Summary of sequencing and assembly for *F. suspensa*.

Type	Unigene
Total number	87,564
Total base	87,582,604
Largest length (bp)	15,856
Smallest length (bp)	201
Average length (bp)	1000.21
GC percent (%)	44
N50 average length (bp)	1851
Q_{30} (%)	94.55
Q_{20} (%)	98

All of the sequencing data were deposited in the National Center for Biotechnology Information (NCBI). The sequence read archive (SRA) accession number is SRP392325. The resulting fastq files were deposited on SRR21066404, SRR21066405, SRR21066406, SRR21066407, SRR21066412, SRR21066413 (leaf) and SRR21066408, SRR21066409, SRR21066410, SRR21066411, SRR21066414, SRR21066415 (fruit).

3.3. Functional Annotation of Unigenes

A total of 37,297, 25,361, 15,161, 30,017, 29,688, and 24,539 unigenes had significant hits (E-value $\leq 10^{-5}$) in the Nr (non-redundant protein) database, Swissprot, Kyoto Encyclopedia of Genes and Genomes (KEGG), the COG database and the Pfam database, respectively (Figure S3). A total of 60,093 CDS unigenes produced by the Illumina Hiseq 2500 platform, and most CDSs, were less than 1800 bp (Figure S4).

For GO annotation, 79,094 unigenes were allocated into three categories ("biological process", "cellular component", and "molecular function"). Among the biological processes, the unigenes were mainly involved in metabolic processes (8180, GO: 0008152), cellular processes (9058, GO: 0009987), and biological regulation (2432, GO: 0065007). Among the cellular components category, the unigenes were primarily associated with cellular fraction (7846, GO: 0044464), membrane fraction (7269, GO: 0044425), and organelles (4424, GO: 0043226). Among the molecular function categories, unigenes with catalytic activity (10,376, GO: 0003824), binding (12,606, GO: 0005488), and transporter protein activity (1323, GO: 0005215), accounted for a larger proportion (Figure 2a).

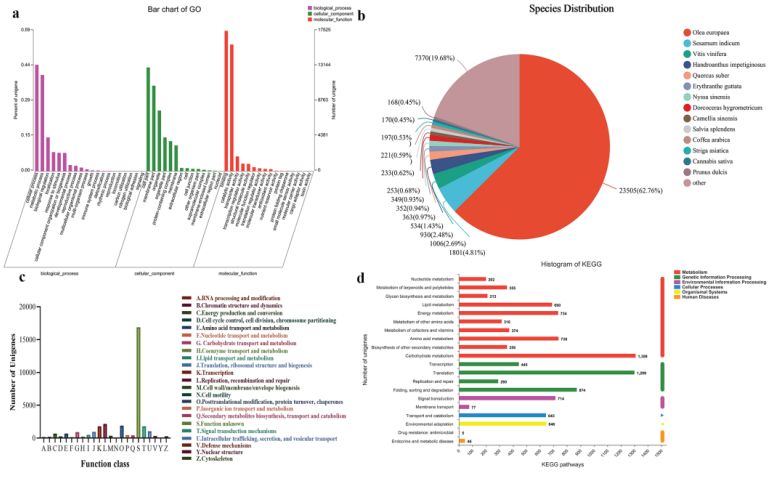


Figure 2. Functional annotation of *F. suspensa* transcriptome. (**a**) Gene Ontology (GO) functional classification of assembled unigenes. (**b**) Homologous species distribution of *F. suspensa* annotated in the NR database. (**c**) Clusters of Orthologous Groups (COG) functional classification of assembled unigenes. (**d**) Functional classification and pathway assignment of assembled unigenes by Kyoto Encyclopedia of Genes and Genomes (KEGG).

Unigene sequences were searched against the NR database for annotation and revealed 23,505 unigenes (62.76%) matched to *Olea europaea* and 1801 unigenes (4.81%) matched to *Sesamum indicum* (Figure 2b).

To further evaluate the completeness of the transcriptome library and the validity of the annotation, the COG classification of unigenes was performed. Among the 23 COG categories, "L: Replication, recombination and repair" was the largest group, followed by "O: Posttranslational modification, protein turnover, chaperones and ion, protein turnover, chaperones" and "T: Signal transduction mechanisms". The smallest group was "Y: Nuclear structure" with only two unigenes annotated to this category (Figure 2c).

The KEGG database was used to identify the biochemical pathways assigned to unigene sequences. In our results, a total of 10,325 unigenes were annotated to six metabolic pathways: metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, and human diseases. Metabolism had the largest number of pathways and human diseases had the smallest number of pathways.

There were 356 unigenes matching to the biosynthesis of other secondary metabolism, and 224 unigenes of phenylpropanoid biosynthesis pathway (Figure 2d).

3.4. Analysis of Differentially Expressed Unigenes (DEGs)

To obtain a comprehensive view of the gene expression profile associated with the response of phillyrin to MeJA, we used edgeR to identify the DEGs. A total of 4211 (2348 up-regulated and 1863 down-regulated) DEGs and 8433 (4134 up-regulated and 4299 down-regulated) DEGs were identified in fruits and leaves of *F. suspensa* treated 12 h with MeJA, respectively (Figure 3); this finding indicated that there were more upregulated than downregulated unigenes in *F. suspensa* fruits and more downregulated than upregulated unigenes in *F. suspensa* leaves under MeJA treatment. Furthermore, we used principal component analysis (PCA) to display the overall relationships between the transcriptome from the different samples (Figure S5), indicating the leaves and fruits of *F. suspensa* before and after MeJA treatment showed mutual aggregation within the group. Meanwhile, two different transcriptional variation tendencies were observed before and after the MeJA treatment of *F. suspensa* fruits, and *F. suspensa* leaves also showed the same trend.

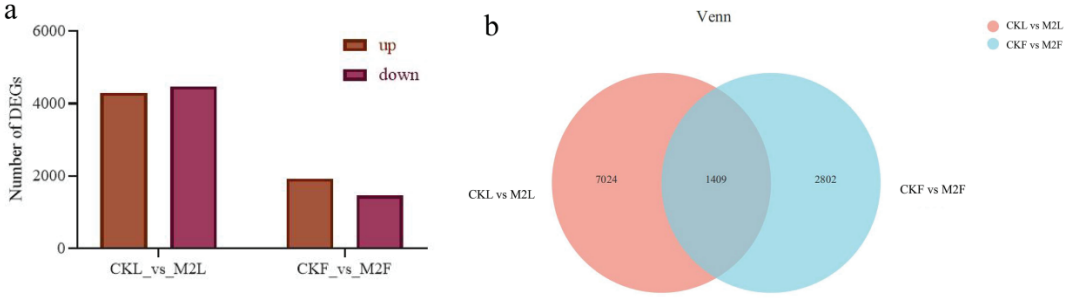


Figure 3. Statistics of DEGs in *F. suspensa* transcriptomes. (**a**) DEG numbers in comparisons of the CKF-vs.-M2F and CKL-vs.-M2L. (**b**) Venn diagram showing significantly DEGs at different tissue. CKL: *F. suspensa* leaf treated with control solution for 0 h; M2L: *F. suspensa* leaf treated with 200 μM MeJA; CKF: *F. suspensa* fruit treated with control solution for 0 h; M2F: *F. suspensa* fruit treated with 200 μM MeJA.

Functional enrichment analysis revealed that 3754 and 1822 DEGs were associated with GO terms in *F. suspensa* leaves and fruits, respectively (Figure 4). DEGs were mainly enriched in oxidoreductase activity (GO: 0016682), and phenylpropanoid metabolic process (GO: 0009698) (Figure 4a,b). The KEGG pathways analysis indicated DEGs were mainly involved in the metabolism of phenylpropanoid biosynthesis, plant-pathogen interaction, MAPK signaling pathway-plant, phenylalanine metabolism, and plant hormone signal transduction (Figure 4c,d).

Previous study showed that phillyrin as a kind of lignan was tightly related to phenylpropanoid biosynthesis [41]. In the phillyrin biosynthesis pathway, three *FsCCR*, four *FsCAD*, six *FsDIR*, five *FsOMT*, one *FsPAL*, and thirteen *FsUGT* genes were significant different after MeJA treatment in *F. suspensa* leaves, respectively (Figures 3 and 5). Meanwhile, two *Fs4CL*, two *FsCAD*, one *FsCCR*, four *FsDIR*, three *FsOMT*, two *FsPAL*, and twelve *FsUGT* were significant different in *F. suspensa* fruits. Furthermore, one *FsUGT*, three *FsOMT*, two *FsDIR*, one *FsPAL*, and one *FsCCR* were significant different both in *F. suspensa* leaves and fruits (Figures 3 and 5). These results suggest that MeJA may induce the accumulation of phillyrin by regulating the expression of genes associated with phillyrin biosynthesis in *F. suspensa*.

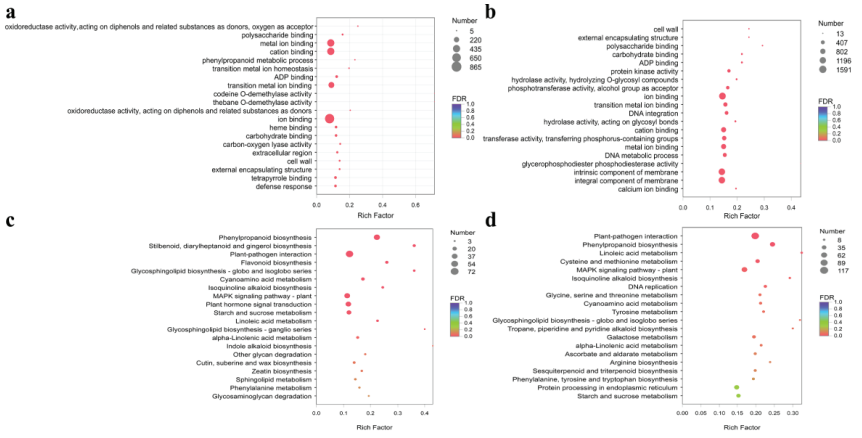


Figure 4. Enrichment analysis of differentially expressed genes in the GO and KEGG pathways. (**a**) GO pathway enrichment analysis in CKF-vs.-M2F. (**b**) GO pathway enrichment analysis in CKL-vs.-M2L. (**c**) KEGG pathway enrichment analysis in CKF-vs.-M2F. (**d**) KEGG pathway enrichment analysis in CKL-vs.-M2L. CKL: *F. suspensa* leaf treated with control solution for 0 h; M2L: *F. suspensa* leaf treated with 200 μM MeJA; CKF: *F. suspensa* fruit treated with control solution for 0 h; M2F: *F. suspensa* fruit treated with 200 μM MeJA.

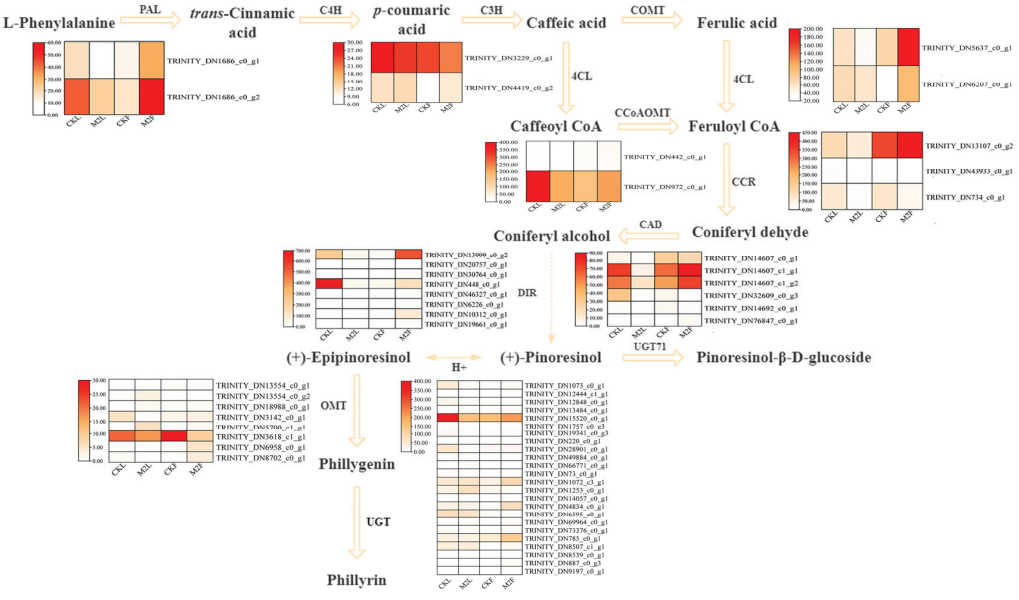


Figure 5. Putative pathway for phillyrin synthesis in *F. suspensa*. PAL, phenylalanine ammonia-lyase; C3H, Coumarate 3-hydroxylase; 4CL, 4-coumarate CoA ligase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; DIR, dirgent protein; OMT, O-methyltransferase; UGT, UDP-glycosyltransferase. CKL: *F. suspensa* leaf treated with control solution for 0 h; M2L: *F. suspensa* leaf treated with 200 μM MeJA; CKF: *F. suspensa* fruit treated with control solution for 0 h; M2F: *F. suspensa* fruit treated with 200 μM MeJA.

3.5. Analysis of DEGs Related to Phillyrin Biosynthetic Pathway

To verify the expression level of genes related to phillyrin biosynthesis exhibited by the identified DEGs in response to MeJA, the expression levels of nine genes (*FsUGT*, *Fs4CL*, *FsCCoAOMT*, *FsOMT*, *FsPAL*, *FsDIR*, *FsCAD*, *FsCCR*, and *FsC3H*) were detected by qRT-PCR. The expression profiles were generally consistent with the RNA-seq results (Figure 6a). Correlation analysis based on the qRT-PCR and transcriptome data also showed a significant correlation with a Spearman correlation coefficient of 0.819, which reflects the accuracy of transcriptome data (Figure 6b). To analyze the dynamic expression profiles of genes putatively related to phillyrin synthesis under MeJA treatment, genes expression level (*FsUGT*, *Fs4CL*, *FsCCoAOMT*, *FsOMT*, *FsPAL*, *FsDIR*, *FsCAD*, *FsCCR*, and *FsC3H*) in *F. suspensa* fruits and leaves after MeJA treatment in different time (0 h, 7 h, 12 h, and 24 h) were studied. The results revealed that MeJA activated or suppressed the expression of genes to variant degrees at different time points and in different tissues.

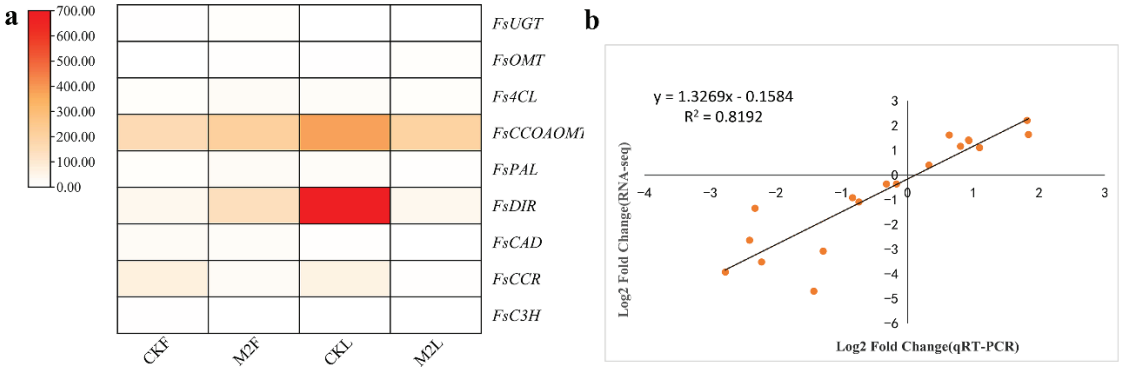


Figure 6. Expression patterns of genes related to phillyrin biosynthesis. (**a**). Transcriptional profiles of DEGs related to phillyrin biosynthetic pathway. (**b**). Scatter plot and linear regression based on qRT-PCR and RNA-seq. The correlation coefficient was calculated using the Spearman correlation method. *FsPAL*, *phenylalanine ammonia-lyase gene*; *FsC3H*, *Coumarate 3-hydroxylase gene*; *Fs4CL*, *4-coumarate CoA ligase gene*; *FsCCoAOMT*, *caffeoyl-CoA O-methyltransferase gene*; *FsCCR*, *cinnamoyl-CoA reductase gene*; *FsCAD*, *cinnamyl alcohol dehydrogenase gene*; *FsDIR*, *dirgent protein gene*; *FsOMT*, *O-methyltransferase gene*; *FsUGT*, *UDP-glycosyltransferase gene.* CKL: *F. suspensa* leaf treated with control solution for 0 h; M2L: *F. suspensa* leaf treated with 200 μM MeJA; CKF: *F. suspensa* fruit treated with control solution for 0 h; M2F: *F. suspensa* fruit treated with 200 μM MeJA.

The expression of *FsDIR*, *FsUGT*, and *FsOMT* genes were up-regulated at all time points in 200 μM MeJA-treated *F. suspensa* fruit. In the presence of MeJA, the expression levels of *Fs4CL*, *FsPAL*, *FsCCR*, and *FsCAD* genes, were inhibited at the earlier time points and activated at the later time points. The expression level of *FsCCoAOMT* gene reached a maximum at 12 h, and then its expression decreased compared with the 0 h control. The expression levels of *FsC3H* gene reached a maximum at 7 h and were repressed thereafter compared with the 0 h control (Figure 7a). In *F. suspensa* leaves, *Fs4CL*, *FsPAL*, and *FsCCoAOMT* genes expression levels were down-regulated at all time points, and *FsOMT* and *FsUGT* genes expression were up-regulated within 24 h of MeJA treatment compared with the 0 h control. *FsCCR*, *FsC3H*, and *FsCAD* genes expression were repressed at early time points, afterwards, expression gradually increased (Figure 7b).

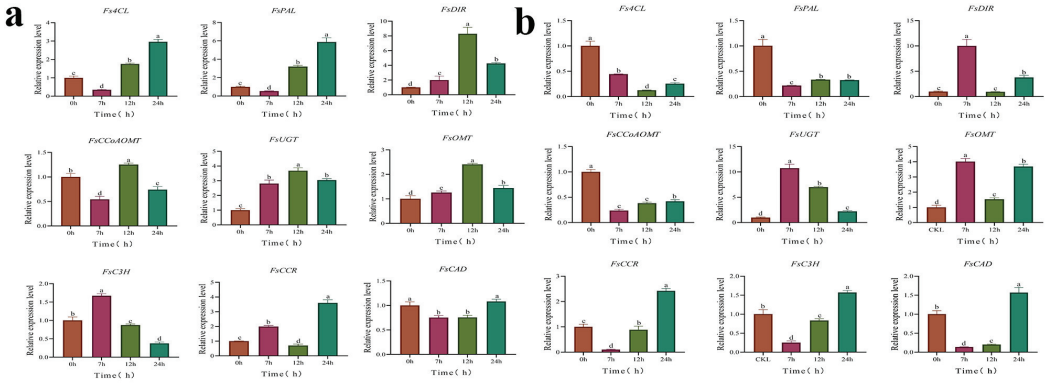


Figure 7. Analysis of the expression profiles of nine genes related to phillyrin synthesis after 200 μM MeJA treatment for 0, 7, 12, and 24 h. (**a**). Relative expression levels of phillyrin biosynthesis genes in *F. suspensa* fruits are detected by qRT-PCR (0 h as 1). (**b**). Relative expression levels of phillyrin biosynthesis genes in *F. suspensa* leaves are detected by qRT-PCR (0 h as 1). The *UKN1* + *SDH* + *G6PD* genes are used as the reference gene group to standardize the RNA samples for each reaction. All Statistical analyses of data had three biological repeats. Data are analyzed by SPSS, followed by Duncan's honestly significant difference test at $p < 0.05$. All Statistical analyses of data had three biological repeats. *FsPAL*, *phenylalanine ammonia-lyase gene*; *FsC3H*, *Coumarate 3-hydroxylase gene*; *Fs4CL*, *4-coumarate CoA ligase gene*; *FsCCoAOMT*, *caffeoyl-CoA O-methyltransferase gene*; *FsCCR*, *cinnamoyl-CoA reductase gene*; *FsCAD*, *cinnamyl alcohol dehydrogenase gene*; *FsDIR*, *dirgent protein gene*; *FsOMT*, *O-methyltransferase gene*; *FsUGT*, *UDP-glycosyltransferase gene*.

4. Discussion

The content of phillyrin is an important index for evaluating the quality of *F. suspensa*. Phillyrin are effective in clearing heat and detoxifying, reducing swelling, and dispersing nodules [42]. It is important to further understand the phillyrin biosynthesis pathway and provide robust candidate genes for further functional investigations aimed at the improvement of phillyrin content and quality.

MeJA is a cyclopentanone derivative-like signaling substance commonly found in the plant kingdom that regulates plant growth and development, triggers cells to initiate protective mechanisms, and stimulates the expression of key enzymes of metabolic pathways [43]. Plant defense systems are initiated upon injury or by signals from exogenous plant hormones, such as MeJA. Moreover, lignans, at least in part, are believed to be involved in host defense systems [44–46]. In combination, the hormone is expected to enhance lignan biosynthesis [45,47]. We also found that MeJA could induce the accumulation of phillyrin in *F. suspensa* fruit (Figure 1a), but a certain concentration of MeJA could inhibit the synthesis of phillyrin in *F. suspensa* leaves (Figure 1b). The content of phillyrin in leaves without MeJA treatment was higher than that in fruits without MeJA treatment, which was consistent with the results of previous studies [48]. The efficacy of *F. suspensa* leaves recorded in traditional Chinese medicine classics is very similar to that of *F. suspensa* fruits, and the chemical composition is also similar [49]. According to the theory of traditional Chinese medicine, the effects of different parts of *F. suspensa* should be completely different, therefore, we performed RNA-seq detection on *F. suspensa* fruits and leaves to compare the molecular mechanism of phillyrin biosynthesis in *F. suspensa* fruits and leaves, laying a foundation for the comprehensive development and utilization of *F. suspensa* resources.

GO and KEGG results showed DEGs were mainly enriched in phenylpropanoid metabolic process in *F. suspensa* after MeJA treatment (Figure 4), the phenylpropanoid pathway, a common pathway of phillyrin biosynthesis pathway, can produce coniferyl alcohol, which are precursors of phillyrin [50], including *Fs4CL*, *FsCCR*, *FsCAD*, *FsC3H*, *FsCCoAOMT*, and *FsPAL* genes, which were significantly different in *F. suspensa* after

MeJA treatment compared with control (Figures 3 and 5). Research shows that during phillyrin biosynthesis, two coniferyl alcohol molecules are coupled to produce compounds of different configurations, and then converted to (+)-epipinoresinol. Furthermore, the (+)-epipinoresinol is converted to phillygenin, eventually forming phillyrin [23], including the *FsDIR, FsOMT,* and *FsUGT* gene, which were noticeably different in *F. suspensa* after MeJA treatment compared with control. When *F. suspensa* fruits were treated with different concentrations of MeJA, the content of phillyrin in the fruits increased, and most of the genes in its synthesis pathway showed an increasing trend; the expression of *Fs4CL*, *FsCCR*, *FsCAD*, *FsC3H*, *FsCCoAOMT*, and *FsPAL* in MeJA-treated *F. suspensa* leaves was suppressed for a certain period of time, while the expression of *FsOMT* and *FsUGT* increased, and phillyrin was also decreased (Figures 1 and 7). Different parts of *F. suspensa* have different responses to MeJA induction, which has been reported in other plants [51]. This may indicate that genes, during the synthesis of the lignan precursor substance coniferyl alcohol, also have an important influence on the synthesis of the end product phillyrin; and that there are complex regulatory mechanisms involved in the transcription of genes to translation into proteins, from the processing and modification of proteins into active enzymes to secondary metabolites and accumulation.

The fruit of *F. suspensa* is the main part of the medicine; our study showed that MeJA treatment could effectively increase the content of phillyrin in *F. suspensa* fruits (Figure 1). The expression of *FsOMT* and *FsUGT* was up-regulated in both the fruits and leaves of *F. suspensa* after MeJA treatment. Previous study showed that *Lycoris aurea LaOMT1* transcripts were significantly increased after MeJA treatment [52]; *Bupleurum chinense DC.* glycosyltransferase genes *BcUGT1*, *BcUGT3,* and *BcUGT6* could be induced by MeJA with different degrees of up-regulation of expression [53]. The expression of *FsCAD*, *FsCCR,* and *FsC3H* did not correspond to the phillyrin content in *F. suspensa* fruit, and the expression of *FsOMT* and *FsUGT* did not correspond to the phillyrin content in *F. suspensa* leaves. The result was not a surprise because the formation of end products is collaborative with all genes in the synthesis pathway. The present study cannot adequately account for the role of these individual genes in phillyrin biosynthesis. Therefore, it is crucial to establish a transgenic system for *F. suspensa* in order to further confirm the function of genes in the phillyrin synthesis pathway. Recently, our team has been making a concerted effort to establish a system for transgenic *F. suspensa* with an emphasis on hairy roots and callus. In this transgenic system, a deeper understanding of the molecular mechanisms of phillyrin will be gained, and the biotechnological improvement of *F. suspensa*.

5. Conclusions

In this study, using de novo sequencing, a *F. suspensa* dataset containing 87,564 unigenes was constructed to molecular mechanism underlying the response of phillyrin to MeJA in *F. suspensa* leaves and fruits. A total of 4211 and 8433 DEGs were identified in the fruits and leaves of *F. suspensa* treated with MeJA for 12 h, respectively. GO and KEGG analyses indicated that DEGs were mainly enriched in phenylpropanoid metabolic process in *F. suspensa*, which was tightly associated with phillyrin biosynthesis. The qRT-PCR results showed that *FsUGT*, *Fs4CL*, *FsCCoAOMT*, *FsOMT*, *FsPAL*, and *FsDIR* were significantly up-regulated in *F. suspensa* fruits after MeJA treatment; the corresponding *F. suspensa* fruit also increased the content of phillyrin. While *FsCAD Fs4CL*, *FsCCoAOMT*, *FsCCR*, *FsPAL*, *FsC3H*, and *FsDIR* were significantly down-regulated in *F. suspensa* leaves, the content of phillyrin in *F. suspensa* leaves with the same treatment was reduced compared with that of the control. These qRT-PCR data were consistent with RNA-seq data. Taken together, the analysis of transcriptome results of *F. suspensa* leaves and fruits lay a foundation for the comprehensive development and utilization of *F. suspensa* resources. These results provide a robust theoretical basis for accelerating the study of the regulatory mechanism of MeJA-induced phillyrin biosynthesis in *F. suspensa*.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12111143/s1, Table S1: qRT-PCR primer information. Figure S1: Melt curve of representative qRT-PCR amplicons. Figure S2: The sequence length distribution of assembled unigenes. Figure S3: Gene function annotations in six databases (GO, KEGG, COG, NR, Swiss-Prot, Pfam) of *F. suspensa*. Figure S4: Length distribution of CDS in unigenes. Figure S5: Principal component analysis of expressed gene. CKL: *F. suspensa* leaf treated with control solution for 0 h; M2L: *F. suspensa* leaf treated with 200 μM MeJA; CKF: *F. suspensa* fruit treated with control solution for 0 h; M2F: *F. suspensa* fruit treated with 200 μM MeJA.

Author Contributions: D.H. is the project administration and designed the experiments. X.L. and J.Z. performed the experiments and wrote the original draft. X.Z. prepared the samples for RNA extraction. H.L. used the software. H.X. and H.Z. conducted the RNA extraction and qRT-PCR analysis. X.Z., H.L. and H.S. reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (U1404829), a key project at central government level: The ability establishment of sustainable use for valuable Chinese medicine resources (2060302), Natural Science Foundation of Henan Province (202300410151). This research was also supported by a special fund for the construction of technical systems of the traditional Chinese medicine industry in Henan Province. Henan Province Chinese medicine industry science and technology commissioner service group.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this study can be found in online repositories. The sequencing data can be found in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under accession number PRJNA869193.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Xia, E.Q.; Ai, X.; Zang, S.; Guan, T.; Xu, X.; Li, H. Ultrasound-assisted extraction of phillyrin from *Forsythia suspensa*. *Ultrason. Sonochem* **2011**, *18*, 549–552. [CrossRef]
2. Li, H.; Chen, F. Preparative isolation and purification of phillyrin from the medicinal plant *Forsythia suspensa* by high-speed counter-current chromatography. *J. Chromatogr. A* **2005**, *1083*, 102–105. [CrossRef]
3. Zhou, W.; Di, L.; Shan, J.; Bi, X.; Chen, L.; Wang, L. Intestinal absorption of forsythoside A in different compositions of Shuang–Huang–Lian. *Fitoterapia* **2010**, *82*, 375–382. [CrossRef]
4. Hu, K.; Guan, W.; Bi, Y.; Zhang, W.; Li, L.; Zhang, B.; Liu, Q.; Song, Y.; Li, X.; Duan, Z.; et al. Efficacy and safety of Lianhua Qingwen capsules, a repurposed Chinese herb, in patients with Coronavirus disease 2019: A multicenter, prospective, randomized controlled trial [Phytomedicine 85 (2021) 153242]. *Phytomedicine* **2022**, *94*, 153800. [CrossRef] [PubMed]
5. Liu, M.; Gao, Y.; Yuan, Y.; Liu, M.; Gao, Y.; Yuan, Y.; Yang, K.; Shi, S.; Tian, J.; Zhang, J. Efficacy and safety of herbal medicine (Lianhuaqingwen) for treating COVID-19: A systematic review and meta-analysis. *Integr. Med. Res.* **2021**, *10*, 100644. [CrossRef] [PubMed]
6. Bu, Y.; Shi, T.; Meng, M.; Bu, Y.; Shi, T.; Meng, M.; Kong, G.; Tian, Y.; Chen, Q.; Yao, X.; et al. A novel screening model for the molecular drug for diabetes and obesity based on tyrosine phosphatase Shp2. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 874–878. [CrossRef]
7. Xue, H.; Yuan, W. Review on Pharmacological Research of Forsythia Suspensa Leaves. *Lishizhen Med. Mater. Med. Res.* **2009**, *20*, 1149–1150.
8. Yuan, J.; Qiu, Z.; Liu, J.L.; Liu, H.; Zhang, S.L. Processing techniques of Forsythia Suspensa leaves green tea and analysis of nutrient active ingredients. *J. Henan Univ. Sci. Technol.* **2015**, *36*, 78–82.
9. Dong, Z.; Lu, X.; Tong, X.; Dong, Y.; Tang, L.; Liu, M. Forsythiae Fructus: A Review on its Phytochemistry, Quality Control, Pharmacology and Pharmacokinetics. *Molecules* **2017**, *22*, 1466. [CrossRef] [PubMed]
10. Satake, H.; Koyama, T.; Bahabadi, S.E. Essences in metabolic engineering of lignan biosynthesis. *Metabolites* **2015**, *5*, 270–290. [CrossRef]
11. Du, Y.; You, L.; Ni, B.; Sai, N.; Wang, W.; Sun, M.; Xu, R.; Yao, Y.; Zhang, Z.; Qu, C.; et al. Phillyrin Mitigates Apoptosis and Oxidative Stress in Hydrogen Peroxide-Treated RPE Cells through Activation of the Nrf2 Signaling Pathway. *Oxidative Med. Cell. Longev.* **2020**, *2020*, 2684672. [CrossRef] [PubMed]

12. Ma, Q.; Li, R.; Pan, W.; Huang, W.; Liu, B.; Xie, Y.; Wang, Z.; Li, C.; Jiang, H.; Huang, J.; et al. Phillyrin (KD-1) exerts anti-viral and anti-inflammatory activities against novel coronavirus (SARS-CoV-2) and human coronavirus 229E (HCoV-229E) by suppressing the nuclear factor kappa B (NF-kappaB) signaling pathway. *Phytomedicine* **2020**, *78*, 153296. [CrossRef] [PubMed]
13. Xu, X.; Saadeldeen, F.S.A.; Xu, L.; Zhao, Y.; Wei, J.; Wang, D.H.M.; Liu, Z.; Kang, W. The Mechanism of Phillyrin from the Leaves of *Forsythia suspensa* for Improving Insulin Resistance. *Biomed. Res. Int.* **2019**, *2019*, 3176483. [CrossRef] [PubMed]
14. Liu, J.; Yu, S.; Ma, Y.; Zhang, T.; Zhao, G. Research progress in metabolic engineering and synthetic biology for natural lignan production. *Chin. Tradit. Herb. Drugs* **2016**, *47*, 2556–2562.
15. Rahman, M.M.A.; Dewick, P.M.; Jackson, D.E.; Lucas John, A. Biosynthesis of lignans in *Forsythia intermedia*. *Phytochemistry* **1990**, *29*, 1841–1846. [CrossRef]
16. Fukasawa, A.T.; Kung, S.D.; Watson, J.C. Phenylalanine ammonia-lyase gene structure, expression, and evolution in Nicotiana. *Plant Mol. Biol.* **1996**, *30*, 711–722. [CrossRef]
17. Gertsch, J.; Tobler, R.T.; Brun, R.; Sticher, O.; Heilmann, J. Antifungal, Antiprotozoal, Cytotoxic and Piscicidal Properties of Justicidin B and a New Arylnaphthalide Lignan from *Phyllanthus piscatorum*. *Planta Med.* **2003**, *69*, 420–424.
18. Kumar, A.; Ellis, B.E. 4-Coumarate:CoA ligase gene family in *Rubus idaeus*: cDNA structures, evolution, and expression. *Plant Mol. Biol.* **2003**, *31*, 327–340. [CrossRef]
19. Anterola, A.; Lewis, N. Trends in lignin modification: A comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. *Phytochemistry* **2002**, *61*, 221–294. [CrossRef]
20. Guo, D.; Chen, F.; Inoue, K.; Blount, J.W.; Dixon, R.A. Downregulation of caffeic acid 3-0-methyltransferase and caffeoyl CoA 3-0-methyltransferase in transgenic alfalfa: Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell Online* **2001**, *13*, 73. [CrossRef]
21. Luderitz, T.; Grisebach, H. Enzymic Synthesis of Lignin Precursors Comparison of Cinnamoyl-CoA Reductase and Cinnamyl Alcohol: NADP iDehydrogenase from Spruce (*Picea abies* L.) and Soybean (*GZycine max* L.). *Eur. J. Biochem.* **1981**, *119*, 115–124. [CrossRef]
22. Davin, L.B.; Wang, H.B.; Crowell, A.L.; Davin, L.B.; Wang, H.B.; Crowell, A.L.; Bedgar, D.L.; Martin, D.M.; Sarkanen, S.; Lewis, N.G. Stereoselective bimo-lecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. *Science* **1997**, *275*, 362–367. [CrossRef] [PubMed]
23. Yuan, W.; Zhang, S.; He, Z.; He, y.; He, S.; Liu, L.; Sun, X.; Li, Q. Comparative transcriptome analyses identify genes involved into the biosynthesis of forsythin and forsythoside A in *Forsythia suspensa*. *Funct. Integr. Genom.* **2022**, *22*, 731–741. [CrossRef] [PubMed]
24. Berim, A.; Ebel, R.; Schneider, B.; Petersen, M. UDP-glucose:(6-methoxy)podophyllotoxin 7-O-glucosyltransferase from suspension cultures of *Linum nodiflorum*. *Phytochemistry* **2008**, *69*, 374–381. [CrossRef]
25. Wasternack, C. Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann. Bot.* **2007**, *100*, 681–697. [CrossRef] [PubMed]
26. Verpoorte, R.; Heijden, R.V.D.; Hoge, J.H.C.; Hoopen, T. Plant cell biotechnology for the production of secondary metabolites. *Pure Appl. Chem.* **1994**, *66*, 2307–2310. [CrossRef]
27. Zhao, J.; Zhu, W.H.; Hu, Q. Selection of fungal elicitors to increase indole alkaloid accumulation in *Catharanthus roseus* suspension cell culture. *Enzym. Microb. Technol.* **2001**, *28*, 666–672. [CrossRef]
28. Diallo, A.O.; Agharbaoui, Z.; Badawi, M.A.; Ali-Benali, M.A.; Moheb, A.; Houde, M.; Sarhan, F. Transcriptome analysis of an mvp mutant reveals important changes in global gene expression and a role for methyl jasmonate in vernalization and flowering in wheat. *J. Exp. Bot.* **2014**, *65*, 2271–2286. [CrossRef]
29. Misra, R.C.; Maiti, P.; Chanotiya, C.S.; Shanker, K.; Ghosh, S. Methyl jasmonate-elicited transcriptional responses and pentacyclic triterpene biosynthesis in sweet basil. *Plant Physiol.* **2014**, *164*, 1028–1044. [CrossRef]
30. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef]
31. Pruitt, K.D.; Tatusova, T.; Maglott, D.R. NCBI Reference Sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic. Acids Res.* **2005**, *33*, D501–D504. [CrossRef] [PubMed]
32. Tatusov, R.L.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Kiryutin, B.; Koonin, E.V.; Krylov, D.M.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; et al. The COG database: An updated version includes eukaryotes. *BMC Bioinform.* **2003**, *4*, 41. [CrossRef] [PubMed]
33. Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; et al. KEGG for linking genomes to life and the environment. *Nucleic. Acids Res.* **2008**, *36*, D480–D484. [CrossRef] [PubMed]
34. Conesa, A.; Gotz, S.; Garcia-Gomez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [CrossRef] [PubMed]
35. Jin, J.P.; Tian, F.; Yang, D.C.; Meng, Y.; Kong, L.; Luo, J.; Gao, G. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic. Acids Res.* **2017**, *45*, D1040–D1045. [CrossRef]
36. Dewey, C.N.; Bo, L. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323.

37. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]
38. Deng, W.; Wang, Y.; Liu, Z.; Cheng, H.; Xue, Y. HemI: A toolkit for illustrating heatmaps. *PLoS ONE* **2014**, *9*, e111988. [CrossRef]
39. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [CrossRef]
40. Shen, J.; Wu, Y.; Jiang, Z.; Xu, Y.; Zheng, T.; Wang, J.; Cheng, T.; Zhang, Q.; Pan, H. Selection and validation of appropriate reference genes for gene expression studies in Forsythia. *Physiol. Mol. Biol. Plants* **2020**, *26*, 173–188. [CrossRef]
41. Lewis, N.G.; Davin, L.B.; Sarkanen, S. Lignin and Lignan Biosynthesis: Distinctions and Reconciliations. *ACS Symp.* **1998**, *697*, 1–27.
42. Zhou, C.; Lu, M.; Cheng, J.; Rohani, E.R.; Hamezah, H.S.; Han, R.; Tong, X. Review on the Pharmacological Properties of Phillyrin. *Molecules* **2022**, *27*, 3670. [CrossRef] [PubMed]
43. Moharramnejad, S.; Azam, A.T.; Panahandeh, J.; Dehghanian, Z.; Ashraf, M. Effect of Methyl Jasmonate and Salicylic Acid on In Vitro Growth, Stevioside Production, and Oxidative Defense System in *Stevia rebaudiana*. *Sugar Tech* **2019**, *21*, 1031–1038.
44. Satake, H.; Ono, E.; Murata, J. Recent advances in the metabolic engineering of lignan biosynthesis pathways for the production of transgenic plant-based foods and supplements. *J. Agric. Food Chem.* **2013**, *61*, 11721–11729. [CrossRef]
45. Malik, S.; Biba, O.; Grúz, J.; Arroo, R.R.J.; Strnad, M. Biotechnological approaches for producing aryltetralin lignans from Linum species. *Phytochem. Rev.* **2014**, *13*, 893–913. [CrossRef]
46. Zhao, J.; Davis, L.C.; Verpoorte, R. Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnol. Adv.* **2005**, *23*, 283–333. [CrossRef] [PubMed]
47. Berim, A.; Spring, O.; Conrad, J.; Maitrejean, M.; Boland, W.; Petersen, M. Enhancement of lignan biosynthesis in suspension cultures of *Linum nodiflorum* by coronalon, indanoyl-isoleucine and methyl jasmonate. *Planta* **2005**, *222*, 769–776. [CrossRef]
48. Wang, J.; Fan, S.; Li, A.; Chen, T.; Sun, B.; Yang, X. Analysis of the Forsythin and Forsythiaside A in different parts of *Forsythia suspensa* and its medicinal use discussion. *Mod. Chin. Med.* **2013**, *15*, 556–559.
49. Ye, H.; Wu, Z.; Li, F.R. Problems and thoughts of folium Forsythia used for medical purpose. *Chin. J. Clin. Ration. Drug Use* **2011**, *4*, 74–75.
50. Vanholme, R.; Demedts, B.; Morreel, K.; John, R.; Wout, B. (2010) Lignin biosynthesis and structure. *Plant Physiol.* **2010**, *153*, 895–905. [CrossRef]
51. Wang, H.; Yang, J.; Deng, K.; He, X.; Zhan, R.; Tang, L. Methyl Jasmonate affects metabolism and gene transcription of volatile terpenoids from *Amomum villosum* Lour. *Mod. Tradit. Chin. Med. Mater. Med.-World Sci. Technol.* **2014**, *16*, 1528–1536.
52. Sun, B.; Wang, P.; Wang, R.; Li, Y.; Xu, S. Molecular Cloning and Characterization of a meta/para-O-Methyltransferase from *Lycoris aurea*. *Int. J. Mol. Sci.* **2018**, *19*, 1911. [CrossRef] [PubMed]
53. Tao, Y.; Xu, J.; Wei, J.; Sun, J.; Xu, Y.; Yang, X.; Zhang, Y.; Liu, J.; Sui, C. Expression analysis of glycosyltransferase BcUGT1 from *Bupleurum chinense* DC. and its expression in E. coli and the target protein purification. *Yao Xue Xue Bao = Acta Pharm. Sin.* **2013**, *48*, 1345–1352.

metabolites

MDPI

Article

Integrated Analysis of Metabolome and Transcriptome Reveals the Difference in Flavonoid Biosynthesis between the Red- and White-Sarcocarp Pomelo Fruits

Chenxu Zhao [1,†], Jiajia Wang [1,†], Yuxia Li [1,†], Lei Zhang [1], Ghazala Nawaz [2], Shaoyuan Wu [1,*] and Tao Xu [1,*]

[1] Jiangsu Key Laboratory of Phylogenomics and Comparative Genomics, Jiangsu Joint International Center of Genomics, School of Life Sciences, Jiangsu Normal University, Xuzhou 221116, China
[2] Department of Botany, Kohat University of Science and Technology, Kohat 26000, Pakistan
* Correspondence: shaoyuanwu@outlook.com (S.W.); xutao_yr@126.com (T.X.)
† These authors contributed equally to this work.

Abstract: Flavonoids are bioactive secondary metabolites that play multiple roles in plants. However, studies on the flavonoid accumulation of the pomelo fruit are rare. In this study, we conducted a widely targeted metabolome analysis by using ultra-performance liquid chromatography and tandem mass spectrometry and identified 550 metabolites in the sarcocarp from red (*C. maxima* Merr. var. Tubtim Siam) and white pomelos (*C. maxima* (Burm.) Osbeck). A total of 263 significantly changed metabolites were detected from the 550 metabolites. Content analysis of the significantly changed metabolites (SCMs) showed that 138 SCMs were highly accumulated, whereas 125 SCMs were observed with lower content in red-sarcocarp pomelo. Importantly, 103 of the 263 SCMs were flavonoids, including 34 flavonoids, 29 flavonols, 18 flavonoid carbonosides, 9 dihydroflavones, 6 isoflavones, 5 anthocyanins, 1 dihydroflavonol, and 1 chalcone. Gene ontology analysis indicated that upregulated genes in red-sarcocarp pomelo were significantly enriched in GO terms related to flavonoids including flavonoid biosynthetic processes. Several important differentially expressed genes were detected in the correlation network, especially Cg2g009540 which is an orthologous gene of *AtCHS*, also detected in flavonoid biosynthesis networks, and which could be related to the high level of total flavonoids in the red-sarcocarp pomelo. Our study demonstrated the fluctuation of flavonoid biosynthesis in the two pomelo cultivars and laid a theoretical foundation for pomelo breeding to generate fruits with a high flavonoid content.

Keywords: pomelo; transcriptome; metabolome; flavonoid; biosynthetic difference

Citation: Zhao, C.; Wang, J.; Li, Y.; Zhang, L.; Nawaz, G.; Wu, S.; Xu, T. Integrated Analysis of Metabolome and Transcriptome Reveals the Difference in Flavonoid Biosynthesis between the Red- and White-Sarcocarp Pomelo Fruits. *Metabolites* **2022**, *12*, 1161. https://doi.org/10.3390/metabo12121161

Academic Editors: Yanjie Zhang, Yan Li and Jose Lorenzo Rodriguez

Received: 4 November 2022
Accepted: 18 November 2022
Published: 23 November 2022

1. Introduction

Flavonoids are ubiquitous secondary metabolites with diverse roles in plants. Flavonoids constitute a class of aromatic bioactive compounds, including sub-classes of flavones, flavonols, isoflavonoids, and anthocyanins. Studies have found that flavonoids are involved in the interactions between the plant and the environment including stress response, color formation of the plant organ, and so on [1–4]. Furthermore, flavonoids have also been demonstrated to have antioxidant and anti-inflammatory effects in animals [5–7]. Dietary flavonoids are important for reverse cholesterol transport, HDL metabolism, and function [8]. As a result, flavonoids have become one of the most attractive metabolites in plants.

With the increase in knowledge in the field of flavonoid metabolism in plants, the unraveling of complexities in the regulation of flavonoid biosynthetic pathways has opened a hotspot for plant scientists. Biosynthesis of flavonoids is a significant event in plant life because it determines the bio-accumulation and structure of bioactive compounds, which greatly affect the physiological activities in plants, particularly citrus fruits. A

recent review has summarized the biosynthesis pathways, cell localization, and factors affecting the biosynthesis of flavonoids [9]. In general, it is speculated that citrus-derived flavonoid biosynthesis pathways occur through the phenylpropanoid pathway. Initially, phenylalanine is transformed into p-coumaroyl-CoA with the catalysis of phenylalanine ammonia-lyase, cinnamate 4-hydroxylase, and 4-coumarate-CoA ligase. Then, chalcone synthase (CHS) catalyzes malonyl-CoA and p-coumaroyl-CoA to form chalcone, which transforms into naringenin under the control of chalcone isomerase (CHI). Naringenin is a key intermediate metabolite in the biosynthetic pathway of flavonoids and further converts into flavanones, anthocyanins, flavones, and flavanols [10]. Flavanones are transformed from naringenin. Under the catalysis of flavone synthase (FNS), naringenin is transformed into flavones. Meanwhile, dihydro-flavonols, which play vital roles in the formation of other flavonoids (such as flavonols and anthocyanins), are formed by flavanones under the regulation of flavanone 3-hydroxylase. Several external and internal factors, including transcription factors (TFs), precursors, intermediate enzymes, UV, and root zone salinity, affect the regulatory pathway of the flavonoid pathway [11–13].

Citruses are the most widely grown and popular fruits because of their palatable taste and nutritional benefits for human health. Citrus fruits are highly rich in bioactive compounds, including ascorbic acid, phenolic compounds, carotenoids, and flavonoids [14]. Pomelo (*Citrus grandis* (L.) Osbeck) belongs to the genus Citrus of the rutaceae family and is one of the most widely cultivated fruits in the world [15,16]. In China, the pomelo (such as *C. grandis* 'Tomentosa') can be used to produce traditional Chinese medicine to relieve several symptoms, such as coughs and phlegm [17,18]. Pomelo contains diverse flavonoids, such as narirutin, naringin, hesperidin, and neo-hesperidin [19,20]. Therefore, the variation of flavonoid profiles in different pomelo cultivars must be studied. In the present study, the total metabolites of two cultivars (*Citrus maxima* (Burm.) Osbeck and *Citrus maxima* Merr. var. Tubtim Siam) were studied using ultra-performance liquid chromatography and tandem mass spectrometry (UPLC-MS/MS) and high-throughput transcriptomic technologies to analyze the biosynthesis regulation and accumulation of flavonoids between pomelo cultivars and their influences on the qualities of pomelo.

2. Materials and Methods

2.1. Plant Materials

Two pomelo cultivars, comprising white- (*Citrus maxima* (Burm.) Osbeck) and red-sarcocarp pomelos (*Citrus maxima* Merr. var. Tubtim Siam), were cultivated in orchards in Xishuangbanna, Yunnan, China. An approval to collect the plant samples was acquired. Formal identification of the plant material was undertaken by Chunfen Xiao, Hui Teng, and Li Wang from Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences (CAS). Voucher specimens of these pomelo cultivars were deposited in the herbarium of Xishuangbanna Tropical Botanical Garden, CAS (Deposition No. 136042 and 162222). Healthy ripe fruits of uniform size were collected in August 2019. Sarcocarp samples at the equatorial plate (0.5 cm thickness) of every five globose fruits were collected, mixed, frozen in liquid nitrogen, and stored at -80 °C in a refrigerator for subsequent analysis. Three biological repeats were carried out for the experiments in this study. Experimental research and studies on the plants complied with relevant institutional, national, and international guidelines and legislation.

2.2. UPLC and Electrospray Ionization-Triple Quadrupole Linear Ion Trap (ESI-Q TRAP)-MS/MS

The pulp samples were freeze-dried in a lyophilizer (Scientz-100F) and ground to powder by a grinding machine (MM 400; Retsch, Germany) for 1.5 min at 30 Hz. A total of 100 mg of the powder was dissolved in 0.6 mL 70% aqueous methanol. The solution was kept at 4 °C overnight and then centrifuged at 10,000× *g* for 10 min before the supernatant was collected. A CNWBOND Carbon-GCB SPE Cartridge (250 mg, 3 mL; ANPEL, Shanghai, China) and SCAA-104 (0.22 μm pore size; ANPEL, Shanghai, China)

were applied to absorb and filter the extracts for further analysis of total metabolites on UPLC-MS/MS, respectively [21].

The filtrates were analyzed by a UPLC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system, Kyoto, Japan; MS, Applied Biosystems 4500 Q TRAP, Waltham, MA, USA). Further, UPLC analysis was performed with a Waters ACQUITY UPLC HSS T3 C18 chromatographic column (1.8 μm, 2.1 mm × 100 mm). The solvent for mobile phase A consisted of ultrapure water with 0.04% acetic acid, and that for mobile phase B was acetonitrile with 0.04% acetic acid. A 0.35 mL/min sample flow rate was used, and the oven temperature was 40 °C. The elution gradient of A:B was 95% A and 5% B at 0 min. Within the first 10 min, B was linearly increased to 95% and held at 95% for 1 min. The proportion of B fell to 5% from 11.00 min to 11.10 min and then remained at 5% for 2.90 min. Then, the effluent was alternatively channeled into an ESI-QTRAP-MS [22].

An API 4500 Q TRAP UPLC/MS/MS system equipped with a linear ion trap (LIT), a triple quadrupole (QQQ), and an ESI turbo IS (ion-spray) interface was operated in controlled mode with positive and negative ions via Analyst 1.6.3. The temperature of ESI was maintained at 550 °C, the MS voltage (or IS voltage) was 5500 V, curtain gas was set at 30 psi, and ion source gases I and II were maintained at 50 and 60 psi, respectively. The parameters of collision-activated dissociation were high. The mass calibration and equipment tuning were carried out with different concentrations of polypropylene glycol, such as 10 μmol/L in the QQQ model and 100 μmol/L in the LIT model. The scans of the QQQ were obtained from the multiple reaction monitoring (MRM) model using nitrogen as the collision gas set at 5 psi. Based on optimized declustering potential and collision energy, each ion pair was scanned following a method described earlier [23].

2.3. RNA-Seq Analysis

For transcriptomic analysis of pomelo fruits, total RNA was extracted with the RNAprep pure plant plus kit (Tiangen, Beijing, China) as previously described [24]. The extracted total RNA was checked on a Nano Photometer® spectrophotometer (IMPLEN, CA, USA) and an Agilent 2100 bio analyzer (Agilent Technologies, Clara, CA, USA) to analyze the purity and quality of the RNA, respectively. The degree of stability of the RNA was measured by running the RNA on 1% agarose gel electrophoresis. The cDNA library for each sample was constructed using the NEBNext UltraTM RNA Library Prep Kit (New England Biolabs, NEB, Ipswich, MA, USA), and was sequenced with an Illumina novaseq 6000 platform to generate 150 bp paired-end reads (Illumina, San Diego, CA, USA).

RNA-Seq analysis was performed as previously described [25]. Firstly, clean data for each sample were obtained from the raw data using Trimmomatic (version 0. 39) with default parameters. Secondly, the obtained clean data were mapped to the *C. maxima* genome (v1.0) using STAR (version 2.7.1a) under the 2-pass mapping mode [26,27]. Thirdly, the read counts mapped to each gene were calculated using featureCounts (v2.0.1), and the FPKM values for each gene were obtained using RSEM (v1.3.1). A gene was considered expressed if its average expression level in FPKM was lower than 1. Finally, the DEseq2 package was employed to perform DEG analysis between white- and red-sarcocarp pomelo [28]. Genes with $|\log2\text{FoldChange}| \geq 2$ and adjusted p-value ≤ 0.01 were considered to be significantly differentially expressed.

The GO annotation of pomelo genes was assigned using the Trinotate pipeline with default parameters, and was used to build the OrgDb database [29]. The functional enrichment analysis of the above identified DEGs was performed using clusterProfiler at adjusted p-values < 0.05 [30].

2.4. Gene Expression Analysis by Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

The expression of 3 DEGs which encoded CHS, FNS II (flavone synthase II), and SHT (shikimate O-hydroxycinnamoyl transferase) was analyzed by qRT-PCR. The total RNA was extracted from pomelo fruits by using OminiPlant RNA Kit (Kangwei, Taizhou, China). cDNA was synthesized from the total RNA using a reverse transcription kit (TaKaRa,

Dalian, China) as previously described [31,32]. β-actin was used as the internal reference gene [26]. Online software Primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/, accessed on 4 November 2022) was used to design qRT-PCR primers (Table S11), and TB Green^{TM} Premix Ex Taq^{TM} II (TaKaRa, Shiga, Japan) was used to perform qRT-PCR analysis on the qPCR instrument (Analytik Jena, Jena, Germany) as described in a previous study [33,34]. The experiments were performed with three repetitions, and the expression was calculated with the $2^{-\Delta\Delta Ct}$ method [35].

2.5. Statistical Analysis

Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS–DA) were used to simplify the high-dimensional data obtained from various experiments. The MRM of the QQQ was used to quantify metabolites. The core program in R was used to calculate the Pearson correlation coefficient (PCC) between the DEGs and significantly changed metabolites (SCMs). The DEGs and SCMs ($|PCC| > 0.9$) were selected to draw a correlation network diagram using the software Cytoscape. Significance analysis was performed by using *t*-test in the qRT-PCR experiment.

3. Results

3.1. PCA and OPLS–DA Analysis in the Fruit of White and Red Pomelo Cultivars

Two pomelo fruits with white and red sarcocarp were collected to compare their flavonoid biosynthesis (Figure 1). To gain insights into the significantly changed flavonoids (SCFs) between the red- and white-sarcocarp pomelos, we performed UPLC–MS/MS assay on the pulp of the pomelos with red or white sarcocarp and detected about 550 metabolites in total (Table S1 (Supplementary Materials)). PCA, a tool that analyzes a small number of principal components to show the structure of multiple variables, was used to analyze the metabolic profile difference between the two pomelo cultivars. The plots for white and red sarcocarp were distributed on the left and right sides, respectively (Figure 2A), implying that significant differences in the metabolic profiles existed between red and white pomelo sarcocarps at principal component 1 (70.71%).

Figure 1. Phenotype of two pomelo cultivars: white- (*Citrus maxima* (Burm.) Osbeck) and red-sarcocarp pomelos (*Citrus maxima* Merr. var. Tubtim Siam). Bars indicate 3 cm.

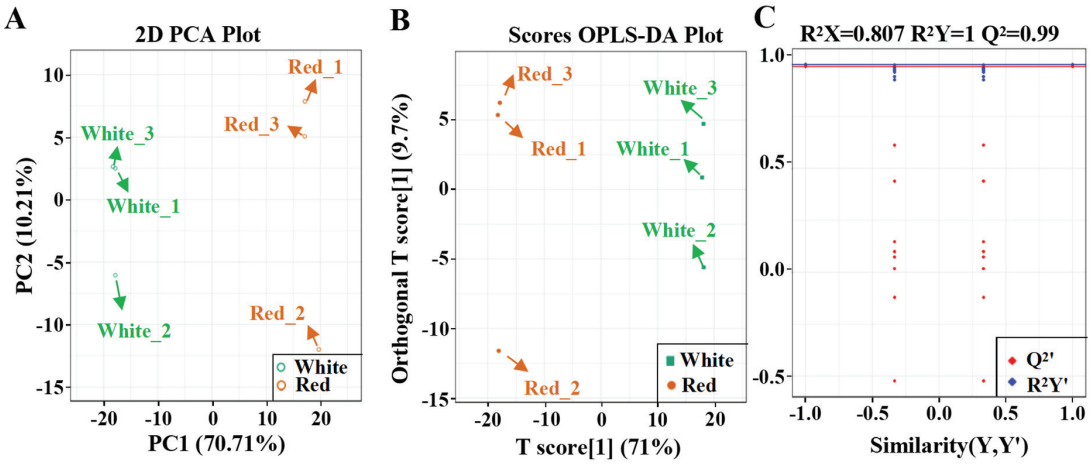


Figure 2. PCA and OPLS-DA on metabolites of pomelo fruit. (**A**) PCA score plot. (**B**) OPLS–DA score plot. (**C**) OPLS–DA permutation analysis model validation graph.

We used OPLS-DA to exclude the small-correlation variables and to achieve the maximum significant differences of the SCMs between the two cultivars. The OPLS–DA score plot results displayed a distinct metabolic difference between the red- and white-sarcocarp cultivars (Figure 2B). A total of 200 alignment experiments were conducted to further verify the OPLS–DA model (R^2X = 0.807, R^2Y = 1, and Q^2 = 0.99). The blue and red lines denote the R^2Y and Q^2 of the original model, respectively. The blue and red dots correspond to the R^2Y' and Q^2' after replacement, respectively. The results indicate that the OPLS-DA model was reliable, and that the SCMs can be selected based on variable importance in projection (VIP) (Figure 2C).

3.2. Screening of Significantly Changed Flavonoids

The SCMs in the two cultivars were screened using fold change (FC, $\geq$2 or $\leq$0.5) and VIP value (VIP $\geq$ 1) as the threshold. Among the 550 metabolites, 263 were SCMs between the white- and red-sarcocarp pomelos. A total of 52.5% SCMs (138 in 263) had a higher content, and the rest had a lower content in the red-sarcocarp pomelo compared with those in the white-sarcocarp pomelo (Figure 3A). To further analyze the composition of the SCMs with the greatest difference in content, the absolute log2FoldChange analysis displayed the top 20 SCMs. As shown in Figure S1, 12 SCMs were flavonoids, only 3 of which were found to be highly accumulated in the red-sarcocarp pomelo including 1 anthocyanin and 2 flavonoids. The remaining nine SCMs with lower content in the red-sarcocarp pomelo consisted of three flavonoid carbonosides, two flavonoids, two flavonols, one dihydroflavonol, and one anthocyanin (Figure S1).

To fully comprehend the type and abundance of metabolites present in the two cultivars, we generated a heatmap of 263 SCMs, and the results showed that 103 SCMs were flavonoids, including 34 flavonoids, 29 flavonols, 18 flavonoid carbonosides, 9 dihydroflavones, 6 isoflavones, 5 anthocyanins, 1 dihydroflavonol, and 1 chalcone (Figure 3B; Table S1). Among these SCFs, 49 exhibited higher levels in the red-sarcocarp pomelo compared with the white-sarcocarp pomelo. Then, the SCMs were annotated to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to analyze their metabolic pathways [36]. Meanwhile, several detected metabolites that were not annotated in KEGG pathways were added to the corresponding pathways based on their structure. The results of the KEGG enrichment indicated that the SCMs of the two cultivars were significantly enriched in the flavonoid, isoflavonoid, flavone and flavonol biosynthesis pathways (Figure 3C).

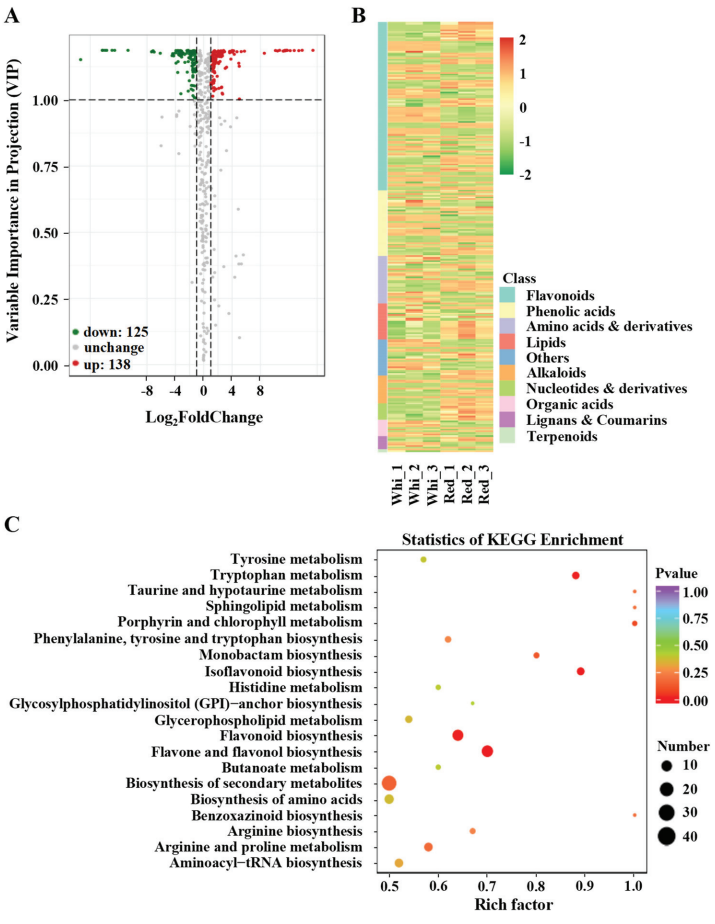


Figure 3. Analysis of SCMs in pomelo fruits. (**A**) Volcano plot of metabolites. (**B**) Heatmap of SCM expression. (**C**) KEGG enrichment analysis of SCMs.

3.3. Analysis of DEGs and Gene Ontology (GO) Enrichment in Red- and White-Sarcocarp Pomelo

RNA libraries were constructed and sequenced by using the Illumina platform to investigate gene expressions in white- and red-sarcocarp pomelo. After filtering the raw data, checking the sequencing error rate, and inspecting the GC content distribution, approximately 44.91 Gb (149.96 million paired-end reads) of clean bases were obtained, ranging from 6.02 to 9.76 Gb per library (Table S2). The clean data of each library were then mapped to the *C. maxima* genome with an average mapping rate of 98.67% [26]. The read counts and FPKM values for all the annotated genes were calculated. A total of 14,158 genes were found to be expressed in our transcriptome datasets, accounting for 47.00% of all genes (14158/30123) in the *C. maxima* genome [26]. The similarity of the six samples were further assessed by using the above expressed genes. The results showed that the three biological replicates for each cultivar were highly correlated (Figure S2).

To obtain the significantly differential expression of genes between the two cultivars, we analyzed gene expressions by using the DESeq2 package [28]. A Venn diagram showed that 12,742 genes were expressed both in the white and red pomelo cultivars, whereas 710 and 706 genes were expressed specifically in red- and white-sarcocarp pomelo, respectively (Figure 4A). In the white- vs. red-sarcocarp pomelo comparison, 1593 differently expressed genes (DEGs) were detected based on FC ($|\log2FoldChange| \geq 2$) and adjusted *p*-value

(≤0.01), in which the expressions of 735 genes were upregulated, and those of 858 genes were significantly downregulated (Figure 4B; Table S3).

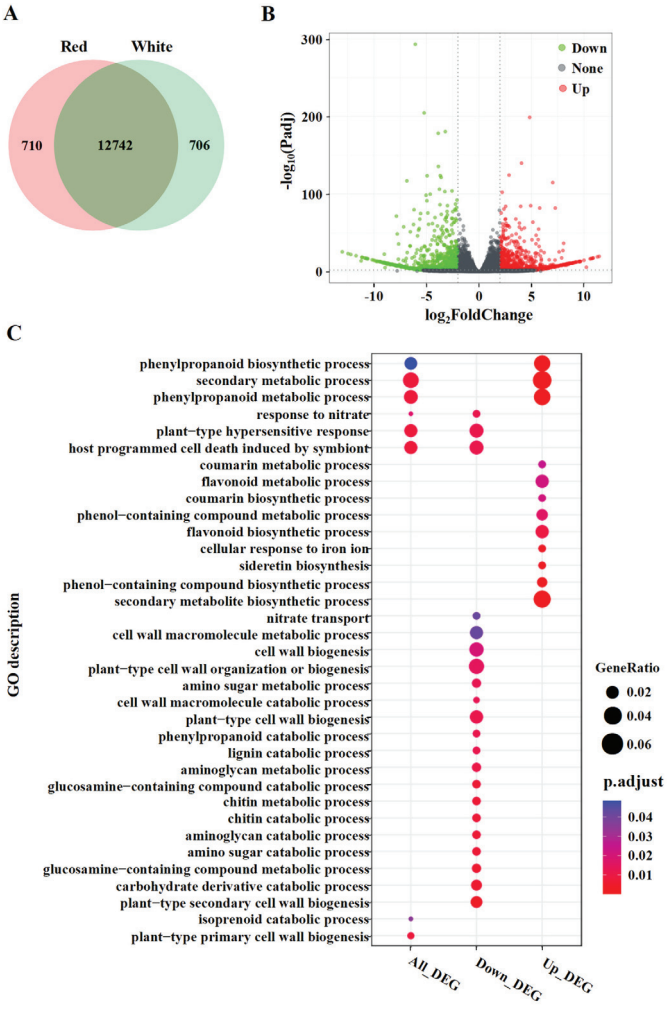


Figure 4. Gene expression analysis and GO annotation of DEGs in pomelo fruits. (**A**) Venn diagram of genes. (**B**) Volcano plot of genes. (**C**) GO annotation of DEGs.

To understand the role of TFs involved in the metabolism of pomelo, we annotated 95 DEGs in 31 TF families via iTAK (1.7a) (Table S4). A total of 17 DEGs were detected in the MYB (11 DEGs) and bHLH (6 DEGs) TF families, which are related to flavonoid biosynthesis. Among these DEGs, four MYBs and three bHLHs showed higher expression in the red pomelo than the white pomelo (Table S5).

GO enrichment analysis was performed to categorize the biological functions of the DEGs. For all the DEGs, a total of 35 GO terms were enriched, including 3 for cellular component, 24 for molecular function, and 8 for biological process (Figure 4C; Table S6). For example, GO terms related to "secondary metabolic process" and "phenylpropanoid biosynthetic process" were enriched. To better understand the biological functions of the DEGs, enrichment analysis was conducted for the up- and downregulated DEGs, respectively (Figure 4C, Table S6). By comparing with the white-sarcocarp pomelo, upregulated

DEGs in the red-sarcocarp pomelo were significantly enriched in GO terms related to the "phenylpropanoid biosynthetic process" and the "flavonoid biosynthetic process", etc. The downregulated DEGs were mainly over-represented with GO terms related to "cell wall biogenesis". These findings were consistent with the characteristics of the red-sarcocarp pomelo, which are enriched in flavonoid metabolites.

3.4. Correlation Network Analysis of DEGs and SCMs Related to Flavonoid Biosynthesis

To study the relationship between genes and metabolites in flavonoid biosynthesis, correlation network analysis was conducted. The 19 DEGs and 17 SCMs related to flavonoid biosynthesis were selected to calculate the PCC, and based on the PCC ($|PCC| > 0.9$), a correlation network including 18 DEGs and 17 SCMs was drawn (Figure 5; Table S7). In this network, 11 DEGs and 12 SCMs were highly correlated with 5 or more SCMs and DEGs, respectively (Figure 5; Table S8). In the network, Cg2g009540, an orthologous gene of *AtCHS*, was highly related to as many as 12 SCMs. Cg5g035630, highly related to 10 SCMs, could be an orthologous gene of *AtbHLH42*. Cg5g043310, involved in regulation of brassinosteroid metabolism, was related to 12 SCMs, while the other plant hormone-related gene Cg8g004960, involved in ABA transport, was only related to 2 SCMs. Cg4g006870, an orthologous gene of *AtCYP711A1*, was only highly related to three SCMs in the network. The above results indicated the complexity of the flavonoid biosynthetic regulatory network.

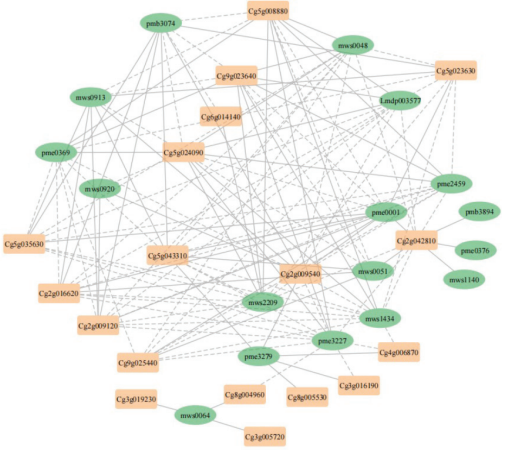


Figure 5. Correlation network of DEGs and SCMs related to flavonoid biosynthesis. Green ellipse, SCMs; light orange squares, DEGs; solid line, positive correlation; dashed line, negative correlation. Lmdp003577: Genistein 7–O–Glucoside; mws0048: Vitexin; mws0051: Acacetin; mws0913: Trifolin; mws1434: Isovitexin; mws2209: Kaempferol-3-O-glucoside; pmb3074: 3–O–p–Coumaroyl quinic acid; pme0001: Neohesperidin; pme0369: Kaempferol–3–O–rutinoside; pme2459: Luteolin–7–O–glucoside; pme3227: Vitexin 2″–O–β–L–rhamnoside; pme3279: 2′-Hydroxygenistein; mws0064: Eriodictyol; mws0920: Tricetin; mws1140: Naringenin chalcone; pmb3894: Di–O–methylquercetin; pme0376: Naringenin. Cg5g043310: *CmBEN1*; Cg5g035630: *CmbHLH42*; Cg2g016620: *Cm5MAT*; Cg2g009120: *CmTT12*; Cg4g006870: *CmCYP711A1*; Cg2g009540: *CmCHS*; Cg3g019230: *CmTT10*.

3.5. Analysis of Flavonoid Biosynthesis Network

The flavonoid biosynthesis network was analyzed and graphically illustrated based on the analysis of KEGG pathway annotation (Figure 6A). A total of 17 SCMs and 11 relevant genes were detected in the network (Tables S9 and S10). The levels of SCM contents and gene expressions are shown in Figure 6B,C, respectively. The content of 10 metabolites in red-sarcocarp pomelo were higher than those in white-sarcocarp pomelo (Figure 6B). Among 11 genes, only three DEGs were detected: two upregulated genes Cg2g009540 and Cg5g023630, and one downregulated gene Cg2g025640, which encoded CHS (chalcone

synthase), FNS II and SHT, respectively (Figure 6C). However, among substrate content, enzyme gene expression and product content, they were not perfectly matched with each other in this network. In general, these results suggested the complexity of the regulatory network in flavonoid biosynthesis.

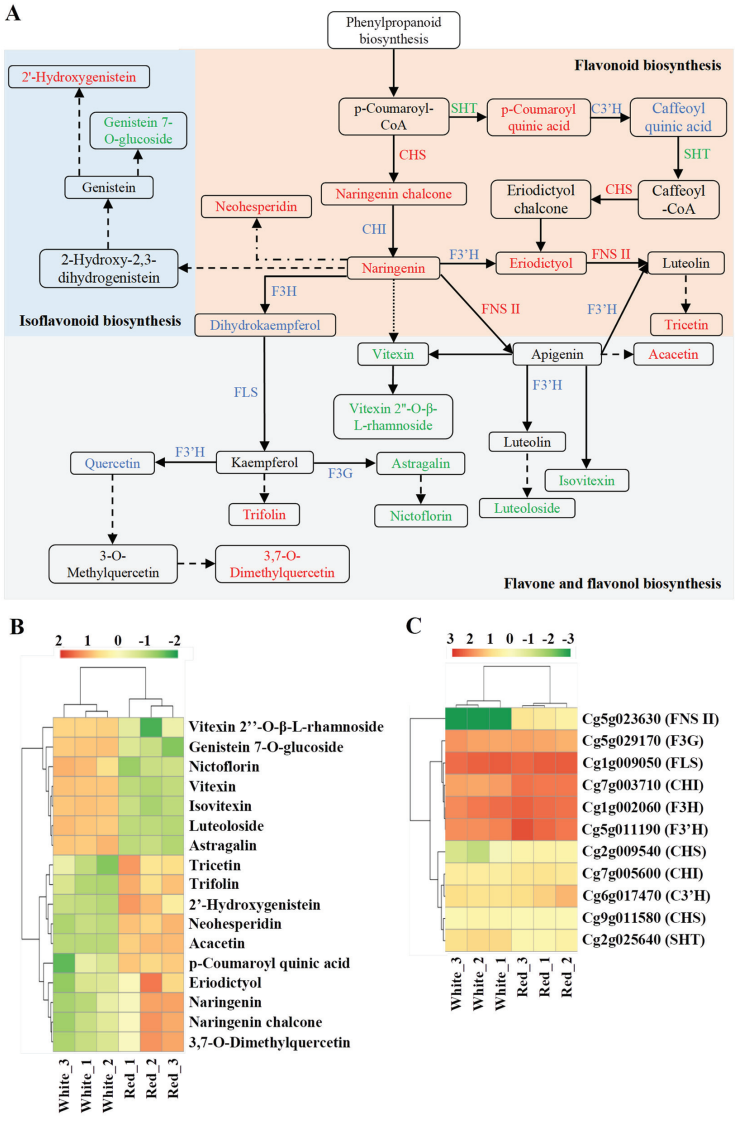


Figure 6. Pathway analysis of flavonoid biosynthesis. (**A**) Flavonoid biosynthesis pathway in pomelo. Solid arrow: molecular interaction or relation; dashed arrow, no genes were detected; dotted arrow, no metabolites were detected; point-line arrow, no genes or metabolites were detected. CHI: chalcone isomerase; CHS: chalcone synthase; C3′H: 5–O–(4-coumaroyl)–D–quinate/shikimate 3′–hydroxylase; F3G: flavonol 3–O–glucosyltransferase; F3H: flavanone 3–hydroxylase; F3′H: flavonoid 3′–hydroxylase; FLS: flavonol synthase; FNS II: flavone synthase II; SHT: shikimate O–hydroxycinnamoyl transferase. (**B**) Heatmap of SCMs in flavonoid biosynthesis. (**C**) Heatmap of gene expressions in flavonoid biosynthesis.

3.6. qRT-PCR Analysis

To confirm the gene expression data obtained from RNA-seq, the transcript levels of the three DEGs in the flavonoid biosynthesis network which encoded *CHS* (Cg2g009540), *FNS II* (Cg5g023630), and *SHT* (Cg2g025640) were analyzed by qRT-PCR using gene-specific primers in the pomelo samples. Comparing with white-sarcocarp pomelo, red-sarcocarp pomelo expressed significantly high levels of *CHS* and *FNS II* and a low level of *SHT* (Figure S3). The expressions of all the DEGs from the qPCR experiment were consistent with those from RNA-seq.

4. Discussion

Members of Citrus are globally known for their nutritional and medicinal benefits. However, the presence of flavonoids in different Citrus cultivars and their roles in flavonoid biosynthesis are largely unknown. In this study, widely targeted metabolome and transcriptome sequencing were employed to investigate the differences in flavonoid accumulation between the two pomelo cultivars (*Citrus maxima* (Burm.) Osbeck and *Citrus maxima* Merr. var. Tubtim Siam). The metabolome data suggested that almost half of the metabolites (263 in 550) detected were differentially accumulated between the white-and red-sarcocarp pomelos (Table S1). Flavonoids were the top SCMs between the two cultivars, implying that flavonoids showed the greatest variation and the most important metabolites in pomelo. The differences in metabolites and transcripts of the flavonoid pathway affect petal color in *Nelumbo nucifera* [4]. Recent studies indicated that flavonoids enhance plant tolerance to abiotic stresses [37,38]. The involvement of flavonoids in symbiotic nitrogen fixation in Rhizobia [39] has been investigated. Altogether, our results suggested that flavonoids were one of the most important bioactive components and their abundance might affect characteristic differences of pomelo fruits.

Transcriptome sequencing was carried out to investigate the DEGs in two pomelo cultivars with white and red sarcocarp. GO terms related to flavonoid biosynthesis including "phenylpropanoid biosynthetic process" and "flavonoid biosynthetic process" were found in upregulated DEGs. TFs, such as the MYB, bHLH, and WD40 families, played important roles in flavonoid biosynthetic regulation by forming the MYB-bHLH-WD repeat complex [40]. Overexpression of the *MYB* gene was an important approach to obtain the high content of flavonoids in tobacco [41]. In *Arabidopsis* seedlings, pigment production depended on Transparent Testa Glabra 1 (TTG1, a WD40 TF) and bHLH [42]. In this study, 11 DEGs in the MYB and 6 DEGs in the bHLH family were annotated. Among them, four MYBs and three bHLHs were highly expressed (Table S5). The function of these transcription factors needs to be demonstrated in further studies.

To understand the mechanism of flavonoid biosynthesis in pomelo, we evaluated the PCC between DEGs and SCMs. In the correlation network, Cg5g035630 encoding a TF in the bHLH family was detected, which could participate in the transcriptional regulation of pomelo flavonoid biosynthesis. Phytohormone and stress can also affect flavonoid biosynthesis [43–45]. In the network, the detection of plant hormone-related genes (Cg5g043310 and Cg8g004960) and a defense-associated gene (Cg5g024090) indicated potential regulation of plant hormone and stress in flavonoid biosynthesis. *AtCYP711A1* was related to 11 genes (*CHS*, *CHI*, etc.) in flavonoid biosynthesis [46]. We detected that an orthologous gene of *AtCYP711A1* in pomelo (Cg4g006870) was highly related to three SCMs. *CmMAX1* would be an entry point to further investigate flavonoid biosynthetic regulation in pomelo.

Naringenin was a pivotal intermediate product which could be converted into other flavonoids [10]. As the results suggested in Figure 6, the higher expression of the *CHS* gene (Cg2g009540) could lead to a higher content of naringenin chalcone, and then naringenin was highly accumulated in red-sarcocarp pomelo due to the increase in substrate content, which could be a reasonable interpretation of the higher eriodictyol content and even the high level of total flavonoid in red-sarcocarp pomelo. Naringenin can be converted into 8-C-glucosylnaringenin and then form vitexin, all the steps occurring without regulation

by enzymes. Interestingly, we found vitexin content was lower in red-sarcocarp pomelo. In general, the production of apigenin from naringenin was catalyzed by FNS (FNS I and FNS II) [47]. A study in 2019 also found that overexpression of *ZmFNS I* and *ZmFNS II* in *Arabidopsis* plants resulted in apigenin formation and remitted damage caused by UV-B [48]. However, only *FNS II* was detected to be occasionally involved in apigenin formation [49]. In our study, the process of apigenin formation might be catalyzed only by FNS II because just the *FNS II* gene (Cg5g023630) was detected in red-sarcocarp pomelo. Given that only the substrate content had changed significantly, we also speculated that saturation of enzyme activity restricted the synthesis of caffeoyl quinic acid and dihydrokaempferol. Additionally, enzyme genes and direct precursors were not detected; further exploration needs to be conducted on the formation of several flavonoids including 2′-hydroxygenistein, trifolin, and tricetin.

5. Conclusions

Metabolome analysis based on UPLC-MS/MS and transcriptome was performed in red- and white-sarcocarp pomelo fruits. DEG analysis revealed that flavonoids were the SCMs between the two kinds of pomelo fruits. GO enrichment analysis showed that specific DEGs were involved in flavonoid biosynthesis-related processes, including the phenylpropanoid biosynthetic process and the flavonoid biosynthetic process. The increase in the *CHS* gene (Cg2g009540) expression could be related to the high level of total flavonoid in red-sarcocarp pomelo. Conclusively, the combined analysis of SCMs and DEGs provided molecular evidence for the flavonoid biosynthesis between red- and white-sarcocarp pomelo and laid a theoretical foundation for the breeding of flavonoid-rich pomelo.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12121161/s1, Figure S1: Foldchange of top 20 SCMs. SCFs are marked by red boxes, Figure S2: Inter-sample correlation diagram, Figure S3: qRT-PCR analysis of candidate DEGs related to flavonoid biosynthesis, Table S1: List of metabolites in two pomelo cultivars, Table S2: Summary of sequenced RNA libraries, Table S3: List of DEGs in two pomelo cultivars, Table S4: Transcription factor annotation, Table S5: Statistics of transcription factor annotation results, Table S6: GO terms of all the DEGs, Table S7: Pearson correlation coefficient between DEGs and SCMs related to flavonoid biosynthesis, Table S8: Correlation network of DEGs and SCMs related to flavonoid biosynthesis, Table S9: SCMs detected in flavonoid biosynthesis based on KEGG pathway annotation, Table S10: Genes related to SCMs in flavonoid biosynthesis, Table S11: Primers used in qRT-PCR analysis. Citrus β-actin was selected as internal reference. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Author Contributions: Conceptualization, S.W. and T.X.; Formal analysis, C.Z., J.W., Y.L. and L.Z.; Writing—original draft, C.Z., J.W. and Y.L.; Writing—review and editing, G.N. and T.X.; Project administration, T.X.; Funding acquisition, S.W. and T.X.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported jointly by the projects of the Natural Science Foundation of Jiangsu Higher Education Institutions of China (19KJA510010), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The funding bodies provided monetary support to cover the transcriptome, metabolome, and qPCR analysis. The funding bodies were not involved in the design, analysis and reporting of the study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and analyzed during the current study are available in the NCBI database repository (Public on 1 February 2021, Accession No. GSE162849, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE162849, accessed on 4 November 2022).

Acknowledgments: We thank MetWare Biological Science and Technology Co., Ltd., (Wuhan, China) for supplying transcriptome and metabolome assay services.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Brunetti, C.; Fini, A.; Sebastiani, F.; Gori, A.; Tattini, M. Modulation of Phytohormone Signaling: A Primary Function of Flavonoids in Plant-Environment Interactions. *Front. Plant Sci.* **2018**, *9*, 1042. [CrossRef]
2. Li, J.; Hossain, M.S.; Ma, H.; Yang, Q.; Gong, X.; Yang, P.; Feng, B. Comparative metabolomics reveals differences in flavonoid metabolites among different coloured buckwheat flowers. *J. Food Compos. Anal.* **2020**, *85*, 103335. [CrossRef]
3. Liu, Y.; Lv, J.; Liu, Z.; Wang, J.; Yang, B.; Chen, W.; Ou, L.; Dai, X.; Zhang, Z.; Zou, X. Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (*Capsicum annuum* L.). *Food Chem.* **2020**, *306*, 125629. [CrossRef]
4. Zhu, H.H.; Yang, J.X.; Xiao, C.H.; Mao, T.Y.; Zhang, J.; Zhang, H.Y. Differences in flavonoid pathway metabolites and transcripts affect yellow petal colouration in the aquatic plant *Nelumbo nucifera*. *BMC Plant Biol.* **2019**, *19*, 277. [CrossRef] [PubMed]
5. Ohkatsu, Y.; Sakurai, T.; Sato, T. Relationship between Chemical Structure and Antioxidant Function of Flavonoids. *J. Jpn. Pet. Inst.* **2010**, *53*, 213–221. [CrossRef]
6. Fardoun, M.M.; Maaliki, D.; Halabi, N.; Iratni, R.; Bitto, A.; Baydoun, E.; Eid, A.H. Flavonoids in adipose tissue inflammation and atherosclerosis: One arrow, two targets. *Clin. Sci.* **2020**, *134*, 1403–1432. [CrossRef] [PubMed]
7. Jian, T.; Chen, J.; Ding, X.; Lv, H.; Li, J.; Wu, Y.; Ren, B.; Tong, B.; Zuo, Y.; Su, K. Flavonoids isolated from loquat (*Eriobotrya japonica*) leaves inhibit oxidative stress and inflammation induced by cigarette smoke in COPD mice: The role of TRPV1 signaling pathways. *Food Funct.* **2020**, *11*, 3516–3526. [CrossRef] [PubMed]
8. Millar, C.L.; Duclos, Q.; Blesso, C.N. Effects of dietary flavonoids on reverse cholesterol transport, HDL metabolism, and HDL function. *Adv. Nutr.* **2017**, *8*, 226–239. [CrossRef] [PubMed]
9. Zhao, C.; Wang, F.; Lian, Y.; Xiao, H.; Zheng, J. Biosynthesis of citrus flavonoids and their health effects. *Crit. Rev. Food Sci. Nutr.* **2020**, *60*, 566–583. [CrossRef]
10. Pfeiffer, P.; Hegedűs, A. Review of the molecular genetics of flavonoid biosynthesis in fruits. *Acta Aliment.* **2011**, *40*, 150–163. [CrossRef]
11. Matsui, K.; Walker, A.R. Biosynthesis and regulation of flavonoids in buckwheat. *Breed. Sci.* **2020**, *70*, 74–84. [CrossRef] [PubMed]
12. Falcone Ferreyra, M.L.; Rius, S.P.; Casati, P. Flavonoids: Biosynthesis, biological functions, and biotechnological applications. *Front. Plant Sci.* **2012**, *3*, 222. [CrossRef]
13. Agati, G.; Biricolti, S.; Guidi, L.; Ferrini, F.; Fini, A.; Tattini, M. The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in *L. vulgare* leaves. *J. Plant Physiol.* **2011**, *168*, 204–212. [CrossRef]
14. Guimarães, R.; Barros, L.; Barreira, J.C.; Sousa, M.J.; Carvalho, A.M.; Ferreira, I.C. Targeting excessive free radicals with peels and juices of citrus fruits: Grapefruit, lemon, lime and orange. *Food Chem. Toxicol.* **2010**, *48*, 99–106. [CrossRef] [PubMed]
15. Gaikwad, K.A.; Haldavanekar, P.C.; Parulekar, Y.R.; Haldankar, P.M. Survey and characterization of pummelo genotypes (*Citrus grandis* L. Osbeck) grown in coastal region of Maharashtra. *Ecoscan* **2015**, *8*, 371–380.
16. Liu, Y.; Heying, E.; Tanumihardjo, S.A. History, global distribution, and nutritional importance of citrus fruits. *Compr. Rev. Food Sci. Food Saf.* **2012**, *11*, 530–545. [CrossRef]
17. Jiang, K.; Song, Q.; Wang, L.; Xie, T.; Wu, X.; Wang, P.; Yin, G.; Ye, W.; Wang, T. Antitussive, expectorant and anti-inflammatory activities of different extracts from Exocarpium *Citri grandis*. *J. Ethnopharmacol.* **2014**, *156*, 97–101. [CrossRef] [PubMed]
18. Zhu, Z.; Wu, H.; Su, W.; Shi, R.; Li, P.; Liao, Y.; Wang, Y.; Li, P. Effects of Total flavonoids from exocarpium citri grandis on air pollution particle-induced pulmonary inflammation and oxidative stress in mice. *J. Food Sci.* **2019**, *84*, 3843–3849. [CrossRef]
19. Xu, G.; Liu, D.; Chen, J.; Ye, X.; Ma, Y.; Shi, J. Juice components and antioxidant capacity of citrus varieties cultivated in China. *Food Chem.* **2008**, *106*, 545–551. [CrossRef]
20. Zhang, M.; Duan, C.; Zang, Y.; Huang, Z.; Liu, G. The flavonoid composition of flavedo and juice from the pummelo cultivar (*Citrus grandis* (L.) Osbeck) and the grapefruit cultivar (*Citrus paradisi*) from China. *Food Chem.* **2011**, *129*, 1530–1536. [CrossRef]
21. Wang, F.; Chen, L.; Chen, H.; Chen, S.; Liu, Y. Analysis of flavonoid metabolites in citrus peels (*Citrus reticulata* "Dahongpao") using UPLC-ESI-MS/MS. *Molecules* **2019**, *24*, 2680. [CrossRef] [PubMed]
22. Li, Q.; Song, J. Analysis of widely targeted metabolites of the euhalophyte Suaeda salsa under saline conditions provides new insights into salt tolerance and nutritional value in halophytic species. *BMC Plant Biol.* **2019**, *19*, 388. [CrossRef] [PubMed]
23. Chen, W.; Gong, L.; Guo, Z.; Wang, W.; Zhang, H.; Liu, X.; Yu, S.; Xiong, L.; Luo, J. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: Application in the study of rice metabolomics. *Mol. Plant* **2013**, *6*, 1769–1780. [CrossRef] [PubMed]
24. Lu, Y.; Sun, J.; Yang, Z.; Zhao, C.; Zhu, M.; Ma, D.; Dong, T.; Zhou, Z.; Liu, M.; Yang, D. Genome-wide identification and expression analysis of glycine-rich RNA-binding protein family in sweet potato wild relative *Ipomoea trifida*. *Gene* **2019**, *686*, 177–186. [CrossRef] [PubMed]

25. Zhang, L.; Yu, Y.; Shi, T.; Kou, M.; Sun, J.; Xu, T.; Li, Q.; Wu, S.; Cao, Q.; Hou, W. Genome-wide analysis of expression quantitative trait loci (eQTLs) reveals the regulatory architecture of gene expression variation in the storage roots of sweet potato. *Hortic. Res.* **2020**, *7*, 90. [CrossRef] [PubMed]
26. Wang, X.; Xu, Y.; Zhang, S.; Cao, L.; Huang, Y.; Cheng, J.; Wu, G.; Tian, S.; Chen, C.; Liu, Y. Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nat. Genet.* **2017**, *49*, 765–772. [CrossRef]
27. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [CrossRef] [PubMed]
28. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
29. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [CrossRef] [PubMed]
30. Yu, G.; Wang, L.; Han, Y.; He, Q. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **2012**, *16*, 284–287. [CrossRef] [PubMed]
31. Zhu, P.; Dong, T.; Xu, T.; Kang, H. Identification, characterisation and expression analysis of MADS-box genes in sweetpotato wild relative Ipomoea trifida. *Acta Physiol. Plant* **2020**, *42*, 163. [CrossRef]
32. Chen, Y.; Zhang, Y.; Nawaz, G.; Zhao, C.; Li, Y.; Dong, T.; Zhu, M.; Du, X.; Zhang, L.; Li, Z. Exogenous melatonin attenuates post-harvest decay by increasing antioxidant activity in wax apple (*Syzygium samarangense*). *Front. Plant Sci.* **2020**, *11*, 569779. [CrossRef]
33. Yang, Z.; Zhu, P.; Kang, H.; Liu, L.; Cao, Q.; Sun, J.; Dong, T.; Zhu, M.; Li, Z.; Xu, T. High-throughput deep sequencing reveals the important role that microRNAs play in the salt response in sweet potato (*Ipomoea batatas* L.). *BMC Genom.* **2020**, *21*, 164. [CrossRef]
34. Li, Y.; Zhang, L.; Zhang, L.; Nawaz, G.; Zhao, C.; Zhang, J.; Cao, Q.; Dong, T.; Xu, T. Exogenous melatonin alleviates browning of fresh-cut sweetpotato by enhancing anti-oxidative process. *Sci. Hortic.* **2022**, *297*, 110937. [CrossRef]
35. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* **2001**, *25*, 402–408. [CrossRef]
36. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef] [PubMed]
37. Sun, J.; Qiu, C.; Ding, Y.; Wang, Y.; Sun, L.; Fan, K.; Gai, Z.; Dong, G.; Wang, J.; Li, X. Fulvic acid ameliorates drought stress-induced damage in tea plants by regulating the ascorbate metabolism and flavonoids biosynthesis. *BMC Genom.* **2020**, *21*, 411. [CrossRef] [PubMed]
38. Dou, X.; Zhou, Z.; Ren, R.; Xu, M. Apigenin, flavonoid component isolated from Gentiana veitchiorum flower suppresses the oxidative stress through LDLR-LCAT signaling pathway. *Biomed. Pharmacother.* **2020**, *128*, 110298. [CrossRef]
39. Wink, M. Evolution of secondary metabolites in legumes (*Fabaceae*). *S. Afr. J. Bot.* **2013**, *89*, 164–175. [CrossRef]
40. Xu, W.; Dubos, C.; Lepiniec, L. Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* **2015**, *20*, 176–185. [CrossRef] [PubMed]
41. Zong, Y.; Li, S.; Xi, X.; Cao, D.; Wang, Z.; Wang, R.; Liu, B. Comprehensive influences of overexpression of a MYB transcriptor regulating anthocyanin biosynthesis on transcriptome and metabolome of tobacco leaves. *Int. J. Mol. Sci.* **2019**, *20*, 5123. [CrossRef] [PubMed]
42. Gonzalez, A.; Zhao, M.; Leavitt, J.M.; Lloyd, A.M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* **2008**, *53*, 814–827. [CrossRef] [PubMed]
43. Das, P.K.; Shin, D.H.; Choi, S.; Park, Y. Sugar-hormone cross-talk in anthocyanin biosynthesis. *Mol. Cells* **2012**, *34*, 501–507. [CrossRef] [PubMed]
44. Koskimäki, J.J.; Hokkanen, J.; Jaakola, L.; Suorsa, M.; Tolonen, A.; Mattila, S.; Pirttilä, A.M.; Hohtola, A. Flavonoid biosynthesis and degradation play a role in early defence responses of bilberry (*Vaccinium myrtillus*) against biotic stress. *Eur. J. Plant Pathol.* **2009**, *125*, 629–640. [CrossRef]
45. Gharibi, S.; Tabatabaei, B.E.S.; Saeidi, G.; Talebi, M.; Matkowski, A. The effect of drought stress on polyphenolic compounds and expression of flavonoid biosynthesis related genes in *Achillea pachycephala* Rech. f. *Phytochemistry* **2019**, *162*, 90–98. [CrossRef] [PubMed]
46. Lazar, G.; Goodman, H.M. MAX1, a regulator of the flavonoid pathway, controls vegetative axillary bud outgrowth in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 472–476. [CrossRef]
47. Tan, G.; Ma, J.; Zhang, X.; Xu, Z.; Xiong, A. AgFNS overexpression increase apigenin and decrease anthocyanins in petioles of transgenic celery. *Plant Sci.* **2017**, *263*, 31–38. [CrossRef]
48. Righini, S.; Rodriguez, E.J.; Berosich, C.; Grotewold, E.; Casati, P.; Falcone Ferreyra, M.L. Apigenin produced by maize flavone synthase I and II protects plants against UV-B-induced damage. *Plant, Cell Environ.* **2019**, *42*, 495–508. [CrossRef] [PubMed]
49. Mizuno, H.; Yazawa, T.; Kasuga, S.; Sawada, Y.; Kanamori, H.; Ogo, Y.; Hirai, M.Y.; Matsumoto, T.; Kawahigashi, H. Expression of flavone synthase II and flavonoid 3′-hydroxylase is associated with color variation in tan-colored injured leaves of sorghum. *Front. Plant Sci.* **2016**, *7*, 1718. [CrossRef] [PubMed]

metabolites

MDPI

Article

Identification and Dissipation of Chlorpyrifos and Its Main Metabolite 3,5,6-TCP during Wheat Growth with UPLC-QTOF/MS

Lili Yu [1], Jia Li [2], Meiqin Feng [1], Qian Tang [1], Zejun Jiang [2], Hui Chen [1], Tingting Shan [1] and Junhui Li [3,*]

[1] College of Animal Science and Food Engineering, Jinling Institute of Technology, Nanjing 210038, China
[2] College of Life Sciences, China Jiliang University, Hangzhou 310018, China
[3] Shandong (Linyi) Institute of Modern Agriculture, Zhejiang University, Linyi 276000, China
* Correspondence: 18511581374@163.com; Tel./Fax: +86-0571-8898-2151

Abstract: Ultrahigh-performance liquid chromatography system coupled to a hybrid quadrupole time-of-flight mass spectrometer (UPLC-QTOF/MS) technology was used to investigate the degradation and metabolism of chlorpyrifos during wheat growth by spraying plants with different doses of chlorpyrifos 7 days after the flowering and filling stage. We analyzed and identified chlorpyrifos metabolites in different parts of wheat in full-scan MS^E mode, and established a chlorpyrifos metabolite screening library using UNIFI software. The results show that the residues of chlorpyrifos in wheat ears, leaves, and stems exhibited a decreasing trend with the prolongation of application time, and the degradation kinetics could be fitted with the first-order kinetic equation $C_t = C_0\, e^{-kt}$. The initial residues of chlorpyrifos in different parts of the wheat were different, in the order of leaves > wheat ears > stems. The degradation rate of chlorpyrifos under field conditions is relatively fast, and the half-life value is 2.33–5.05 days. Chlorpyrifos can undergo a nucleophilic addition substitution reaction under the action of hydrolase to generate secondary metabolite 3,5,6-trichloro-2-pyridinol (3,5,6-TCP). The residual amount of 3,5,6-TCP in each part of wheat first showed an increasing trend and then decreased over time. It reached the maximum on the 3rd, 7th, or 11th day after application, and then gradually degraded. Considering that 3,5,6-TCP is a biomarker with potential threats to humans and animals, it is recommended that 3,5,6-TCP be included in the relevant regulations for dietary exposure risk assessment.

Keywords: wheat; chlorpyrifos; 3,5,6-TCP; dissipation; metabolite; UPLC-QTOF/MS

Citation: Yu, L.; Li, J.; Feng, M.; Tang, Q.; Jiang, Z.; Chen, H.; Shan, T.; Li, J. Identification and Dissipation of Chlorpyrifos and Its Main Metabolite 3,5,6-TCP during Wheat Growth with UPLC-QTOF/MS. *Metabolites* **2022**, *12*, 1162. https://doi.org/10.3390/metabo12121162

Academic Editors: Yanjie Zhang and Yan Li

Received: 12 October 2022
Accepted: 18 November 2022
Published: 23 November 2022

1. Introduction

Chlorpyrifos is an effective organophosphorus pesticide that is widely used worldwide [1,2]. Its degradation behavior in the environment and organisms, the types of metabolites, metabolic pathways, and the fate of metabolites in the environment affect the ecological environment and the safety of organisms [3,4]. There are many reports on the residual dynamics of chlorpyrifos in plants. In field trials, the rate of chlorpyrifos degradation in plants is fast, and the residual amount of chlorpyrifos is closely related to the number of applications, dosage, and weather conditions after application. The research of Shen Yan (2007) showed that spraying plants with the recommended doses of chlorpyrifos at different periods after flowering resulted in a downward trend of chlorpyrifos residues in wheat ears, decreasing in the order of 35 days after flowering > 14 days after flowering > 21 days after flowering > 7 days after flowering; the degradation half-life was relatively short, between 1.20 and 3.36 days. During the maturity period, chlorpyrifos residues were detected in different parts of wheat, mainly distributed in the bran and glumes, but not in flour [5]. Wu et al. (2012) found that the degradation trend of chlorpyrifos in green beans was in line with a first-order reaction kinetics equation, and the initial residual concentration of chlorpyrifos in green beans was low, ranging from 0.571 to 1.737 mg/kg. The degradation half-lives of chlorpyrifos in beans treated one- and twofold the recommended dosage were

1.6 and 1.5 days, respectively. During the safety interval (14 days), the residual amount of chlorpyrifos in green beans was lower than the safety limits of the European Union, South Korea, Japan, and other countries, further indicating that chlorpyrifos is an easily degradable pesticide [6].

Chlorpyrifos undergoes metabolic transformation in organisms, mainly through oxidation and hydrolysis, to generate corresponding metabolites [7–9]. The oxidative reaction of chlorpyrifos in the organism is mainly the metabolic reaction that first occurs under the catalysis of cellular P450s to form the unstable intermediate oxysulfide phosphate [10], which further undergoes oxidative desulfurization or dearylation to generate corresponding metabolites [11]. The oxidative desulfurization of chlorpyrifos is a toxicity-enhancing metabolic reaction that is mainly caused by the exchange of phosphorus atoms and oxygen atoms in the molecular structure to form the metabolite chlorpyrifos-oxon (CPO) [12]. A study by the United States Geological Survey found that the main oxidative metabolite of chlorpyrifos in the environment, CPO, is more toxic to animals than the parent pesticide is [13]. Antonious et al. (2017) reported that, under field conditions, after spraying chlorpyrifos on two kale varieties, the CPO metabolite was found. Although the residual concentration of CPO was lower than that of the parent pesticide chlorpyrifos, the degradation half-life (1.15–18.02 days) of CPO was longer than that of the parent pesticide chlorpyrifos (2.21–5.10 days), and its persistence and biological toxicity in plants were much higher than those of chlorpyrifos [14]. The dearylation of chlorpyrifos is a detoxification metabolic reaction through which chlorpyrifos can form metabolites diethylphosphate (DEP), diethylphosphorothioate (DETP) and 3,5,6-trichloro-2-pyridinol (3,5,6-TCP) [11]. In addition, chlorpyrifos can break the covalent bond between phosphorus atoms and leave groups under the catalytic action of various hydrolases, resulting in the less-toxic metabolites DEP, DETP, and 3,5,6-TCP [15–17].

QTOF/MS has the characteristics of high resolution, a wide mass range, high accuracy, and fast analytical speed, combined with the structural information obtained in MS/MS mode. It can obtain the accurate mass information of compounds through high sensitivity in the full-scan acquisition mode. In addition, QTOF/MS can optimize collision energy to obtain accurate mass precursor ions and fragment ions. Recently, QTOF/MS has been used as a powerful tool for multiple pesticide analysis and metabolite screening. Saito-Shida et al. (2016) developed a multiresidue method for the identification and quantification of 149 pesticides in four vegetables and fruits at a spiking level of 0.01 mg/kg using LC-QTOF/MS [18]. Similarly, Yang et al. (2018) developed a multiresidue detection method that used UPLC-QTOF/MS to identify and quantify 50 pesticides in fruits with high sensitivity and accuracy [19]. In another study, a multicomponent analytical method for the screening and quantification of pesticide residues in paprika samples via UPLC-QTOF/MS was developed and validated [20]. These results showed that QTOF/MS has high accuracy, sensitivity, and selectivity for the qualitative and quantitative aspects of compounds, and can be widely used in the analysis of pesticide residues, the identification of metabolites, and metabolomics research. However, there is still a lack of knowledge about the nature of the metabolite screening of chlorpyrifos during wheat growth.

The maximal residue limits (MRLs) of chlorpyrifos in wheat are 0.5 mg/kg in the Codex Alimentarius Commission (CAC), United States, Japan, and China [21–24]. Although MRLs are a feasible method to ensure the safe and effective implementation of pesticides, some metabolites in pesticides may be more toxic than the parent compound. However, few studies have focused on determining the MRLs of pesticide metabolites in food. It is, therefore, of great research significance to study the metabolic mechanism of chlorpyrifos in wheat.

The overall aim of the present study was to analyze and identify chlorpyrifos metabolites in different tissues of wheat through UPLC-QTOF/MS in full-scan MS^E mode, and to establish a chlorpyrifos metabolite screening library using UNIFI software. On the basis of accurate mass, retention time, and diagnostic ions, we found chlorpyrifos metabolite 3,5,6-TCP in different tissues of wheat in the open field. This research provides information to develop strategies for the safe use of chlorpyrifos under open-field conditions.

2. Materials and Methods

2.1. Reagents and Chemicals

A standard solution of chlorpyrifos (1000 mg/L) was purchased from the Agro-Environment Protection Institute, Ministry of Agriculture and Rural Affairs of China (Beijing, China). Commercial chlorpyrifos (45% emulsifiable concentrate (EC)) was obtained from Suzhou Jiahui Chemical Co., Ltd. (Suzhou, China). The 3,5,6-TCP standard (purity 99.6%) was supplied by Sigma-Aldrich (Shanghai) Trading Co., Ltd. (Shanghai, China). Organic solvents, including HPLC/MS-grade methanol, acetonitrile, formic acid, ethyl acetate, and ammonium acetate, were purchased from Thermo Fisher Scientific Corporation (Shanghai, China). Ultrapure water was collected from a Milli-Q system (Millipore, Billerica, MA, USA).

Standard stock solutions of chlorpyrifos and 3,5,6-TCP (100 mg/L) were prepared with LC-grade methanol and stored at −20 °C. These working standard solutions (0.001, 0.02, 0.05, 0.1, 0.2, 0.5, and 1 mg/L) were prepared by serially diluting the stock solution. Correspondingly, matrix-matched standard solutions of chlorpyrifos and 3,5,6-TCP (0.001, 0.02, 0.05, 0.1, 0.2, 0.5, and 1 mg/L) were prepared by serially diluting the working matrix standard solutions. These solutions were stored in the dark at 4 °C.

2.2. Field Experiment Design

Field trials were conducted from May to June in 2017 and 2018 at the Experimental Base of the Chinese Academy of Agricultural Sciences in Shunyi Farm, Beijing (116°33′ E, 40°13′ N). Weather data represent a monsoon climate of medium latitudes and were provided by the China Meteorological Administration. In 2017, the average daily temperature was 18.0–30.5 °C, the average daily sunshine hours were 14.38–14.93 h, and the average daily precipitation was 1.03–2.52 mm. In 2018, the average daily temperature was 18.5–30.5 °C, the average daily sunshine hours were 14.37–14.98 h, and the average daily precipitation was 1.09–2.65 mm. Chlorpyrifos was sprayed 7 days after the flowering and filling stage of wheat; plants were sprayed with clean water as a control. Treatment plots were sprayed with four different concentrations of chlorpyrifos. Plot 1 was used as the control. Plots 2–5 were sprayed with commercial 45% EC chlorpyrifos at the recommended dosage (450 mL/hectare), 2-fold the recommended dosage (900 mL/hectare), 5-fold the recommended dosage (2250 mL/hectare), and 10-fold the recommended dosage (4500 mL/hectare), respectively. Plots were randomly arranged, and each treatment was replicated three times. Wheat ear, leaf, and stem samples were taken on Days 0, 1, 3, 7, 11, 14, 21, 28, and 35 (harvest period) after application, and 50 plants were randomly sampled in each plot. These samples were placed in polyethylene bags and stored at −40 °C until analysis.

2.3. Sample Treatment

The homogenized sample (5 g) was placed in a 50 mL polypropylene centrifuge tube and extracted with 20 mL acetonitrile–water (50:50, *v*/*v*) for 30 min using an automatic shaker. Afterwards, a total of 4 g of $MgSO_4$, 1 g of NaCl, 1 g of sodium citrate dihydrate, and 0.5 g of sodium hydrogen citrate sesquihydrate were added and shaken vigorously for 2 min, and then all the samples were centrifuged for 5 min at 6000 rpm. Then, 5 mL of the upper layer (acetonitrile) was transferred to a 15 mL centrifuge tube containing 150 mg of PSA, 150 mg of C_{18}, and 900 mg of $MgSO_4$. The tube was immediately vortexed for 1 min and centrifuged for 5 min at 6000 rpm. Then, 2 mL of supernatant-cleaned extract was evaporated to dryness at 60 °C. The dry extract was dissolved with 1 mL methanol, and the resulting solution was then filtered through a 0.22 μm nylon syringe filter (Jinteng, Tianjin, China) and analyzed with UPLC-QTOF/MS.

The extraction and clean-up of 3,5,6-TCP were performed as follows. The homogenized sample (5 g) was placed in a 50 mL polypropylene centrifuge tube and extracted with 20 mL of a mixture of acetonitrile and ethyl acetate (50:50, *v*/*v*), 10 mL of saturated sodium chloride solution, and 0.5 mL of 12 mol/L hydrochloric acid. The sample was vigorously shaken for 30 min. Subsequently, a centrifugation step (10,000 rpm, 5 min) was performed.

Following this, 4 mL of the supernatant-cleaned extract was collected and transferred to an Oasis Prime HLB cartridge (Waters Corp, Milford, MA, USA) for purification. Then, the extract was evaporated to dryness at 45 °C and 200 mbar. The dry extract was reconstituted with 1 mL of methanol and the resulting solution was passed through a 0.22 μm nylon syringe filter (Jinteng, Tianjin, China) and analyzed using UPLC-QTOF/MS.

2.4. Instrumentation and Conditions

An ultrahigh-performance liquid chromatography system (ACQUITY UPLC I-CLASS, Waters Corp., Milford, MA, USA) coupled to a hybrid quadrupole time-of-flight mass spectrometer (VION IMS QTOF, Waters Corp., Milford, MA, USA) was used for this study. Sample separation was performed using a Waters ACQUITY UPLC HSS T3 analytical column (particle size 1.8 μm, 2.1 mm (i.d.) × 100 mm (length)). The mobile phase consisted of methanol (Solvent A) and 10 mmol ammonium acetate in water (Solvent B) applied at a flow rate of 0.45 mL/min in the following gradient mode: (i) 0 min (A–B, 2:98, *v*/*v*), (ii) 0.25 min (A–B, 2:98, *v*/*v*), (iii) 12.25 min (A–B, 99:1, *v*/*v*), (iv) 13 min (A–B, 99:1, *v*/*v*), (v) 13.01 min (A–B, 2:98, *v*/*v*) and (vi) 17 min (A–B, 2:98, *v*/*v*). The injection volume was set at 5 μL, and the column temperature was held at 45 °C.

The mass spectrometer in the positive electrospray ionization (ESI^+) mode was used for the analysis of chlorpyrifos and its metabolites. The MS source conditions were as follows: capillary voltage, 1.0 kV; sample cone voltage, 40 V; nitrogen gas-flow of the nebulizer, 50 L/h; desolvation gas-flow, 1000 L/h; desolvation temperature, 550 °C; and source temperature, 120 °C. Spectral data was acquired in a mass range of *m*/*z* 50–1000 using full scan and the mass spectrometer elevated (MS^E) experiment mode with an acquisition speed of 0.2 scan/s. In the MS^E function, the LE spectrum was recorded at 6.0 eV, and the HE spectrum was recorded with a collision energy ramped from 10 to 45 eV. Real-time calibration was performed with 200 pg/μL leucine–enkephalin (*m*/*z* 556.2766 in positive mode).

2.5. Method Validation

Recovery experiments were conducted to investigate the accuracy and precision of the method. Five replicates of spiked untreated wheat samples at five different levels (20, 50, 100, 200, and 500 μg/kg) were prepared with chlorpyrifos and 3,5,6-TCP working solutions in methanol. Then, the recoveries obtained with the spiked samples were compared with those of the matrix-matched calibration solutions. The limits of detection (LODs) for chlorpyrifos and 3,5,6-TCP were considered to be the concentration that produced a signal-to-noise (S/N) ratio of 3, and the limits of quantitation (LOQs) were assessed at an S/N ratio of 10. The linearity (R^2) was evaluated using a matrix-matched calibration.

2.6. Statistical Analysis

The data were collected and processed using UNIFI™ 1.8.1 software (Waters Corp., Milford, MA, USA). Data were processed with a scientific library that was created inhouse containing a suspected database of chlorpyrifos metabolites (four library entities) with information about molecular structures, the exact mass analysis of precursor ions, fragment ions, adducts, and retention time in the database. In addition, in this study, the method conditions for pesticide screening to establish the scientific compound library were set according to Waters Corp. (Waters, Milford, MA, USA) [25]. The dissipation studies of residues of chlorpyrifos and its metabolite were performed using linear regression. Statistical analyses were performed using the PSAW Statistic 19.0 statistical software package (SPSS, Chicago, IL, USA). All data were statistically evaluated with one-way analysis of variance (ANOVA) followed by Duncan's multiple-range test. The data are shown and were analyzed as micrograms per kilogram of matrix (mg/kg) on a dry-matter basis. All the values are reported as the means ± the standard deviation (SD) of five replicates.

3. Results and Discussion

3.1. Identification and Confirmation of Chlorpyrifos and Its Metabolites with UPLC-QTOF/MS

Table 1 shows the molecular formula, accurate mass, fragment ions, retention time and adducts of chlorpyrifos and its metabolite 3,5,6-TCP obtained in the full-scan MS^E mode of UPLC-QTOF/MS. The retention time of chlorpyrifos was 11.20 min; the characteristic fragment ions of chlorpyrifos with the same retention time obtained by QTOF/MS analysis were m/z 311.2577, m/z 197.9273 and m/z 96.9505. The retention time of the metabolite 3,5,6-TCP was 6.00 min, and the characteristic fragment ions of 3,5,6-TCP with the same retention time obtained by QTOF/MS analysis were m/z 184.0724, m/z 116.0693, and m/z 70.0646.

Table 1. Accurate UPLC-QTOF/MS mass measurements of chlorpyrifos and 3,5,6-TCP in wheat samples.

Compound	Retention Time (min)	Formula	Observed Neutral Mass (Da)	Fragments	Adducts
Chlorpyrifos	11.20	$C_9H_{11}Cl_3NO_3PS$	349.9335	311.2577 197.9273 96.9505	+H, +Na
3,5,6-TCP	6.00	C_5H_2 Cl_3NO	197.9268	184.0724 116.0693 70.0646	+H

For all the possible metabolites of chlorpyrifos reported in Table 2, we established a metabolite screening library for all suspected metabolites of chlorpyrifos in different plant species to detect and identify the possible metabolites of chlorpyrifos during wheat growth. Then, we analyzed and identified chlorpyrifos metabolites in different parts of wheat through UPLC-QTOF/MS in full-scan MS^E mode, and established a pesticide metabolite screening library using UNIFI software. Table 2 shows the molecular formula and calculated exact mass of all possible metabolites of chlorpyrifos reported in the references. The identification and diagnostic proposal of metabolites were performed in accordance with the following criteria: (i) unique peaks in the processed sample compared with the blank sample; (ii) accurate mass deviation of precursor ion < 2 ppm mass error; (iii) retention time error of all samples < 0.1 min; (iv) at least $\geq$ 1 characteristic fragment ion by UNIFI™ [26,27].

Table 2. List of reported metabolites of chlorpyrifos in references, their formula, and calculated exact masses.

Compound	Formula	Exact Mass (Da)
3,5,6-TCP	$C_5H_2Cl_3NO$	198.4345
DEP	$C_4H_{11}O_4P$	154.1015
DETP	$C_4H_{11}O_3PS$	170.1671
CPO	$C_9H_{11}Cl_3NO_4P$	332.9491

We analyzed and identified chlorpyrifos metabolite 3,5,6-TCP in different parts of wheat on the basis of accurate mass, retention time, adducts, diagnostic ions, and other standards. Since the reference substance of 3,5,6-TCP is commercially available, the further quantitative determination of 3,5,6-TCP in wheat samples was performed with UPLC-QTOF/MS. Figure 1A summarizes the UPLC-QTOF/MS-extracted ion chromatogram and MS^E spectrum of chlorpyrifos and its metabolite 3,5,6-TCP. Figure 1B shows the total ion chromatogram (TIC) and the low collision energy channel data for chlorpyrifos and 3,5,6-TCP. Chlorpyrifos presented an $[M + H]^+$ peak as a base peak in the spectrum, and 3,5,6-TCP presented the sodium adduct $[M + H]^+$ and $[M + Na]^+$ as a base peak in which the hydrogen adduct might be the main form of chlorpyrifos and 3,5,6-TCP in mass spectrometry.

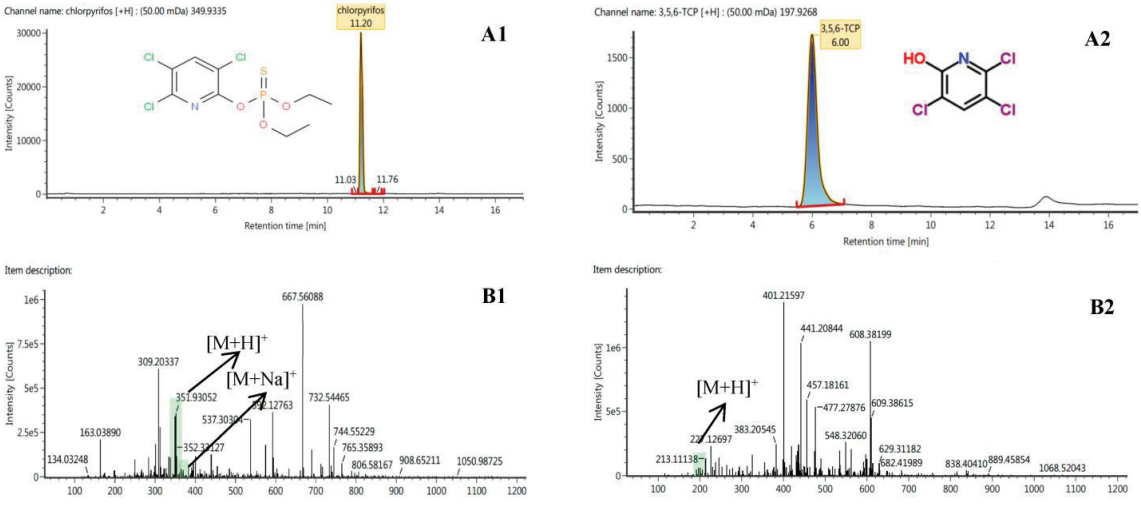


Figure 1. UPLC-QTOF/MS-extracted ion chromatogram and MS^E spectra of chlorpyrifos and 3,5,6-TCP: (**A1**) chlorpyrifos standard at 100 ug/kg in wheat sample; (**A2**) 3,5,6-TCP standard at 100 ug/kg in wheat sample; (**B1**) the low collision energy adducts of H^+, Na^+ chlorpyrifos; (**B2**) low-collision-energy adducts of H^+, 3,5,6-TCP.

3.2. Method Validation for Chlorpyrifos and 3,5,6-TCP in Different Parts of Wheat

Five levels of chlorpyrifos and 3,5,6-TCP standard solutions were added to the blank wheat samples to verify the reliability of the method. The results are shown in Table 3. Within the range of 20–500 μg/kg, the concentrations of chlorpyrifos and 3,5,6-TCP exhibited excellent linearity with the peak area of the quantitative ion and the correlation coefficients (R^2) ranged from 0.9958 to 0.9995. The mean recovery of chlorpyrifos and 3,5,6-TCP in different wheat matrices ranged from 63.38% to 102.13%, and the relative standard deviation (RSD) ranged from 1.76% to 8.99%. The LOD range of chlorpyrifos was 0.38–0.85 μg/kg, and the LOQ range was 1.30–2.70 μg/kg, which is lower than the MRLs established by the United States, the European Union, Australia, Japan, and China. The LOD range of 3,5,6-TCP was 2.90–4.50 μg/kg and the LOQ range was 11.00–14.00 μg/kg. The recovery rate and accuracy of this method are suitable for the determination of chlorpyrifos and 3,5,6-TCP content in different wheat matrices.

Table 3. R^2, recoveries, LOD and LOQ of chlorpyrifos and its metabolite 3, 5, 6-TCP in different matrices (n = 5). Values (mean ± SD) in the same row.

Compound	Matrix	R^2	Average Recovery and Standard Deviations (%)					LOD (μg/kg)	LOQ (μg/kg)
			Spiking Level (μg/kg)						
			20	50	100	200	500		
Chlorpyrifos	Wheatear	0.9995	72.00 ± 2.18	71.33 ± 6.43	72.67 ± 7.57	82.83 ± 1.76	102.13 ± 3.35	0.38	1.30
	Leaf	0.9974	76.67 ± 5.77	83.33 ± 5.03	85.00 ± 2.00	85.50 ± 2.78	76.73 ± 4.27	0.85	2.70
	Stem	0.9983	77.00 ± 3.29	78.46 ± 5.87	79.16 ± 7.22	85.45 ± 2.77	89.53 ± 4.88	0.70	2.20
3, 5, 6-TCP	Wheatear	0.9977	66.33 ± 6.24	70.01 ± 7.35	72.37 ± 8.35	83.57 ± 7.37	83.36 ± 8.99	2.90	11.00
	Leaf	0.9958	65.79 ± 5.34	63.38 ± 4.26	69.24 ± 6.38	72.43 ± 8.46	77.84 ± 6.59	4.50	14.00
	Stem	0.9983	67.86 ± 5.76	71.47 ± 7.58	73.15 ± 5.37	81.38 ± 3.58	87.67 ± 5.68	4.25	13.50

LOD: The limit of detection; LOQ: The limit of quantitation; 3,5,6-TCP: 3,5,6-trichloro-2-pyridinol; SD: Standard Deviation.

3.3. Dynamic Distribution of Chlorpyrifos in Different Parts of Wheat

The dynamic distribution of different concentrations of chlorpyrifos in different parts of wheat was fitted with the first-order reaction kinetics equation: $C_t = C_0\ e^{-kt}$ (where C_t is the concentration of the pesticide at time t, C_0 is the initial concentration of the pesticide, k is the degradation rate constant (1/d) and t is the number of days after applica-

tion) [28,29]. $t_{1/2}$ (degradation half-life) refers to the time required for the applied pesticide to be degraded by 50%. As presented in Table 4, in the field trials of 2017 and 2018, after spraying plants with different concentrations of chlorpyrifos pesticides, the degradation trends of chlorpyrifos in wheat ears, leaves, and stems all accorded with the first-order reaction kinetics equation. Here, the fitting degree was high, and the correlation coefficient ranged from 0.8646 to 0.9923. The field trial results (Figures 2, 3 and 4A) showed that, in 2017, the initial residues of chlorpyrifos in plants sprayed with 1-, 2-, 5- and 10-fold the recommended dosage were 2.649, 6.621, 18.316, and 46.653 mg/kg, respectively, in wheat ears (Figure 2(A-1)); 9.989, 12.656, 32.600 and 77.520 mg/kg, respectively, in the leaves (Figure 3(A-1)); and 1.515, 2.160, 7.033 and 19.333 mg/kg, respectively, in the stems (Figure 4(A-1)). In 2018, the initial residues of chlorpyrifos in plants sprayed with 1-, 2-, 5- and 10-fold the recommended dosage were 2.569, 5.860, 21.619 and 51.173 mg/kg, respectively, in wheat ears (Figure 2(A-2)); 10.288, 16.181, 34.887 and 81.440 mg/kg, respectively, in the leaves (Figure 3(A-2)); and 1.579, 2.701, 8.540 and 22.960 mg/kg, respectively, in the stems (Figure 4(A-2)). Field trial results in two localities for 2 years showed that the initial residues of chlorpyrifos in different parts of the wheat were different, and the distribution of initial residues for 2 years was in the order of leaves > wheat ears > stems. In addition, after application, the residual amount of chlorpyrifos in each part of the wheat plant showed a gradual degradation trend over time, dropping to a minimum at the maturity period (35 days after application). Our results are consistent with the report of Shen (2007), who indicated that, after spraying wheat plants with chlorpyrifos at different periods after flowering, the residual amount of chlorpyrifos in wheat ears, leaves, and stems showed a downward trend, with the initial deposition of chlorpyrifos in the different organs of wheat in the order of leaves > wheat ears > stems. The research results of Yang et al. (2018) showed that, after spraying rice plants with chlorpyrifos, the residual amount of chlorpyrifos in the various organs of the plants showed a single-peak curve and gradually degraded over time, and the initial deposits of chlorpyrifos were in the order of leaves > grains > stems > roots. Similarly, Zhang et al. (2012) reported that, after spraying rice plants with chlorpyrifos, the residual amount of chlorpyrifos in rice, husks, and straws dissipated gradually over time. The initial residues of chlorpyrifos in different tissues of rice were in the order of straw > husks > rice [30]. These results show that the degradation rule of chlorpyrifos in different parts of wheat is consistent and unaffected by changes in application dose and plant species.

Table 4. Degradation kinetics of chlorpyrifos at different applied dosages in different parts of wheat.

Matrix	Time	Treatment	First-Order Kinetic Equation	C_0 (mg/kg)	R^2	K (1/d)	$t_{1/2}$ (d)
Wheatear	2017	1×	$Ct = 2.4668\ e^{-0.2765t}$	2.4668	0.9201	0.2765	2.51
		2×	$Ct = 4.5245\ e^{-0.2609t}$	4.5245	0.9638	0.2609	2.66
		5×	$Ct = 12.3650\ e^{-0.2510t}$	12.3650	0.9653	0.2510	2.76
		10×	$Ct = 24.3582\ e^{-0.2080t}$	24.3582	0.9265	0.2080	3.33
	2018	1×	$Ct = 2.4479\ e^{-0.1921t}$	2.4479	0.9693	0.1921	3.61
		2×	$Ct = 2.9560\ e^{-0.1653t}$	2.9560	0.8646	0.1653	4.19
		5×	$Ct = 9.3164\ e^{-0.1775t}$	9.3164	0.8839	0.1775	3.90
		10×	$Ct = 34.7860\ e^{-0.1562t}$	34.7860	0.9595	0.1562	4.44
Leaf	2017	1×	$Ct = 7.6341\ e^{-0.2199t}$	7.6341	0.9771	0.2199	3.15
		2×	$Ct = 10.8377\ e^{-0.1580t}$	10.8377	0.9726	0.1580	4.39
		5×	$Ct = 32.4195\ e^{-0.1954t}$	32.4195	0.9381	0.1954	3.55
		10×	$Ct = 56.8813\ e^{-0.1494t}$	56.8813	0.9898	0.1494	4.64
	2018	1×	$Ct = 8.6738\ e^{-0.2117t}$	8.6738	0.9863	0.2117	3.27
		2×	$Ct = 10.5375\ e^{-0.1516t}$	10.5375	0.9778	0.1516	4.57
		5×	$Ct = 25.7491\ e^{-0.1512t}$	25.7491	0.9909	0.1512	4.58
		10×	$Ct = 62.3189\ e^{-0.1372t}$	62.3189	0.9919	0.1372	5.05

Table 4. *Cont.*

Matrix	Time	Treatment	First-Order Kinetic Equation	C_0 (mg/kg)	R^2	K (1/d)	$t_{1/2}$ (d)
Stem	2017	1×	$Ct = 1.0919\ e^{-0.2687t}$	1.0919	0.8979	0.2687	2.58
		2×	$Ct = 1.8643\ e^{-0.2954t}$	1.8643	0.9253	0.2954	2.35
		5×	$Ct = 6.8317\ e^{-0.3142t}$	6.8317	0.9685	0.3142	2.21
		10×	$Ct = 27.7730\ e^{-0.2979t}$	27.7730	0.9742	0.2979	2.33
	2018	1×	$Ct = 1.8867\ e^{-0.2950t}$	1.8867	0.9893	0.2950	2.35
		2×	$Ct = 2.3523\ e^{-0.2311t}$	2.3523	0.9799	0.2311	3.00
		5×	$Ct = 8.0059\ e^{-0.2213t}$	8.0059	0.9843	0.2213	3.13
		10×	$Ct = 20.6520\ e^{-0.2002t}$	20.6520	0.9923	0.2002	3.46

1×: recommended dosage. 2×: twofold recommended dosage. 5×: fivefold recommended dosage. 10×: tenfold recommended dosage.

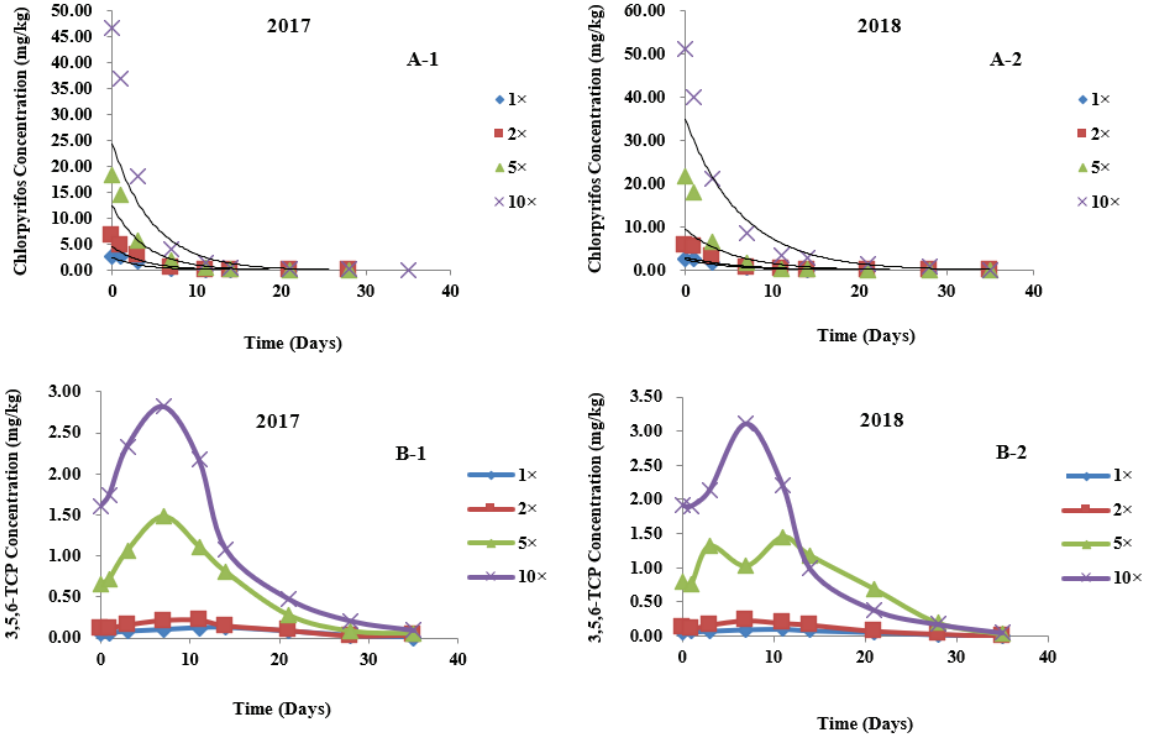


Figure 2. Degradation curves of chlorpyrifos and its metabolite 3,5,6-TCP at different applied dosages in wheat ears. (**A-1**) Degradation curves of Chlorpyrifos at different applied dosages in wheat ears in 2017; (**A-2**) Degradation curves of Chlorpyrifos at different applied dosages in wheat ears in 2018; (**B-1**) Degradation curves of 3,5,6-TCP at different applied dosages in wheat ears in 2017; (**B-2**) Degradation curves of 3,5,6-TCP at different applied dosages in wheat ears in 2018. 1×: recommended dosage. 2×: twofold recommended dosage. 5×: fivefold recommended dosage. 10×: tenfold recommended dosage.

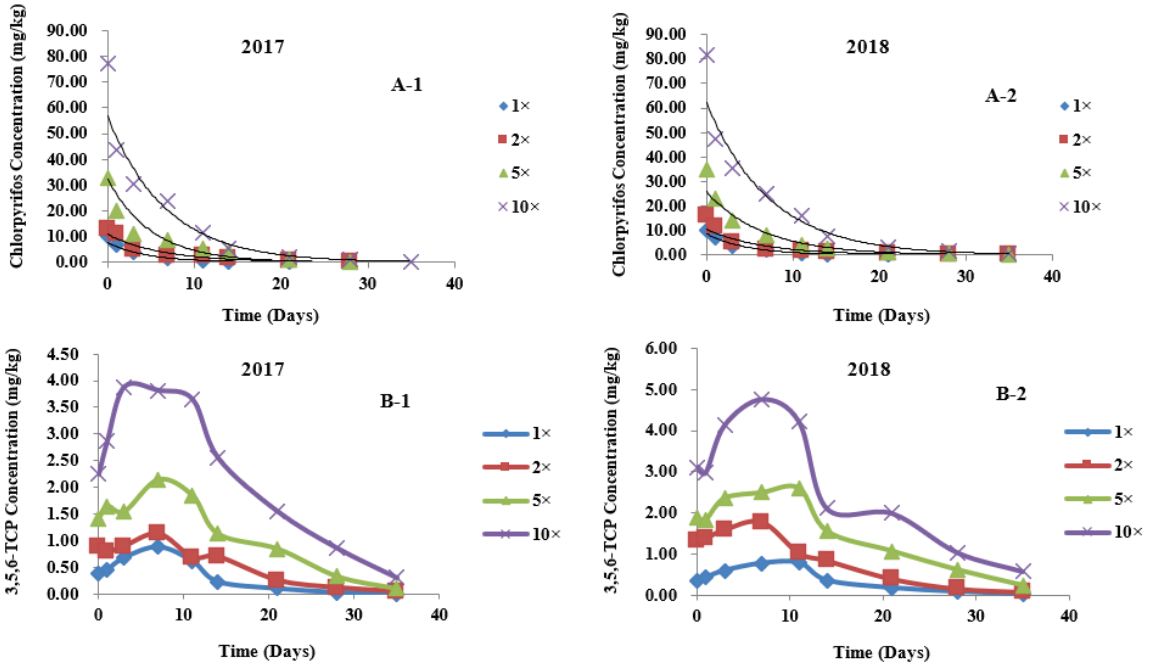


Figure 3. Degradation curves of chlorpyrifos and its metabolite 3,5,6-TCP at different applied dosages in leaves. (**A-1**) Degradation curves of Chlorpyrifos at different applied dosages in leaves in 2017; (**A-2**) Degradation curves of Chlorpyrifos at different applied dosages in leaves in 2018; (**B-1**) Degradation curves of 3,5,6-TCP at different applied dosages in leaves in 2017; (**B-2**) Degradation curves of 3,5,6-TCP at different applied dosages in leaves in 2018. 1×: recommended dosage. 2×: twofold recommended dosage. 5×: fivefold recommended dosage. 10×: tenfold recommended dosage.

Table 4 shows that the rate of chlorpyrifos degradation in each part of wheat was relatively fast, and the degradation half-lives of chlorpyrifos in wheat ears, leaves, and stems were 2.51–4.44, 3.15–5.05, and 2.21–3.46 days, respectively. The rate of chlorpyrifos degradation was fastest within 2–5 days after application, indicating that chlorpyrifos is an easily degradable pesticide, which may be related to its chemical properties, the characteristics of the applied plants, and environmental factors. In addition, in the early stage of application, the rate of chlorpyrifos degradation in each part of wheat was fast, and with the passage of time, it tended to be flat. The rule of chlorpyrifos degradation in various parts of wheat is not only affected by external environmental factors; the growth rate of the crop itself also greatly affects the persistence of pesticide residues in various parts of wheat. With the growth of plants, the pesticide is diluted by a larger surface area, and the pesticide content gradually decreases. However, as the plants grow, the cuticle becomes thicker in each part of the plant, especially the stems, as they contain more lignin, which means that pesticides are retained more strongly by mature plants than by young ones, and the pesticides that penetrate into the wheat tissue are less affected by external environmental factors; therefore, the degradation rate slows down. During the wheat harvest period (35 days after application), the residual amount of chlorpyrifos in wheat ears was below the national safety limit. Among treatments, in the group where chlorpyrifos was applied at the recommended dose, it was completely degraded, indicating that chlorpyrifos is an easily degradable and low-residue insecticide. This shows that spraying different doses of chlorpyrifos during the flowering and filling period of wheat is safe for wheat.

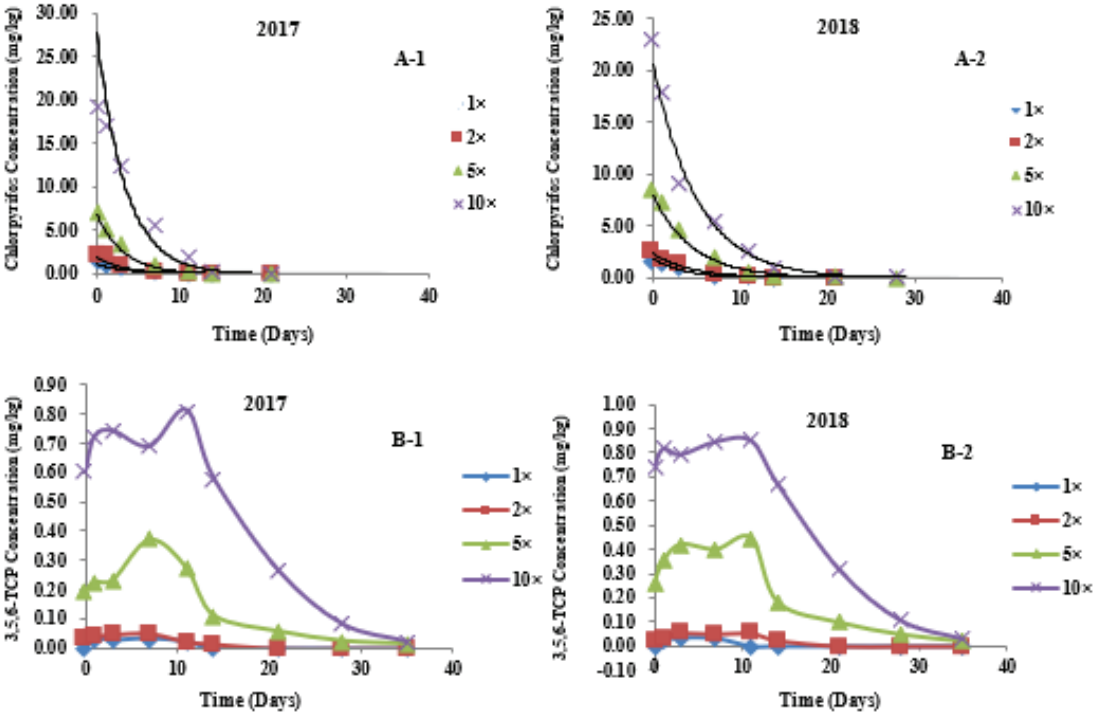


Figure 4. Degradation curves of chlorpyrifos and its metabolite 3,5,6-TCP at different applied dosages in stems. (**A-1**) Degradation curves of Chlorpyrifos at different applied dosages in stems in 2017; (**A-2**) Degradation curves of Chlorpyrifos at different applied dosages in stems in 2018; (**B-1**) Degradation curves of 3,5,6-TCP at different applied dosages in stems in 2017; (**B-2**) Degradation curves of 3,5,6-TCP at different applied dosages in stems in 2018. 1×: recommended dosage. 2×: twofold recommended dosage. 5×: fivefold recommended dosage. 10×: tenfold recommended dosage.

3.4. Metabolic Kinetics of 3,5,6-TCP in Different Parts of Wheat

In this study, UPLC-QTOF/MS and pesticide metabolite screening library analysis showed that, after chlorpyrifos was applied in the field, the 3,5,6-TCP metabolite could be detected in all parts of wheat. The dynamic relationship between 3,5,6-TCP residue and application time is shown in Figures 2, 3 and 4B. In 2017, the initial residues of 3,5,6-TCP in plants sprayed with 1-, 2-, 5-, and 10-fold the recommended dosage of chlorpyrifos were 0.057, 0.116, 0.657, and 1.604 mg/kg, respectively, in wheat ears (Figure 2(B-1)); 0.385, 0.889, 1.418 and 2.246 mg/kg, respectively, in the leaves (Figure 3(B-1)); and 0.000, 0.034, 0.193 and 0.604 mg/kg, respectively, in the stems (Figure 4(B-1)). In 2018, the initial residues of 3,5,6-TCP in plants sprayed with 1-, 2-, 5- and 10-fold the recommended dosage of chlorpyrifos were 0.048, 0.134, 0.807 and 1.909 mg/kg, respectively, in wheat ears (Figure 2(B-2)); 0.357, 1.336, 1.889 and 3.112 mg/kg, respectively, in the leaves (Figure 3(B-2)); and 0.000, 0.018, 0.262 and 0.740 mg/kg, respectively, in the stems (Figure 4(B-2)). The 2-year field data show that the initial residual concentrations of the 3,5,6-TCP metabolite in various parts of wheat were much lower than those of the parent pesticide chlorpyrifos, and the distribution of 3,5,6-TCP content in different parts of wheat in different dose groups was basically the same as that of the parent pesticide chlorpyrifos, i.e., in the order of leaves > wheat ears > stems. This may be because, under the same application conditions, the different structures of plant organs lead to differences in pesticide residues of plants. The critical surface tension, stratum corneum structure, and thickness of different plant organs are different

and affect the adhesion and diffusion of pesticides on the plant surface [31,32]. As the main plant target organ for biotransformation and metabolism, leaves have a large surface area and a thin natural stratum corneum, which facilitate absorbing pesticides. Therefore, the initial residues of 3,5,6-TCP in leaves (0.357–3.112 mg/kg) and the persistence of pesticides in leaves were relatively high. During the wheat harvest period (35 days after application), the residues of 3,5,6-TCP in the leaves of plants sprayed with the recommended dosage were degraded, while the residues in the leaves of plants treated with 2–10-fold the recommended dosage were 0.038–0.588 mg/kg (Figure 3B). Compared with the leaves, stems have many similarities in the absorption of pesticides. However, the stem has a small surface area and thick cuticle, so its ability to absorb pesticides is relatively weak. The initial residual amount of 3,5,6-TCP in the stems was low (0.000–1.604 mg/kg), and the degradation rate was fast: 21 days after application, it was completely degraded in groups treated with 1- and 2-fold the recommended dose. During the harvest period, only a small amount of 3,5,6-TCP residues (0.012–0.033 mg/kg) could be detected in the groups treated with 5- and 10-fold the recommended dose (Figure 4B). As the edible part of wheat, wheat ears have a large surface area, but due to the enveloping glumes, the stratum corneum is thick, and the absorption capacity of both the parent pesticide and metabolites is relatively weak compared with that of the leaves. During the wheat harvest period, 3,5,6-TCP in wheat ears was completely degraded in the group sprayed with the recommended dose, and a small amount of 3,5,6-TCP was detected in the groups treated with 2–10-fold the recommended dose (0.031–0.093 mg/kg). The variation pattern of 3,5,6-TCP in each part of wheat was that it increased first and then decreased with degradation of the parent pesticide chlorpyrifos, with a slight fluctuation in the middle; it reached the maximal value on the 7th or 11th day after application, and then gradually degraded with the extension of time (Figure 2B).

Pesticides applied in the field can undergo metabolic transformation through oxidation, reduction, hydrolysis, photolysis, biodegradation, and other forms in plants through the action of enzymes or the influence of external environmental factors (temperature, precipitation, humidity, etc.), to produce specific metabolites [1,14,15,28,33]. In plant, soil, water and biological reaction systems, especially during photolysis, hydrolysis, and microbial metabolism, chlorpyrifos can be converted into more than 20 metabolites with a slow kinetic process and a half-life ranging from weeks to months [34]. Zhang et al. (2015) reported that, after spraying rice plants with chlorpyrifos, its 3,5,6-TCP metabolite was found in rice and rice stalks at a significantly lower concentration than that of the parent pesticide chlorpyrifos [35]. Antonious et al. (2017) reported that, after applying chlorpyrifos under field conditions, metabolites CPO and 3,5,6-TCP were found in two different kale varieties. Although the concentrations of the two metabolites were much lower than those of the parent pesticide chlorpyrifos, the degradation half-life of CPO (1.15–18.02 days) was much longer than that of chlorpyrifos (2.21–5.10 days), and that of 3,5,6-TCP (3.33–3.34 days) was similar to that of chlorpyrifos. The content of 3,5,6-TCP in kale also increased first and then decreased with the degradation of chlorpyrifos [14]. However, the results of Chai et al. (2009) showed that, after applying chlorpyrifos to green mustard and soil, no metabolites were found in green mustard, while 3,5,6-TCP was detected in soil. The degradation half-life of 3,5,6-TCP in soil was 3.6–9.4 days. The degradation rate was fast in the early stage and gradually slowed down after 21 days until the 77th day, at which point the 3,5,6-TCP in soil was basically degraded [36]. Similarly, in this study, we found that, after applying chlorpyrifos to wheat, chlorpyrifos would degrade under the action of natural conditions such as temperature, light, microorganisms, and rainwater, and produce the 3,5,6-TCP metabolite, which is highly toxic to the environment, water quality, animals, and the human body. Although the residual concentration of 3,5,6-TCP was lower than that of the parent pesticide chlorpyrifos, 3,5,6-TCP is a water-soluble compound with stronger polarity, greater toxicity, and greater mobility and durability [9,37,38]. Therefore, in order to reduce the risk of exposing chlorpyrifos metabolites to the ecological environment, and

ensure the quality and safety of agricultural products, it is necessary to further clarify the metabolic pathways of chlorpyrifos in animals and plants.

Some studies have reported the metabolic pathway of chlorpyrifos in organisms. Racke (1993) found that chlorpyrifos was mainly metabolized into 3,5,6-TCP in the soil using isotope-labeled ^{36}Cl-chlorpyrifos, and further verified that 3,5,6-TCP was mineralized and forms carbon oxides as secondary metabolites using ^{14}C markers [34]. Lu et al. (2013) found that chlorpyrifos underwent hydrolysis under the action of strain DT-1 to produce 3,5,6-TCP and speculated that 3,5,6-TCP would then undergo three steps of dehydrogenation and produce secondary metabolites 2-hydroxypyridine, 5,6-dichloro-2-pyridinol and 6-chloro-2-pyridinol; lastly, the pyridine ring may be broken and degraded into small molecular compounds [39]. These research results are consistent with our predictions. After spraying wheat with chlorpyrifos, chlorpyrifos undergoes a nucleophilic addition substitution reaction under the action of organophosphorus hydrolase to generate the metabolite 3,5,6-TCP (Figure 5) [14,35]. With a continuous increase in OH concentration in the medium, the rate of hydrolysis accelerates, and the concentration of 3,5,6-TCP continues to increase. As time goes on, 3,5,6-TCP is affected by multiple factors such as its own properties, natural conditions, and crop characteristics, and may be further mineralized to produce small molecular compounds such as water, CO_2, and ammonium ions [40].

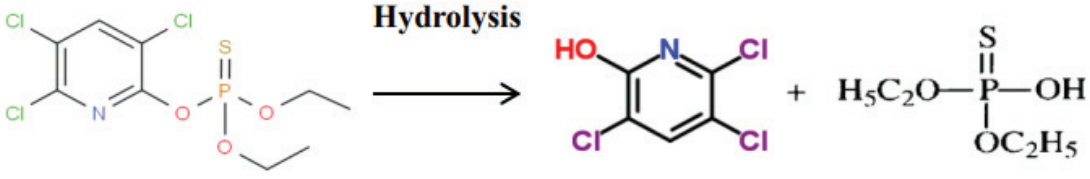


Figure 5. Hydrolyzed metabolic pathway of chlorpyrifos in wheat samples.

4. Conclusions

In conclusion, our study is the first to reveal the metabolic mechanism of chlorpyrifos during wheat growth in addition to the dissipation kinetics in wheat of the parent chlorpyrifos and its metabolite 3,5,6-TCP. The degradation dynamics of chlorpyrifos could be fitted by the first-order kinetic equation $C_t = C_0\, e^{-kt}$. The initial residues of chlorpyrifos in different organs of wheat were different at the early stage of application, which was in the order of leaves > wheat ears > stems. Chlorpyrifos was hydrolyzed to 3,5,6-TCP which was further dissipated. 3,5,6-TCP first showed an increasing trend and then a decreasing trend over time. In addition, it reached the maximum on the 3rd, 7th, or 11th day after application, and then gradually dissipated.

The primary metabolic pathway of chlorpyrifos was initially verified, but studies on the secondary product's metabolic pathway have not been further verified. Further research is still needed to reveal the metabolic pathway of chlorpyrifos in plants. Considering that 3,5,6-TCP is a biomarker with potential threats to humans and animals, the establishment of MRLs for 3,5,6-TCP should be considered during dietary exposure risk assessment.

Author Contributions: L.Y., J.L. (Jia Li), M.F., Z.J. and Q.T. conducted technical support on the pesticide detection, and L.Y., H.C., T.S. and L.Y., J.L. (Junhui Li) reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Research Foundation for Talented Scholars of Jinling Institute of Technology (no. JIT6202216), the Zhejiang Basic Public Welfare Research Program (no. LGF19C200001), the Zhejiang Provincial Natural Science Foundation of China (no. LQ20B070008), the Fundamental Research Funds for the Provincial Universities of Zhejiang, China (no. 2020YW34), the National Natural Science Foundation of China Project (no. 32101931), and the Fundamental Research Funds for the Provincial Universities of Zhejiang, China (no. 2021YW47).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in a publicly accessible repository The data presented in this study are openly available in [repository name e.g., FigShare] at [doi], reference number [reference number].

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

References

1. Bose, S.; Kumar, P.S.; Dai-Viet, N.V. A review on the microbial degradation of chlorpyrifos and its metabolite TCP. *Chemosphere* **2021**, *283*, 131447. [CrossRef] [PubMed]
2. Sud, D.; Kumar, J.; Kaur, P.; Bansal, P. Toxicity, natural and induced degradation of chlorpyrifos. *J. Chil. Chem. Soc.* **2020**, *65*, 4807–4816. [CrossRef]
3. Akbar, S.; Sultan, S. Soil bacteria showing a potential of chlorpyrifos degradation and plant growth enhancement. *Braz. J. Microbiol.* **2016**, *47*, 563–570. [CrossRef]
4. Sanchez-Hernandez, J.C.; Sandoval, M. Effects of chlorpyrifos on soil carboxylesterase activity at an aggregate-size scale. *Ecotoxicol. Environ. Saf.* **2017**, *142*, 303–311. [CrossRef] [PubMed]
5. Shen, Y. *Residue Characteristics and Degradation Mechanisms of Common Pesticides in Wheat [D]*; University of Yangzhou: Yangzhou, China, 2007; pp. 51–57.
6. Wu, Y.J.; Lv, C.Q.; Li, Y.; Chen, Y.P.; Yao, W.M.; Pan, Y.K.; Zhen, Y.K. Study on resides and degradation dynamics of chlorpyrifos in green bean. *Chin. J. Insp. Quar.* **2012**, *6*, 44–47.
7. Cao, L.; Xu, J.H.; Wu, G.; Li, M.M.; Jiang, J.D.; He, J.; Li, S.P.; Hong, Q. Identification of two combined genes responsible for dechlorination of 3,5,6-trichloro-2-pyridinol (TCP) in *Cupriavidus pauculus* P2. *J. Hazard. Mater.* **2013**, *260*, 700–706. [CrossRef]
8. Haque, M.A.; Hong, S.Y.; Hwang, C.E.; Kim, S.C.; Cho, K.M. Cloning of an organophosphorus hydrolase (opdD) gene of *Lactobacillus sakei* WCP904 isolated from chlorpyrifos-impregnated *kimchi* and hydrolysis activities of its gene product for organophosphorus pesticides. *Appl. Biol. Chem.* **2018**, *61*, 643–651. [CrossRef]
9. Watts, M. *Chlorpyrifos as a Possible Global Persistent Organic Pollutant*; Pesticide Action Network North America: Berkeley, CA, USA, 2012; pp. 1–34.
10. Furnes, B.; Schlenk, D. Extrahepatic metabolism of carbamate and organophosphate thioether compounds by the flavin-containing monooxygenase and cytochrome P450 systems. *Drug. Metab. Dispos. Biol. Fate. Chem.* **2005**, *33*, 214–218. [CrossRef]
11. Forsyth, C.S.; Chambers, J.E. Activation and degradation of the phosphorothionate insecticides parathion and EPN by rat brain. *Biochem. Pharmacol.* **1989**, *38*, 1597–1603. [CrossRef]
12. Zhao, Y.Y.; Wendling, L.A.; Wang, C.H.; Pei, Y.S. Behavior of chlorpyrifos and its major metabolite TCP (3,5,6-trichloro-2-pyridinol) in agricultural soils amended with drinking water treatment residuals. *J. Soils. Sediments* **2017**, *17*, 889–900. [CrossRef]
13. Sparling, D.W.; Fellers, G. Comparative toxicity of chlorpyrifos, diazinon, malathion and their oxon derivatives to larval Rana boylii. *Environ. Pollut.* **2007**, *147*, 535–539. [CrossRef] [PubMed]
14. Antonious, G.F.; Turley, E.T.; Abubakari, M.; Snyder, J.C. Dissipation, half-lives, and mass spectrometric identification of chlorpyrifos and its two metabolites on field-grown collard and kale. *J. Environ. Sci. Health. B* **2017**, *52*, 251–255. [CrossRef] [PubMed]
15. Briceño, G.; Fuentes, M.S.; Palma, G.; Jorquera, M.A.; Amoroso, M.J.; Diez, M.C. Chlorpyrifos biodegradation and 3,5,6-trichloro-2-pyridinol production by actinobacteria isolated from soil. *Int. Biodeterior. Biodegrad.* **2012**, *73*, 1–7. [CrossRef]
16. Das, S.; Adhya, T.K. Degradation of chlorpyrifos in tropical rice soils. *J. Environ. Manag.* **2015**, *152*, 36–42. [CrossRef] [PubMed]
17. Nabil, S.A.; Hussein, H.S.; Salah, E.M.A.A.; Reda, A.B. Isolation, characterization and fingerprinting of some chlorpyrifos-degrading bacterial strains isolated from Egyptian pesticides-polluted soils. *Afr. J. Microbiol. Res.* **2011**, *5*, 2855–2862. [CrossRef]
18. Saito-Shida, S.; Nemoto, S.; Teshima, R.; Akiyama, H. Quantitative analysis of pesticide residues in vegetables and fruits by liquid chromatography quadrupole time-of-flight mass spectrometry. *Food. Addit. Contam. Part. A.* **2016**, *33*, 119–127. [CrossRef]
19. Yang, X.F.; Luo, J.H.; Duan, Y.; Li, S.H.; Liu, C.H. Simultaneous analysis of multiple pesticide residues in minor fruits by ultrahigh-performance liquid chromatography/hybrid quadrupole time-of-fight mass spectrometry. *Food Chem.* **2018**, *241*, 188–198. [CrossRef]
20. Liu, X.; Fan, Y.; Chang, C.P.; Lo, C.K.; Wang, X. High-throughput screening and quantification of pesticides in paprika by UHPLC-Q-TOF/MS. *Food. Anal. Methods* **2021**, *14*, 2186–2198. [CrossRef]
21. Codex Alimentarius Commission. Codex Pesticide Residues Limits in Food and Feed Database. 2003. Available online: https://www.fao.org/fao-who-codexalimentarius/codex-texts/dbs/pestres/pesticide-detail/en/?p_id=17 (accessed on 12 June 2021).
22. United States Environmental Protection Agency. Office of Pesticide Programs. Index to Pesticide Chemical Names, Part 180 Tolerance Information, and Food and Feed Commodities (by Commodity). 2012. Available online: https://www.epa.gov/sites/default/files/2015-01/documents/tolerances-commodity.pdf (accessed on 20 November 2021).
23. Japan Food Chemical Research Foundation. The Japanese Positive List System for Agricultural Chemical Residues in Foods. 2021. Available online: http://db.ffcr.or.jp/front/pesticide_detail?id=21800 (accessed on 25 April 2022).
24. National Standard of the People's Republic of China (NSPRC). *GB 2763-2021*; Maximum Residue Limits for Pesticides in Food. China Standard Publishing House: Beijing, China, 2016; pp. 95–97.

25. López, M.G.; Fussell, R.J.; Stead, S.L.; Roberts, D.; McCullagh, M.; Rao, R. Evaluation and validation of an accurate mass screening method for the analysis of pesticides in fruits and vegetables using liquid chromatography-quadrupole-time of flight-mass spectrometry with automated detection. *J. Chromatogr. A.* **2014**, *1373*, 40–50. [CrossRef]
26. Bauer, A.; Luetjohann, J.; Hanschen, F.S.; Schreiner, M.; Kuballa, J.; Jantzen, E.; Rohn, S. Identification and characterization of pesticide metabolites in brassica species by liquid chromatography travelling wave ion mobility quadrupole time-of-flight mass spectrometry (UPLC-TWIMS-QTOF-MS). *Food Chem.* **2018**, *244*, 292–303. [CrossRef]
27. Sánchez-Hernández, L.; Hernández-Domínguez, D.; Martín, M.T.; Nozal, M.J.; Higes, M.; Bernal Yagüe, J.L. Residues of neonicotinoids and their metabolites in honey and pollen from sunflower and maize seed dressing crops. *J. Chromatogr. A.* **2016**, *1428*, 220–227. [CrossRef] [PubMed]
28. González-Curbelo, M.Á.; Socas-Rodríguez, B.; Herrero, M.; Herrera-Herrera, A.V.; Hernández-Borges, J. Dissipation kinetics of organophosphorus pesticides in milled toasted maize and wheat flour (gofio) during storage. *Food Chem.* **2017**, *229*, 854–859. [CrossRef]
29. Kong, Z.Q.; Li, M.M.; Chen, J.Y.; Gui, Y.J.; Bao, Y.M.; Fan, B.; Jian, Q.; Francis, F.; Dai, X.F. Behavior of field-applied triadimefon, malathion, dichlorvos, and their main metabolites during barley storage and beer processing. *Food Chem.* **2016**, *211*, 679–686. [CrossRef] [PubMed]
30. Zhang, X.; Shen, Y.; Yu, X.Y.; Liu, X.J. Dissipation of chlorpyrifos and residue analysis in rice, soil and water under paddy field conditions. *Ecotoxicol. Environ. Saf.* **2012**, *78*, 276–280. [CrossRef]
31. Mahdavi, V.; Farimani, M.M.; Fathi, F.; Ghassempour, A. A targeted metabolomics approach toward understanding metabolic variations in rice under pesticide stress. *Anal. Biochem.* **2015**, *478*, 65–72. [CrossRef] [PubMed]
32. Zhao, H.Z.; Jiang, M.; Xie, C.F.; Song, S.Y.; Wang, J.; Bai, G.; Luo, G.A. Metabolic fingerprinting of acs7 mutant and wild-type arabidopsis thaliana under salt stress by ultra performance liquid chromatography coupled with quadrupole/time of flight mass spectrometry. *Anal. Lett.* **2012**, *45*, 1786–1798. [CrossRef]
33. Regueiro, J.; López-Fernández, O.; Rial-Otero, R.; Cancho-Grande, B.; Simal-Gándara, J. A review on the fermentation of foods and the residues of pesticides—Biotransformation of pesticides and effects on fermentation and food quality. *Crit. Rev. Food. Sci. Nutr.* **2015**, *55*, 839–863. [CrossRef]
34. Racke, K.D. Environmental fate of chlorpyrifos. *Rev. Environ. Contam. Toxicol.* **1993**, *131*, 1–150.
35. Zhang, Z.Y.; Jiang, W.W.; Jian, Q.; Song, W.C.; Zheng, Z.T.; Wang, D.L.; Liu, X.J. Changes of field incurred chlorpyrifos and its toxic metabolite residues in rice during food processing from-RAC-to-consumption. *PLoS ONE* **2015**, *10*, e0116467. [CrossRef]
36. Chai, L.K.; Mohd-Tahir, N.; Bruun Hansen, H.C. Dissipation of acephate, chlorpyrifos, cypermethrin and their metabolites in a humid-tropical vegetable production system. *Pest. Manage. Sci.* **2009**, *65*, 189–196. [CrossRef]
37. Abraham, J.; Silambarasan, S. Biodegradation of chlorpyrifos and its hydrolyzing metabolite 3,5,6-trichloro-2-pyridinol by *Sphingobacterium* sp. JAS3. *Process. Biochem.* **2013**, *48*, 1559–1564. [CrossRef]
38. Rathi, B.S.; Kumar, P.S. Application of adsorption process for effective removal of emerging contaminants from water and wastewater. *Environ. Pollut.* **2021**, *280*, 116995. [CrossRef] [PubMed]
39. Lu, P.; Li, Q.F.; Liu, H.M.; Feng, Z.Z.; Yan, X.; Hong, Q.; Li, S.P. Biodegradation of chlorpyrifos and 3,5,6-trichloro-2-pyridinol by *Cupriavidus* sp. DT-1. *Bioresour. Technol.* **2013**, *127*, 337–342. [CrossRef] [PubMed]
40. Feng, Y.C.; Minard, R.D.; Bollag, J.M. Microbial and photolytic degradation of 3,5,6-trichloro-2-pyridinol. *Environ. Toxicol. Chem.* **2010**, *17*, 814–819. [CrossRef]

metabolites

MDPI

Article

Transcription and Metabolism Pathways of Anthocyanin in Purple Shamrock (*Oxalis triangularis* A.St.-Hil.)

Baobing Luo [1], Liujun Chen [1], Guoping Chen [1], Yunshu Wang [1], Qiaoli Xie [1], Xuqing Chen [2,*] and Zongli Hu [1,*]

[1] Laboratory of Molecular Biology of Tomato, Bioengineering College, Chongqing University, Chongqing 400044, China
[2] Institute of Grassland, Flowers and Ecology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
* Correspondence: chenxuqing@baafs.net.cn (X.C.); huzongli71@163.com (Z.H.); Tel.: +86-1051503868 (X.C.); +86-13996265017 (Z.H.)

Abstract: Anthocyanins are water-soluble pigments that can impart various colors to plants. Purple shamrock (*Oxalis triangularis*) possesses unique ornamental value due to its purple leaves. In this study, three anthocyanins, including malvidin 3-O-(4-O-(6-O-malonyl-glucopyranoside)-rhamnopyranosyl)-5-O-(6-O-malonyl-glucopyranoside), delphinidin-3-O-rutinoside and malvidin-3,5-di-O-glucoside, were characterized with ultra-performance liquid chromatography-electrospray ionization tandem mass spectrometry (UPLC-ESI-MS/MS) in purple shamrock. To investigate the molecular mechanism of anthocyanin biosynthesis in green shamrock (*Oxalis corymbosa*) and purple shamrock, RNA-seq and qRT-PCR were performed, and the results showed that most of the anthocyanin biosynthetic and regulatory genes were up-regulated in purple shamrock. Then, dark treatment and low temperature treatment experiments in purple shamrock showed that both light and low temperature can induce the biosynthesis of anthocyanins.

Keywords: anthocyanin biosynthesis; *Oxalis triangularis*; UPLC-ESI-MS/MS; RNA-seq; stress resistance

Citation: Luo, B.; Chen, L.; Chen, G.; Wang, Y.; Xie, Q.; Chen, X.; Hu, Z. Transcription and Metabolism Pathways of Anthocyanin in Purple Shamrock (*Oxalis triangularis* A.St.-Hil.). *Metabolites* **2022**, *12*, 1290. https://doi.org/10.3390/metabo12121290

Academic Editor: Marijana Zovko Končić

Received: 28 November 2022
Accepted: 12 December 2022
Published: 19 December 2022

1. Introduction

Anthocyanins, a group of water-soluble pigments, are widely spread in higher plants. They are flavonoid compounds derived from the phenylpropanoid biosynthesis pathway [1,2]. While flavonoids, as the most important secondary metabolites in plants, can produce the colors of purple, blue and red in flowers, leaves and fruits [3]. Anthocyanins confer a variety of colors to plants, which contributes to the completion of pollination and seed dispersal via attracting insects or animals [4]. In addition, they also play important roles in resisting biotic and abiotic stress [4], such as cold, drought, salt and UV irradiation. Furthermore, more and more evidence show that intaking foods rich in anthocyanins can effectively reduce the risk of suffering from obesity [4], cancers [5,6], and cardiovascular disease [7,8]. For ornamental plants, cultivars that contain an amount of anthocyanins are more eye-catching than others.

So far, the anthocyanin biosynthetic pathway has been studied extensively in higher plants, including tree peony (*Paeonia suffruticosa Andrews*) [9], tomato (*Solanum lycopersicum* L.) [10], rugosa rose (*Rosa rugosa* Thunb.) [11], and others. Anthocyanin biosynthesis is classified as a branch of the phenylpropanoid pathway, which requires the participation of multiple enzymes. The genes encoding these enzymes can be divided into three categories, namely beginning, early, and late biosynthetic genes [12,13]. The beginning biosynthetic genes consist of PAL, C4H, and 4CL, which are in the phenylpropanoid pathway. The early biosynthetic genes (EBGs), CHS, CHI and F3H, are involved in a common flavonoid pathway. While the late biosynthetic genes (LBGs), including F3′H, F3′5′H, DFR, ANS and UFGT, lead to the production of the different flavonoid pigments in many plants.

Over the past years, many transcription factors (TFs), including MYBs, bHLHs and WDRs, were identified in many plant species [14–16]. R2R3-MYBs have been shown to be involved in the anthocyanin biosynthetic pathway via activating the enzyme genes in many plants. Previous studies showed that these TFs can form MYB-bHLH-WD40 (MBW) complex binding to the promoter of enzyme genes and promoting the transcription of these genes [1,17,18]. Recent reports indicated that WRKY TFs are also involved in the regulation of anthocyanin biosynthesis [1,12,19]. In this study, genes coding for six TFs, including MYB113, TT8, TTG1, TTG2, GL3 and CPC, were characterized. Among these genes, CPC, a MYB TF, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis thaliana* L. [20,21].

Environmental conditions, such as light and temperature, also affect anthocyanin biosynthesis in many plants. Light could induce the expression of light-responsive genes, such as HY5 and COP1, which regulate the expression of MYB TF genes that are related to anthocyanin biosynthesis [22,23]. Similarly, low temperature can also induce the expression of low temperature response factor genes to affect the anthocyanin levels in plants. For example, CBFs can interact with MYB113 to regulate anthocyanin biosynthesis in eggplant (*Solanum melongena* L.) [24]. In kiwifruit (*Actinidia chinensis* Planch.), light and low temperature both can induce the expression of AcMYB10, which leads to the increase of anthocyanin content [15]. In contrast, anthocyanin accumulation in kiwifruit decreased under high temperature conditions [25]. An exposure to light induced PpbHLH64 expression and anthocyanin accumulation in pear fruit (*Pyrus pyrifolia* Nakai) [26]. In purple head Chinese cabbage (*Brassica rapa* L.), BrMYB2, and BrTT8 were highly up-regulated after low temperature treatment, which promotes the biosynthesis of anthocyanins [27].

Oxalis corymbosa DC., the green shamrock, belongs to the family Oxalidaceae and is widely distributed all over the world. Groom et al. [28] demonstrated that *Oxalis corniculata* probably comes from southeastern Asia. In addition, *Oxalis corniculata* is regarded as an herb with antibacterial, antifungal and anticancer potentials [29–31] and has the third largest distribution in vascular plants [32]. *Oxalis triangularis* A. St.-Hil., also called purple shamrock or false shamrock [33], is an important traditional ornamental and medicinal plant. As an ornamental plant, the purple leaves of this plant are the most outstanding feature. The trifoliate leaves are subdivided into three obtriangular to obovate-triangular leaflets and resemble a clover in shape. The leaves of purple shamrock move in response to light levels, opening in high ambient light (in the day) and closing at low light levels (at night).

So far, the molecular mechanisms underlying the different leaf colors between purple and green shamrocks are still unclear. In this study, the anthocyanins in purple shamrock were characterized with ultra-performance liquid chromatography-electrospray ionization tandem mass spectrometry (UPLC-ESI-MS/MS). The anthocyanin biosynthetic and regulatory genes were analyzed by real-time quantitative reverse transcription PCR (qRT-PCR) in the two species of shamrocks. Meanwhile, the expression levels of anthocyanin biosynthetic and regulatory genes were further determined in both species under dark/light conditions. RNA-seq was also used to analyze the differentially expressed genes (DEGs) in the leaves of green and purple shamrocks. Based on the above, the biosynthetic pathway of anthocyanins in purple shamrock was summarized (Figure 1). These results further our understanding about the molecular mechanisms of anthocyanin biosynthesis in purple shamrock.

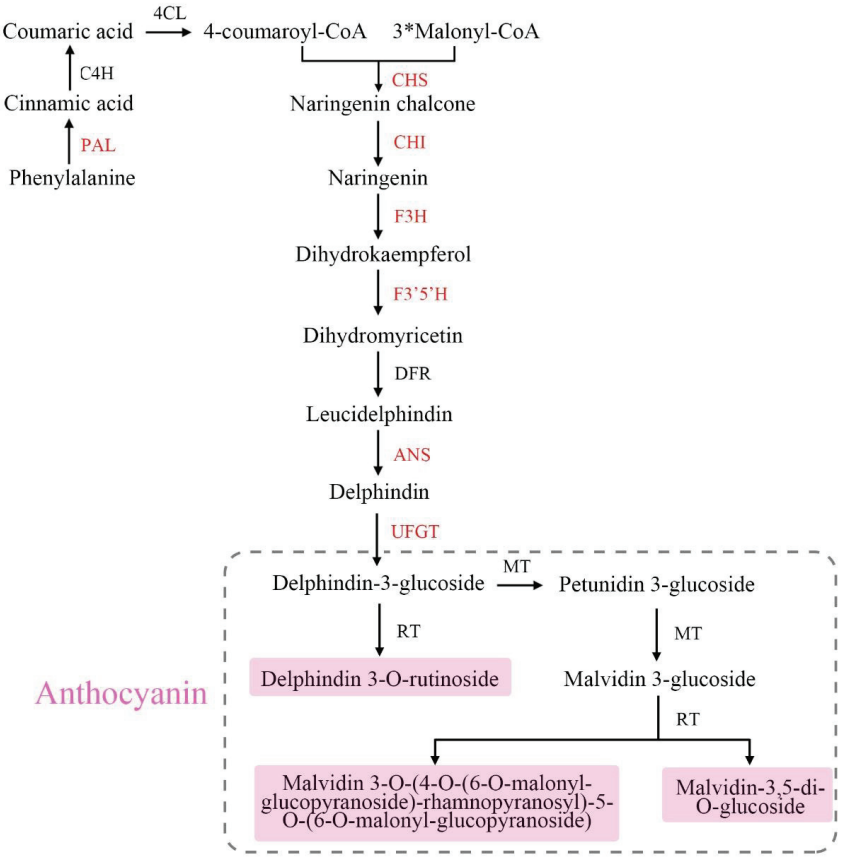


Figure 1. Anthocyanin biosynthetic pathway in purple shamrock. Abbreviations: PAL, phenylalanine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarateCoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′5′-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; UFGT, flavonoid 3-O-glucosyltransferase; RT, rhamnosyltransferase; MT, methyltransferase.

2. Materials and Methods

2.1. Plant Materials and Growth Conditions

The bulbs of green shamrock (*Oxalis corymbosa* DC.) and purple shamrock (*Oxalis triangularis* A.St.-Hil.) were used as the original experimental materials. The two plants reached maturity (Figure 2A) after 30 days of culture under the following conditions: 28 °C/light for 16 h (intensity of 250 μmol m-2 s-1), 18 °C/dark for 8 h, 80% relative humidity (RH). The mature leaves (Figure 2B) of green and purple shamrock were used for UPLC-ESI-MS/MS analysis. The samples used for qRT-PCR analysis were collected from mature leaves and flowers of green and purple shamrocks. For dark treatment, the purple shamrocks plants were grown in the dark (28 °C/dark for 16 h, 18 °C/dark for 8 h, 80% RH.)/light (28 °C/light for 16 h, 18 °C/dark for 8 h, 80% RH.) conditions, respectively, for 30 days. For low temperature treatment, the purple shamrocks were cultivated in normal (28 °C/light for 16 h, 18 °C/dark for 8 h, 80% RH.)/low (4 °C/light for 16 h, 4 °C/dark for 8 h, 80% RH.) temperature conditions for 0, 3, 6, and 12 days, respectively. All samples were frozen in liquid nitrogen and stored at −80 °C until further analysis.

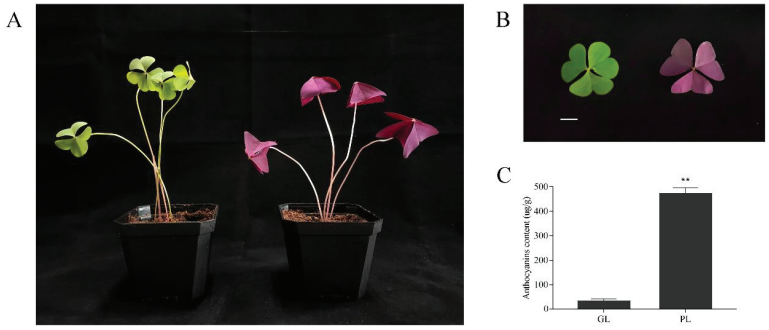


Figure 2. (**A**) Phenotypes of green and purple shamrocks. (**B**) The leaf of green and purple shamrocks. Bar = 1 cm. (**C**) Total contents of anthocyanins in leaves of green and purple shamrocks. Abbreviations: GL, leaves of green shamrock; PL, leaves of purple shamrock. ** indicates a significant difference ($p < 0.01$).

2.2. Extraction and Total Concentration of Anthocyanins

The method of Rapisarda et al. [34] was used to determine the total anthocyanin content in plant tissues. Frozen samples (100 mg) were ground into powder in liquid nitrogen and extracted separately with 1 mL of pH 1.0 buffer (50 mM KCl and 150 mM HCl) and 1 mL of pH 4.5 buffer (400 mM CH3COONa and 240 mM HCl). Then, the mixtures were centrifuged at 14,000× *g* for 5 min at 4 °C. The supernatants were gathered for measurement of absorbance at 510 nm. Concentration of anthocyanins was calculated using the following equation:

$$\text{concentration } (\mu g\ g^{-1}\ FW) = (A1 - A2) \times 484.8 \times 1000/24{,}825 \times \text{dilution factor}$$

A1 represents the absorbance of supernatants gathered from pH 1.0 buffer solution at 510 nm, while A2 represents the absorbance of supernatants gathered from pH 4.5 buffer solution at 510 nm. The value 484.8 represents the molecular mass of cyanidin-3-glucoside chloride, while 24,825 (L g^{-1} cm^{-1}) reflects its molar absorption coefficient at 510 nm in the pH 1.0 solution. Three biological replicates were used for total anthocyanin extraction, and each sample was from a different plant.

2.3. UPLC and ESI-MS/MS Analysis of Anthocyanins

The sample extracts were analyzed using an UPLC-ESI-MS/MS system (UPLC, Shimpack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 4500 Q TRAP). The analytical conditions were as follows, UPLC: column, Agilent SB-C18 (1.8 μm, 2.1 mm × 100 mm); The mobile phase consisted of solvent A, pure water with 0.1% formic acid, and solvent B, acetonitrile. Sample measurements were performed with a gradient program that employed the starting conditions of 95% A, 5% B. Within 9 min, a linear gradient to 5% A, 95% B was programmed, and a composition of 5% A, 95% B was kept for 1 min. Subsequently, a composition of 95% A, 5.0% B was adjusted within 1.1 min and kept for 2.9 min. The column oven was set to 40 °C; the injection volume was 4 μL. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-M.

The ESI source operation parameters were as follows: ion source, turbo spray; source temperature 550 °C; ion spray voltage (IS) 5500 V (positive ion mode)/−4500 V (negative ion mode); ion source gas I (GSI), gas II (GSII), curtain gas (CUR) was set at 50, 60, and 30.0 psi, respectively; the collision gas (CAD) was high. Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM experiments with collision gas (nitrogen) set to 5 psi. DP and CE for individual MRM transitions was performed with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period. Three

biological replicates were used for UPLC and ESI-MS/MS analysis, and each sample was from a different plant.

2.4. RNA Extraction and qRT-PCR Analysis

The frozen samples were crushed into powder in liquid nitrogen. Total RNA was isolated from the samples with RNAiso (Takara, Otsu, Japan), and 1–2 μL of total RNA was used to synthesize the complementary DNA (cDNA) with reverse transcriptase (Promega, Beijing, China) and an oligo (dT)20 primer. The qRT-PCR was carried out using the CFX96 real-time system (Bio-Rad, Hercules, CA, USA). Three biological replicates were used for qRT-PCR analysis, and each sample was from a different plant. The primers used for qRT-PCR analysis were designed by Primer Premier 5 (Table S1).

2.5. Library Preparation and Transcriptome Sequencing

Total RNA was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Waltham, MA, USA) following the manufacturer's recommendations. Briefly, mRNA was purified from total RNA using oligo (dT)-attached magnetic beads. The first strand cDNA was synthesized using random hexamer primer and M-MLV Reverse Transcriptase. Second strand cDNA synthesis was subsequently performed using DNA Polymerase I. The cDNA fragments were end-repaired and ligated to NEBNext adaptor. Then, the library fragments were purified with AMPure XP system to select cDNA fragments of preferentially 250–300 bp in length, and PCR was performed to enrich the cDNA fragments. Finally, the library preparations were sequenced on an Illumina Hiseq platform and paired-end reads were generated. Three biological replicates were used for RNA-seq analysis, and each sample was from a different plant.

2.6. Establishment of Local mRNA and Protein Database

Since there is no genome information of *Oxalis triangularis* online, the result of RNA-seq analysis was used to establish the local mRNA and protein database for further research. BioEdit, a sequence analysis software, was used to build the database by importing mRNA and protein sequence information. Based on a BLASTP search using the amino acid sequences of anthocyanin biosynthetic genes in Arabidopsis as queries, single orthologs for PAL, CHS, CHI, F3H, F3′5′H, ANS, and UFGT were obtained from purple shamrock contigs. In addition, the regulatory genes, including MYB113, TT8, TTG1, TTG2, GL3 and CPC, were also identified.

2.7. Differential Expression Analysis

Differential expression analysis of the two groups was performed using the DESeq R package (1.10.1). DESeq provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate [35]. Genes with an adjusted p-value < 0.05 found by DESeq were assigned as differentially expressed.

2.8. Gene Functional Annotation and Enrichment Analysis

To acquire comprehensive gene function information, seven databases, including Nr (NCBI non-redundant protein sequences), Nt (NCBI non-redundant nucleotide sequences), Pfam (Protein family), KOG/COG (Clusters of Orthologous Groups of proteins), Swiss-Prot (A manually annotated and reviewed protein sequence database), KO (KEGG Ortholog database) and GO (Gene Ontology), were used to annotate the assembled unigenes. Expression level of each transcript was estimated by FPKM (fragments per kilobase of transcript per million fragments mapped) method [36].

3. Results

3.1. Identification of Anthocyanins in Green and Purple shamrocks

By analyzing the extracts from leaves of green and purple shamrocks with the method of UPLC-ESI-MS/MS, three anthocyanins were separated and identified (Figure 3). As shown in Table 1, three anthocyanins, including malvidin 3-O-(4-O-(6-O-malonyl-glucopyranoside)-rhamnopyranosyl)-5-O-(6-O-malonyl-glucopyranoside), delphinidin-3-O-rutinoside and malvidin-3,5-di-O-glucoside, were characterized. Through KEGG enrichment analysis, these three anthocyanins are derivatives of delphinidin-3-O-glucoside that is the main anthocyanin for purple plants (Figure S1). To further verify the effect of anthocyanins on purple shamrock's phenotype, total content of anthocyanins was determined. As expected, the total content of anthocyanins in purple shamrock was about 15 times that of green shamrock (Figure 2C).

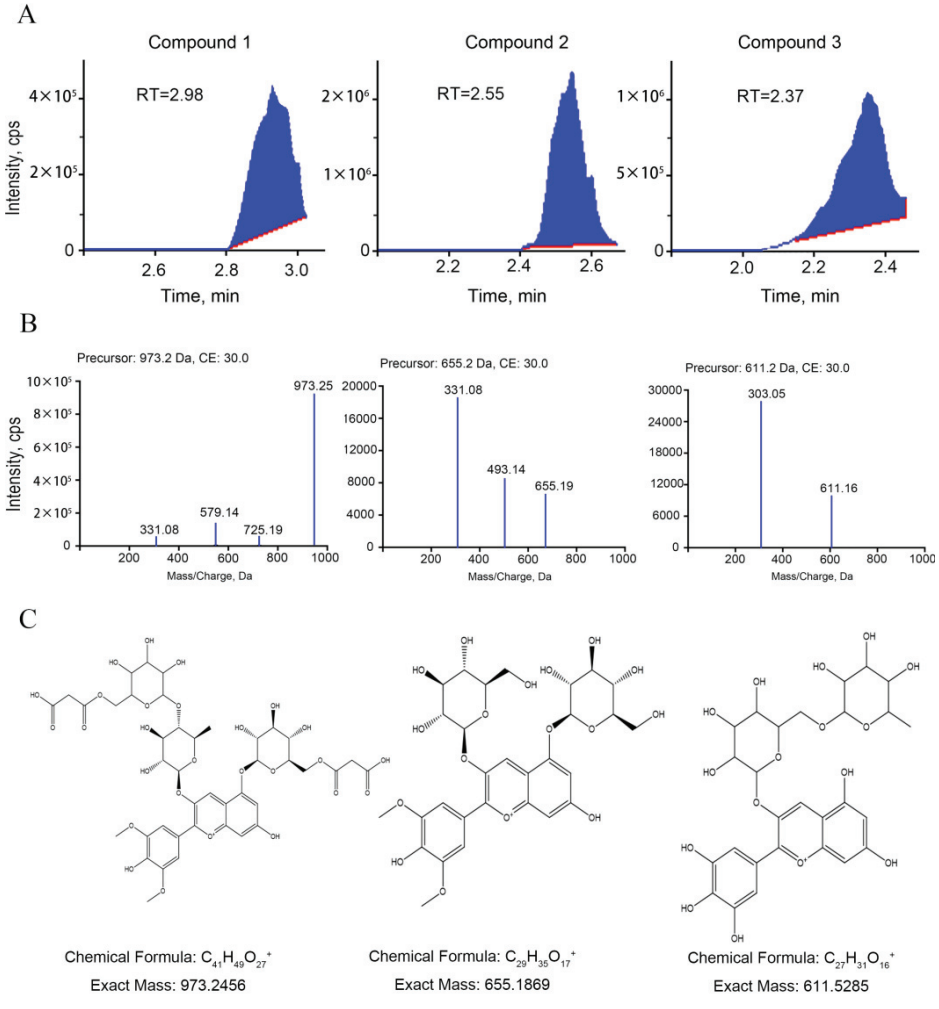


Figure 3. UPLC and ESI-MS/MS Analysis of Anthocyanins. (**A**) Chromatograms of three anthocyanins detected in green and purple shamrocks. (**B**) Mass spectrogram of three anthocyanins. (**C**) Structural formulas and chemical formulas of three anthocyanins.

Table 1. Anthocyanin Contents in Leaves of Green and Purple Shamrock.

NO.	Compound	Sample	
		Green	Purple
1	Malvidin 3-O-(4-O-(6-O-malonyl-glucopyranoside)-rhamnopyranosyl)-5-O-(6-O-malonyl-glucopyranoside)	9	4,290,000
2	malvidin-3,5-di-O-glucoside (Malvin)	856,000	22,200,000
3	delphinidin-3-O-rutinoside	416,000	15,000,000

Note: the values in the table represent the area of the chromatographic peak (cps min).

3.2. Transcriptome Analysis of Green and Purple Shamrocks by RNA-seq

The mature leaf samples of green and purple shamrocks were collected for RNA-seq analysis to investigate the molecular mechanisms of anthocyanin accumulation in purple leaves. To identify the differentially expressed genes (DEGs) involved in purple leaf coloration, the fragment per kilobase of exon per million fragments mapped (FPKM) values were analyzed for each gene in leaves of green and purple shamrocks.

As shown in the volcanic plot map, a total of 39,356 DEGs were upregulated and 19,638 DEGs were downregulated (Figure 4A). To classify the genes involved in the anthocyanin biosynthetic pathway, all DEGs were subjected to KEGG database. According to the different functions, the DEGs were divided into five categories and thirty-four subcategories. The five categories including cellular processes, environmental information processing, genetic information processing, and metabolism and organismal systems, and the DEGs mainly belong to metabolism. (Figure 4B). For further research, GO enrichment was performed. All of the DEGs based on their functions were divided into three main categories, including four biological processes (BP) terms; 10 cellular components (CC) terms and 15 molecular function (MF) terms. For CC, the majority of DEGs were associated with nuclear and membrane-bound organelle functions. For MF, most of DEGs were related to binding and catalytic activity (Figure 4C). Above all, although there are many DEGs between green and purple shamrocks, only a small part of them are involved in anthocyanin biosynthesis (Supporting Information 1, Supporting Information 2 and Figure S5).

3.3. Expression of Anthocyanin Biosynthetic and Regulatory Genes in Green and Purple Shamrocks

Through qRT-PCR technology, the expression levels of anthocyanin biosynthetic genes PAL, CHS, CHI, F3H, F3′5′H, ANS and UFGT were analyzed. Compared with the green shamrock, all the anthocyanin pathway genes were significantly upregulated in purple shamrock, which is consistent with the results of RNA-seq (Figure S2). Moreover, the expression levels of CHS and CHI were especially increased at least 700-fold in the purple as compared with the green shamrock (Figure 5A). Meanwhile, the transcripts of some anthocyanin biosynthesis regulatory orthologous genes of Arabidopsis, MYB113, TT8, TTG1, TTG2, GL3 and CPC, were detected in shamrocks. As shown in Figure 5B, all the regulatory genes were upregulated in purple shamrock, especially MYB113 and TT8. These results indicated that the expression of anthocyanin pathway genes in purple shamrock is activated.

3.4. Effects of Dark Treatment on Anthocyanin Accumulation in Leaves of Purple Shamrock

According to previous reports [22,23], light could induce anthocyanin accumulation in plants. In this study, purple shamrocks were cultured for 30 days under light and dark conditions. After 30 days of dark treatment, the leaf area of purple shamrocks was much smaller than that grown under normal light, and the color changed to pale pink (Figure S4). In addition, the stem became slender and easy to lodge. The anthocyanin content of purple shamrock leaves grown under dark condition was significantly less than that under light condition (Figure 6A). Further, the expression of anthocyanin biosynthetic and regulatory genes was analyzed by qRT-PCR. As shown in Figure 6B, most of the anthocyanin biosynthetic genes in purple shamrock under dark condition were downregulated compared to that under light condition except PAL and CHI. Moreover, the expression level of CHI was

significantly up-regulated after dark treatment, indicating that expression of OtCHI may not be induced by light. Meanwhile, MYB113 and TT8, the key regulators of anthocyanin biosynthesis, were downregulated under dark condition (Figure 6C), although there are no obvious changes in other regulatory genes. These results showed that dark treatment inhibits the expression of some anthocyanin biosynthetic and regulatory genes, thereby reducing the accumulation of anthocyanins in purple shamrock.

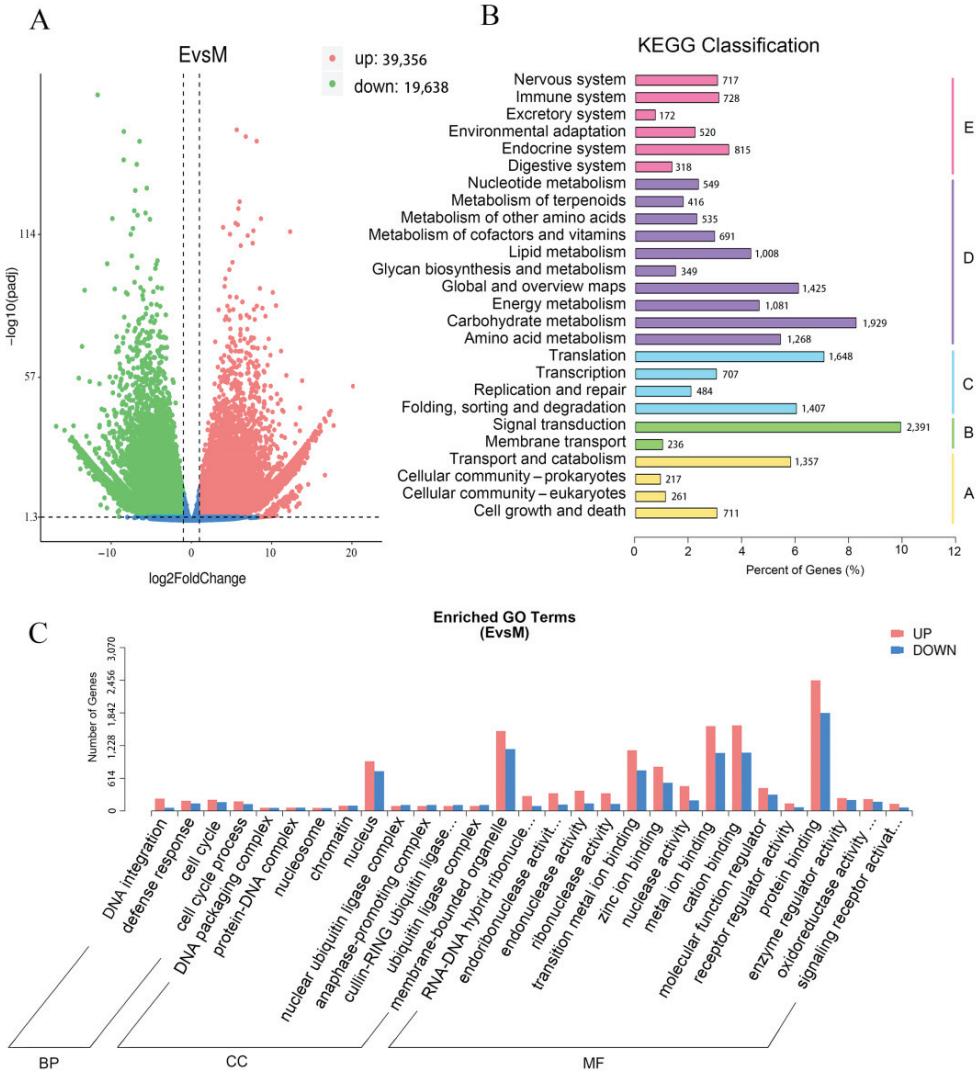


Figure 4. DEGs, KEGG classification and GO Enrichment analysis. (**A**) Volcano plot of differentially expressed genes. (**B**) KEGG classification of assembled differentially expressed genes. Abbreviations: A, cellular processes; B, environmental information processing; C, genetic information processing; D, metabolism; E, organismal systems. (**C**) GO Enrichment of assembled differentially expressed genes. Abbreviations: BP, biological processes; CC, cellular components; MF, molecular functions.

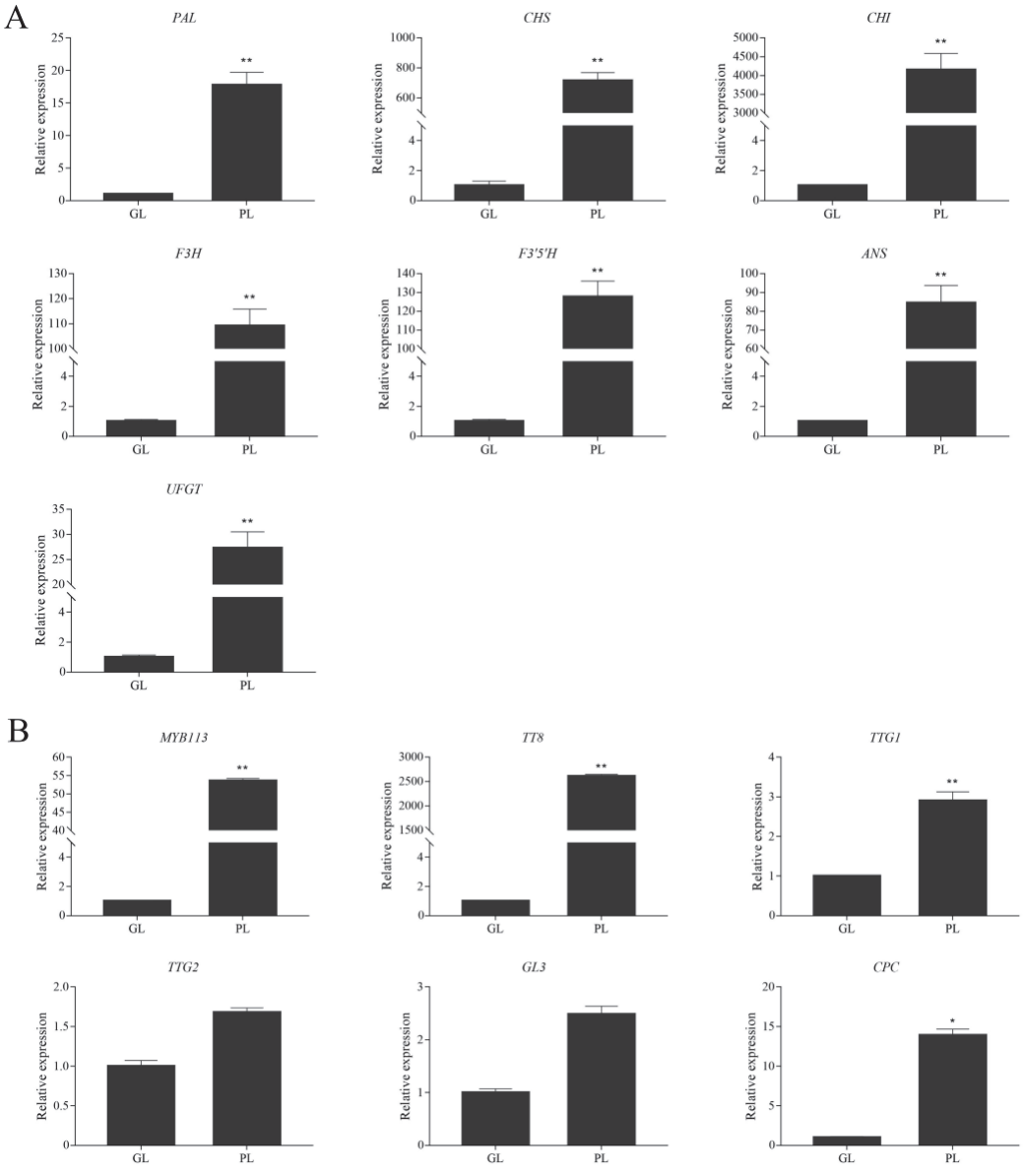


Figure 5. Relative expression level analysis of anthocyanin biosynthetic and regulatory genes in leaves of green and purple shamrocks. (**A**) Relative expression level analysis of anthocyanin biosynthetic genes. (**B**) Relative expression level analysis of regulatory genes. Abbreviations: GL, leaves of green shamrock; PL, leaves of purple shamrock. Error bars represent the standard error of the mean ($n = 3$). Significant differences (* $p < 0.05$, ** $p < 0.01$).

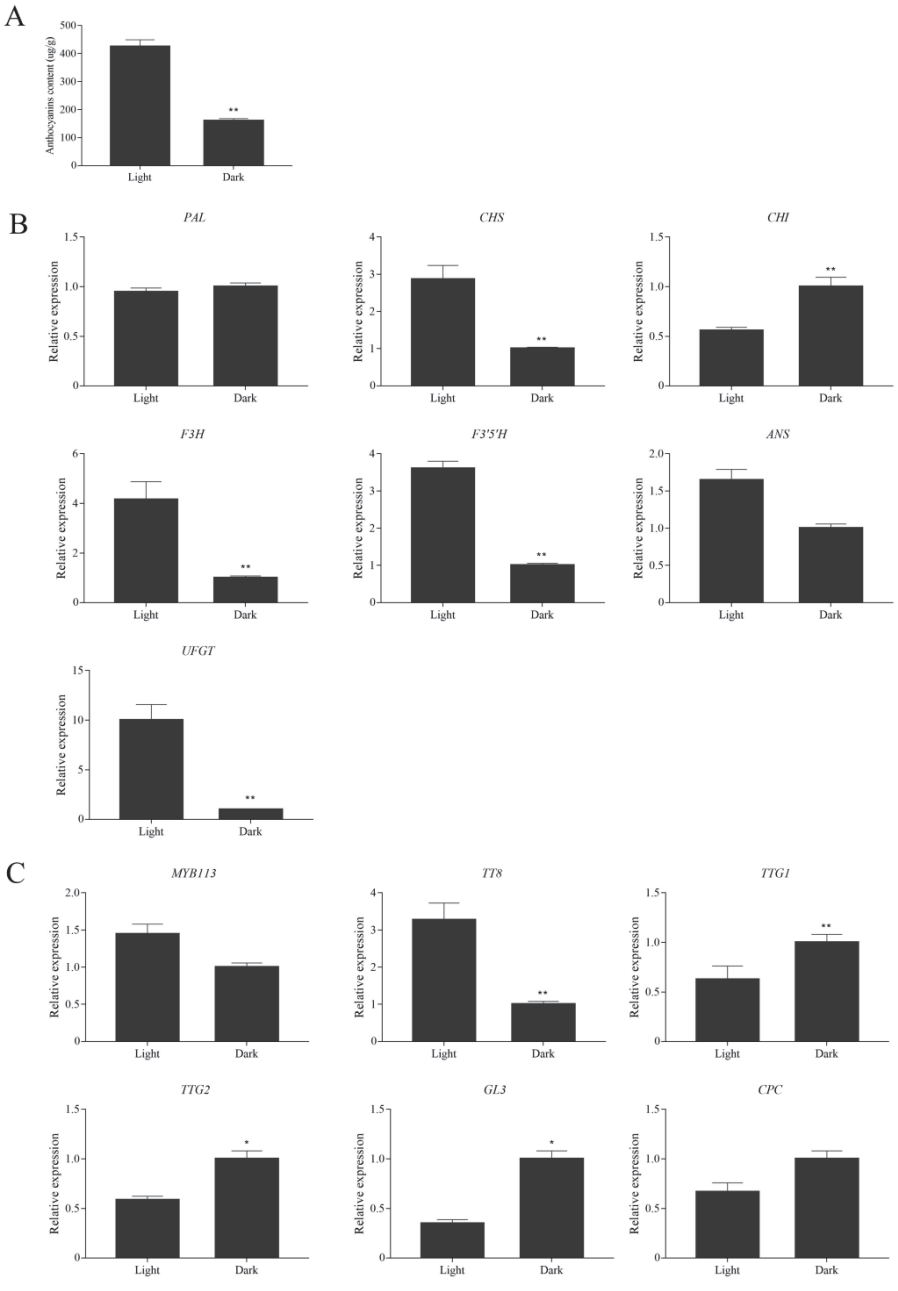


Figure 6. Relative expression level analysis of anthocyanin biosynthetic and regulatory genes in purple shamrock under light/dark conditions. (**A**) Total contents of anthocyanins in light/dark conditions of purple shamrock. (**B**) Relative expression level analysis of anthocyanin biosynthetic genes. (**C**) Relative expression level analysis of regulatory genes. Error bars represent the standard error of the mean ($n = 3$). Significant differences (* $p < 0.05$, ** $p < 0.01$).

3.5. Effect of Low Temperature Treatment on Anthocyanin Accumulation in Leaves of Purple Shamrock

Previous studies showed that temperature can affect the biosynthesis of anthocyanins [27,37]. In this study, purple shamrocks were cultured at 4 °C and normal temperature for 0, 3, 6, and 12 days. After 12 days of low temperature treatment, the color of the purple shamrock leaves is a little darker than that growth under normal temperature, and the anthocyanin content of leaves after low temperature treatment is higher (Figure S3). To further reveal the molecular mechanism of the effect of low temperature on anthocyanin biosynthesis, the expression levels of anthocyanin biosynthetic and regulatory genes were analyzed by qRT-PCR. The results showed that the transcription level of CHS in low temperatures was significantly higher than that in normal temperatures, but there were no significant differences in other anthocyanin biosynthetic genes (Figure 7).

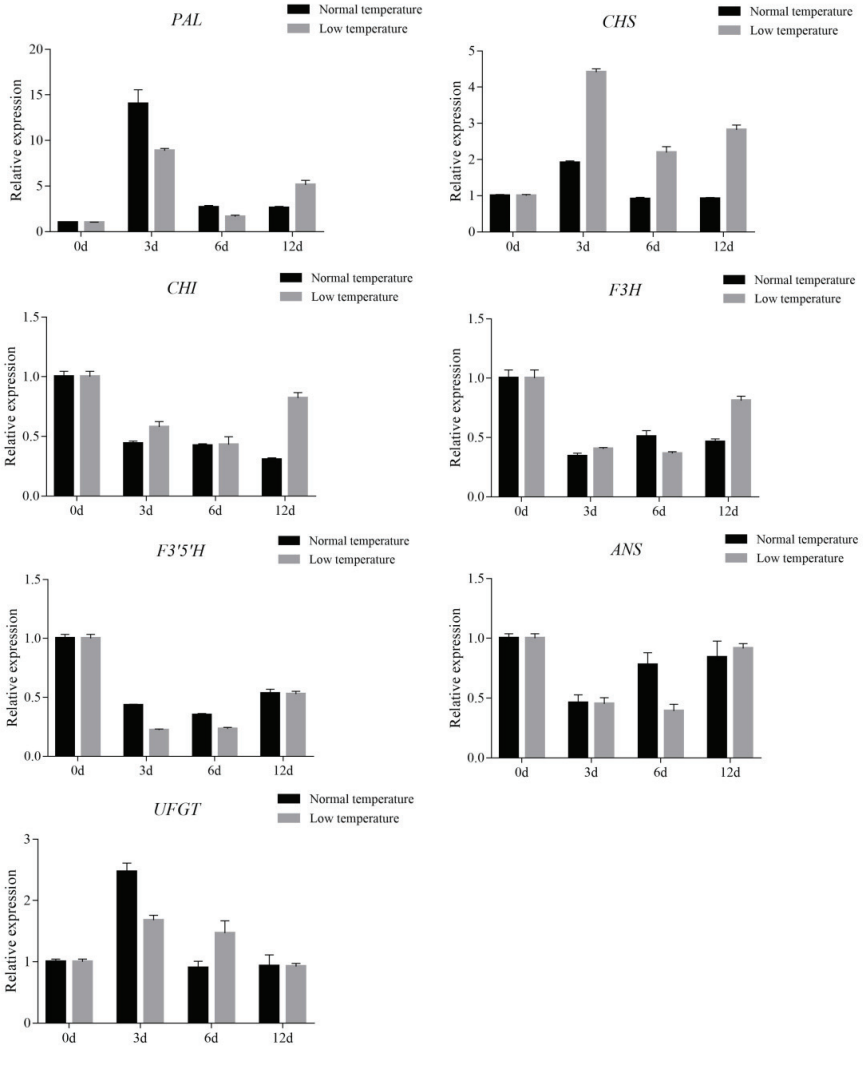


Figure 7. Relative expression level analysis of anthocyanin biosynthetic genes in purple shamrock under normal/low temperature conditions. Error bars represent the standard error of the mean (n = 3).

Moreover, most of the regulatory genes were upregulated in purple shamrock leaves exposure to low temperature, except MYB113 (Figure 8). These results indicated that low temperature treatment increases the accumulation of anthocyanins in purple shamrock leaves by promoting the expression of some anthocyanin biosynthetic and regulatory genes.

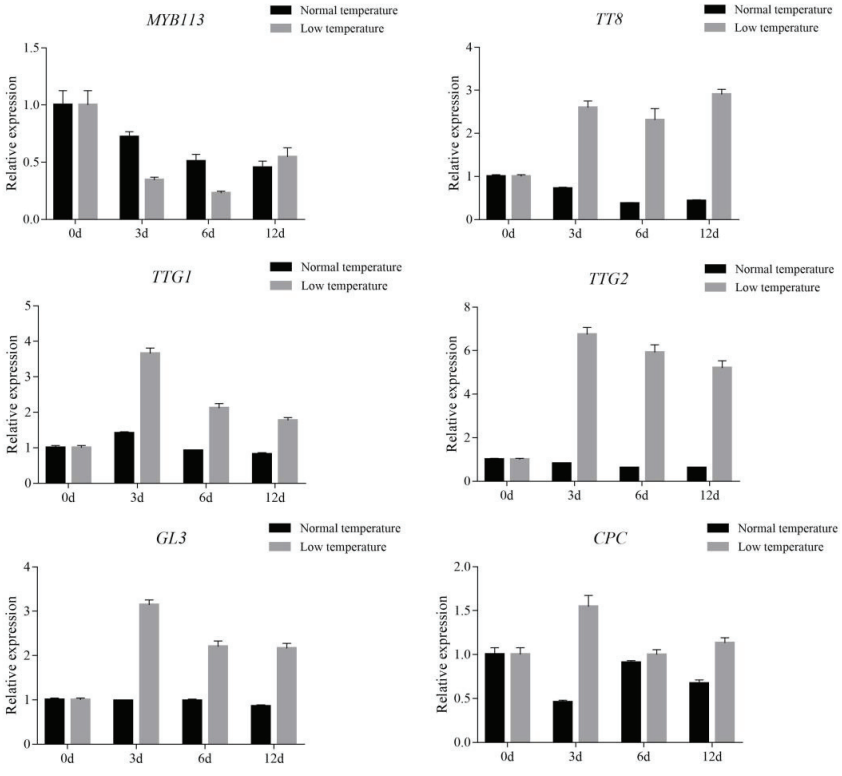


Figure 8. Relative expression level analysis of regulatory genes in purple shamrocks under normal/low temperature conditions. Error bars represent the standard error of the mean (n = 3).

3.6. Expression Levels of Anthocyanin Biosynthetic and Regulatory Genes in Flowers of Green and Purple Shamrocks

Interestingly, the flowers of green shamrock are darker in appearance than those of purple shamrock (Figure 9A). As shown in Figure 9B, the anthocyanin content in flowers of purple shamrock was markedly lower than that of green shamrock. To investigate the molecular mechanism of anthocyanin accumulation in flowers, qRT-PCR was performed. In flowers of green shamrock, just one regulatory gene, TTG2, has a higher expression level than in purple shamrock (Figure 9D). Besides, among the anthocyanin biosynthetic genes, only F3H and F3′5′H were upregulated in green shamrock (Figure 9C). According to the results mentioned above, in the flowers and leaves of shamrocks, anthocyanin biosynthesis may have different regulatory mechanisms.

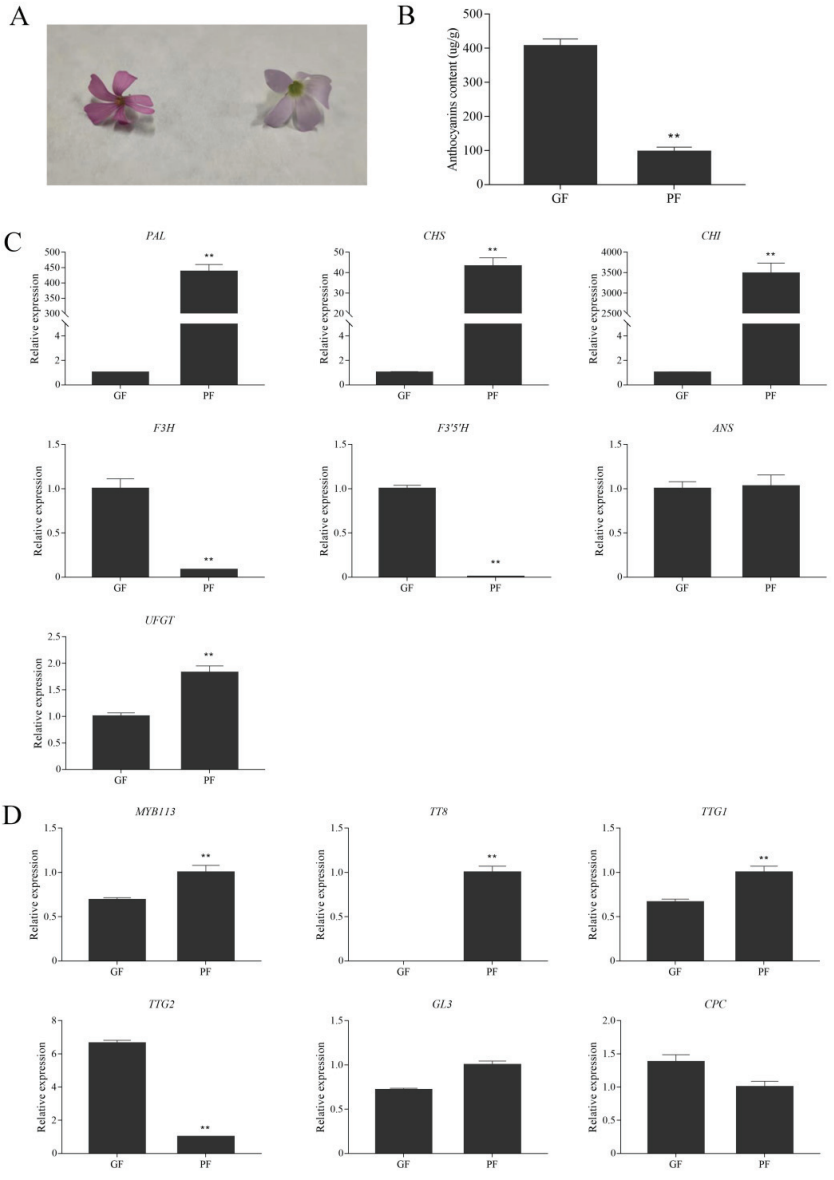


Figure 9. Relative expression level analysis of anthocyanin biosynthetic and regulatory genes in flowers of green and purple shamrocks. (**A**) The flower of green (left) and purple (right) cultivars. (**B**) Total contents of anthocyanins in flowers of green and purple shamrocks. (**C**) Relative expression level of anthocyanin biosynthetic genes in flowers of green and purple shamrocks. (**D**) Relative expression level of regulatory genes in flowers of green and purple shamrocks. Abbreviations: GF, flowers of green shamrock; PF, flowers of purple shamrock. Error bars represent the standard error of the mean ($n = 3$). Significant differences (** $p < 0.01$).

4. Discussion

Anthocyanins are flavonoid compounds found in many plants, which is the main reason why so many plants have various colors, ranging from red to blue [38]. The biosynthesis of anthocyanins is affected by many factors; therefore, its regulation is quite complex.

The color of purple shamrock leaves is the most charming feature. However, as an ornamental plant, this species has not been investigated at the molecular level. To investigate the molecular mechanisms of anthocyanin biosynthesis, UPLC-ESI-MS/MS was performed first to analyze the types of anthocyanins in purple shamrock, and three anthocyanins were identified. According to KEGG pathway, interestingly, all these three anthocyanins are derivatives of delphinidin-3-O-glucoside that makes plants appear purple [39]. Consistently, F3′5′H is highly expressed in purple shamrock, which commits the pathway to delphinidin biosynthesis. As reported by Fossen et al. [40], seven identified anthocyanins also belonged to delphinidin-type anthocyanins, consistent with our results. Anthocyanin biosynthesis is an enzyme cascade reaction process. In this pathway, all anthocyanin biosynthesis genes in purple shamrock that were identified are significantly up-regulated in comparison with green shamrock. This result may be caused by inhibition of the upstream genes of anthocyanin biosynthesis in green shamrock.

In plants, anthocyanins are compartmented in specific locations. For example, stable anthocyanins formed by different modifications, such as glycosylation, methylation and acylation, are transported to the vacuole where they accumulate [41]. Anthocyanins function in plants as protection against environmental stresses, including light, cold, salt and drought stresses [42]. Light is a key environmental factor that can affect anthocyanin biosynthesis in many plants. In litchi (*Litchi chinensis* Sonn.), anthocyanins accumulated rapidly in the pericarp under light condition [43]. The same phenomenon also occurs in teinturier grape (*Vitis vinifera* L.) [44], apple (*Malus domestica* Borkh.) [45], and pear [46]. In this report, the anthocyanin content in the leaves of purple shamrock significantly decreased after dark treatment. Under dark conditions, although the anthocyanin biosynthesis was inhibited, there was still a small amount of anthocyanin accumulation, which indicated that anthocyanin biosynthesis in purple shamrock has two ways: light-dependent and light-independent. However, the molecular mechanism of light-independent anthocyanin biosynthesis remains unclear. Temperature also affects the biosynthesis of anthocyanins. High temperature inhibits anthocyanin biosynthesis, e.g., in Cymbidium hybrids [47], while low temperature induces anthocyanin biosynthesis, as in blood oranges (*Citrus sinensis* L. Osbeck) [37], and Arabidopsis [48]. According to the qRT-PCR results, the expression of CHS was upregulated after low temperature treatment, which is consistent with Naing's report [49]. Meanwhile, most of the regulatory genes were significantly upregulated in low temperature, except for MYB113, suggesting that MYB113 in purple shamrock may not be affected by low temperature.

Interestingly, during the research, we found that flowers of green shamrock contain more anthocyanins than flowers of purple shamrock. In flowers of green shamrock, the expression of F3H, F3′5′H, was higher than that of flowers of purple shamrock, while the expression of PAL, CHS, CHI, was evidently lower. It is worth noting that TT8 is almost not expressed, while TTG2 is highly expressed in flowers of green shamrock. This result indicated that the biosynthesis of anthocyanins in flowers may be regulated by another mechanism. For example, TTG2 and ARF8 control flower coloration by regulating anthocyanin biosynthesis in tobacco (*Nicotiana tabacum* L.) [50]. According to previous studies, DFR can catalyze the conversion of dihydroquercetin to leucoanthocyanidins. However, FLS catalyzes dihydroflavonols to flavonols, which competes with DFR in the anthocyanin pathway at a key branch point [51]. In tea (*Camellia sinensis* L.), three DFR gene variants were upregulated in pink flowers relative to white flowers. On the contrary, FLS has a higher level of expression in white flowers [52]. Above all, although the key genes for anthocyanin biosynthesis have been identified, there are still many potential genes involved in anthocyanin biosynthesis.

In summary, this research clarified the molecular mechanism of anthocyanin biosynthesis in purple shamrock and explored the response of anthocyanin biosynthesis under different external environmental conditions, such as light and low temperature. Meanwhile, the anthocyanin biosynthesis mechanism of the flowers of the two species was studied, which is in contrast with the anthocyanin biosynthesis in leaves. These results

expand our understanding about the molecular mechanism of anthocyanin biosynthesis in purple shamrock.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo12121290/s1, Table S1: The primers used for qRT-PCR analysis; Figure S1: KEGG enrichment analysis; Figure S2: The expression of anthocyanin pathway genes in RNA-seq; Figure S3: The anthocyanin content of leaves after low temperature treatment; Figure S4: The phenotype of purple shamrock after dark treatment; Figure S5: DEGs in Flavonoid Biosynthesis; Supporting Information 1: Expression levels of genes involved in anthocyanin biosynthesis in the transcriptome; Supporting Information 2: DEGs involved in anthocyanin biosynthesis between green and purple shamrocks.

Author Contributions: Conceptualization, L.C.; methodology, G.C.; resources, Y.W.; writing—original draft preparation, B.L.; writing—review and editing, Z.H.; project administration, X.C.; funding acquisition, Q.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Innovation Project of People Returned from Studying Abroad of Chongqing, grant number CX2019158.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the main article and the supplementary materials.

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

References

1. Lloyd, A.; Brockman, A.; Aguirre, L.; Campbell, A.; Bean, A.; Cantero, A.; Gonzalez, A. Advances in the MYB-bHLH-WD Repeat (MBW) Pigment Regulatory Model: Addition of a WRKY Factor and Co-option of an Anthocyanin MYB for Betalain Regulation. *Plant Cell Physiol.* **2017**, *58*, 1431–1441. [CrossRef] [PubMed]
2. Zhang, Y.; Chen, G.; Dong, T.; Pan, Y.; Zhao, Z.; Tian, S.; Hu, Z. Anthocyanin Accumulation and Transcriptional Regulation of Anthocyanin Biosynthesis in Purple Bok Choy (*Brassica rapa* var. *chinensis*). *J. Agric. Food Chem.* **2014**, *62*, 12366–12376. [CrossRef]
3. Tanaka, Y.; Ohmiya, A. Seeing is believing: Engineering anthocyanin and carotenoid biosynthetic pathways. *Curr. Opin. Biotechnol.* **2008**, *19*, 190–197. [CrossRef] [PubMed]
4. Harborne, J.B.; Williams, C.A. Advances in flavonoid research since 1992. *Phytochemistry* **2000**, *55*, 481–504. [CrossRef]
5. Hazafa, A.; Rehman, K.U.; Jahan, N.; Jabeen, Z. The Role of Polyphenol (Flavonoids) Compounds in the Treatment of Cancer Cells. *Nutr. Cancer* **2020**, *72*, 386–397. [CrossRef] [PubMed]
6. Lin, B.W.; Gong, C.C.; Song, H.F.; Cui, Y.Y. Effects of anthocyanins on the prevention and treatment of cancer. *Br. J. Pharmacol.* **2017**, *174*, 1226–1243. [CrossRef]
7. Cassidy, A. Berry anthocyanin intake and cardiovascular health. *Mol. Asp. Med.* **2018**, *61*, 76–82. [CrossRef]
8. Krga, I.; Milenkovic, D. Anthocyanins: From Sources and Bioavailability to Cardiovascular-Health Benefits and Molecular Mechanisms of Action. *J. Agric. Food Chem.* **2019**, *67*, 1771–1783. [CrossRef]
9. Qi, Y.; Zhou, L.; Han, L.; Zou, H.; Miao, K.; Wang, Y. PsbHLH1, a novel transcription factor involved in regulating anthocyanin biosynthesis in tree peony (*Paeonia suffruticosa*). *Plant Physiol. Biochem.* **2020**, *154*, 396–408. [CrossRef]
10. Sun, C.; Deng, L.; Du, M.; Zhao, J.; Chen, Q.; Huang, T.; Jiang, H.; Li, C.; Li, C. A Transcriptional Network Promotes Anthocyanin Biosynthesis in Tomato Flesh. *Mol. Plant* **2020**, *13*, 42–58. [CrossRef]
11. Li, Z.; Zhao, M.; Jin, J.; Zhao, L.; Xu, Z. Anthocyanins and their biosynthetic genes in three novel-colored Rosa rugosa cultivars and their parents. *Plant Physiol. Biochem.* **2018**, *129*, 421–428. [CrossRef]
12. Liu, Y.; Tikunov, Y.; Schouten, R.E.; Marcelis, L.F.M.; Visser, R.G.F.; Bovy, A. Anthocyanin Biosynthesis and Degradation Mechanisms in Solanaceous Vegetables: A Review. *Front. Chem.* **2018**, *6*, 52. [CrossRef] [PubMed]
13. Xie, Q.; Yan, F.; Hu, Z.; Wei, S.; Lai, J.; Chen, G. Accumulation of Anthocyanin and Its Associated Gene Expression in Purple Tumorous Stem Mustard (*Brassica juncea* var. *tumida* Tsen et Lee) Sprouts When Exposed to Light, Dark, Sugar, and Methyl Jasmonate. *J. Agric. Food Chem.* **2018**, *67*, 856–866. [CrossRef]
14. Yan, S.; Chen, N.; Huang, Z.; Li, D.; Zhi, J.; Yu, B.; Liu, X.; Cao, B.; Qiu, Z. Anthocyanin Fruit encodes an R2R3-MYB transcription factor, SlAN2-like, activating the transcription of SlMYBATV to fine-tune anthocyanin content in tomato fruit. *New Phytol.* **2020**, *225*, 2048–2063. [CrossRef]
15. Yu, M.; Man, Y.; Wang, Y. Light- and Temperature-Induced Expression of an R2R3-MYB Gene Regulates Anthocyanin Biosynthesis in Red-Fleshed Kiwifruit. *Int. J. Mol. Sci.* **2019**, *20*, 5228. [CrossRef] [PubMed]

16. Zhou, H.; Peng, Q.; Zhao, J.; Owiti, A.; Ren, F.; Liao, L.; Wang, L.; Deng, X.; Jiang, Q.; Han, Y. Multiple R2R3-MYB Transcription Factors Involved in the Regulation of Anthocyanin Accumulation in Peach Flower. *Front. Plant Sci.* **2016**, *7*, 1557. [CrossRef] [PubMed]
17. Gil-Munoz, F.; Sanchez-Navarro, J.A.; Besada, C.; Salvador, A.; Badenes, M.L.; Naval, M.; Rios, G. MBW complexes impinge on anthocyanidin reductase gene regulation for proanthocyanidin biosynthesis in persimmon fruit. *Sci. Rep.* **2020**, *10*, 3543. [CrossRef]
18. Gonzalez, A.; Zhao, M.; Leavitt, J.M.; Lloyd, A.M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* **2008**, *53*, 814–827. [CrossRef]
19. Tang, B.; Li, L.; Hu, Z.; Chen, Y.; Tan, T.; Jia, Y.; Xie, Q.; Chen, G. Anthocyanin Accumulation and Transcriptional Regulation of Anthocyanin Biosynthesis in Purple Pepper. *J. Agric. Food Chem.* **2020**, *68*, 12152–12163. [CrossRef]
20. Zhu, H.F.; Fitzsimmons, K.; Khandelwal, A.; Kranz, R.G. CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Mol. Plant* **2009**, *2*, 790–802. [CrossRef]
21. Zhang, W.; Ning, G.; Lv, H.; Liao, L.; Bao, M. Single MYB-type transcription factor AtCAPRICE: A new efficient tool to engineer the production of anthocyanin in tobacco. *Biochem. Biophys. Res. Commun.* **2009**, *388*, 742–747. [CrossRef] [PubMed]
22. Maier, A.; Schrader, A.; Kokkelink, L.; Falke, C.; Welter, B.; Iniesto, E.; Rubio, V.; Uhrig, J.F.; Hulskamp, M.; Hoecker, U. Light and the E3 ubiquitin ligase COP1/SPA control the protein stability of the MYB transcription factors PAP1 and PAP2 involved in anthocyanin accumulation in Arabidopsis. *Plant J.* **2013**, *74*, 638–651. [CrossRef] [PubMed]
23. Shin, D.H.; Choi, M.; Kim, K.; Bang, G.; Cho, M.; Choi, S.B.; Choi, G.; Park, Y.I. HY5 regulates anthocyanin biosynthesis by inducing the transcriptional activation of the MYB75/PAP1 transcription factor in Arabidopsis. *FEBS Lett.* **2013**, *587*, 1543–1547. [CrossRef] [PubMed]
24. Zhou, L.; He, Y.; Li, J.; Liu, Y.; Chen, H. CBFs Function in Anthocyanin Biosynthesis by Interacting with MYB113 in Eggplant (*Solanum melongena* L.). *Plant Cell Physiol.* **2020**, *61*, 416–426. [CrossRef]
25. Man, Y.P.; Wang, Y.C.; Li, Z.Z.; Jiang, Z.W.; Yang, H.L.; Gong, J.J.; He, S.S.; Wu, S.Q.; Yang, Z.Q.; Zheng, J.; et al. High-temperature inhibition of biosynthesis and transportation of anthocyanins results in the poor red coloration in red-fleshed Actinidia chinensis. *Physiol. Plant* **2015**, *153*, 565–583. [CrossRef]
26. Tao, R.; Yu, W.; Gao, Y.; Ni, J.; Yin, L.; Zhang, X.; Li, H.; Wang, D.; Bai, S.; Teng, Y. Light-Induced Basic/Helix-Loop-Helix64 Enhances Anthocyanin Biosynthesis and Undergoes CONSTITUTIVELY PHOTOMORPHOGENIC1-Mediated Degradation in Pear. *Plant Physiol.* **2020**, *184*, 1684–1701. [CrossRef]
27. He, Q.; Ren, Y.; Zhao, W.; Li, R.; Zhang, L. Low Temperature Promotes Anthocyanin Biosynthesis and Related Gene Expression in the Seedlings of Purple Head Chinese Cabbage (*Brassica rapa* L.). *Genes* **2020**, *11*, 81. [CrossRef]
28. Groom, Q.J.; Van der Straeten, J.; Hoste, I. The origin of *Oxalis corniculata* L. *PeerJ* **2019**, *7*, e6384. [CrossRef]
29. Rehman, A.; Rehman, A.; Ahmad, I. Antibacterial, Antifungal, and Insecticidal Potentials of *Oxalis corniculata* and Its Isolated Compounds. *Int. J. Anal. Chem.* **2015**, *2015*, 842468. [CrossRef]
30. Salahuddin, H.; Mansoor, Q.; Batool, R.; Farooqi, A.A.; Mahmood, T.; Ismail, M. Anticancer activity of Cynodon dactylon and *Oxalis corniculata* on Hep2 cell line. *Cell Mol. Biol.* **2016**, *62*, 60–63. [CrossRef]
31. Satish, S.; Raghavendra, M.P.; Raveesha, K.A. Evaluation of the antibacterial potential of some plants against human pathogenic bacteria. *Adv. Biol. Res.* **2008**, *2*, 44–48.
32. Pyšek, P.; Pergl, J.; Essl, F.; Lenzner, B.; Dawson, W.; Kreft, H.; Weigelt, P.; Winter, M.; Kartesz, J.; Nishino, M.; et al. Naturalized alien flora of the world: Species diversity, taxonomic and phylogenetic patterns, geographic distribution and global hotspots of plant invasion. *Preslia* **2017**, *89*, 203–274. [CrossRef]
33. Rayyan, S.; Fossen, T.; Andersen, O.M. Flavone C-glycosides from leaves of *Oxalis triangularis*. *J. Agric. Food Chem.* **2005**, *53*, 10057–10060. [CrossRef] [PubMed]
34. Rapisarda, P.; Fanella, F.; Maccarone, E. Reliability of Analytical Methods for Determining Anthocyanins in Blood Orange Juices. *J. Agric. Food Chem.* **2000**, *48*, 2249–2252. [CrossRef] [PubMed]
35. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **1995**, *57*, 289–300. [CrossRef]
36. Trapnell, C.; Williams, B.A.; Pertea, G.; Mortazavi, A.; Kwan, G.; van Baren, M.J.; Salzberg, S.L.; Wold, B.J.; Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **2010**, *28*, 511–515. [CrossRef]
37. Carmona, L.; Alquezar, B.; Diretto, G.; Sevi, F.; Malara, T.; Lafuente, M.T.; Pena, L. Curing and low-temperature combined post-harvest storage enhances anthocyanin biosynthesis in blood oranges. *Food Chem.* **2021**, *342*, 128334. [CrossRef]
38. Naing, A.H.; Kim, C.K. Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. *Plant Mol. Biol.* **2018**, *98*, 1–18. [CrossRef]
39. Yan, H.; Pei, X.; Zhang, H.; Li, X.; Zhang, X.; Zhao, M.; Chiang, V.L.; Sederoff, R.R.; Zhao, X. MYB-Mediated Regulation of Anthocyanin Biosynthesis. *Int. J. Mol. Sci.* **2021**, *22*, 3103. [CrossRef]
40. Fossen, T.; Rayyan, S.; Holmberg, M.H.; Nateland, H.V.S.; Andersen, Ø.M. Acylated anthocyanins from leaves of *Oxalis triangularis*. *Phytochemistry* **2005**, *66*, 1133–1140. [CrossRef]
41. Liu, H.; Liu, Z.; Wu, Y.; Zheng, L.; Zhang, G. Regulatory Mechanisms of Anthocyanin Biosynthesis in Apple and Pear. *Int. J. Mol. Sci.* **2021**, *22*, 8441. [CrossRef] [PubMed]

42. 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] [PubMed]
43. Zhang, H.N.; Li, W.C.; Wang, H.C.; Shi, S.Y.; Shu, B.; Liu, L.Q.; Wei, Y.Z.; Xie, J.H. Transcriptome Profiling of Light-Regulated Anthocyanin Biosynthesis in the Pericarp of Litchi. *Front. Plant Sci.* **2016**, *7*, 963. [CrossRef]
44. Guan, L.; Dai, Z.; Wu, B.H.; Wu, J.; Merlin, I.; Hilbert, G.; Renaud, C.; Gomes, E.; Edwards, E.; Li, S.H.; et al. Anthocyanin biosynthesis is differentially regulated by light in the skin and flesh of white-fleshed and teinturier grape berries. *Planta* **2016**, *243*, 23–41. [CrossRef] [PubMed]
45. Vimolmangkang, S.; Zheng, D.; Han, Y.; Khan, M.A.; Soria-Guerra, R.E.; Korban, S.S. Transcriptome analysis of the exocarp of apple fruit identifies light-induced genes involved in red color pigmentation. *Gene* **2014**, *534*, 78–87. [CrossRef]
46. Yu, B.; Zhang, D.; Huang, C.; Qian, M.; Zheng, X.; Teng, Y.; Su, J.; Shu, Q. Isolation of anthocyanin biosynthetic genes in red Chinese sand pear (Pyrus pyrifolia Nakai) and their expression as affected by organ/tissue, cultivar, bagging and fruit side. *Sci. Hortic.* **2012**, *136*, 29–37. [CrossRef]
47. Nakatsuka, T.; Suzuki, T.; Harada, K.; Kobayashi, Y.; Dohra, H.; Ohno, H. Floral organ- and temperature-dependent regulation of anthocyanin biosynthesis in *Cymbidium hybrid* flowers. *Plant Sci.* **2019**, *287*, 110173. [CrossRef]
48. Ilk, N.; Ding, J.; Ihnatowicz, A.; Koornneef, M.; Reymond, M. Natural variation for anthocyanin accumulation under high-light and low-temperature stress is attributable to the ENHANCER OF AG-4 2 (HUA2) locus in combination with PRODUCTION OF ANTHOCYANIN PIGMENT1 (PAP1) and PAP2. *New Phytol.* **2015**, *206*, 422–435. [CrossRef]
49. Naing, A.H.; Park, D.Y.; Park, K.I.; Kim, C.K. Differential expression of anthocyanin structural genes and transcription factors determines coloration patterns in gerbera flowers. *3 Biotech.* **2018**, *8*, 393. [CrossRef]
50. Li, P.; Chen, X.; Sun, F.; Dong, H. Tobacco TTG2 and ARF8 function concomitantly to control flower colouring by regulating anthocyanin synthesis genes. *Plant Biol.* **2017**, *19*, 525–532. [CrossRef]
51. Martens, S.; Preuss, A.; Matern, U. Multifunctional flavonoid dioxygenases: Flavonol and anthocyanin biosynthesis in *Arabidopsis thaliana* L. *Phytochemistry* **2010**, *71*, 1040–1049. [CrossRef] [PubMed]
52. Zhou, C.; Mei, X.; Rothenberg, D.O.N.; Yang, Z.; Zhang, W.; Wan, S.; Yang, H.; Zhang, L. Metabolome and Transcriptome Analysis Reveals Putative Genes Involved in Anthocyanin Accumulation and Coloration in White and Pink Tea (*Camellia sinensis*) Flower. *Molecules* **2020**, *25*, 190. [CrossRef] [PubMed]

metabolites

MDPI

Article

Comparative Metabolomics Profiling Reveals Key Metabolites and Associated Pathways Regulating Tuber Dormancy in White Yam (*Dioscorea rotundata* Poir.)

Jeremiah S. Nwogha [1,2,3], Abtew G. Wosene [1], Muthurajan Raveendran [2], Jude E. Obidiegwu [3], Happiness O. Oselebe [4], Rohit Kambale [2], Cynthia A. Chilaka [5,*] and Veera Ranjani Rajagopalan [2]

1 Department of Horticulture and Plant Sciences, College of Agriculture and Veterinary Medicine, Jimma University, Jimma P.O. Box 307, Ethiopia; wosene.gebreselassie@ju.edu.et (A.G.W.)
2 Centre for Plant Molecular Biology & Biotechnology, Departments of Plant Biotechnology and Biochemistry, Tamil Nadu Agricultural University, Coimbatore 641003, India; raveendrantnau@gmail.com (M.R.); rohitkamble568@gmail.com (R.K.)
3 Yam Research Programme, National Root Crops Research Institute, Umudike 440001, Nigeria; ejikemeobidiegwu@gmail.com
4 Department of Crop Production and Landscape Management, Ebonyi State University, Abakaliki 480282, Nigeria; h.oselebe@gmail.com
5 Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, 19 Chlorine Gardens, Belfast BT9 5DL, UK
* Correspondence: c.chilaka@qub.ac.uk

Abstract: Yams are economic and medicinal crops with a long growth cycle, spanning between 9–11 months due to their prolonged tuber dormancy. Tuber dormancy has constituted a major constraint in yam production and genetic improvement. In this study, we performed non-targeted comparative metabolomic profiling of tubers of two white yam genotypes, (*Obiaoturugo* and *TDr1100873*), to identify metabolites and associated pathways that regulate yam tuber dormancy using gas chromatography–mass spectrometry (GC–MS). Yam tubers were sampled between 42 days after physiological maturity (DAPM) till tuber sprouting. The sampling points include 42-DAPM, 56-DAPM, 87DAPM, 101-DAPM, 115-DAPM, and 143-DAPM. A total of 949 metabolites were annotated, 559 in *TDr1100873* and 390 in *Obiaoturugo*. A total of 39 differentially accumulated metabolites (DAMs) were identified across the studied tuber dormancy stages in the two genotypes. A total of 27 DAMs were conserved between the two genotypes, whereas 5 DAMs were unique in the tubers of *TDr1100873* and 7 DAMs were in the tubers of *Obiaoturugo*. The differentially accumulated metabolites (DAMs) spread across 14 major functional chemical groups. Amines and biogenic polyamines, amino acids and derivatives, alcohols, flavonoids, alkaloids, phenols, esters, coumarins, and phytohormone positively regulated yam tuber dormancy induction and maintenance, whereas fatty acids, lipids, nucleotides, carboxylic acids, sugars, terpenoids, benzoquinones, and benzene derivatives positively regulated dormancy breaking and sprouting in tubers of both yam genotypes. Metabolite set enrichment analysis (MSEA) revealed that 12 metabolisms were significantly enriched during yam tuber dormancy stages. Metabolic pathway topology analysis further revealed that six metabolic pathways (linoleic acid metabolic pathway, phenylalanine metabolic pathway, galactose metabolic pathway, starch and sucrose metabolic pathway, alanine-aspartate-glutamine metabolic pathways, and purine metabolic pathway) exerted significant impact on yam tuber dormancy regulation. This result provides vital insights into molecular mechanisms regulating yam tuber dormancy.

Keywords: metabolites; differential–accumulation; energy; antioxidant; metabolism; molecular–mechanism; dormancy; regulation

Citation: Nwogha, J.S.; Wosene, A.G.; Raveendran, M.; Obidiegwu, J.E.; Oselebe, H.O.; Kambale, R.; Chilaka, C.A.; Rajagopalan, V.R. Comparative Metabolomics Profiling Reveals Key Metabolites and Associated Pathways Regulating Tuber Dormancy in White Yam (*Dioscorea rotundata* Poir.). *Metabolites* **2023**, *13*, 610. https://doi.org/10.3390/metabo13050610

Academic Editors: Yanjie Zhang and Yan Li

Received: 15 March 2023
Revised: 11 April 2023
Accepted: 21 April 2023
Published: 28 April 2023

1. Introduction

Yam is a multi-species crop that has about 613 known species that produce tubers, bulbils, or rhizomes [1]. Of these, about 10 are cultivated over a larger area and serve as a staple food crop, and about 50 other species are also eaten as wild-harvested staple famine food; thus, this genus occupies a prominent position in global food insecurity combat [2,3]. Yam is placed fourth position among the utilized root and tuber crops globally, after potatoes (*Solanum* spp.), cassava (*Manihot esculenta*), and sweet potatoes (*Ipomoea* spp.), and second in West Africa after cassava [4,5]. However, the sensorial preference for yam coupled with its better organoleptic properties and storability compared to cassava, potatoes, sweet potatoes, and plantain has led to its high demand as a major cash crop in Sub-Sahara Africa [1]. Its potential as a source of food is attributed to its high levels of carbohydrates including fiber, starch, and sugar, contributing about 200 dietary calories per person per day to more than 300 million people in the tropics [6], in addition to other nutritional benefits such as proteins, lipids, vitamins, and minerals [2]. Yam tuber also contains an abundance of steroidal C_{27} saponins, and diosgenin. the aglycone portion of the abundant saponin dioscin has been industrially exploited as the starting material for the synthesis of pregnenolone-derived steroids, thus making yam an industrial crop [7].

Yam tuber dormancy is the temporary growth arrest of the continuous underlying meristematic cells beneath the dead brown skin of the tuber [8]. It is an inherent mechanism that regulates tuber sprouting. The dormancy phase has been divided into three categories according to Lang, et al. [8], including: endodormancy (deep dormancy during which growth arrest is influenced by internal physiological and genetic factors within the meristem), para-dormancy (this occurs when growth is arrested by physiological factors external to the meristem), and eco-dormancy (growth is stopped by unfavorable external or environmental factors).

Prolonged tuber dormancy after physiological maturity has constituted a great constraint in yam crop production and genetic improvement [9]. Tuber dormancy is the major cause of ware and seed tubers' inability to sprout for a prolonged period, during which tubers remain dormant; incapable of developing internal or external shoot buds/sprouts for about 150 to 210 days depending on: species, genotype, and growing and storage environmental conditions [10–12]. Thus, this makes it impossible to have more than one crop cycle per year, thereby limiting yam crop productivity, tuber availability for industrial use, and slowing down the rate of genetic improvement [13]. Dormancy induction and release is associated with numerous physiological and biochemical processes that are regulated by gene expression, protein synthesis, hormonal signaling, and energy metabolism [14]. In these processes, both the decomposition and synthesis of metabolites, conversion, and consumption of energy occur, thus generating a genotype and dormancy stage-specific metabolites [15,16]. Some of these metabolites have been found to regulate dormancy in seeds of various plant species. The typical example is phytohormones, whose involvement in seed dormancy regulation has always been a popular topic in seed physiology and biochemistry research [17]. The most reported phytohormone influencing dormancy induction in crops is abscisic acid (ABA), which inhibits the process of seed germination, though complex interactions exist between ABA and other phytohormones [18]. Gibberellic acid (GA) is another important phytohormone known to be involved in dormancy regulation. It acts antagonistically with ABA to promote seed germination (dormancy breaking), while jasmonic acid (JA) shows a synergetic effect with ABA in regulating dormancy induction and maintenance in plants [7]. Exogenous treatment with auxin in soybean was found to induce dormancy by decreasing the GA/ABA ratio [19]. Therefore, the roles of these phytohormones in dormancy regulation are complex and subtle.

Polyamine has also been reported to play a key role as a modulator of plant growth and development, and its involvement in seed dormancy regulation has attracted research attention recently. Dormancy-breaking led to the accumulation of the total polyamine content in the seeds [20]. Polyamine was found to be distributed compartmentally in dormancy-breaking seeds, which may suggest its involvement in the process [21]; presoak-

ing with polyamine improved seed germinability [22]. This effect has been suspected to involve starch metabolism, phytohormones interactions, and antioxidant defenses [23,24]. In addition, the involvement of amino acids in dormancy regulation has also attracted research attention. For instance, it has been reported that essential amino acids such as lysin, methionine, leucine, isoleucine, threonine, phenylalanine, and valine increased significantly during the dormancy breaking process in wheat, brown rice, triticale, and Arabidopsis [25,26]. Asparagine, arginine, and γ-aminobutyric (GABA) are the three most common amino compounds in the mobilization of nitrogen reserves in dormancy breaking chestnut seeds [27]. Therefore, as the major transporter of all forms of nitrogen and the key factor of carbon/nitrogen balance in the plants, the role of amino acids in dormancy regulation should not be ignored [18].

Carbohydrates play multiple roles in plant growth and development, besides acting as the primary source of carbon and energy; hence, energy metabolism has also been implicated in dormancy regulation. It has been reported that Embden Meyerhof Parnas (EMP) pathway, tricarboxylic acid (TCA) pathway, and pentose phosphate pathway (PPP) strictly regulated dormancy in different plants [28]. For instance, the TCA cycle was enhanced while the PPP pathway was slowly decreased during apple bud dormancy breaking [29]. In grape, dormancy breaking was induced by chemicals and low temperature was regulated by PPP, EMP, and TCA cycles [30,31]. Also, other carbohydrates generating energy pathways has been reported to be involved in dormancy regulation in herbaceous plants; for example, sucrose metabolism, lipid metabolism, and amino acid metabolism were all downregulated in imbibed dormant seeds of herbaceous plants [32], thus demonstrating that the ability of imbibed seeds to synthesize protein and generate adenosine triphosphate (ATP) decreases during the dormancy period [33,34]. Also, two studies have reported that the imbibed dormant seeds of *Picea glauca* and *Julans regia* exhibited an inactivation of sucrose and amino acids due to wet cold induced dormancy [32]. Furthermore, an accumulation the galactinol and raffinose family of oligosaccharides (RFOs) during seed maturation stages in some legumes has indicated their roles in desiccation tolerance and longevity [35]. Galactinol content in dry seeds has also been used as a biomarker for seed longevity in *Brassiceae* and tomato [36]. The sucrose-non-fermenting 1-related kinase1 (SnRK1) transcription family has been reported to play the role of energy metabolism sensors in plant systems; they detect when energy is below optimum in the plant system and inhibit plant growth and developmental processes that are energy-consuming, such as dormancy-breaking (sprouting), in order to maintain energy homeostasis during a low-energy economy, which characterizes the dormant state [37,38]. There are also reports that trehalose-6-phospate (T6P) initiates the process of growth in any plant part undergoing growth arrest (dormancy) by inhibiting the action of (SnRK1) [39,40].

The effective reduction of yam tuber dormancy duration through genetic manipulation is considered a reliable solution to the challenge of prolonged tuber dormancy and ensures more than one yam crop cycle per annual. It is important to note that the paucity of genetic resource information on factors controlling the dormancy trait has limited efforts towards overcoming the barriers mentioned above. Our study presents an alternative route in understanding implicated metabolites and their associated pathways at the biochemical and genetic level which could be utilized in genetic manipulation yam tuber dormancy.

The development of metabolomics technologies has improved our understanding of the complex molecular interaction of biological systems. Due to its high throughput and efficiency, metabolomics has been applied to research on the mechanism of dormancy in several plant species [18]. However, in yam crop only few metabolomics studies exist, and to the best of our knowledge this is the first time in which the metabolomics approach is being applied to study yam tuber dormancy. In this study, an untargeted GC-MS approach was used to profile tubers of two white yam genotypes (*Obiaoturugo* and *TDr1100873*) for differentially accumulated metabolites from dormancy induction to dormancy breaking in order to determine the metabolites and their associated pathways that regulate yam tuber dormancy.

2. Experimental Design

2.1. Genetic Material

Two yam genotypes (*Obiaoturugo* and *TDr1100873*) of white yam specie (*D. rotundata*,) were used for the study. *Obiaoturugo* is a popular landrace while *TDr1100873* is an advanced breeding clone. Both individuals were sourced from the Yam breeding program of the National Root Crops Research Institute (NRCRI), Umudike, Nigeria.

2.2. Field Study Area

The field was established at an Eastern research farm of the National Root Crops Research Institute (NRCRI) in Umudike, Nigeria. Umudike is a rainforest agroecology located at; 5.4729° N, 7.5480° E, 152 m ASL with a mean annual rainfall of 2093 mm, mean annual temperature of 27.3 °C, mean annual relative humidity of 82%, and sandy loam soil with a pH range of 4.3 to 5.27.

2.3. Planting and Agronomic Management

A uniform set size (200 g) of only proximal and distal regions of the tubers was planted in order to maintain relatively uniform germination time. A total of 30 plant stands of each genotype were planted in randomized complete block design, replicated three times, with 10 plant stands of each genotype per replicate. The field was adequately maintained under a rainfed condition, and all cultural practices including application of NPK fertilizer at the recommended dose of 80:60:100 kg/ha [41] were carried out. Tubers were harvested at 50% senescence, which is the tuber physiological maturity stage. The Figure 1 below shows the vegetative growth and tuber storage stages.

Figure 1. Howing (**a**); established field of the two yam genotypes, (**b**) tubers of *Obiaoturugo* in post-harvest study facility, and (**c**) tubers of *TDr1100873* in post-harvest study facility.

2.4. Postharvest Study and Sampling

After harvesting, tubers were freighted to the Centre of Plant Molecular Biology and Bioinformatics, Department of Plant Biotechnology Tamil Nadu Agricultural University, Coimbatore, India, where the tubers were stored in a greenhouse facility with natural light conditions and sunshine hours per day between 10 to 12 h, a relative humidity range of 38% to 78%, and a maximum and minimum temperature range from 26 to 36.5 °C and 17.5 to 26.5 °C, respectively, during the study. Triplicate samples of tubers were collected at the following days after physiological maturity (DAPM): 42-DAPM, 56-DAPM, 87-DAPM, 101-DAPM, 115-DAPM and sprouted tuber (143-DAPM). The two genotypes (*Obiaoturugo* and *TDr1100873*) exhibited different dormant phenotypes. *Obiaoturugo* exhibited a shorter dormant phenotype and sprouted at 101 days after tuber physiological maturity, and was sampled only four times within this period at 42, 56, 87, and 101 days. On the other hand, *TDr1100873* exhibited longer dormant phenotype and sprouted at 143 days after tuber physiological maturity, and was sampled six times within the period at 42, 56, 87,

101, 115, and 143 days after tuber physiological maturity. The samples were frozen in liquid nitrogen immediately after collection, and stored at −80 °C until the sampling was completed. At the completion of sampling, freeze-dried samples were microwaved for 5 min and then oven-dried in an oven at 72 °C for 48 h, and grinded into powder using a commercial grinder.

2.5. Metabolites Extraction

A total of 10 g of powdered samples were wrapped in filter papers, and dropped in the cylindrical tubes of Soxhlet apparatus and filled with 400 mL of 100% methanol solvent produced by Sigma–Aldrich. The temperature of the Soxhlet apparatus was set to 70 °C. Each sample was subjected to nine extraction cycles, and 50 mL of the extract were collected in 50 mL centrifuging tubes at the end of the nine extraction cycles, then air-dried to 7.5 mL before further processing.

Sample Processing

A total of 500 µL of the extracts was pipetted into 2.0 Eppendoff tubes and incubated in a water bath for 15 min at 37 °C, before centrifuging at 10,000 rpm and 4 °C for 10 min. Samples were filtered using 0.2 mm filter paper and syringe. After filtering, samples were diluted with methanol at a ratio of 1:9. A total of 500 mL of the diluted samples were concentrated at 20,000 rpm for 2 h. The concentrated solid extracts were derivatized according to [1] with slight modifications by dissolving it in 50 mL solution of (20 mg methoxylamine hydrochloride + 1 mL of pyrimidine), and incubated for 1.5 h at 37 °C. Afterward, 80 µL of N-Methyl-N-(trimethylsily)trifluoroacetamide was also added and incubated for 30 min at 37 °C. The derivatization chemicals were sourced from Sigma–Aldrich (Bengaluru, Karnataka 560099, India).

2.6. GC-MS Analysis

Untargeted metabolomics data were acquired using two auto-injecting GC-MS machines (Turbomass and 8040NX) by injecting 1ml of each sample. The GC oven was held for 2 min at 705 °C, ramp for 10 °C/min to 150 °C, then held for 2 min, ramp at 5 °C/min to 220 °C, held for 1 min, ramp 10 °C/min to 250 °C, held for 1 min. Auto-injection was conducted at 220 °C. Helium was the carrier gas at a flow rate of 1.3 mL/min. The interface with the MS was set at 290 °C and MS performed in full scan mode using 70 eV EI + and scanned from 50 to 500 Da. Retention time locking to ribitol was used according to Enfissi, et al. [42]. A mixture of n-alkanes, ranging from 8 to 32 carbons, was used for retention index external calibration. One milliliter of the extraction solvent (methanol) was used as the internal standard. Sample sets were run in batches of three replicates, Solvent Delay = 3.00 min, Transfer Temp = 250 °C, Source Temp = 220 °C, Scan mass range: 50 to 500 Da. Column was set at 30.0 m × 250 µm.

Data Processing

To identify chromatogram components found in white yam tuber profiles, a mass spectral library was constructed from machine in-built libraries: NIST '20M&R MS lib, Main lib, rep lib, OA_TMS_DB5_67min_V3 lib and FA_ME_SP2560_V3 lib. Component peak identification and spectral deconvolution was performed using the Automated Mass Spectral Deconvolution and Identification System (AMDIS v2.71, NIST) using Kovat's retention indices (RI) and MS for identification according to the metabolomics reporting guidelines [1]. Each compound was assigned a representative ion and response areas were integrated and expressed relative to internal standard. Linearity was assessed following the standard extraction procedure on amounts of material (10 mg) and the range further enhanced by taking the quantity of aliquot used (400 µL). Recovery was assessed by expressing measurements from each extract of sample (10 mg) as a percentage of the total following three sequential extractions. All measurements for the developmental stages were conducted in triplicate and relative to internal standard.

2.7. Data Statistical Analysis and Visualization

Data were further processed in metaboAnalyst 5.0 platform. Row-wise normalization by the sum, logarithmic transformation and scaling by mean centering were performed to normalize the data (Figure S1a,b). All statistical analyses were perform using metaboanalyst software. Venn tool (Version 2.1) [43] was used to determine the relationship between differential accumulated metabolites (DAMs) in the tubers of the two genotypes. One-way univariate ANOVA analyses were performed to determine the significant differentially accumulated metabolites across different data points at a *p*-value threshold of $p \leq 0.05$, and Turkey's HSD was used to compare the means. Principal component analysis (PCA) was performed to determine the overall pattern of metabolites distribution in the two genotypes along the data points. In order to filter out the variables in the metabolites that were not related to categorical variables and to obtain more reliable metabolite information on the structure of different data points, partial least squares–discriminant analysis (PLS–DA) was performed using variable importance in projection (VIP) at a threshold of VIP ≤ 1. Agglomerative hierarchical cluster analysis was performed to classify the differentially accumulated metabolites (DAMs) across the data point using euclidean and ward.D distance algorithms and visualized it in a clustering heatmap. The DAMs were mapped into PubChem, Kyoto encyclopedia genes and genomes (KEGG) and human metabolome data base (HMDB) to identify their functional groups. Metabolites set enrichment analysis (MSEA) and pathway annotation were performed to determine the significantly induced metabolisms that DAMs are involved in, during the studied yam dormancy stages. Pathway topology analysis was performed to visualize the significantly induced metabolic pathways.

3. Results

3.1. Metabolic Profiling of White Yam Tuber during Dormancy

A total of 949 metabolites were annotated in tubers of the two genotypes during the study period; 390 metabolites were annotated in the tubers of *Obiaoturugo* and 559 metabolites annotated in the tubers of *TDr1100873*. Using Venn interactive tool (accessed on 1 November 2022) [43], we compared the tuber metabolomes of the two genotypes and it showed that 158 metabolites were common between the genotypes (Figure 2a). Differentially accumulated metabolites (DAMs) in tubers of the two genotypes were also compared, and Figure 2b showed that 27 DAMs were conserved in the tubers of both genotypes, while 7 unique DAMs were observed *Obiaoturugo* and 5 unique DAMs were observed *TDr1100873*. It is believed that the phenotypic variation between the two yam genotypes with respect to dormancy duration was driven by these unique DAMs.

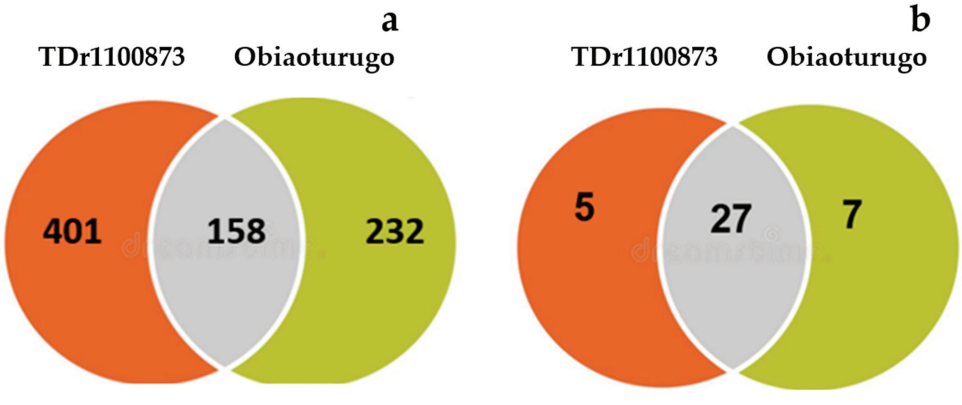


Figure 2. (**a**) Comparison of total annotated metabolites in the tubers of the two yam genotypes. (**b**) Comparison of significantly differentially accumulated metabolites during dormancy in the tubers of the two genotypes.

Analysis of variance (ANOVA) with Bonferroni correction identified 28 and 30 significantly differentially accumulated metabolites at ($p \leq 0.05$ and FDR ≤ 0.05) across the data points in the tubers of *Obiaoturugo* and *TDr1100873* respectively (Figure 3a,b).

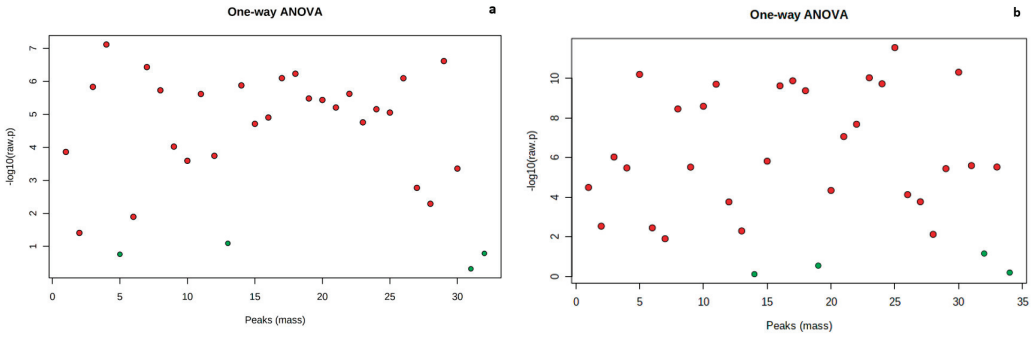


Figure 3. Differential accumulated metabolites identified by ANOVA plot with *p*-value threshold 0.05, FDR $\leq$ 0.05 and logFC 1.5. Means were compared using Turkey's Post-hoc test. (**a**) shows DAMs in the tubers of *TDr1100873* across the studied tuber dormancy stages. (**b**) Shows DAMs in the tubers of *Obiaoturugo* across the studied tuber dormancy stages. Red dots are significant differential accumulated metabolites, while the green dots are the metabolites that were differentially accumulated but are not significant. The vertical axis indicates the metabolites log-fold-change, while the horizontal axis shows the mass peak of the metabolites.

3.2. Principal Component Analysis (PCA) and Partial Least Square–Discriminant Analysis (PLS–DA) of Metabolites

Principal component analysis (PCA) was performed to determine the overview of yam tuber metabolome at the six data points of *TDr1100873* and four data points of *Obiaoturugo*. Figure 4a shows the discriminating pattern of *TDr1100873* tuber metabolome across the studied dormancy stages. The first principal component (PCI) clearly separated *TDr1100873* tuber metabolome at 42-DAPM, 115-DAPM and 143-DAPM separated from the rest of the data points and it accounts for 28.4% of total variation, while PC2 accounts for 21.4% of the total variation and it drives the clustering together of the tuber metabolome at 56-DAPM, 87-DAPM, and 101-DAPM. This indicates that the later three dormancy stages shared similar metabolome, while the former three dormancy stages were distinct. Although 115-DAPM and 143-DAPM together constitute the dormancy-breaking stage and were expected to share similarities in their metabolomes, PCA did not detect any similarity between them. The tuber metabolome at 115-DAPM, which marks the appearance shoot buds (the first physical sign that tuber is getting ready for sprouting), differed substantially from the metabolome at 143-DAPM (sprouted tuber). We used PCA loading to identify the major metabolites contributing to the tuber metabolome variations at these dormancy stages (Table S1). The metabolites playing key roles in tuber metabolome discrimination at different dormancy stages include: four esters, four amines, two alcohols, an alkaloid, fatty acid, lipid, coumarin, terpenoid, and one flavonoid. Similarly, Figure 4b shows an *Obiaoturugo* tuber metabolome discriminating pattern across its four studied dormancy stages. The tuber metabolomes at 87-DAPM and 101-DAPM were clearly separated from that at 42-DAPM and 56-DAPM, respectively, by PC1, and it accounted for 50.5% total variations. PC2 accounts for 32.4% of total variations and drives the clustering together of 42-DAPM and 56-DAPM together. This is an indication that *Obiaoturugo* tuber metabolome at these dormancy stages are similar, implying that during the 14 days between the two sampling points no substantive metabolic change occurred. However, the distance between the tuber metabolome at 87-DAPM and 101-DAPM indicates that tuber metabolomes at these two stages are clearly different from each other. 87-DAPM marked the appearance of shoot buds in the tubers of *Obiaoturugo*, while 101-DAPM is its sprouted

tuber metabolome. PCA loading revealed that the metabolites driving the discrimination of *Obiaoturugo* tuber metabolome at 87-DAPM and 101-DAPM were: five esters, three amines, three alcohols, lipid, fatty acid, coumarin, flavonoid, and a carbohydrate (Table S2).

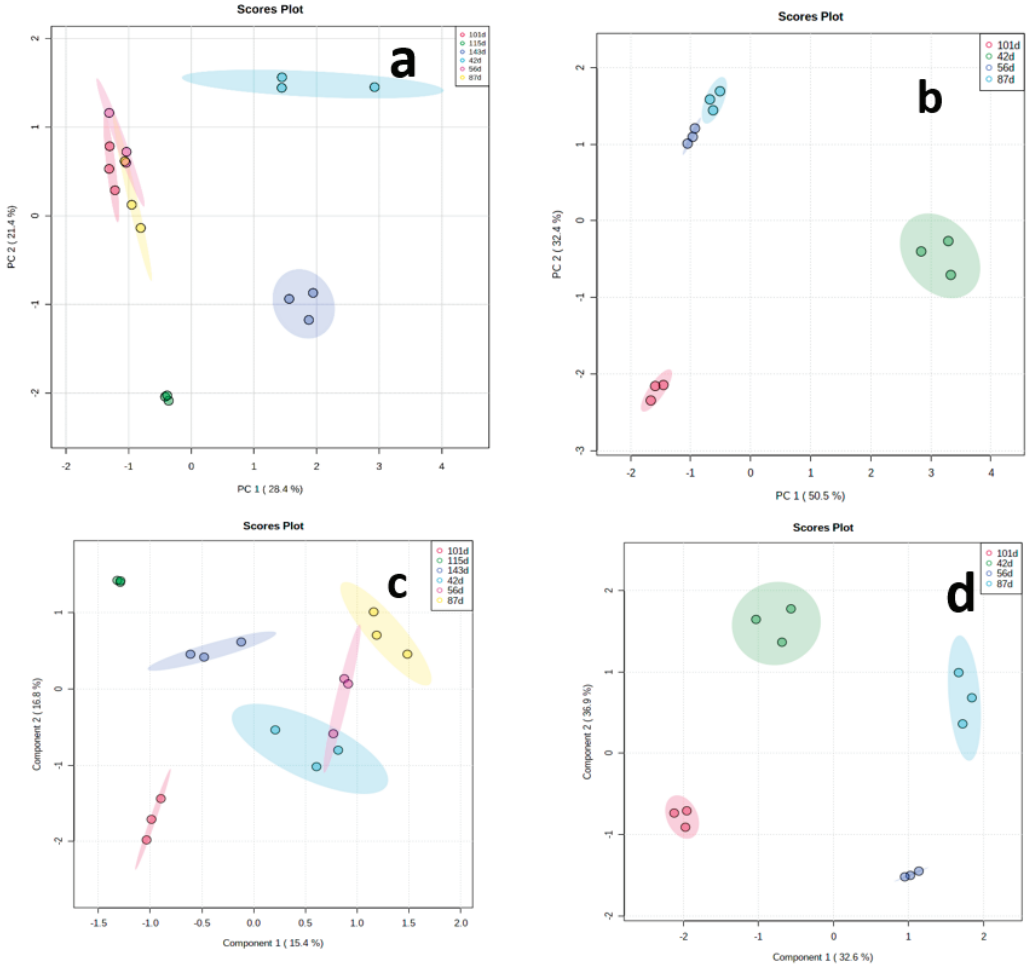


Figure 4. Score plots between the selected Principal components of PCA and components of PLS-DA showing the discrimination of tuber metabolomes of the two genotypes across the dormancy stages studied, and the explained variances are shown in brackets; (**a**) shows the PCA of *TDr1100873*, (**b**) shows the PCA of *Obiaoturugo*, (**c**) shows the PLS-DA of TDr110873, while (**d**) shows the PLS-DA of *Obiaoturugo*.

We performed partial least squares–discriminant analysis (PLS–DA) to further access the performance of the discrimination of the tuber metabolome under a supervised technique across the studied stages, using the very importance in projection (VIP) scores to evaluate the pattern of differential accumulation of important metabolites identified by PLS–DA across the dormancy stages (Figures S2 and S3). Figure 4c shows the discriminating of *TDr110873* tuber metabolome under supervised classification technique across the investigated dormancy stages. Component 2 separated tuber metabolome at 115-DAPM and 143-DAPM from the metabolome in the rest of the dormancy stages, and it accounts for 16.8% of the total variations, while component 1, which accounts for 15.4% variation, separated the tuber metabolome at 101-DAPM from 42-DAPM, 56-DAPM, and 87-DAPM.

Figure 4d shows the discriminating pattern of *Obiaoturugo* tuber metabolome across the studied dormancy stages under supervised classification. It indicates that the *Obiaoturugo* metabolome at each stage represents a distinct dormancy progression stage unlike what was observed with PCA. The separations were driven by components 1 and 2 of the PLS–DA scores plot which accounts for 32.6% and 36.9% of the total variations, respectively. The studied dormancy stages exhibited a consistent pattern in their separation into groups of clusters. The separation into distinct metabolome group or clustering into groups of similar metabolomes was determined mainly by the time duration between dormancy stages and dormancy phenotype, which necessitate change in metabolome composition.

3.3. Identification of Differentially Accumulated Metabolites across the Studied Tuber Dormancy Stages

To identify the differential accumulated metabolites (DAMs) across tuber dormancy stages studied, we performed agglomerative hierarchical clustering (AHC) analysis which partitioned DAMs earlier identified by ANOVA and PLS–DA VIP scores into the pattern of the exact change in their concentrations across the different dormancy stages using Euclidean distance measure and ward. D clustering algorithm. If a metabolite plays a role in the biological process, its concentration normally fluctuates during the process to adapt to the physiological changes in the process. First, AHC grouped the studied dormancy stages of *TDr1100873* tuber into two main clusters; cluster 1 comprises of 143-DAPM, 115DAPM, and 42-DAPM, cluster 2 comprises of 101-DAPM, 87-DAPM, and 56-DAPM (Figure 5a). In *Obiaoturugo*, the only distinct cluster comprises of 56-DAPM and 87-DAPM, while 42-DAPM and 101-DAPM were separately linked to it (Figure 5b). This was a deviation from the normal three curve stages in the germination process of botanic seed. However, this was expected because the physical characteristics of dormant botanic seeds that influence its growth curve exist in yam tuber. While botanical seeds are in the desiccation state at maturity, necessitating their requirement of imbibition during the germination process, yam tubers at maturity are in the hydrated state, and do not require imbibition during germination. Instead, as we reported earlier [44], it requires moisture loss to a 40% threshold before the dormancy-breaking process can commence. This might be linked to this sharp deviation of yam tuber germination stages from well-established three curve stages of the germination process in botanical seeds.

Agglomerative hierarchical clustering of the DAMs revealed the dynamic changes in the tuber metabolomes of the two genotypes (*TDr1100873 and Obiaoturugo*) across the studied dormancy stages. It identified 32 DAMs in tubers of *TDr1100873*, and 34 DAMs in tubers of *Obiaoturugo* (Figure 5a,b). Figure 5a shows that in *TDr1100873*, eight DAMs which comprised of aromatic amine, (two) alcohols and polyols, organosiloxane, amino acid and derivatives, alkaloid, biogenic polyamine, coumarin were upregulated, and 15 DAMs that comprised of nucleotide (adenine), (two) alkaloids, amino acid and derivative, galactonolactone (sugar acid), flavonoids, fatty acid, (three) carboxylic acids, coumarin, (two) alcohols and polyols, ester, benzene derivatives and phytohormone, were down-regulated at 42-DAPM. At 56-DAPM, nine DAMs which comprised of aromatic amine, (two) amino acids and derivatives, (two) alcohols, flavonoid, organosiloxane, biogenic polyamines and ester were up-regulated, and seven DAMs that comprised of benzene derivatives, phenol, sugar (sucrose), benzoquinone, alcohol and polyol, ester, and carboxylic acid were down-regulated. At 87-DAPM, 10 DAMs which comprised of fatty ester, biogenic polyamine, flavonoid, amines, alcohol, phenol, lipid, coumarin, fatty acid, benzene derivatives were up-regulated, and seven DAMs that comprised of, alkaloids, (two) nucleotides (adenine and nucleic acid), aromatic amine, alcohol and polyol, amino acid and derivatives, organosiloxane were down-regulated. At 101-DAPM, six DAMs which comprised of fatty ester, terpene, alcohol, (two) nucleotides (adenine and nucleotide sugar acid) and alkaloid, were up-regulated, alkaloids, and seven DAMs that comprised of alcohol, flavonoid, amine, biogenic polyamine, alkaloid, phenol, phytohormone (salicyluric acid) down-regulated. At 115-DAPM, seven DAMs which comprised of sugar (sucrose),

benzoquinone, (two) carboxylic acids, fatty acid, lipid, phytohormone (salicyluric acid) were up-regulated, and 12 DAMs that comprised of fatty ester, biogenic polyamine, terpene, (two) amino acid and derivatives, (three) alcohols and polyols, (two) flavonoids, fatty acid, organosiloxane were down-regulated. At 143-DAPM, 12DAMs which comprised of (two) flavonoids, sugar (sucrose), (two) alkaloids, carboxylic acid, phytohormone (salicyluric acid), biogenic Polyamines, benzene derivatives, amino acid, benzoquinone were up-regulated, whereas seven DAMs that comprised of aromatic amine, alcohol, amino acids and derivatives, sugar acid, flavonoid, fatty acid, lipid were down-regulated.

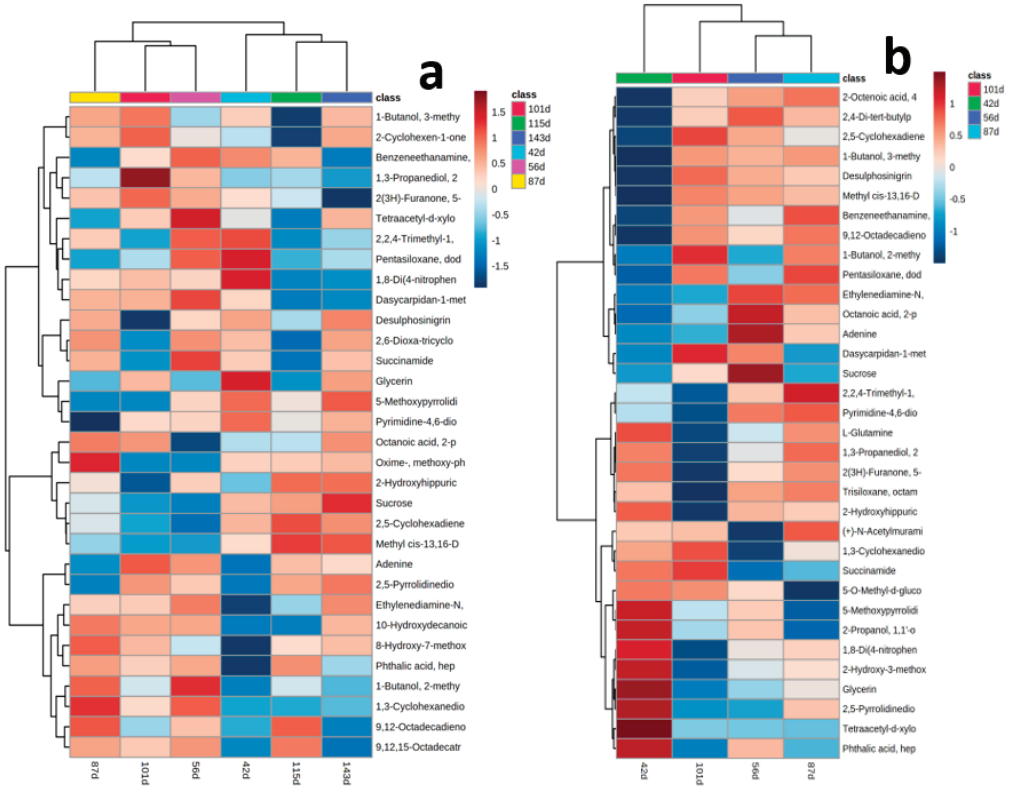


Figure 5. Agglomerative Hierarchical Cluster heatmap of DAMs across the studied dormancy stages in each of the genotypes. Metabolites and dormancy stages were separated into clusters using distance measure Euclidean distance measure, and clustering algorithms ward. linkage. The relative concentration of each metabolite at different studied dormancy stages was standardized, ranked using Spearman's rank correlation, and shown in the box according to the color criterion indicating their log fold change in concentration at each dormancy stage. The log fold change values according to the color criterion and dormancy stages are shown on the legends at the right of the figures. (**a**) shows the hierarchical cluster heatmap of DAMs in the tubers of *TDr1100873* across the studied dormancy stages, while (**b**) shows the hierarchical cluster heatmap of DAMs in the tubers of *Obiaoturugo* across the studied dormancy stages.

Figure 5b shows that in *Obiaoturugo*, 15 DAMs which comprised of (three) amino acids and derivatives, (three) alcohols and polyols, (two) alkaloids, (two) carboxylic acids, (two) phytohormones (salicyluric and 3-methylsalicylic acids), sugar acid, amino alcohol, biogenic polyamine were up-regulated, and 14 DAMs that comprised of (two) carboxylic acids, (two) esters, (two) flavonoids, phenol, fatty acid, aromatic amine, organosiloxane,

benzoquinone, down-regulated at 42-DAPM. At 56-DAPM, seven DAMs which comprised of amine, (two) flavonoids, benzene derivatives, (two) nucleotides (adenine and nucleic acid), sugar (sucrose) were up-regulated, and five DAMs that comprised of ester, nucleotide (nucleotide sugar), alcohol, biogenic polyamine, and alkaloid were down-regulated. At 87-DAPM, 11 DAMs which comprised of (two) nucleotides (nucleic acid and nucleotide sugar), benzoquinone, (two) esters, aromatic amine, fatty acid, carboxylic acid, flavonoid, alcohol, organosiloxane were up-regulated, and four DAMs that comprised of flavonoids, phytohormone (methylsalicylic acid), alkaloids, amino alcohol were down-regulated. At 101-DAPM, eight DAMs which comprised of (two) flavonoids, benzoquinone, ester, carboxylic acid, alcohol organosiloxane and biogenic polyamines were up-regulated, and nine DAMs that comprised of phytohormone (methysalicylic acid), amino acid and derivatives, alcohols and polyols, sugar acid, inorganic oxide, (two) carboxylic acids were down-regulated. This result demonstrates that amines, biogenic polyamines, amino acids and derivatives, alcohols, flavonoids, alkaloids, organosiloxane, phenols, esters, coumarins and phytohormone dominated the metabolic activities during the dormancy period, whereas, fatty acids, lipids, nucleotides, carboxylic acids, sugars, amino acid (glutamine), terpenoids, benzoquinones, and benzene derivatives dominated the metabolic activities during tuber dormancy breaking and sprouted tubers of both yam genotypes. However, polyamine accumulated in the tubers of both genotypes during the deep dormant stages and in sprouted tubers, indicating that it might be playing a role in both dormancy induction and yam vine development.

3.4. Metabolites Mapping and Chemical Functional Groups Identification

To perform further analysis with the identified DAMs the compound labels, or identity need to be standardized. The raw metabolites labels were used as queries in mapping DAMs into HMDB, PubChem, and KEGG databases to identify their standardized compound names, chemical IDs, and functional groups. It was observed that the DAMs were distributed across sixteen (13) major chemical functional groups, comprising primary and secondary metabolisms. Alcohols and polyols were the most abundant DAMs, and it constitutes 12.82% proportion of the total DAMs, followed by fatty acids and esters, and amines and polyamines as each of them constitute 10.26% (Figure 6). Table 1 shows the conversion of the raw DAMs labels to standard compound names, their chemical functional groups, and IDs in HMDB, PubChem, and KEGG.

Table 1. Result of mapping of DAMs in HMDB, PubChem, and KEGG for standard name, chemical functional group and IDs determination.

Query	Match	Chem-Fun Group	HMDB	Pub Chem	KEGG
9,12,15-Octadecatrienoic acid, (Z,Z,Z)-	Alpha-Linolenic acid	Fatty acids	HMDB0001388	5280934	C06427
Ethylenediamine-N,N′-dipropionic acid	Edetic Acid	Flavonoids	HMDB0015109	6049	C00284
1-Butanol, 2-methyl-, acetate	Methyl methacrylate	Esters	HMDB0032385	6658	C19504
Tetraacetyl-d-xylonic nitrile	Sarcosine	Amino Acids and Derivatives	HMDB0000271	1088	C00213
(+)-N-Acetylmuramic acid	UDP-N-acetylmuraminate	nucleotide-sugar	HMDB0011720	24755495	C01050
1-Butanol, 3-methyl-, formate	3-Methylbutyl formate	Fatty esters	HMDB0034163	8052	C12293
1,3-Propanediol, 2-ethyl-2-(hydroxymethyl)-	1,3-Butanediol	Alcohols and Polyols	HMDB0031320	7896	C20335
1,8-Di(4-nitrophenylmethyl)-3,6-diazahomoadamantan-9-one	Tyramine	Amino Acids and Derivatives	HMDB0000306	5610	C00483
Benzeneethanamine, 2-fluoro-á,3,4-trihydroxy-N-isopropyl-	Phenylethylamine	Amines, Aromatic	HMDB0012275	1001	C05332
Desulphosinigrin	Delphinidin 3-O-sophoroside	Flavonoids		47205615	C16307
Glycerin	Glycerol	Alcohols and Polyols	HMDB0000131	753	C00116
9,12-Octadecadienoic acid (Z,Z)-	Linoleic acid	Lipids	HMDB0000673	5280450	C01595
Dasycarpidan-1-methanol, acetate (ester)	Acetaminophen	Flavonoids	HMDB0001859	1983	C06804
2,2,4-Trimethyl-1,3-pentanediol diisobutyrate	3-Phenylpropyl 2-methylpropanoate	Alcohols	HMDB0034472	7662	C02008
Sucrose	Sucrose	Disaccharides	HMDB0000258	5988	C00089
5-O-Methyl-d-gluconic acid dimethylamide	3-Cresotinic acid	Phytohormone	HMDB0002390	6738	C14088

Table 1. *Cont.*

Query	Match	Chem-Fun Group	HMDB	Pub Chem	KEGG
2,6-Dioxa-tricyclo [3.3.2.0(3,7)] decan-9-one	Ibuprofen	Amines	HMDB0001925	3672	C01588
1,3-Cyclohexanedione, 2,5,5-trimethyl-	1,4-Cyclohexanedione	Alcohols		10263	C08063
5-Methoxypyrrolidin-2-one	Galactosylglycerol	Alkaloids	HMDB0006790	656504	C05401
2(3H)-Furanone, 5-heptyldihydro	Galactonolactone	Sugar Acids	HMDB0002541	5640	C03383
2-Propanol, 1,1′-oxybis-	Propranolol	Amino Alcohols	HMDB0001849	4946	C07407
10-Hydroxydecanoic acid	12-Hydroxydodecanoic acid	Fatty Acids	HMDB0002059	79034	C08317
Succinamide	Putrescine	Biogenic Polyamines	HMDB0001414	1045	C00134
2-Octenoic acid, 4,5,7-trhydroxy	Methyl acrylate	Carboxylic Acids	HMDB0033977	7294	C19443
2-Hydroxy-3-methoxy-succinic acid, dimethyl ester	L-Malic acid	carboxylic acid	HMDB0000156	222656	C00149
Octanoic acid, 2-phenylethyl ester	Styrene	Benzene Derivatives	HMDB0034240	7501	C19506
2-Cyclohexen-1-one, 4-(3-hydroxy-1-butenyl)-3,5,5-trimethyl-	Amoxycillin	Terpenoids	HMDB0030500	2171	C06827
8-Hydroxy-7-methoxycoumarin	Oxybenzone	Coumarins	HMDB0015497	4632	C14285
Pyrimidine-4,6-diol, 5-methyl-	Uracil	Nucleic Acids and Derivatives	HMDB0000300	1174	C00106
Oxime-, methoxy-phenyl-_	3,4-Dihydroxyphenylacetaldehyde	phenols	HMDB0003791	119219	C04043
Adenine	Adenine	Nucleotides	HMDB0000034	190	C00147
2,5-Cyclohexadiene-1,4-dione, dioxime	Quinone	Benzoquinones	HMDB0003364	4650	C00472
2,5-Pyrrolidinedione, 3-(1-aminoethylidene)-4-methyl-	Isopyridoxal	Alkaloids	HMDB0004290	440899	C06051
Phthalic acid, hept-4-yl isobutyl ester	Thalidomide	Carboxylic Acids	HMDB0015175	92142	C07910
L-Glutamine	L-Glutamine	Amino Acids and Derivatives	HMDB0000641	5961	C00064
Methyl cis-13,16-Docosadienate	Methylitaconate	carboxylic acid	METPA0268		C02295
2-Hydroxyhippuric acid-3TMS	Salicyluric acid	Phytohormone	HMDB0000840	10253	C07588
2,4-Di-tert-butylphenoxytrimethylsilane	Butenafine	Amines	HMDB0015223	2484	C08067
Pentasiloxane, dodecamethyl-	Dodecamethylpentasiloxane	organosiloxane	HMDB0062731	8853	

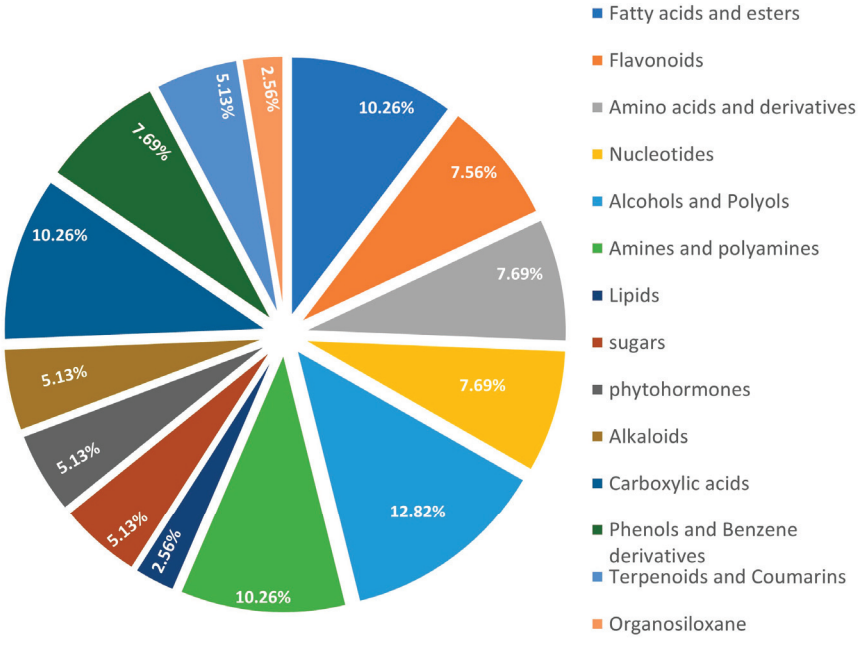


Figure 6. Proportional composition of chemical functional groups of the DAMs. The color and area of each component of the pie chart show the percentage composition of a particular chemical functional group.

3.5. Metabolites Sets Enrichment Analysis and Pathways Annotation

We performed metabolite set enrichment analysis using a hypergeometric test to determine whether the metabolism of any set of DAMs is represented more than expected by chance during the studied yam tuber dormancy stages. It was found that 12 metabolisms

were significantly enriched at ($p \leq 0.05$) during the studied yam tuber dormancy stages (Figure 7). The enrichment ratio of each of the induced metabolism denoted by the length of its bar length indicates the fold change in a particular metabolism from normal, while the *p*-value indicates determines whether a particular metabolism has a significant impact on the biological process under consideration. Linoleic acid metabolism recorded the highest enrichment ratio of over 20 at $p = 0.02$, an indication that linoleic acid metabolism plays the most important role in yam tuber dormancy regulation. On the other hand, other metabolisms, such as galactose metabolism, biosynthesis of unsaturated fatty acids, pyrimidine metabolism, nitrogen metabolism, purine metabolism, D-glutamine and D-glutamate metabolism, phenylalanine metabolism, alpha-linoleic acid metabolism, Arginine biosynthesis, Glycerolipid metabolism, and starch and sucrose metabolism play some roles in yam tuber dormancy regulation.

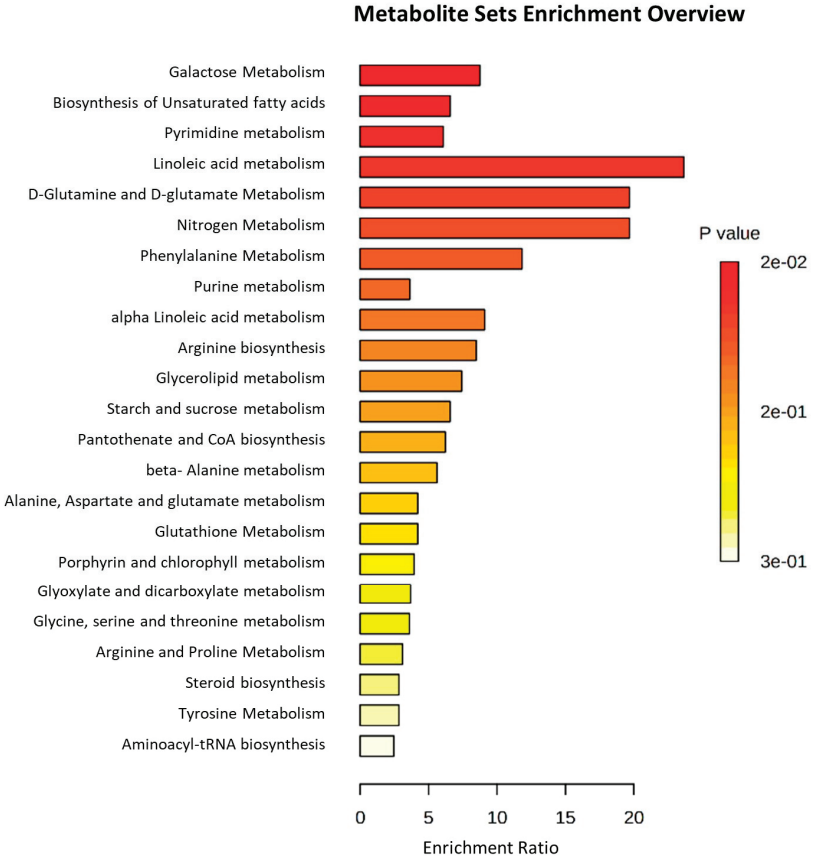


Figure 7. Bar plots for metabolites set enrichment analysis showing the top 23 metabolisms of enriched metabolites. The enrichment ratio values are the ratio between the fold change in the metabolisms across the studied dormancy stages and their expected values under normal conditions. The color criterion indicates the *p*-value at which each metabolism was induced to its enrichment ratio, and it determines whether enrichment is significant or not.

3.6. Pathway Topology Analysis

Since the metabolic network is directed, we used the relative betweenness centrality of the metabolites in a node to determine the impact of the pathway on yam tuber dormancy regulation. The result indicates that seven metabolic pathways (linoleic acid metabolic path-

way, phenylalanine metabolic pathway, galactose metabolic pathway, starch and sucrose metabolic, alanine, aspartate, and glutamate metabolic pathway, purine metabolic pathway, and lysine biosynthesis) had a significant impact on yam tuber dormancy regulation, with linoleic acid metabolic pathway exacting maximum impact of 1 (Figure 8). Two pathways (Aminoacy-tRNA biosynthesis, and porphyrin and chlorophyll metabolism) were also observed to be suppressed during tuber endodormancy dormancy stages. Generally, the seven significantly induced pathways can be grouped into three major functional categories, which include: lipids metabolism, energy and nucleotides metabolism, and amino acids metabolism. Lipids and energy nucleotide metabolisms were majorly involved in yam tuber dormancy-breaking activities and their metabolites were preponderance during tuber dormancy-breaking stages. While amino acids metabolisms dominated the activities during yam tuber endodormancy dormancy stages. Table 2 presents the summary of all induced pathways, their log-fold change and impacts on yam tuber dormancy regulation.

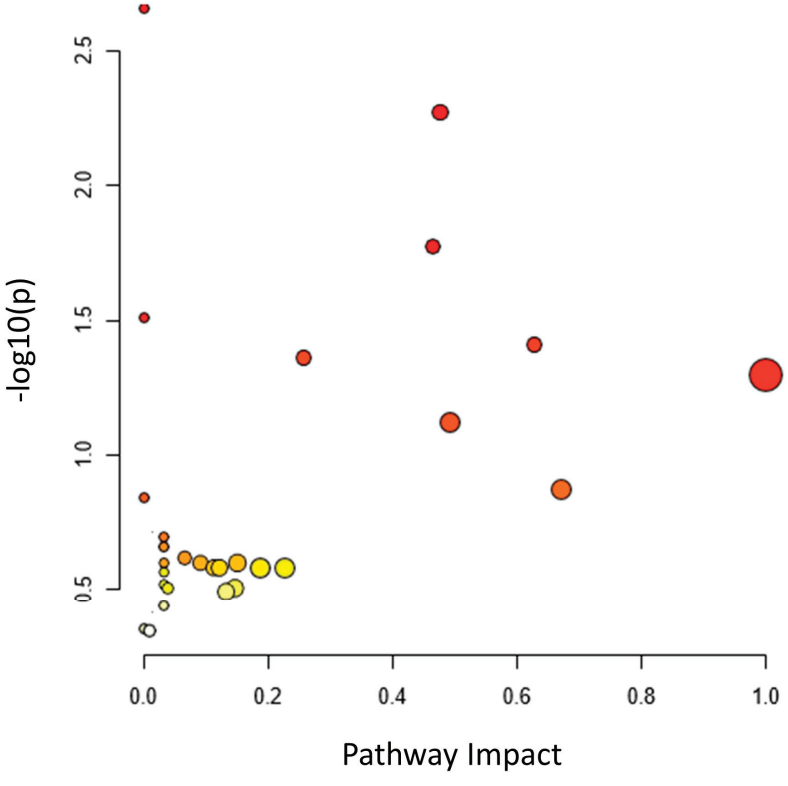


Figure 8. Metabolome view of pathway topology showing the pathways that significantly impacted yam tuber dormancy regulation. The dots denote metabolic pathways, size of the dots indicates the value of the impact of the pathway on yam tuber dormancy regulation on a scale of 0 to 1, whereas, the color intensity of the dots indicates the fold change of the pathway.

Table 2. Result of pathway topology analysis, showing the significantly induced pathways and their impact on yam tuber dormancy regulation.

	Total	Expected	Hits	Raw *p*	−log10 (*p*)	Holm Adjust	FDR	Impact
Isoquinoline alkaloid biosynthesis	6	0.08	2	2.26×10^{-3}	2.65×10	2.15×10^{-1}	2.03×10^{-1}	0.00
Galactose metabolism	27	0.35	3	4.28×10^{-3}	2.37×10	4.02×10^{-1}	2.03×10^{-1}	0.76
Tyrosine metabolism	18	0.23	2	2.11×10^{-2}	1.68×10	1.00×10	6.67×10^{-1}	0.04
Biosynthesis of unsaturated fatty acids	22	0.28	2	3.09×10^{-2}	1.51×10	1.00×10	7.33×10^{-1}	0.00
Linoleic acid metabolism	4	0.05	1	5.04×10^{-2}	1.30×10	1.00×10	8.15×10^{-1}	1.00
Glyoxylate and dicarboxylate metabolism	29	0.37	2	5.14×10^{-2}	1.29×10	1.00×10	8.15×10^{-1}	0.06
Pyrimidine metabolism	38	0.49	2	8.32×10^{-2}	1.08×10	1.00×10	1.00×10	0.06
Lysine biosynthesis	9	0.12	1	1.10×10^{-1}	9.59×10^{-1}	1.00×10	1.00×10	0.47
Nitrogen metabolism	12	0.15	1	1.44×10^{-1}	8.41×10^{-1}	1.00×10	1.00×10	0.54
Phenylalanine metabolism	12	0.15	1	1.44×10^{-1}	8.41×10^{-1}	1.00×10	1.00×10	0.79
Purine metabolism	63	0.81	2	1.92×10^{-1}	7.17×10^{-1}	1.00×10	1.00×10	0.26
Arginine biosynthesis	18	0.23	1	2.09×10^{-1}	6.81×10^{-1}	1.00×10	1.00×10	0.00
beta-Alanine metabolism	18	0.23	1	2.09×10^{-1}	6.81×10^{-1}	1.00×10	1.00×10	0.00
Citrate cycle (TCA cycle)	20	0.26	1	2.29×10^{-1}	6.40×10^{-1}	1.00×10	1.00×10	0.33
Zeatin biosynthesis	21	0.27	1	2.39×10^{-1}	6.22×10^{-1}	1.00×10	1.00×10	0.00
Carbon fixation in photosynthetic organisms	21	0.27	1	2.39×10^{-1}	6.22×10^{-1}	1.00×10	1.00×10	0.06
Glycerolipid metabolism	21	0.27	1	2.39×10^{-1}	6.22×10^{-1}	1.00×10	1.00×10	0.26
Phenylalanine, tyrosine, and tryptophan biosynthesis	22	0.28	1	2.49×10^{-1}	6.04×10^{-1}	1.00×10	1.00×10	0.08
Starch and sucrose metabolism	22	0.28	1	2.49×10^{-1}	6.04×10^{-1}	1.00×10	1.00×10	0.59
Pyruvate metabolism	22	0.28	1	2.49×10^{-1}	6.04×10^{-1}	1.00×10	1.00×10	0.24
Alanine, aspartate and glutamate metabolism	22	0.28	1	2.49×10^{-1}	6.04×10^{-1}	1.00×10	1.00×10	0.57
Pantothenate and CoA biosynthesis	23	0.30	1	2.59×10^{-1}	5.87×10^{-1}	1.00×10	1.00×10	0.00
Cyanoamino acid metabolism	26	0.33	1	2.87×10^{-1}	5.42×10^{-1}	1.00×10	1.00×10	0.00
Glutathione metabolism	27	0.35	1	2.97×10^{-1}	5.28×10^{-1}	1.00×10	1.00×10	0.01
alpha-Linolenic acid metabolism	27	0.35	1	2.97×10^{-1}	5.28×10^{-1}	1.00×10	1.00×10	0.11
Arginine and proline metabolism	28	0.36	1	3.06×10^{-1}	5.14×10^{-1}	1.00×10	1.00×10	0.16
Glycine, serine, and threonine metabolism	33	0.42	1	3.50×10^{-1}	4.56×10^{-1}	1.00×10	1.00×10	0.00
Aminoacyl-tRNA biosynthesis	46	0.59	1	4.53×10^{-1}	3.44×10^{-1}	1.00×10	1.00×10	0.00
Porphyrin and chlorophyll metabolism	47	0.60	1	4.61×10^{-1}	3.37×10^{-1}	1.00×10	1.00×10	0.01

Total is the number of compounds in the pathway, **expected** is the normal ratio value, **hits** are the matched number from the uploaded data, **Raw *p*** is the *p*-value before adjustment, **Holm *p*** is the *p*-value adjusted using the Holm–Bonferroni method. **FDR** is the *p* value adjusted using a false discovering rate. **Impact** is the effect of the pathway on a biological process (yam tuber dormancy regulation).

4. Discussions

Dormancy in yam tuber is a complex process that is synergistically modulated by molecular processes and environmental signals. Physiologically, during the period of dormancy to sprouting, yam tuber transit into three modes of nutrition; from heterotrophic to autotrophic and back to heterotrophic. Metabolites undoubtedly play essential roles during these processes, since they contribute the raw materials for cellular structure composition, energy supply, biochemical reactions, and signaling which are required for the completion of the physiological process of dormancy induction and breaking [18,28]. Usually, upon dormancy induction some specific metabolites are accumulated and some signals activated to maintain low energy conditions in yam tuber that characterize the dormancy period, whereas during dormancy breaking, varieties of metabolites and metabolisms are activated to provide nutrients and energy required for tuber dormancy breaking [7,45]. However, elucidating the key metabolites and metabolic pathways that regulate yam tuber dormancy, to the best of our knowledge has not been done previously, although the functions of some metabolites in botanical seeds have been established [32,46–48]. Our study used untargeted metabolomics profiling to determine some key metabolites and their associated metabolic pathways that regulate yam tuber dormancy.

4.1. Efficiency of Metabolomics in Deciphering Molecular Mechanisms in a Biological System

In recent years, metabolomics has been frequently used in the study of seed dormancy [28,46,49]. In *Davidia involucrate*: a dove tree that is common in China, which has very recalcitrant seeds that stay in a dormant state for about four years, metabolome profiling identified 48 differentially accumulated metabolites (DAMs) which were enriched in purine metabolism, pyrimidine metabolism, arginine, and proline metabolism, flavone and

flavonol biosynthesis, phenylpropanoid biosynthesis, arginine biosynthesis pathways, as well as key phytohormones, abscisic acid, indole-3 acetic acid and sinapic acid that were implicated in the seed dormancy regulation [32]. In rice, comparative metabolic profiling identified 74 candidate metabolites, among which 29 belong to the ornithine-asparagine-polyamine module and the shikimic acid-tyrosine-tryptamine-phenylalamine-flavonoid module. The result revealed that shikimic acid promoted seed dormancy breaking [18]. In Arabidopsis, metabolomic and transcriptomic profiling identified two distinctive profiles involved in the metabolic regulation of seed dormancy breaking and seedling establishment [35]. The distribution of metabolites in two inbred lines (B73 and Mo17) of maize seeds revealed that the two inbred lines were highly differentiated in their metabolite profiles during dormancy breaking, with regard to amino acids, sugar alcohols, and organic acids accumulation, which determined their dormancy phenotype [50]. In the present study, we performed a comparative metabolomics profiling of tubers of two white yam genotypes at different time points during dormancy to dormancy breaking. We observed that a total of 39 metabolites were differentially accumulated in tubers of the two yam genotypes during the study period. The 39 DAMs were dominated by amino acids and derivatives, alkaloids and alcohols, sugars, nucleotides, amines and polyamines, fatty acids and esters, phenols and benzene derivatives, and organic acids. Therefore, this study demonstrated the efficiency of our metabolomic profiling approach in identifying differentially accumulated metabolites, which revealed the dynamic physiological processes in white yam tubers from dormancy to sprouting.

4.2. Differentially Accumulated Metabolites Determined the Phenotypic Variation in Dormancy Duration of the Two White Yam Genotypes

Our data showed genotype-specific variation in differentially accumulated metabolites during tuber dormancy and dormancy break of the two white yam genotypes (*TDr1100873* and *Obiaoturugo*). The tubers of *TDr1100873* had 5 unique DAMs that spread between dormancy and dormancy break periods, and exhibited long dormancy phenotype. The 5 unique DAMs observed in the tubers of *TDr1100873* comprised of terpenoids, amines, phenol, and 2 fatty acids. Among the 5 unique DAMs observed in tubers of *TDr110873*, terpenoid was accumulated at 101-DAPM, amine accumulated at 56-DAPM, while phenol and one of the fatty acids accumulated at 87-DAPM, all these time points, except 101-DAPM constitute the dormant period in *TDr1100873*. *Obiaoturugo* had 7 unique DAMs (nucleotide sugar, L-malic acid, amino alcohol, amine, phytohormone, carboxylic acid, and amino acid and derivatives) that also spread between dormancy and dormancy break periods and exhibited short dormant phenotype. Among the 7 unique DAMs observed in the tubers of *Obiaoturugo*, nucleotide sugar, amino acid, and carboxylic acid were all upregulated at 87-DAPM which is the dormancy breaking point in *Obiaoturugo*. The role of amines in dormancy regulation has been shown to be species-specific. In some plants, they promote dormancy breaking (germination), and in others, they induce dormancy, [46,51,52]. Its accumulation in the tuber of long-dormant phenotype is indicative that it induced or at least functioned to prolong yam tuber dormancy. The study has shown phenol as an antioxidant that participates in the scavenging of reactive oxygen species (ROS) thus catalyzing oxygenation reactions through the formation of metallic complexes, and inhibiting the activities of oxidizing enzymes [53]. Its accumulation at the point when the tubers of short dormant genotype-initiated dormancy breaking this suggests it plays a vital role in the process. Conversely, the presence of unique nucleotide sugar and amino acid (glutamine) that accumulated at 87-DAPM which marked the endodormancy breaking point suggests that these metabolites play some roles in determining early dormancy breaking. It has been demonstrated that nucleotide sugar plays a regulatory role in cell division. A strong correlation was observed between the supply of Glc and the expression of cyclins (*cycD2;1, D3;2, A3;2,* and *B1;2*) [54,55], that the accumulation of nucleotide sugar in the tubers of short dormant phenotype might have triggered the early reactivation of cell division and consequently dormancy break. Similarly, it has been reported that L-Glutamine is a di-

rect precursor of Glutamate, and Glutamate occupies a central position in amino acid metabolism in plants and as well plays key role initiation of dormancy breaking process in seeds [56]. We infer that 2 metabolites (amine and phenol) out of the 5 unique DAMs in *TDr1100873* tubers might be playing some role in maintaining prolonged dormancy duration, whereas, 2 metabolites (nucleotide sugar and amino acid) out of the 7 unique DAMs in *Obiaoturugo* tubers might be playing some roles in early dormancy break. The amine and phenol might have induced and maintained dormancy through their antioxidative ability. Since dormancy breaking process generally generates ROS, and ROS has been reported to promote dormancy breaking by oxidation of NADP and release of energy (ATP) required for dormancy breaking process [57] It can be hypothesized that any process that suppresses any process that generates energy in the tuber during dormancy will enhance dormancy maintenance. Conversely, nucleotide sugar and amino acid might promote dormancy break by enhancing energy availability and increasing the nitrogen/carbon ratio which are the primary requirement for dormancy break. Our results agreed with the finding of [58] who reported that composition and changes in concentration of metabolites in the seeds of two wild barley. Our study buttresses the fact that variation in the duration of tuber dormancy could be driven by the absence of nucleotide sugar and nucleic amino acid (L-Glutamine) in the tubers of long dormancy duration genotype (*TDr1100873*).

4.3. Metabolic Regulation of Yam Tuber Dormancy

Using three typical methods; ANOVA, PLS-DA, and AHC, for comparative metabolome analysis at different time points during tuber dormancy of the two white yam genotypes, we identified 39 candidates metabolites associated with tuber dormancy regulation, distributed across 14 well-known dormancy regulating chemical functional groups such as: amines and biogenic polyamines, amino acids and derivatives, carboxylic acids, sugars, Alkaloids, alcohols and polyols, flavonoids, phenols, and benzene derivatives, nucleotides, fatty acids and esters, lipids, coumarins and terpenoids, and phytohormones [7,26,59,60]. We did not identify any of the two established dormancy-regulating phytohormones; Abscisic acid (ABA) and Gibberellic acids (GA) directly, although, their precursors were observed in some of the metabolic pathways enriched. Our data showed the dynamic changes in the concentrations of the identified DAMs across the studied dormancy stages of the tubers of the two white yam genotypes in response to different physiological processes that characterized each of the studied dormancy progression stages. Metabolites set enrichment analysis revealed that these DAMs were involved in 12 significantly induced metabolisms during the white yam tuber dormancy-to-dormancy break period. These observations can be further grouped into amines and amino acids metabolisms, secondary metabolites metabolisms, energy metabolisms, and nucleotides metabolisms.

4.4. Amines and Amino Acids Metabolisms

Generally, most the amines and amino acids accumulated in the tubers of both yam genotypes during dormant stages, except for L-Glutamine that accumulated during the dormancy breaking stage in the tubers of the short dormant genotype. The preponderance of amines and amino acids during the yam tuber dormancy stage is an indication that the cells of yam tubers during dormancy were in the post-replication phase, implying that replication which is the hallmark of the cell cycle and consequently growth process was halted during this period. It has been reported that active dormancy induction is characterized by the arrest of the cell cycle process at the G1 phase [54]. Studies have also revealed that amino acids have prominent functions in plant growth and development, besides their role as a protein building block, they as well play pivotal roles in the biosynthesis of secondary metabolites, and during signaling processes in plant stress response [61,62]. These functions of amino acids, amines, and polyamines have been shown to be specie-specific, in some plant they promote dormancy breaking (germination), and in others, they induce dormancy [52,63]. The molecular mechanism of exerting any of the dormancy regulatory action (either induction or inhibition) differs too. In dormancy

induction, it seems that amino acids and amines utilize the same molecular module that they use when they act as osmo-protectant during plant abiotic and biotic stress response and biosynthesis of secondary metabolites. It has been demonstrated that amino acids have antioxidant properties and act as reactive oxygen species (ROS) scavenger during plant response to stressful conditions [64]. ROS also serve as signaling molecules during plant growth and in the development and responses to biotic and abiotic stresses [65]. ROS had been shown to directly act on plant cell-wall polysaccharides by breaking their glycosidic bonds, which cause the cell wall to relax, thereby promoting cell division and elongation, which marks the breaking of endodormancy, [66]. During seed germination of dicotyledonous plants, such as cress and lettuce, ROS accumulation in their micropylar endosperm and radicle, led to endosperm weakening and radicle elongation, thereby promoting seed germination [67]. Certain cell-wall hydrolases and non-enzymatic substances, such as ROS and expansion, are needed for these processes [68]. Li et al. [57] have shown that ROS accumulated in the radicle and coleorhiza of rice seeds during germination and that ROS production and the number of germinated seeds decreased in the presence of ROS inhibitors, i.e., diphenyleneiodonium chloride and guazatine. We therefore infer that the ROS inhibition action of amino acids in yam tuber led to dormancy induction and maintenance, and hence they accumulated during tuber dormancy in the two-genotype studied. Previous studies have also made similar observations, in soybean (*Glaycin max*), high seed-specific expression of feedback-insensitive genes encoding dihydrodipicolinate synthase (DHPS) and aspartate kinase (AK) resulted in wrinkled seeds and low germination rates [69], while grain yield reduction was observed in high lysine crops [70]. Studies suggest that lysine also affect starch synthesis and endosperm development in maize (*Zea mays* Linn) and rice (*Orayze sative*) [70,71]. It mobilizes sucrose and other energy sources in a cell thereby leading to a state of low-energy condition that characterizes dormancy period. Also, seed mobilization of minor proteins such as seed maturation and dormancy-associated proteins that are degraded during imbibition to provide primary amino acids during germination have been reported [72,73]. In red rice, pyruvate in arginine-proline metabolism was induced in dormant seeds, and implicated in seed dormancy induction and maintenance [74]. In the Dove tree (*D. involucrata Baill*), the seeds that do not germinate after stratification treatment had the strongest pyruvate– arginine– proline metabolism induction, suggesting that this metabolic pathway might be a key contributor to dormancy induction [32]. In strawberries, polyamine in ABA-dominated, IAA, and ethylene-participating crosstalk inhibited seed ripening and germination [75]. In the present study, our pathway topology analysis also revealed that arginine-proline was significantly induced during yam tuber dormancy (Table 2). Our pathway topology analysis also revealed that lysine metabolism leads to the biosynthesis of secondary metabolites such as; tropane, piperidine, and pyridine alkaloids, which are known for dormancy induction and maintenance in the seed.

Polyamine (putrescine) has also been implicated in other plant physiological processes such as senescence induction, and stress tolerance [76,77]. It has been reported that senescence induction in yam crops is the physiological marker of dormancy onset [7], hence, it is reasonable to hypothesize that putrescine can induce tuber dormancy through the same molecular module that it uses to induce senescence. In the current study, putrescine accumulated in the tubers of both yam genotypes during the dormancy stage, and was suppressed as dormancy progresses towards breaking (germination). Phenylalanine is another important amine that accumulated during dormancy in the present study. Although, the aromatic amine acts as a nitrogen source participates in the activation of cellular processes that induce germination, but its metabolic pathway is also a hub for the biosynthesis of diverse secondary metabolites, most of which counter its own growth-inducing action [78,79]. The study has also demonstrated that upregulation of alanine aminotransferase (*AlaAT*) isoenzymes encoded by long and short dormancy alleles differ in barley seed embryos led to the exhibition of different dormant phenotypes [80] The reduced dormancy allele *Qsd1* mutant enhanced early and uniform germination, whereas long dormancy allele *qsd1* mutant exhibited prolonged dormancy. In the current study, pathway topology analysis

revealed that phenylalanine accumulated during tuber dormancy stages, suggesting that the metabolite plays role in tuber dormancy induction. Our pathway topology analysis also revealed that phenylalanine was significantly induced during yam tuber dormancy, and it exerted a significant impact on yam tuber dormancy regulation.

Our result contrasted several studies that have demonstrated that the majority of amino acids and amines promote dormancy break (germination), and inhibit dormancy induction and maintenance in botanical seeds [52,81,82]. Ornithine and arginine are both nitrogenous amino acids that can form polyamines through decarboxylation [18]. In canola (*Brassica napus*), it was demonstrated that tyramine differentially accumulated in osmo-primed seeds under saline stress, and led to early seed germination [83]. Llebres, et al. [84] identified the genes involved in arginine metabolism in *Pinus pinaster* Ait and found that arginine plays important role in seed germination in *Pinus pinaster*. In cotton, nitrogen flow occurs in the seed during seed formation and germination through asparagine —> arginine —> storage protein —> arginine —> asparagine sequence [85]. The shikimic acid–tyrosine–tryptamine–phenylalanine–flavonoid module is closely related to the pathway of phenylalanine, tyrosine, and tryptophan biosynthesis along with biosynthesis of phenylpropanoids. This module is also well known for its involvement in plant defense systems, especially shikimic acid (shikimate). Some of the members of this pathway have been reported to be involved in seed germination. In legume seeds, tryptamine was found to be the main biogenic amine detected, and its concentration considerably increased during the germination process [86], suggesting that it might be playing role in dormancy breaking. These aforementioned findings contradict the present result, because tyramine accumulated during the yam tuber deep dormancy stage in both genotypes, suggesting that it might be playing role in yam tuber dormancy induction. We also identified one amino acid (sarcosine) that accumulated during yam tuber dormancy, but has not been reported to have any role in plant growth regulation.

Our result demonstrated that amines and amino acids' role in dormancy induction and maintenance in yam tuber is due to their antioxidant property, and the action of secondary metabolites synthesized in their biosynthetic pathways (Figure 5). The contrast between our result and some of the previous reports of amines and amino acids actions in botanical seeds dormancy regulation could possibly be explained by two reasons namely; (1) The presence of seed coat and requirement of imbibition for dormancy break in botanical seeds, and non-requirement of imbibition for yam tuber dormancy break. In addition to developmental pathways in zygotic tissues, the seed coat plays a critical role in the control of dormancy in botanical seeds of many species. In some seeds, dormancy is known as coat imposed, and dormancy can be removed or reduced by simple removal of or damage to the seed coat. In various species the seed coat has been shown to restrict germination by regulating permeability to either water [87,88], oxygen [89], or germination inhibitors that leach from the seed coat into the embryo [90]. It has been reported that aromatic amines, and amino acids reduced the thickness and increased seed tegument porosity which helped in water imbibition and boost seed germination rate [91]. whereas in the yam tuber, the reverse is the case, as we earlier reported that yam tuber required moisture reduction to a critical threshold of 40% before the dormancy-breaking process could be initiated [44]. The study has also demonstrated that phenolics, aromatic amines, and amino acids in exerting their antioxidant property inhibit enzymes that generate ROS in a cell system [57], in doing this, they might also inhibit enzymes that catalyze the removal of cell cycle arrest at G1 phase, thereby preventing the resumption of mitotic cell division which mark endodormancy break. (2) The botanical seed has a well-organized embryonic system in which dormancy regulation is processed, whereas, yam tuber has a continuous layer of meristematic cells beneath the dried tuber skin in which tuber dormancy regulation is processed, and the basis of yam tuber dormancy on this structure has been explained by Nwogha, et al. [7]. These postulations however need to be further validated through more

4.5. Energy Metabolism

It is believed that yam tuber dormancy induction is an adaptive response to low energy conditions due to the stoppage of the supply of sugar from photosynthesis as a result of vine senescence at tuber physiological maturity [44]. Since photosynthetic sugar supply and mineral uptake systems are not active during yam tuber dormancy, the energy needs for low and high physiological activities during tuber dormancy are supplied from reserved energy-source metabolites in the tuber [45]. In the present study, several energy source metabolites such as; (sucrose, galactose, carboxylic acids, malic acid, fatty acids, esters, and lipids) accumulated at different stages of tuber dormancy. The role of sugar signaling in dormancy regulation has been reported [92]. Based on changes in carbohydrate metabolism, it was suggested that sugar signaling is involved in regulating bud dormancy in non-woody perennial specie of leafy spurge [93]. In woody perennials, bud dormancy is associated with gene expression patterns in the typical manner of carbon starvation, while bud burst in spring depends on sucrose availability, and carbohydrate import from a distant wood parenchyma [94,95]. Expression of an Arabidopsis sucrose-phosphate synthase gene resulted in earlier bud dormancy break in spring transgenic and hybrid poplar (*Populus alba X Populus grandidentata*) [96]. Wengler [92], reported that carbohydrate mobilization for bud dormancy break may depend on interactions between the clock and sugar signaling. While in rice seeds, the study demonstrated that the amount of total storable sugar does not affect the seed dormancy duration, but a mixture of sucrose and raffinose, and thus hypothesized that the ratio of oligosaccharides to sucrose may be the determinant of seed dormancy duration [49,97]. In our study, sucrose accumulated in the tubers of the short dormant genotype (*Obiaoturugo*) at 56-DAPM, whereas, it accumulated in the tubers of the long dormancy duration genotype (*TDr1100873*) at 115-DAP. The 56-DAPM represents 28 days before the appearance of the shoot bud in the short dormant genotype (*Obiaoturugo*), whereas, 115-DAPM was the point of appearance of the shoot bud in the long-dormant genotype (*TDr1100873*). This is an indication that the stored starch might have been mobilized into sucrose at the onset of the dormancy breaking process to provide the required energy to reactivate cellular activities towards tuber dormancy break. our result is in agreement with the findings of [94,95] and the report of Wengler; [93]. Also, we earlier reported that biochemical profiling of non-structural sugar from yam tuber dormancy to dormancy break revealed that the quantity of nonreducing sugar which includes sucrose increased up to 9 and 10 folds in the tubers of *TDr110073* and *Obiaoturugo* respectively and that this increase started from 56-DAPM in tubers of *Obiaoturugo* and 101-DAPM in tubers of *TDr1100873* [44]. This indicates that sucrose plays a role in the early cellular activities culminating in dormancy break in the tubers of both genotypes.

The tricarboxylic acid (TCA) cycle intermediates, L-malic acid was observed in the tubers of *Obiaoturugo* (the short dormancy duration genotype), but not in the tubers of *TDr110873*, and surprisingly, it accumulated at 42-DAPM only. As dormancy progressed towards breaking L-malic acid was downregulated to an almost non-existence level. However, TCA cycles precursors and intermediates such as, carboxylic acids, esters, lipids and fatty acids accumulated from 28 days before the appearance of shoot bud to dormancy break stage in the tubers of both genotypes. This indicates that the TCA cycle plays an essential role in energy metabolism during yam tuber dormancy breaking. Our pathway topology analysis also revealed that the TCA cycle pathway was induced during the yam tuber dormancy progression stage (Table 2). The TCA cycle is one of the central energy metabolic pathways in a biological system, as well as being the biosynthetic pathway of many metabolites [98]. The plant TCA cycle has a set of eight enzymes primarily linking to the oxidation of pyruvate and malate (generated in the cytosol) to CO_2 with the generation of nicotinamide adenine dinucleotide reduced form (NADH) for oxidation by mitochondrial respiratory chain [99], and subsequently releases the energy required to reactivate cell cycle, which had been arrested at G1 phase during dormancy induction. It has been demonstrated that the accumulation of organic acids, particularly TCA cycle intermediates in plants supports many cellular processes, which are species, tissue, and developmental

stage specific [100]. In plants, lipids are the major component of cell membranes, and are used as a compact energy source for seed germination [101]. This was further supported by our enrichment and pathway topology analyses, which revealed that linoleic acid (lipid) metabolism was significantly induced with a fold change ratio above 20 (Figure 7), and its pathway exerted a maximum impact of 1 on yam tuber dormancy regulation (Table 2). We therefore, infer that lipids, carboxylic acids, fatty acids, and esters were metabolized through Embden–Meyerhof–Parnas (EMP), glycolysis, pyruvate, and TCA cycle pathways to generate the required energy pool for yam tuber dormancy break. It is noteworthy to mention that linoleic acid metabolism is very essential for white yam tuber dormancy break, and therefore can be a target metabolite in the genetic manipulation of white yam tuber dormancy duration. Previous studies have also demonstrated that pathways in carbohydrate metabolism and ATP production such as; glycolysis, TCA cycle, EMP pathway, and oxidative phosphorylation were accumulated at the end of endodormancy to produce energy for bud sprouting and as well as precursors for amino acids biosynthesis [46,53]. Put together our results demonstrated that sucrose, carboxylic acids, esters, lipids fatty acids with exception of (octadecadienoic acid) provided the energy required for yam tuber dormancy breaking, with lipid metabolism and its pathway exerting the strongest regulatory impact on the yam dormancy.

4.6. Nucleotides Metabolism and Cellular Processes

Four nucleotides comprised of L-Glutamine (nucleic amino acid), Adenine, nucleotide-sugar, and sugar acid were identified as part of DAMs in our study. Two of these nucleotides (L-Glutamine and nucleotide sugar) were found only in the short dormancy duration genotype (*Obiaoturugo*), whereas, Adenine and sugar acid were found in both genotypes. Adenine accumulated in the tubers of *Obiaoturugo* at 56-DAPM and in the tubers of *TDr1100873* at 101-DAPM. Note; that due to the short dormancy duration of the tubers of *Obiaoturugo,* 56-DAPM in the tubers of *Obiaoturugo* is physiologically equivalent to 101-DAPM in the tubers of *TDr1100873*. Hence, in physiological timing of Adenine accumulation in the tubers of both genotypes is actually the same time, and it marked the point of commencement of molecular activities towards tuber dormancy break in tubers of both genotypes. Hence, the accumulation of adenine in the tubers of both genotypes at their respective stages suggests that adenine plays a key role in the reactivation of cellular processes that led to dormancy break irrespective of genotypic variation. Nucleotide sugar and nucleic amino acid (L-Glutamine) all accumulated at 87-DAPM which marked the appearance of shoot bud in the short dormancy duration genotype, suggesting that these two metabolites play a key role in a later cellular processes that hastened dormancy breaking process in the tubers of *Obiaoturugo*. A previous study has demonstrated that L-Glutamine is direct precursor of Glutamate, and Glutamate occupies a central position in amino acid metabolism in plants and as well plays key role initiation of dormancy breaking process in seeds [56]. Glutamate, glutamine, aspartate, and asparagine are the central regulators of carbon/nitrogen metabolism and they interact with multiple metabolic networks to reactivate the cell cycle process [102,103]. Glutamine and glutamate are precursors of nitrogenous compounds in plants, and the production of glutamate from glutamine is a key point in the synthesis of a variety of organic molecules, such as nucleic acids, amino acids, and secondary metabolites [102]. We infer that L-glutamine could be a biomarker for early yam tuber dormancy breaking, and can be a target metabolic candidate in the genetic manipulation of yam tuber for short dormancy duration.

Two nucleotide sugars (nucleotide sugar and sugar acid) were identified in this study and both of them accumulated at the onset of dormancy breaking stage. Studies have demonstrated the regulatory role of sugar in cell division. A close correlation was observed between the supply of Glc and the expression of cyclins (*cycD2;1, D3;2, A3;2,* and *B1;2*) [54,56]. The D-type cyclins have been reported to act as a sensor of external conditions within the cellular environment, and are associated with cyclin-dependent kinase (CDKA) to regulate cell cycles [104]. It has also been reported that the regulatory effect of Glc on

the rate of cell division is primarily due to signaling rather than nutrient availability and energy status, as cell proliferating activity positively correlated with endogenous hexose levels, but not their uptake rate [54,55]. Implying that nucleotide sugar is not functioning as an energy source, but rather as signaling to reactivate and mobilize other molecular machinery towards the dormancy-breaking process. These reports are in agreement with our data which shows that nucleotide sugar accumulated in the tubers of the short dormancy duration genotype (*Obiaoturugo*) at the onset of dormancy-breaking activities. Sugar acid (galactonolactone) is a precursor of a nucleotide amino acid (ascorbate). The plant has evolved an efficient multi-component antioxidant defense system to protect vital molecules from the damage of mitochondrial generated ROS, as a consequence of the activity of respiratory chains [105]. This antioxidant defense system includes both enzymatic and non-enzymatic elements [106]. One of the most important non-enzymatic members of this system is the water-soluble antioxidant, ascorbate [105]. Paradiso, et al. [107] demonstrated that Ascorbate biosynthesis improved wheat seed kernel development and delayed programmed cell death (PCD), thereby delaying senescing by feeding the plant with its immediate precursor L-galactone-γ-lactone (GL). We infer that galactose accumulation at an early stage of cellular activities towards dormancy break in the current study suggests that it plays double roles during the yam tuber dormancy breaking process. (1) it acted as a precursor in the biosynthesis of ascorbate that scavenged ROS generated during the reactivation of the cell cycle. (2) Galactonolactone also promoted growth and protected the immature cells generated during cell division from being oxidized by ROS, which culminated in a dormancy break. Our result showed that adenine accumulated in the tubers of both genotypes at the onset of dormancy break, 42 days, and 45 days before physical germination of tubers of *TD1100873* and *Obiaoturugo* respectively (Figure 5a,b). This indicates that non de novo synthesis of purine from adenine might have taken place, and synthesis of purine positively regulated the release of the G1 phase of the cell cycle that was arrested during yam tuber dormancy. The release of the arrested G1 phase of the cell cycle marks the release of yam tuber endodormancy, and this also coincides with the 40 days plus that has been reported to be the duration between yam tuber endodormancy release and eco-dormancy release [7]. We postulate that adenine is the metabolic marker for yam tuber endo-dormancy break, whereas, L-glutamine and nucleotide sugar are the potential candidates' metabolites for early tuber dormancy breaking. In yam, these three candidate metabolites hold the promise of metabolic manipulation of yam crops for a desirable tuber dormancy duration. Therefore, we hypothesize that nucleotide sugar (glucose) might have acted as a signal that triggered the accumulation of L-glutamine, which functions together with the accumulated adenine to increase the nitrogen/carbon ratio (a critical factor in energy metabolism and DNA synthesis) in short dormancy duration genotype (*Obiaoturugo*), thereby leading to its tuber earlier dormancy break.

4.7. Secondary Metabolites and Phytohormones Metabolism

Secondary metabolites play key roles in plant growth and developmental processes such as; cell division, hormonal regulation, photosynthetic activity, radicals' homeostasis, modulation of seeds transition to quiescent state at maturity, environmental signals transduction, nutrient mobilization and mineralization, stress tolerance, plant-soil, and environmental interactions [108–110]. In the current study, diverse secondary metabolites that accumulated at different yam tuber dormancy progression stages include; flavonoids, coumarins, terpenoids, alkaloids and alcohols, benzoquinone and benzoic derivatives, phenols, and salicylic acid (hormone). In the tubers of long-dormant genotype (*TDr1100873*) the 3 flavonoids that were detected accumulated during tuber dormancy at (42-DAPM, 56-DAPM) and on sprouted tuber (143-DAPM), whereas, in the tubers of short dormant genotype (*Obiaoturugo*), they accumulated from the onset of dormancy break (56-DAPM, 87-DAPM), and on sprouted tuber (101-DAPM). This suggests that in the tubers of the long dormant genotype flavonoids play role in dormancy induction and maintenance, whereas, in the short dormant genotype, they promote dormancy break (germination). This also ex-

plained the variation between the two genotypes with respect to tuber dormancy duration. The flavonoids accumulation in the sprouted tubers indicates that, it might be playing a role also in vine elongation. It has been reported that flavonoids play different roles in different plant species, at different plant growth stages, organs, and tissues with regard to growth regulation [111]. Studies have demonstrated that in botanical seeds, flavonoids induce and prolong seed dormancy [28,59,109,112]. It has been shown that flavonoids induce dormancy by inhibiting reactive oxygen species (ROS), through; suppression of singlet oxygen, inhibition of enzymes that generate ROS, such as (cyclooxygenase, lipoxygenase, monooxygenase, and xanthine oxidase), chelation of ions transition which catalyzes ROS production, suppressing the cascade of free-radicals in lipid peroxidation and recycling of other antioxidants such as amino acids and polyamines [113–115]. This agreed with our data on the tubers of the long dormant genotype, but varies with data on the tubers of the short dormant genotype. The opposing role of flavonoids in the tubers of the short dormant genotype compared to the long dormant genotype and botanical seeds presents a frontier for exploration of the new role of flavonoids in dormancy regulation.

Similarly, coumarins also accumulated in the tubers of the long dormant genotype during the dormant stage (87-DAPM), and in the tubers of short dormant during the onset of dormancy breaking stages (56-DAPM and 87-DAPM). This also indicates that coumarins prolonged dormancy in the tubers of the long dormant genotype, but induced dormancy breaking process in the tubers of the short dormant genotype. Studies have demonstrated that coumarins induced both primary and secondary seed dormancy [116–118]. It has been reported that coumarins induce dormancy due to their ability to antagonize gibberellins and auxin's function, while synergizing with abscisic acid (ABA) [117,119,120]. Our result on the tubers of long dormant phenotype with regard to coumarins metabolism is supported by this model. Conversely, it has been reported that plants show different response depending on the species, size, growth stage, and process of cellular respiration and associated-enormous changes in the mitochondria [121]. This may explain the dormancy-breaking effect of coumarin observed on the tubers of the short dormant genotype in the current study. It has been shown that secondary metabolite can also serve as energy source for cellular processes, when degraded [112]. In light of this, we hypothesize that there could be a mechanism in the tubers of the short dormant genotype that inhibits dormancy induction action of secondary metabolites, and subsequently degrade them at the onset of dormancy breaking to provide additional energy source required to fuel the cell cycle reactivation process, thus leading to their accumulation during this period. It also could be a result of differential accumulation quantity, because it has been reported that extremely high quantities of phenolic compounds including coumarins such as; free (1042%), bound (120%), and total phenolic acid content of (741%) in canary grass seed promoted germination [122].

Terpenoid was detected only in the tubers of long dormant phenotype, and it accumulated at the onset of dormancy-breaking activities (101-DAPM). The accumulation of terpenes at 101-DAPM is indicative that it might be playing some role in cellular primary metabolism that is involved in cell cycle reactivation. Although, the direct involvement of terpene in dormancy or plant growth regulation is sparsely reported in the literature. However, terpenes play vital roles in many other plants physiological processes, especially in protection against external invaders, and response to environmental signals [110]. In the recent, it has been reported that terpenes also participate in plant primary metabolism [123]. There is also a report that most dormancy-regulating phytohormones belong to one terpenoid group or the other, for instance, gibberellins belong to diterpenoids, brassinolide belongs to triterpenoids, abscisic acid, and strigolactones belong to tetraterpenoids (carotenoids), cytokinin belongs to isoprene and is also known as primary metabolism terpene [110]. In light of this, it is reasonable to postulate that the terpene that accumulated in the tuber of long dormant genotype at the onset of the dormancy break stage in this study might be a precursor of one of these hormones especially the gibberellins and cytokinin which are known for their role in dormancy breaking and promotion of activities towards germination. Benzoquinone and benzene derivatives also accumulated from the onset

dormancy-breaking activities to sprouted tubers in both genotypes, an indication that these metabolites are implicated in the yam tuber dormancy-breaking process. Phenol accumulation was recorded only in the tubers of long dormant genotype, at its mid-dormancy stage (87-DAPM). Study has shown that phenol as an antioxidant, also participate in the scavenging of reactive oxygen species (ROS), catalyzing oxygenation reactions through the formation of metallic complexes, and inhibiting the activities of oxidizing enzymes [53], thereby inducing prolonged dormancy. The absence of phenol in the tubers of the short dormant genotype, and its cumulation in the tubers of the long dormant genotype at 87-DAPM, when endodormancy was broken in the short dormant genotype is indicative of the phenol role in yam tuber dormancy elongation.

In the tubers of the short dormant genotype, the two alkaloids detected in this study accumulated during the dormancy stage at (42-DAPM), whereas, in the tubers of long dormant genotype one (galactosylglycerol) accumulated at the dormant stage (42-DAPM) and in the sprouted tuber. Isopyridoxal accumulated at the dormancy break stage and in the sprouted tuber. This result demonstrates alkaloids play diverse roles in plant growth and developmental process that depend on the species, plant variety, and growth stage. Hence, in the tubers of the short dormant genotype it induced dormancy, and did not play any role in vine development. In the tubers long dormant genotype, galactosylglycerol played role in dormancy induction, and also show to be playing some role in vine development. Five alcohols detected in this study all accumulated during the early dormancy stage in the tubers of both genotypes, although, two of the alcohols accumulated in dormant and sprouted tubers of both genotypes. Our result shows that alcohols, similar to other osmo-protectants, are also involved in yam tuber dormancy induction, whereas the two that accumulated in sprouted yam tubers might have been mobilized to serve as an energy additional source during the early stage of vine elongation. Studies have demonstrated that alcohols improved crops' tolerance against abiotic stresses, possess antioxidant properties, and promote germination and seedling development of botanical seeds [124–126]. The two hormones (salicylic acid and 3-Cresotinic acid) that were identified as DAMs in this study, all accumulated in the tubers of the short dormant genotype during the dormant stage at (42-DAPM), whereas, 3-Cresotinic acid was not found in the tubers of long dormant genotype. Salicylic acid accumulated in its tubers at the dormancy break stage and in the sprouted tuber. Salicylic acid has been classified as one of the phenolics compounds among the secondary metabolites, and has been reported to act in synergy with ABA to induce dormancy, while it functions alone to promote dormancy break [7,122].

4.8. Pathways Elucidations

Our pathway topology analysis revealed that linoleic acid has the strongest influence on the yam tuber dormancy regulation process (Table 2). The pathway view shows that it is a linoleic acid biosynthetic process regulated by Cytochrome P450 (*CYP*) genes family. *CYPs* have been implicated in Gibberellic acids (GAs) biosynthesis and response to an environmental signal (light and photoperiod) during dormancy breaking in many botanical seeds [7], and GAs is one of the major plant growth regulatory hormones that promote seed dormancy break, by inhibiting the action abscisic acid (ABA). ABA and GAs are the two major phytohormones that act in an antagonistic manner to each other to induce and break dormancy respectively, in all crops. The ratio of ABA quantity to GAs quantity in seeds or any other plant tissue that has the capability of being dormant has been reported to be the determinant factor of the seed or tissue dormant status [7], and the effect of this factor (ABA/GAs ratio) surpasses the effect of any other factors and molecular processes that regulate dormancy in plant system. In maize, *CYP01A26* was reported to exhibit ent-kaurene oxidase activity which led to increased accumulation of bioactive GAs and consequently result in early seed dormancy break [127,128] Similarly, in rice, 13-hydroxylation pathway that involves the activity of 13-hydroxlases which were coded by *CYP714b1* and *CYP714B2* genes led to early seed germination [129]. We therefore infer that the activities of *CYP1A2*, *CYP2C*, *CYP2J*, *CYP2EI*, and *CYP3A4*, found in linoleic acid

metabolic pathway might have led to increasing the biosynthesis of GAs and activities of GAs in the tubers, led to dormancy break. Although, we did not detect GAs in our study, probably because of the difference between metabolomic profiling and hormonal profiling protocols, however, some of its intermediates were detected which indicates its presence. The Cytochrome P450 (*CYP*) genes upregulation of linoleic acid biosynthesis can lead to inhibition of abscisic acids (ABA) dormancy inducing action, and increased the accumulation of gibberellic acids (GAs) which consequently led to yam tuber dormancy break. Hence, we further infer that the linoleic acid pathway is a putative metabolic pathway that can be exploited for genetic manipulation of yam tuber dormancy. Our pathway topology analysis also revealed that the phenylalanine metabolic pathway has a significant impact on the yam tuber dormancy regulation process. The phenylalanine pathway is a phenylalanine, tyrosine, and tryptophan biosynthetic pathway, and was shown to be significantly induced during yam tuber dormancy (Table 2).

The pathway view shows that it is a hub for the biosynthesis of secondary metabolite, amino acid derivatives, and degradation of one ester that accumulated during the yam tuber dormancy breaking stage (Figure 9). The secondary metabolites synthesized in this pathway accumulated during the tuber dormancy stage, thus establishing a crosstalk link between secondary metabolites and amino acids and how their synergetic action leads to induction and prolonging of dormancy in yam tuber. The ester (octanoic acid, 2-phenylethyl ester) that was degraded in this pathway eventually accumulated during the tuber dormancy breaking, implying that in addition to antioxidant property effects of secondary metabolites and amino acids, they might be antagonizing some tuber dormancy break promoting metabolites such as Octanoic acid, 2-phenylethyl ester, whose action promotes dormancy break. This further supports our earlier postulation that dormancy inducing effect of most of the amino acids in yam tuber was partly as a result of secondary metabolites synthesis along their biosynthetic pathways. In plants, β-alanine is important for the synthesis of pantothenate and subsequently coenzyme A, which is an essential coenzyme in lipid and carbohydrate metabolism [130]. Also, glycolysis and TCA cycle have been reported to be linked to alanine aminotransferase during hypoxia induced waterlogging [131], confirming the role alanine metabolic pathway in stress tolerance, which is been linked to dormancy inducing-action. Moreover, in rice, alanine aminotransferase 1 encoded by the *flo12* gene was found to simultaneously regulate carbon and nitrogen metabolism, while the *flo12* mutant exhibited a floury white-core endosperm [132]. A previous study has demonstrated that phenylalanine, tyrosine and tryptophan are not only essential components of protein synthesis, but are also located upstream of a number of growth hormones and secondary metabolites with multiple biological functions and health-promoting properties, such as protection against abiotic and biotic stress [133]. It has been further shown that phenylalanine is required for protein biosynthesis and cell survival during desiccation in botanical seed, and acts as a precursor of a large number of multi-functional secondary metabolites, among which is a principal structural component in the supporting tissues of vascular plants and some algae [134]. Tyrosine has been reported to be the central hub to various specialized metabolic pathways, including vitamin E and plastoquinone which are essential metabolites of plant nutrition and antioxidant synthesis, as well as a precursor of specialized metabolites with diverse physiological functions such as protein, amino acids, attractants, and defense compounds [135]. It has been demonstrated that tryptophan is an essential amino acid in the synthesis of a large number of bioactive molecules, such as auxin, tryptamine derivatives, phytoalexins, indole glucosinolates, and terpenoid indole alkaloids, as well as playing a pivotal role in the regulation of plant growth and development [62,63].

Two energy metabolic pathways (starch and sucrose metabolism and galactose metabolism) were significantly induced during this study (Figure 9b,c). These pathways are biosynthetic pathways of three important metabolites (sucrose, sugar acid, and nucleotide sugar) that were identified as DAMs in this study. Interestingly, these sugars all accumulated at the onset of dormancy breaking stage in the tubers of both genotypes, except for nucleotide

sugar which was detected only in the tubers of the short dormant genotype, and has been earlier highlighted as the possible putative candidate metabolite driving the early dormancy break in that genotype. The significant induction and high impact of these pathways on the yam tuber dormancy regulation process support our earlier postulation that accumulation of these sugars at the onset of tuber dormancy breaking is an indication that they serve as an energy source for the reactivation of the cell cycle which marks the breaking of endodormancy in yam tuber. In addition to the biosynthesis of these sugars, these pathways are also involved in the pentose and glucuronate interconversions through UDP-glucose, pentose phosphate pathway, starch mobilization to amylose, galactosamine, D-glucose-6p, treholse-6p, all of which have been reported to play a pivotal role in cell cycle reactivation, including working in synergy with auxin during the dormancy-breaking process, and, as well, serve as signaling molecules [37,107,136]. Alanine, aspartate, and glutamate metabolism and purine metabolism are other two pathways that were induced during our study period in yam tubers. They are mainly involved in the biosynthesis of nucleotides such as adenine, L-glutamine, DNA precursor, DNA, RNA, histidine proteins and pentose phosphate. Adenine and L-glutamine are the two major metabolites in these pathways that were identified as DAMs in our study, and both accumulated at the onset of the dormancy break stage in both genotypes, which indicates that they play a pivotal role in the array of early cellular events towards dormancy break such as; histidine unpacking, DNA synthesis and replication, synthesis of replicated cells components, including house-keeping RNA during cell division, these required central carbon and nitrogen metabolism. L-glutamine was detected only in the tubers of the short dormant genotype, therefore, the induction and impact of its biosynthesis as revealed by pathway topology analysis support our earlier inference that L-glutamine could be a putative metabolic marker determining the early yam tuber dormancy break phenotype. Put together, our results demonstrate that adenine and galactose (sugar acid) might be conserved putative metabolic for yam tuber dormancy break irrespective of dormancy duration phenotype. Meanwhile, L-glutamine and nucleotide sugar are the two putative metabolic markers for yam tuber short dormant phenotype. Hence, alanine-aspartate-glutamine and purine metabolic pathways are some of the pathways that hold promise around the metabolic manipulation of yam tuber dormancy.

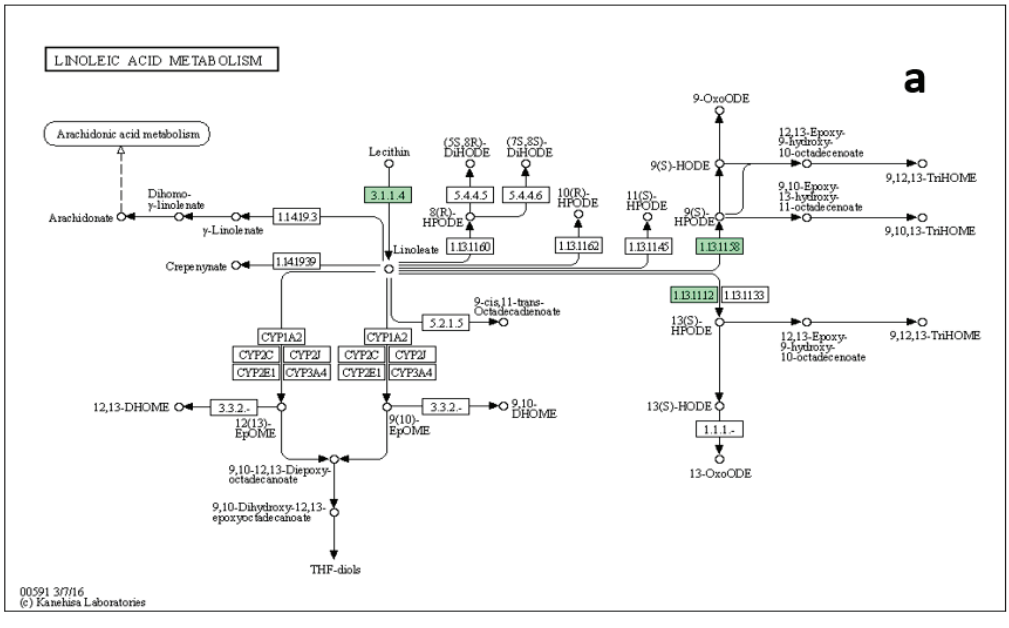


Figure 9. *Cont*.

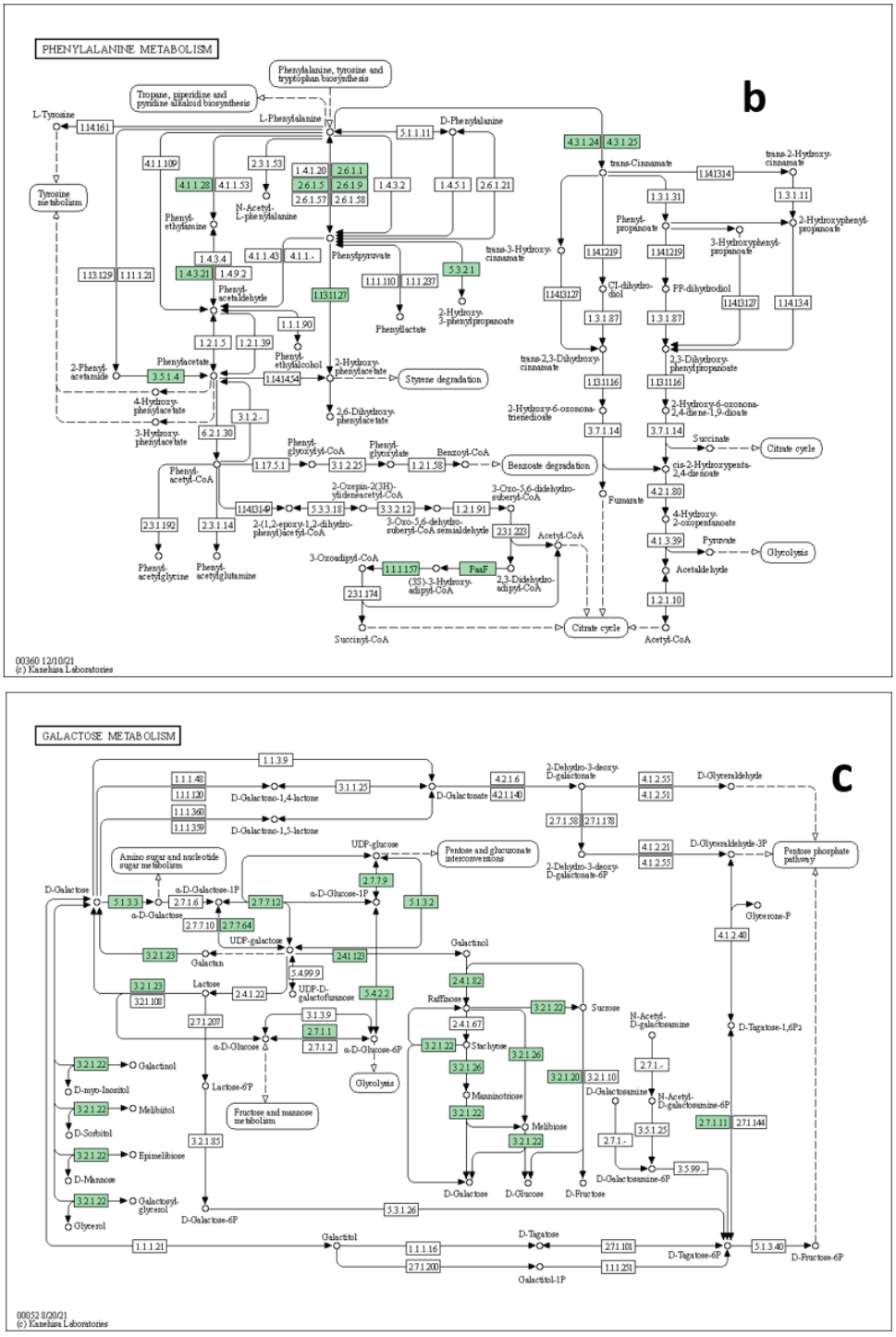
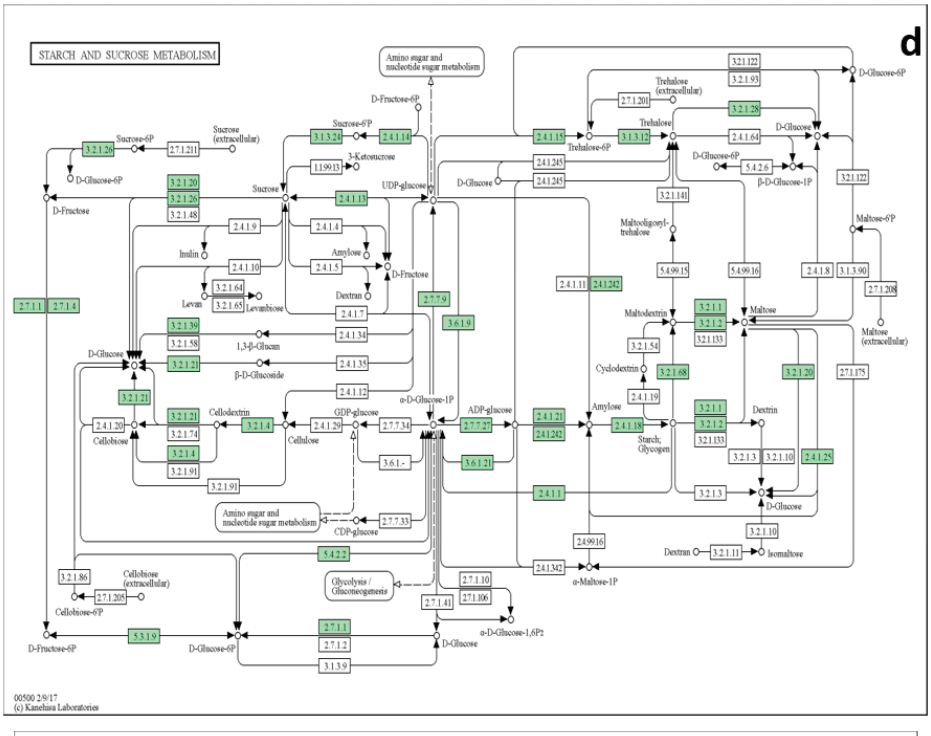


Figure 9. *Cont*.



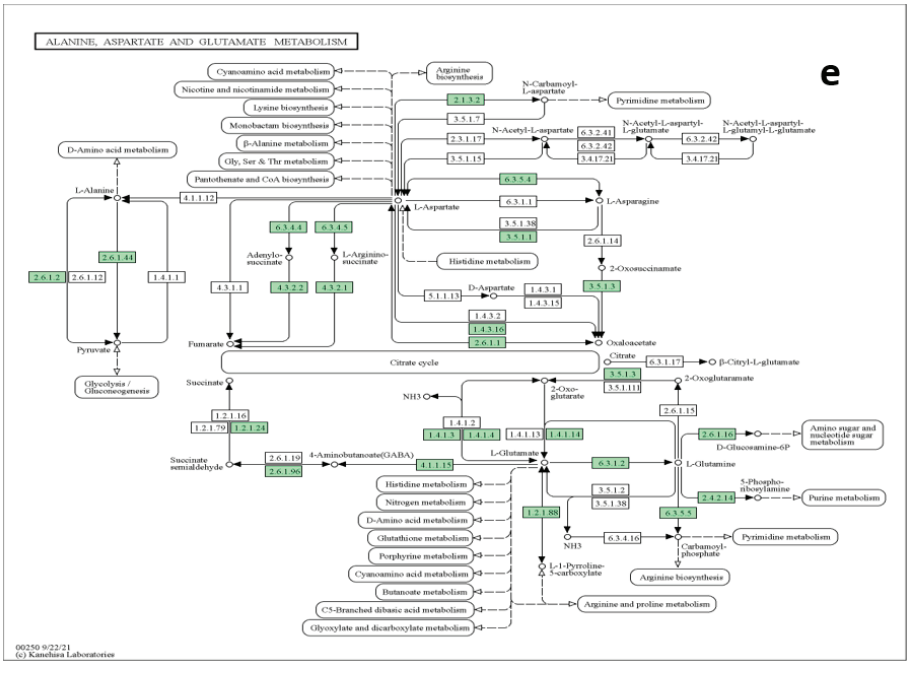


Figure 9. *Cont*.

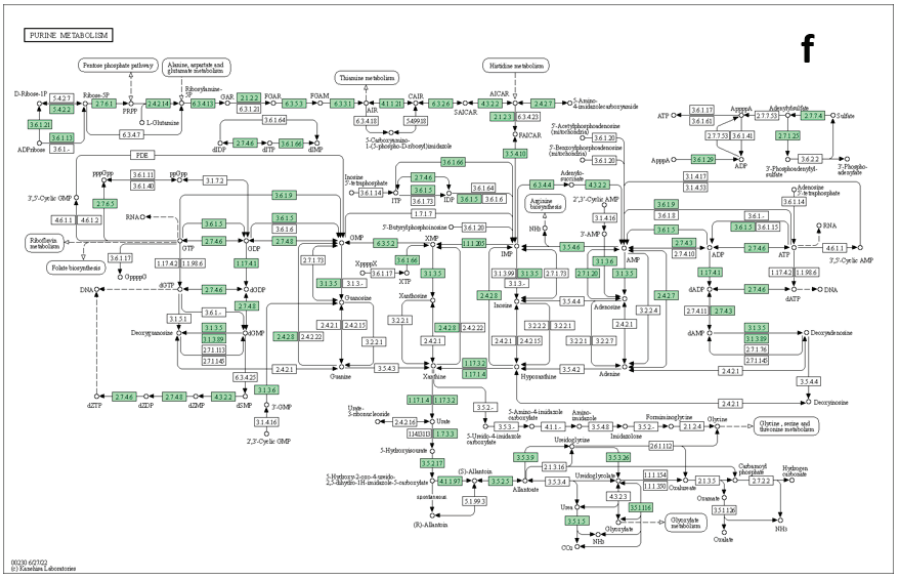


Figure 9. Six candidate pathways identified by pathway topology analysis that have a significant impact on yam tuber dormancy regulation. (**a**) is linoleic acid metabolic pathway, it exerted a maximum impact of 1, on a scale of 0 to 1. (**b**) is phenylalanine metabolic pathway, it exerted an impact of 0.79. (**c**) is galactose metabolic pathway with an impact value of 0.76. (**d**) is starch and sucrose metabolic pathway with an impact value of 0.59. (**e**) is Alanine–aspartate–glutamine metabolic pathway with an impact value of 0.57. (**f**) is the purine metabolic pathway with an impact value of 0.54.

5. Conclusions

This result demonstrates that yam tuber dormancy is a variable trait, and the two genotypes used in this study varied in the duration of their tuber dormancy. This variation was driven by the differential accumulation of terpenoid, phenol, amine, and one fatty acid in the tubers of *TDr110873*, and the differential accumulation of nucleotide sugar, amino acid (L-glutamine), and one carboxylic acid in the tubers of *Obiaoturugo*. The metabolomics changes from dormancy to sprouting in the tubers of both genotypes were systematically revealed, and a total of 39 differentially accumulated metabolites were identified across the investigated yam tuber dormancy stages of both yam genotypes. Amines and biogenic polyamines, amino acids and derivatives, Alcohols, flavonoids, alkaloids, phenols, esters, coumarins, and phytohormone regulated yam tuber dormancy induction and maintenance, whereas, fatty acids, lipids, nucleotides, carboxylic acids, sugars, terpenoids, benzoquinones, and benzene derivatives regulated dormancy breaking and sprouting process in tubers of both yam genotypes. Twelve metabolisms were significantly enriched during yam tuber dormancy and dormancy break process, and these were involved majorly in six metabolic (linoleic acid, phenylalanine, galactose, starch, and sucrose, purine, and alanine-aspartate-glutamate) pathways that exerted significant impacts on yam tuber dormancy regulation. Among these metabolic pathways, only phenylalanine metabolic regulated tuber dormancy induction, whereas, the rest five (linoleic acid, galactose, starch and sucrose, purine and alanine-aspartate-glutamate) pathways regulated dormancy break process, with linoleic acid metabolic pathway exerting strongest regulatory control. Therefore, the linoleic acid pathway holds a promise and the potential of being a putative metabolic pathway to be targeted in the genetic manipulation of yam tuber dormancy. In addition, metabolites such as; adenine, sucrose, and galactose (sugar acid) exhibited the potential of being conserved putative metabolic markers for yam tuber dormancy break irrespective of

genotypic variation with respect to dormancy duration. Our results also demonstrated that nucleotide sugar and glutamine are the two potential metabolic markers for early tuber dormancy break, and therefore determined the dormancy duration phenotypic variation between the tubers of two yam genotypes studied. Comparative to the literature information on botanical seeds' dormancy metabolic regulation mechanisms, our result demonstrated that yam tubers exhibit deviation, especially with regards to the roles of amino acids, amine and polyamines, and some secondary metabolites in dormancy regulations. Our result provides insight into molecular mechanisms regulating yam tuber dormancy, the basis for further metabolomics investigation of yam tuber dormancy regulation mechanisms and future genetic manipulation of yam tuber dormancy.

Supplementary Materials: The following supporting information have also been submitted and can be downloaded at: https://www.mdpi.com/article/10.3390/metabo13050610/s1, Figure S1a. Peak intensity normalization of TDr1100873 data Figure S1b. Peak intensity normalization of Obiaoturugo data, Figure S2. PLS–DA very important in project scores for TDr1100873, Figure S3. PLS–DA very important in project scores for Obiaoturugo; Table S1. PCA loadings of TDr1100873's data, Table S2. PCA loadings of Obiaoturugo's data.

Author Contributions: Conceptualization, J.S.N., J.E.O. and C.A.C.; methodology, M.R., J.S.N. and C.A.C.; software, J.S.N.; validation, M.R., J.E.O. and R.K.; formal analysis, J.S.N.; investigation, J.S.N., R.K. and C.A.C.; resources, J.E.O. and M.R.; data curation, J.S.N., R.K. and V.R.R.; writing—original draft preparation, J.S.N.; writing—review and editing, J.E.O. and A.G.W.; visualization, J.S.N.; supervision A.G.W., M.R. and J.E.O.; project administration, H.O.O. and A.G.W.; funding acquisition, H.O.O. and A.G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by GENES Intra-Africa Academic Mobility Project; "Mobility for plant genomics scholars to accelerate climate-smart adaptation options and food security in Africa (GENES)" funded by European Union. GENES Intra-Africa Academic Mobility Project. grant number EACEA/2017/2552. The APC was funded by Queen's University Belfast, Northern Ireland, United Kingdom.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the supplementary material here.

Acknowledgments: The authors acknowledge the field staff of yam research programme of NRCRI, Umudike Nigeria, especially Anthony Udeagbara, and staff of department Plant biotech-nology Tamil Nadu Agricultural University, Coimbatore, India for their technical support in field and laboratory experimentations respectively.

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

References

1. Price, E.J.; Wilkin, P.; Sarasan, V.; Fraser, P.D. Metabolite profiling of Dioscorea (yam) species reveals underutilised biodiversity and renewable sources for high-value compounds. *Sci. Rep.* **2016**, *6*, 29136. [CrossRef]
2. Obidiegwu, J.E.; Lyons, J.B.; Chilaka, C.A. The Dioscorea Genus (Yam)—An appraisal of nutritional and therapeutic potentials. *Foods* **2020**, *9*, 1304. [CrossRef]
3. Mignouna, H.D.; Abang, M.M.; Asiedu, R. Harnessing modern biotechnology for tropical tuber crop improvement: Yam (*Dioscorea* spp.) molecular breeding. *Afr. J. Biotechnol.* **2003**, *2*, 478–485.
4. Lev, L.S.; Shriver, A.L. A trend analysis of yam production, area, yield, and trade (1961–1996). *Colloq.-CIRAD* **1998**, 11–20.
5. Dansi, A.; Dantsey-Barry, H.; Dossou-Aminon, I.; N'kpenu, E.; Agré, A.; Sunu, Y.; Kombaté, K.; Loko, Y.; Dansi, M.; Assogba, P. Varietal diversity and genetic erosion of cultivated yams (*Dioscorea cayenensis* Poir-*D. rotundata* Lam complex and *D. alata* L.) in Togo. *Int. J. Biodivers. Conserv.* **2013**, *5*, 223–239.
6. FAOSTAT. Food and Agriculture Organization of the United Nations. 2020. Available online: http://wwwfaoorg/faostat/en/#data/QC (accessed on 23 October 2022).
7. Nwogha, J.S.; Wosene, A.G.; Raveendran, M.; Oselebe, H.O.; Obidiegwu, J.E.; Amirtham, D. Physiological and Molecular basis of dormancy in yam tuber: A way forward towards genetic manipultion of dormancy in yam tubers. *Glob. J. Sci. Front. Res.* **2022**, *XXll*, 47–73.

8. Lang, G.A.; Early, J.D.; Martin, G.C.; Darnell, R.L. Endo-, para-, and ecodormancy: Physiological terminology and classification for dormancy research. *HortScience* **1987**, *22*, 371–377. [CrossRef]
9. Hartmann, H.; Trumbore, S. Understanding the roles of nonstructural carbohydrates in forest trees–from what we can measure to what we want to know. *New Phytol.* **2016**, *211*, 386–403.
10. Mounirou, Y.; Jeanne, Z.; Djakaridia, T.; Sèssi Ida, A.; Elie Idossou, A.; Hibath Audrey Agbékè Mahougnon, A.; Jean Didier, Z.; Akpovi, A. Evaluation of yam (*Dioscorea cayenensis–Dioscorea rotundata*) seed germination grown in Centre Benin. *Int. J.* **2015**, *3*, 277–284.
11. Ile, E.; Craufurd, P.; Battey, N.; Asiedu, R. Phases of dormancy in yam tubers (*Dioscorea rotundata*). *Ann. Bot.* **2006**, *97*, 497–504. [CrossRef] [PubMed]
12. Passam, H. Dormancy of yams in relation to storage. In *Yams. Ignames*; Oxford University Press: Oxford, UK, 1982; pp. 285–293.
13. Hamadina, E.I.; Craufurd, P.Q. Changes in Free Phenolics Contents during Tuber Development, Dormancy and Sprouting in White Yam (*Dioscorea rotundata* Poir.). *Int. J. Plant Res.* **2015**, *5*, 34–41.
14. Shangguan, L.; Chen, M.; Fang, X.; Xie, Z.; Gong, P.; Huang, Y.; Wang, Z.; Fang, J. Comparative transcriptome analysis provides insight into regulation pathways and temporal and spatial expression characteristics of grapevine (*Vitis vinifera*) dormant buds in different nodes. *BMC Plant Biol.* **2020**, *20*, 390. [CrossRef]
15. Kaczmarska, K.T.; Chandra-Hioe, M.V.; Zabaras, D.; Frank, D.; Arcot, J. Effect of germination and fermentation on carbohydrate composition of Australian sweet lupin and soybean seeds and flours. *J. Agric. Food Chem.* **2017**, *65*, 10064–10073. [CrossRef]
16. González-Orenga, S.; Ferrer-Gallego, P.P.; Laguna, E.; López-Gresa, M.P.; Donat-Torres, M.P.; Verdeguer, M.; Vicente, O.; Boscaiu, M. Insights on salt tolerance of two endemic Limonium species from Spain. *Metabolites* **2019**, *9*, 294. [CrossRef] [PubMed]
17. Shu, K.; Liu, X.-d.; Xie, Q.; He, Z.-h. Two faces of one seed: Hormonal regulation of dormancy and germination. *Mol. Plant* **2016**, *9*, 34–45.
18. Guo, H.; Lyv, Y.; Zheng, W.; Yang, C.; Li, Y.; Wang, X.; Chen, R.; Wang, C.; Luo, J.; Qu, L. Comparative metabolomics reveals two metabolic modules affecting seed germination in rice (*Oryza sativa*). *Metabolites* **2021**, *11*, 880. [CrossRef] [PubMed]
19. Shuai, H.; Meng, Y.; Luo, X.; Chen, F.; Zhou, W.; Dai, Y.; Qi, Y.; Du, J.; Yang, F.; Liu, J. Exogenous auxin represses soybean seed germination through decreasing the gibberellin/abscisic acid (GA/ABA) ratio. *Sci. Rep.* **2017**, *7*, 12620. [CrossRef] [PubMed]
20. Kralj Cigić, I.; Rupnik, S.; Rijavec, T.; Poklar Ulrih, N.; Cigić, B. Accumulation of agmatine, spermidine, and spermine in sprouts and microgreens of alfalfa, fenugreek, lentil, and daikon radish. *Foods* **2020**, *9*, 547. [CrossRef] [PubMed]
21. Sepúlveda, G.; de Jiménez, E.S. Polyamine distribution among maize embryonic tissues and its relation to seed germination. *Biochem. Biophys. Res. Commun.* **1988**, *153*, 881–887. [CrossRef]
22. Farooq, M.; Basra, S.M.; Rehman, H.; Hussain, M. Seed priming with polyamines improves the germination and early seedling growth in fine rice. *J. New Seeds* **2008**, *9*, 145–155. [CrossRef]
23. Li, Z.; Peng, Y.; Zhang, X.-Q.; Ma, X.; Huang, L.-K.; Yan, Y.-H. Exogenous spermidine improves seed germination of white clover under water stress via involvement in starch metabolism, antioxidant defenses and relevant gene expression. *Molecules* **2014**, *19*, 18003–18024. [CrossRef] [PubMed]
24. Huang, Y.; Lin, C.; He, F.; Li, Z.; Guan, Y.; Hu, Q.; Hu, J. Exogenous spermidine improves seed germination of sweet corn via involvement in phytohormone interactions, H2O2 and relevant gene expression. *BMC Plant Biol.* **2017**, *17*, 1. [CrossRef] [PubMed]
25. Sibian, M.S.; Saxena, D.C.; Riar, C.S. Effect of germination on chemical, functional and nutritional characteristics of wheat, brown rice and triticale: A comparative study. *J. Sci. Food Agric.* **2017**, *97*, 4643–4651. [CrossRef]
26. Gipson, A.B.; Morton, K.J.; Rhee, R.J.; Simo, S.; Clayton, J.A.; Perrett, M.E.; Binkley, C.G.; Jensen, E.L.; Oakes, D.L.; Rouhier, M.F. Disruptions in valine degradation affect seed development and germination in Arabidopsis. *Plant J.* **2017**, *90*, 1029–1039. [CrossRef]
27. Desmaison, A.M.; Tixier, M. Amino acids content in germinating seeds and seedlings from *Castanea sativa* L. *Plant Physiol.* **1986**, *81*, 692–695. [CrossRef]
28. Zhang, T.; Yuan, Y.; Zhan, Y.; Cao, X.; Liu, C.; Zhang, Y.; Gai, S. Metabolomics analysis reveals Embden Meyerhof Parnas pathway activation and flavonoids accumulation during dormancy transition in tree peony. *BMC Plant Biol.* **2020**, *20*, 484. [CrossRef]
29. Wang, S.Y.; Jiao, H.J.; Faust, M. Changes in metabolic enzyme activities during thidiazuron-induced lateral budbreak of apple. *HortScience* **1991**, *26*, 171–173. [CrossRef]
30. Halaly, T.; Pang, X.; Batikoff, T.; Crane, O.; Keren, A.; Venkateswari, J.; Ogrodovitch, A.; Sadka, A.; Lavee, S.; Or, E. Similar mechanisms might be triggered by alternative external stimuli that induce dormancy release in grape buds. *Planta* **2008**, *228*, 79–88. [CrossRef]
31. Shen, C.; Wang, H.; Wang, X.; Wang, B.; Zheng, X.; Shi, X.; Liu, W.; Liu, F. Respiratory changes during dormancy of grape buds. *Sci. Agric. Sin.* **2013**, *46*, 1201–1207.
32. Deng, S.; Xiao, Q.; Xu, C.; Hong, J.; Deng, Z.; Jiang, D.; Luo, S. Metabolome profiling of stratified seeds provides insight into the regulation of dormancy in *Davidia involucrata*. *Plant Divers.* **2022**, *44*, 417–427. [CrossRef]
33. Arc, E.; Chibani, K.; Grappin, P.; Jullien, M.; Godin, B.; Cueff, G.; Valot, B.; Balliau, T.; Job, D.; Rajjou, L. Cold stratification and exogenous nitrates entail similar functional proteome adjustments during Arabidopsis seed dormancy release. *J. Proteome Res.* **2012**, *11*, 5418–5432. [CrossRef]
34. Szczotka, Z.; Pawłowski, T.; Krawiarz, K. Proteins and polyamines during dormancy breaking of European beech (*Fagus sylvatica* L.) seeds. *Acta Physiol. Plant.* **2003**, *25*, 423–435. [CrossRef]

35. Silva, A.C.D.; Suassuna, J.F.; Melo, A.S.D.; Costa, R.R.; Andrade, W.L.D.; Silva, D.C.D. Salicylic acid as attenuator of drought stress on germination and initial development of sesame. *Rev. Bras. Eng. Agrícola Ambient.* **2017**, *21*, 156–162. [CrossRef]
36. De Souza Vidigal, D.; Willems, L.; van Arkel, J.; Dekkers, B.J.; Hilhorst, H.W.; Bentsink, L. Galactinol as marker for seed longevity. *Plant Sci.* **2016**, *246*, 112–118. [CrossRef]
37. Baena-González, E.; Lunn, J.E. SnRK1 and trehalose 6-phosphate–two ancient pathways converge to regulate plant metabolism and growth. *Curr. Opin. Plant Biol.* **2020**, *55*, 52–59. [CrossRef] [PubMed]
38. Zhang, Z.; Zhu, J.-Y.; Roh, J.; Marchive, C.; Kim, S.-K.; Meyer, C.; Sun, Y.; Wang, W.; Wang, Z.-Y. TOR signaling promotes accumulation of BZR1 to balance growth with carbon availability in Arabidopsis. *Curr. Biol.* **2016**, *26*, 1854–1860. [CrossRef]
39. Bledsoe, S.W.; Henry, C.; Griffiths, C.A.; Paul, M.J.; Feil, R.; Lunn, J.E.; Stitt, M.; Lagrimini, L.M. The role of Tre6P and SnRK1 in maize early kernel development and events leading to stress-induced kernel abortion. *BMC Plant Biol.* **2017**, *17*, 74. [CrossRef]
40. Figueroa, C.M.; Lunn, J.E. A tale of two sugars: Trehalose 6-phosphate and sucrose. *Plant Physiol.* **2016**, *172*, 7–27. [CrossRef]
41. Tiama, D.; Sawadogo, N.; Traore, R.E.; Yolou, M.; Bationo-Kando, P.; Zoundjihekpon, J.; Sawadogo MZongo, J.D. Effect of chemical fertilizers on production of yams (nyù) of passore in farmers' environment. *Agron. Afr.* **2018**, *30*, 99–105.
42. Enfissi, E.M.; Barneche, F.; Ahmed, I.; Lichtlé, C.; Gerrish, C.; McQuinn, R.P.; Giovannoni, J.J.; Lopez-Juez, E.; Bowler, C.; Bramley, P.M. Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. *Plant Cell* **2010**, *22*, 1190–1215. [CrossRef]
43. Oliveros, J.C.; VENNY. An Interactive Tool for Comparing Lists with Venn Diagrams. 2007. Available online: http://bioinfogp.cnb.csic.es/tools/venny/index.html (accessed on 1 November 2022).
44. Nwogha, J.S.; Abtew, W.G.; Raveendran, M.; Oselebe, H.O.; Obidiegwu, J.E.; Chilaka, C.A.; Amirtham, D.D. Role of Non-Structural Sugar Metabolism in Regulating Tuber Dormancy in White Yam (*Dioscorea rotundata*). *Agriculture* **2023**, *13*, 343. [CrossRef]
45. Han, C.; Zhen, S.; Zhu, G.; Bian, Y.; Yan, Y. Comparative metabolome analysis of wheat embryo and endosperm reveals the dynamic changes of metabolites during seed germination. *Plant Physiol. Biochem.* **2017**, *115*, 320–327. [CrossRef] [PubMed]
46. Wang, Y.; Zhang, M.; Dong, S.; Liu, Y.-L.; Li, Z.-H. The Opposite Roles of White Light in Regulating Germination of Fresh and Aged Seed in Tobacco. *Plants* **2021**, *10*, 2457. [CrossRef] [PubMed]
47. Yang, M.; Yang, J.; Su, L.; Sun, K.; Li, D.; Liu, Y.; Wang, H.; Chen, Z.; Guo, T. Metabolic profile analysis and identification of key metabolites during rice seed germination under low-temperature stress. *Plant Sci.* **2019**, *289*, 110282. [CrossRef]
48. Mazlan, O.; Aizat, W.M.; Baharum, S.N.; Azizan, K.A.; Noor, N.M. Metabolomics analysis of developing Garcinia mangostana seed reveals modulated levels of sugars, organic acids and phenylpropanoid compounds. *Sci. Hortic.* **2018**, *233*, 323–330. [CrossRef]
49. Yan, S.; Huang, W.; Gao, J.; Fu, H.; Liu, J. Comparative metabolomic analysis of seed metabolites associated with seed storability in rice (*Oryza sativa* L.) during natural aging. *Plant Physiol. Biochem.* **2018**, *127*, 590–598. [CrossRef]
50. Feenstra, A.D.; Alexander, L.E.; Song, Z.; Korte, A.R.; Yandeau-Nelson, M.D.; Nikolau, B.J.; Lee, Y.J. Spatial mapping and profiling of metabolite distributions during germination. *Plant Physiol.* **2017**, *174*, 2532–2548. [CrossRef]
51. Yang, L.; Liu, S.; Lin, R. The role of light in regulating seed dormancy and germination. *J. Integr. Plant Biol.* **2020**, *62*, 1310–1326. [CrossRef]
52. Horikoshi, H.M.; Sekozawa, Y.; Kobayashi, M.; Saito, K.; Kusano, M.; Sugaya, S. Metabolomics analysis of 'Housui' Japanese pear flower buds during endodormancy reveals metabolic suppression by thermal fluctuation. *Plant Physiol. Biochem.* **2018**, *126*, 134–141. [CrossRef]
53. Amarowicz, R.; Weidner, S. Biological activity of grapevine phenolic compounds. In *Grapevine Molecular Physiology & Biotechnology*; Springer: Dordrecht, The Netherlands, 2009; pp. 389–405.
54. Wang, L.; Ruan, Y.-L. Regulation of cell division and expansion by sugar and auxin signaling. *Front. Plant Sci.* **2013**, *4*, 163. [CrossRef]
55. Hartig, K.; Beck, E. Crosstalk between auxin, cytokinins, and sugars in the plant cell cycle. *Plant Biol.* **2006**, *8*, 389–396. [CrossRef] [PubMed]
56. Forde, B.G.; Lea, P.J. Glutamate in plants: Metabolism, regulation, and signalling. *J. Exp. Bot.* **2007**, *58*, 2339–2358. [CrossRef] [PubMed]
57. Li, W.-Y.; Chen, B.-X.; Chen, Z.-J.; Gao, Y.-T.; Chen, Z.; Liu, J. Reactive oxygen species generated by NADPH oxidases promote radicle protrusion and root elongation during rice seed germination. *Int. J. Mol. Sci.* **2017**, *18*, 110. [CrossRef] [PubMed]
58. Huang, Y.; Cai, S.; Ye, L.; Hu, H.; Li, C.; Zhang, G. The effects of GA and ABA treatments on metabolite profile of germinating barley. *Food Chem.* **2016**, *192*, 928–933. [CrossRef] [PubMed]
59. Chen, C.; Zeng, L.; Zhao, H.; Ye, Q. Proteomic analysis of the early development of the Phalaenopsis amabilis flower bud under low temperature induction using the iTRAQ/MRM approach. *Molecules* **2020**, *25*, 1244. [CrossRef] [PubMed]
60. Tiburcio, A.F.; Alcazar, R. Potential applications of polyamines in agriculture and plant biotechnology. In *Polyamines: Methods in Molecular Biology*; Humana Press: New York, NY, USA, 2018; pp. 489–508.
61. Ma, Z.; Bykova, N.V.; Igamberdiev, A.U. Cell signaling mechanisms and metabolic regulation of germination and dormancy in barley seeds. *Crop J.* **2017**, *5*, 459–477.
62. Hildebrandt, T.M.; Nesi, A.N.; Araújo, W.L.; Braun, H.-P. Amino acid catabolism in plants. *Mol. Plant* **2015**, *8*, 1563–1579.

63. Yang, Q.; Zhao, D.; Liu, Q. Connections between amino acid metabolisms in plants: Lysine as an example. *Front. Plant Sci.* **2020**, *11*, 928. [CrossRef]
64. Hosseinifard, M.; Stefaniak, S.; Ghorbani Javid, M.; Soltani, E.; Wojtyla, Ł.; Garnczarska, M. Contribution of exogenous proline to abiotic stresses tolerance in plants: A review. *Int. J. Mol. Sci.* **2022**, *23*, 5186. [CrossRef]
65. Fujita, Y.; Fujita, M.; Satoh, R.; Maruyama, K.; Parvez, M.M.; Seki, M.; Hiratsu, K.; Ohme-Takagi, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K. AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. *Plant Cell* **2005**, *17*, 3470–3488. [CrossRef] [PubMed]
66. Yadav, U.P.; Ivakov, A.; Feil, R.; Duan, G.Y.; Walther, D.; Giavalisco, P.; Piques, M.; Carillo, P.; Hubberten, H.-M.; Stitt, M. The sucrose–trehalose 6-phosphate (Tre6P) nexus: Specificity and mechanisms of sucrose signalling by Tre6P. *J. Exp. Bot.* **2014**, *65*, 1051–1068. [CrossRef] [PubMed]
67. Muller, K.; Linkies, A.; Vreeburg, R.A.; Fry, S.C.; Krieger-Liszkay, A.; Leubner-Metzger, G. In vivo cell wall loosening by hydroxyl radicals during cress seed germination and elongation growth. *Plant Physiol.* **2009**, *150*, 1855–1865. [CrossRef] [PubMed]
68. Gonzalez-Calle, V.; Barrero-Sicilia, C.; Carbonero, P.; Iglesias-Fernandez, R.R. Mannans and endo-β-mannanases (MAN) in Brachypodium distachyon: Expression profiling and possible role of the BdMAN genes during coleorhiza-limited seed germination. *J. Exp. Bot.* **2015**, *66*, 3753–3764. [CrossRef]
69. Falco, S.; Guida, T.; Locke, M.; Mauvais, J.; Sanders, C.; Ward, R.; Webber, P. Transgenic canola and soybean seeds with increased lysine. *Bio/technology* **1995**, *13*, 577–582. [CrossRef]
70. Liu, X.; Zhang, C.; Wang, X.; Liu, Q.; Yuan, D.; Pan, G.; Sun, S.S.; Tu, J. Development of high-lysine rice via endosperm-specific expression of a foreign *LYSINE RICH PROTEIN* gene. *BMC Plant Biol.* **2016**, *16*, 147. [CrossRef]
71. Jia, M.; Wu, H.; Clay, K.L.; Jung, R.; Larkins, B.A.; Gibbon, B.C. Identification and characterization of lysine-rich proteins and starch biosynthesis genes in the opaque2mutant by transcriptional and proteomic analysis. *BMC Plant Biol.* **2013**, *13*, 60. [CrossRef] [PubMed]
72. Yang, P.; Li, X.; Wang, X.; Chen, H.; Chen, F.; Shen, S. Proteomic analysis of rice (*Oryza sativa*) seeds during germination. *Proteomics* **2007**, *7*, 3358–3368. [PubMed]
73. Nakabayashi, K.; Okamoto, M.; Koshiba, T.; Kamiya, Y.; Nambara, E. Genome-wide profiling of stored mRNA in Arabidopsis thaliana seed germination: Epigenetic and genetic regulation of transcription in seed. *Plant J.* **2005**, *41*, 697–709. [CrossRef]
74. Gianinetti, A.; Finocchiaro, F.; Bagnaresi, P.; Zechini, A.; Faccioli, P.; Cattivelli, L.; Valè, G.; Biselli, C. Seed dormancy involves a transcriptional program that supports early plastid functionality during imbibition. *Plants* **2018**, *7*, 35. [CrossRef]
75. Guo, G.; Liu, X.; Sun, F.; Cao, J.; Huo, N.; Wuda, B.; Xin, M.; Hu, Z.; Du, J.; Xia, R. Wheat miR9678 affects seed germination by generating phased siRNAs and modulating abscisic acid/gibberellin signaling. *Plant Cell* **2018**, *30*, 796–814. [CrossRef]
76. Liu, J.-H.; Honda, C.; Moriguchi, T. Involvement of polyamine in floral and fruit development. *Jpn. Agric. Res. Q. JARQ* **2006**, *40*, 51–58. [CrossRef]
77. Nambeesan, S.; Handa, A.K.; Mattoo, A.K. Polyamines and Regulation of Ripening and Senescence. In *Postharvest Biology and Technology of Fruit, Vegetables and Flowers*; Paliyath, G., Murr, D.P., Handa, A.K., Lurie, S., Eds.; Wiley-Blackwell: Oxford, UK, 2016.
78. Feduraev, P.; Skrypnik, L.; Riabova, A.; Pungin, A.; Tokupova, E.; Maslennikov, P.; Chupakhina, G. Phenylalanine and tyrosine as exogenous precursors of wheat (*Triticum aestivum* L.) secondary metabolism through PAL-associated pathways. *Plants* **2020**, *9*, 476. [CrossRef]
79. Jiao, Y.; Chen, Y.; Ma, C.; Qin, J.; Nguyen, T.H.N.; Liu, D.; Gan, H.; Ding, S.; Luo, Z.-B. Phenylalanine as a nitrogen source induces root growth and nitrogen-use efficiency in Populus× canescens. *Tree Physiol.* **2018**, *38*, 66–82. [CrossRef]
80. Sato, K.; Yamane, M.; Yamaji, N.; Kanamori, H.; Tagiri, A.; Schwerdt, J.G.; Fincher, G.B.; Matsumoto, T.; Takeda, K.; Komatsuda, T. Alanine aminotransferase controls seed dormancy in barley. *Nat. Commun.* **2016**, *7*, 11625. [CrossRef] [PubMed]
81. Das, A.; Kim, D.-W.; Khadka, P.; Rakwal, R.; Rohila, J.S. Unraveling key metabolomic alterations in wheat embryos derived from freshly harvested and water-imbibed seeds of two wheat cultivars with contrasting dormancy status. *Front. Plant Sci.* **2017**, *8*, 1203. [CrossRef]
82. Verslues, P.E.; Sharma, S. Proline Metabolism and Its Implications for Plant-Environment Interaction. *Arab. Book/Am. Soc. Plant Biol.* **2010**, *8*, e0140. [CrossRef] [PubMed]
83. Lechowska, K.; Wojtyla, Ł.; Quinet, M.; Kubala, S.; Lutts, S.; Garnczarska, M. Endogenous Polyamines and Ethylene Biosynthesis in Relation to Germination of *Osmoprimed Brassica napus* Seeds under Salt Stress Katarzyna. *Int. J. Mol. Sci.* **2022**, *23*, 349. [CrossRef] [PubMed]
84. Llebrés, M.-T.; Pascual, M.-B.; Debille, S.; Trontin, J.-F.; Harvengt, L.P.; Avila, C.; Cánovas, F.M. The role of arginine metabolic pathway during embryogenesis and germination in maritime pine (*Pinus pinaster* Ait.). *Tree Physiol.* **2018**, *38*, 471–484. [PubMed]
85. Dilworth, M.F.; Dure, L., III. Developmental biochemistry of cotton seed embryogenesis and germination: X. Nitrogen flow from arginine to asparagine in germination. *Plant Physiol.* **1978**, *61*, 698–702. [CrossRef]
86. Shalaby, A. Changes in biogenic amines in mature and germinating legume seeds and their behavior during cooking. *Food/Nahrung* **2000**, *44*, 23–27. [CrossRef]
87. MacGregor, D.R.; Kendall, S.L.; Florance, H.; Fedi, F.; Moore, K.; Paszkiewicz, K.; Smirnoff, N.; Penfield, S. Seed production temperature regulation of primary dormancy occurs through control of seed coat phenylpropanoid metabolism. *New Phytol.* **2014**, *205*, 642–652. [CrossRef] [PubMed]

88. Wyatt, J.E. Seed coat and water absorption properties of seed of near-isogenic snap bean lines differing in seed coat color. *J. Am. Soc. Hortic. Sci.* **1977**, *102*, 478–480. [CrossRef]
89. Corbineau, F.; Picard, M.; Côme, D. Germinability of leek seeds and its improvement by osmopriming. *Acta Hortic.* **1993**, *371*, 45–52. [CrossRef]
90. Edwards, M. MDormancy in seeds of Charlock: III. Occurrence and mode of action of an inhibitor associated with dormancy. *J. Exp. Bot.* **1968**, *19*, 601–610. [CrossRef]
91. Tobe, K.; Zhang, L.; Qiu, G.Y.; Shimizu, H.; Omasa, K. Characteristics of seed germination in five non-halophytic Chinese desert shrub species. *J. Arid. Environ.* **2001**, *47*, 191–201. [CrossRef]
92. Wingler, A.; Henriques, R. Sugars and the speed of life—Metabolic signals that determine plant growth, development and death. *Physiol. Plant.* **2022**, *174*, e13656. [CrossRef] [PubMed]
93. Anderson, G.H.; Veit, B.; Hanson, M.R. The Arabidopsis AtRaptor genes are essential for post-embryonic plant growth. *BMC Biol.* **2005**, *3*, 12. [CrossRef]
94. Tixier, A.; Gambetta, G.A.; Godfrey, J.; Orozco, J.; Zwieniecki, M.A. Non-structural carbohydrates in dormant woody perennials; the tale of winter survival and spring arrival. *Front. For. Glob. Chang.* **2019**, *2*. [CrossRef]
95. Tarancon, C.; González-Grando, E.; Oliveros, J.C.; Nicolas, M.; Cubas, P. A conserved carbon starvation response underlies bud dormancy in woody and herbaceous species. *Front. Plant Sci.* **2017**, *8*, 788.
96. Park, J.Y.; Canam, T.; Kang, K.Y.; Unda, F.; Mansfield, S.D. Sucrose phosphate synthase expression influences poplar phenology. *Tree Physiol.* **2009**, *29*, 937–946. [CrossRef]
97. Vandecasteele, C.; Teulat-Merah, B.; MorÈRe-LE Paven, M.C.; Leprince, O.; Ly Vu, B.; Viau, L.; Ledroit, L.; Pelletier, S.; Payet, N.; Satour, P.; et al. Quantitative trait loci analysis reveals a correlation between the ratio of sucrose/raffinose family oligosaccharides and seed vigour in Medicago truncatula. *Plant Cell Environ.* **2011**, *34*, 1473–1487. [CrossRef] [PubMed]
98. Zhang, Y.; Fernie, A.R. On the role of the tricarboxylic cid cycle in plant productivity. *J. Integr. Plant Biol.* **2018**, *60*, 1199–1216. [CrossRef]
99. Fernie, A.R.; Carrari, F.; Sweetlove, L.J. Respiratory metabolism: Glycolysis, the TCA cycle and mitochondrial electron transport. *Curr. Opin. Plant Biol.* **2004**, *7*, 254–261. [CrossRef]
100. Fernie, A.R.; Malate, E. Jack of all trades or master of a few? *Phytochemistry* **2009**, *70*, 828–883. [CrossRef]
101. Kim, H.U. Lipid metabolism in plants. *Plants* **2020**, *9*, 871. [CrossRef] [PubMed]
102. Qiu, X.M.; Sun, Y.Y.; Ye, X.Y.; Li, Z.G. Signaling Role of Glutamate in Plants. *Front. Plant Sci.* **2020**, *10*, 1743. [CrossRef] [PubMed]
103. Greco, M.; Chiappetta, A.; Bruno, L.; Bitonti, M.B. In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *J. Exp. Bot.* **2012**, *63*, 695–709. [CrossRef] [PubMed]
104. Nieuwland, J.; Menges, M.; Murray, J.A. The plant cyclins. In *Cell Cycle Control and Plant Development*; Blackwell Publishing: Oxford, UK, 2007; Volume 41.
105. Zsigmond, L.; Tomasskovics, B.; Deák, V.; Rigó, G.; Szabados, L.; Bánhegyi, G.; Szarka, A. Plant Physiology and Biochemistry Enhanced activity of galactono-1, 4-lactone dehydrogenase and ascorbate e glutathione cycle in mitochondria from complex III deficient Arabidopsis. *Plant Physiol. Biochem.* **2011**, *49*, 809–815. [CrossRef]
106. Blokhina, O.; Fagerstedt, K.V. Reactive oxygen species and nitric oxide in plant mitochondria: Origin and redundant regulatory systems. *Physiol. Plant.* **2010**, *138*, 447–462.
107. Paradiso, A.; Pinto, M.C.D.; Locato, V.; Gara, L.D. Galactone-γ-lactone-dependent ascorbate biosynthesis alters wheat kernel maturation. *Plant Biol.* **2012**, *14*, 1435–8603. [CrossRef]
108. Xool-Tamayo, J.; Tamayo-Ordoñez, Y.; Monforte-González, M.; Muñoz-Sánchez, J.A.; Vázquez-Flota, F. Alkaloid Biosynthesis in the Early Stages of the Germination of *Argemone mexicana* L. (Papaveraceae). *Plants* **2021**, *10*, 2226. [CrossRef] [PubMed]
109. Shah, A.; Smith, D.L. Flavonoids in agriculture: Chemistry and roles in, biotic and abiotic stress responses, and microbial associations. *Agronomy* **2020**, *10*, 1209. [CrossRef]
110. Pichersky, E.; Raguso, R.A. Why do plants produce so many terpenoid compounds? *New Phytol.* **2018**, *220*, 692–702. [CrossRef]
111. Mierziak, J.; Kostyn, K.; Kulma, A. Flavonoids as Important Molecules of Plant Interactions with the Environment. *Molecules* **2014**, *19*, 16240–16265. [CrossRef]
112. Corso, M.; Perreau, F.; Mouille, G.; Lepiniec, L. Specialized phenolic compounds in seeds: Structures, functions, and regulations. *Plant Sci.* **2020**, *296*, 110471.
113. Higdon, J.V.; Frei, B. Tea catechins and polyphenols: Health effects, metabolism, and antioxidant functions. *Crit. Rev. Food Sci. Nutr.* **2003**, *43*, 89–143. [CrossRef]
114. Arora, A.; Byrem, T.M.; Nair, M.G.; Strasburg, G.M. Modulation of liposomal membrane fluidity by flavonoids and isoflavonoids. *Arch. Biochem. Biophys.* **2000**, *373*, 102–109. [CrossRef] [PubMed]
115. Cotelle, N.; Bernier, J.L.; Catteau, J.P.; Pommery, J.; Wallet, J.C.; Gaydou, E.M. Antioxidant properties of hydroxy-flavones. *Free Radic. Biol. Med.* **1996**, *20*, 35–43. [CrossRef]
116. Araniti, F.; Prinsi, B. Seed Germination Induced by Coumarin Is Mediated by a Lower Ability to Sustain the Energetic Metabolism. *Plants* **2022**, *11*, 843. [CrossRef] [PubMed]
117. Chattha, F.A.; Munawar, M.A.; Nisa, M.; Kousar, S. Plant growth regulating activity of coumarins. In *Coumarin-Based Heteroaromatics as Plant Growth Regulators*; University of the Punjab: Lahore, Pakistan, 2016; pp. 91–104.

118. Chen, Z.; Ma, Y.; Yang, R.; Gu, Z.; Wang, P. Effects of exogenous $Ca2^{+}$ on phenolic accumulation and physiological changes in germinated wheat (*Triticum aestivum* L.) under UV-B radiation. *Food Chem.* **2019**, *288*, 368–376. [CrossRef]
119. Pergo, É.M.; Abrahim, D.; Soares da Silva, P.C.; Kern, K.A.; Da Silva, L.J.; Voll, E.; Ishii-Iwamoto, E.L. *Bidens pilosa* L. exhibits high sensitivity to coumarin in comparison with three other weed species. *J. Chem. Ecol.* **2008**, *34*, 499–507. [CrossRef] [PubMed]
120. Berrie, A.M.M.; Parker, W.; Knights, B.A.; Hendrie, M.R. Studies on lettuce seed germination-I. Coumarin induced dormancy. *Phytochemistry* **1968**, *7*, 567–573. [CrossRef]
121. Goren, R.; Tomer, E. Effect of seselin and coumarin on growth, indoleacetic acid oxidase and peroxidase with special reference to cucumber radicles. *Plant Physiol.* **1971**, *47*, 312–316. [CrossRef]
122. Sharma, A.; Shahzad, B.; Rehman, A.; Bhardwaj, R.; Landi, M. Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. *Molecules* **2019**, *24*, 2452. [CrossRef] [PubMed]
123. Ninkuu, V.; Zhang, L.; Yan, J.; Fu, Z.; Yang, T.; Zeng, H. Biochemistry of Terpenes and Recent Advances in Plant Protection. *Int. J. Mol. Sci.* **2021**, *22*, 5710. [CrossRef]
124. Das, A.K.; Anik, T.R.; Rahman, M.M.; Keya, S.S.; Islam, M.R.; Rahman, M.A.; Sultana, S.; Ghosh, P.K.; Khan, S.; Ahamed, T.; et al. Ethanol Treatment Enhances Physiological and Biochemical Responses to Mitigate Saline Toxicity in Soybean. *Plants* **2022**, *11*, 272. [CrossRef] [PubMed]
125. Ventura, I.; Brunello, L.; Iacopino, S.; Valeri, M.C.; Novi, G.; Dornbusch, T.; Perata, P.; Loreti, E. Arabidopsis phenotyping reveals the importance of alcohol dehydrogenase and pyruvate decarboxylase for aerobic plant growth. *Sci. Rep.* **2020**, *10*, 16669. [CrossRef]
126. Su, W.; Ren, Y.; Wang, D.; Su, Y.; Feng, J.; Zhang, C.; Tang, H.; Xu, L.; Muhammad, K.; Que, Y. The alcohol dehydrogenase gene family in sugarcane and its involvement in cold stress regulation. *BMC Genom.* **2020**, *21*, 521.
127. Skalák, J.; Vercruyssen, L.; Claeys, H.; Hradilová, J.; Černý, M.; Novák, O.; Plačková, L.; Saiz-Fernández, I.; Skaláková, P.; Coppens, F.; et al. Multifaceted activity of cytokinin in leaf development shapes its size and structure in Arabidopsis. *Plant J.* **2019**, *97*, 805–824. [CrossRef]
128. Mao, J.; Li, W.; Mi, B.; Mujitaba, M.; Ma, Z.; Zhang, Y.; Chen, B.; Caldero, A. Different exogenous sugars affect the hormone signal pathway and sugar metabolism in "Red Globe" (*Vitis vinifera* L.) plantlets grown in vitro as shown by transcriptomic analysis. *Planta* **2017**, *246*, 537–552. [CrossRef]
129. Magome, H.; Nomura, T.; Hanada, A.; Takeda-Kamiya, N.; Ohnishi, T.; Shinma, Y.; Katsumata, T.; Kawaide, H.; Kamiya, Y.; Yamaguchi, S. CYP714B1 and CYP714B2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1947–1952. [CrossRef] [PubMed]
130. Parthasarathy, A.; Savka, M.A.; Hudson, A.O. The synthesis and role of β-alanine in plants. *Front. Plant Sci.* **2019**, *10*, 921. [CrossRef] [PubMed]
131. Rocha, M.; Sodek, L.; Licausi, F.; Hameed, M.W.; Dornelas, M.C.; Van Dongen, J.T. Analysis of alanine aminotransferase in various organs of soybean (*Glycine max*) and in dependence of different nitrogen fertilisers during hypoxic stress. *Amino Acids* **2010**, *39*, 1043–1053. [CrossRef]
132. Zhong, M.; Liu, X.; Liu, F.; Ren, Y.; Wang, Y.; Zhu, J.; Teng, X.; Duan, E.; Wang, F.; Zhang, H. FLOURY ENDOSPERM12 encoding alanine aminotransferase 1 regulates carbon and nitrogen metabolism in rice. *J. Plant Biol.* **2019**, *62*, 61–73. [CrossRef]
133. Tzin, V.; Malitsky, S.; Aharoni, A.; Galili, G. Expression of a bacterial bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in Arabidopsis. *Plant J.* **2009**, *60*, 156–167. [CrossRef]
134. Vanholme, R.; De Meester, B.; Ralph, J.; Boerjan, W. Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* **2019**, *56*, 230–239. [CrossRef]
135. Schenck, C.A.; Maeda, H.A. Tyrosine biosynthesis, metabolism, and catabolism in plants. *Phytochemistry* **2018**, *149*, 82–102. [PubMed]
136. Dennis, R. The Role of Primary Carbohydrate Metabolism in Wheat Grain Dormancy and Germination. Ph.D. Thesis, The Australian National University, Canberra, Australia, 2019.

metabolites

MDPI

Article

SlbHLH22-Induced Hypertrophy Development Is Related to the Salt Stress Response of the *GTgamma* Gene in Tomatoes

Baolu Cui [1,2], Min Yu [1], Jiaojiao Bai [1,*] and Zhiguo Zhu [1,*]

[1] College of Pharmacy and Life Sciences, Jiujiang University, Jiujiang 332005, China; 6090111@jju.edu.cn (B.C.); yuminyumin@whu.edu.cn (M.Y.)
[2] College of Biological Sciences and Agriculture, Qiannan Normal University for Nationalities, Duyun 558000, China
* Correspondence: jjbai@sjziam.ac.cn (J.B.); zhuzhiguo@jju.edu.cn (Z.Z.)

Abstract: Hypertrophy development induced by the overexpression of *SlbHLH22* (also called *SlUPA-like*) was susceptible to *Xanthomonas* in tomatoes. Transcriptome and metabolome analyses were performed on the hypertrophy leaves of a *SlbHLH22*-overexpressed line (OE) and wild type (WT) to investigate the molecular mechanism. Metabolome analysis revealed that six key metabolites were over-accumulated in the OE, including Acetylserine/O-Acetyl-L-serine, Glucono-1,5-lactone, Gluconate, 2-Oxoglutarate, and Loganate, implying that the OE plants increased salt or oxidant resistance under normal growth conditions. The RNA-seq analysis showed the changed expressions of downstream genes involved in high-energy consumption, photosynthesis, and transcription regulation in OE lines, and we hypothesized that these biological processes were related to the GTgamma subfamily of trihelix factors. The RT-PCR results showed that the expressions of the *GTgamma* genes in tomatoes, i.e., *SlGT-7* and *SlGT-36*, were suppressed in the hypertrophy development. The expression of the *GTgamma* gene was downregulated by salinity, indicating a coordinated role of GTgamma in hypertrophy development and salt stress. Further research showed that both *SlGT-7* and *SlGT-36* were highly expressed in leaves and could be significantly induced by abscisic acid (ABA). The GTgamma protein had a putative phosphorylation site at S^{96}. These results suggested GTgamma's role in hypertrophy development by increasing the salt resistance.

Keywords: *SlbHLH22*-induced hypertrophy; metabolome; transcriptome; *GTgamma* gene; salt stress; *Solanum lycopersicum*

Citation: Cui, B.; Yu, M.; Bai, J.; Zhu, Z. *SlbHLH22*-Induced Hypertrophy Development Is Related to the Salt Stress Response of the *GTgamma* Gene in Tomatoes. *Metabolites* **2023**, *13*, 1195. https://doi.org/10.3390/metabo13121195

Academic Editor: Robert D Hall

Received: 31 October 2023
Revised: 7 December 2023
Accepted: 7 December 2023
Published: 11 December 2023

1. Introduction

Xanthomonas causes a broad disease in crop cultivars, such as spot disease. To overcome plant defense, *Xanthomonas* delivers transcription activator-like effectors (TALes) into host cells to suppress immune responses [1]. AvrBs3, one of the TALe families, induces cell enlargement in the host leaf by directly activating a master regulator of cell size, i.e., *UPA20*, a bHLH family gene [2,3]. We also found that *SlUPA-like* (the orthology of *UPA20*, also called *SlbHLH22*) overexpression caused severe hypertrophy and facilitated the infection of *Xanthomonas* in tomato leaves. The experimental evidence proved that the Gibberellin (GA) response was upregulated and that the jasmonic acid (JA) response was downregulated in *SlUPA-like* overexpressed lines (OEs) [4]. Additionally, the mature leaves of OEs curled upward and wilted under normal conditions, and the total chlorophyll decreased remarkably [4]. These phenotypes implied that other factors might be involved in the developmental malformation of OE plants.

Previous reports proved that altering plant development with trihelix factors contributes to pathogen susceptibility or resistance. *GhGT-3b* was strongly induced by *Verticillium dahlia* and the heterologous expression of *GhGT-3b* in *Arabidopsis* enhanced resistance to *Verticillium dahlia* but inhibited the growth of rosette leaves [5]. *ARABIDOPSIS SH4-RELATED 3* (*ASR3*) overexpressed plants were smaller than the control but enhanced

susceptibility to infections of *Pseudomonas syringae* pv *tomato DC3000* and *Pseudomonas syringae* pv *maculicola ES4326* [6]. Meanwhile, a similar mechanism was also found in the over-accumulation of the ASR3-interacting transcriptional factor 1 (AITF1), which negatively regulated *Pseudomonas syringae* resistance in *Arabidopsis* [7]. In maize, the seedlings of *ZmGT-3b* knockdown showed reduced photosynthesis activity but were resistant to the *Fusarium graminearum* challenge [8]. However, few data verified the role of the trihelix gene in hypertrophy developments.

Most studies focus on trihelix factor functions in abiotic stress. The overexpression of *ShCIGT* (GT-1) improved cold and drought tolerance in tomatoes [9]. In cotton, *GhGT26* (GT-1)-overexpressed lines had higher salt tolerance than the control via the ABA independent pathway, which was partially similar to the SIP1subfamily gene *GhGT23* [10]. In rice, the experimental data proved that *OsGTgamma-1* and *OsGTgamma-2* have specific roles in promoting salt tolerance when directly regulating salinity transporter genes [11,12]. Interestingly, *SlbHLH22* enhanced plant tolerance to salinity in MicroTom (one dwarf cultivar of tomato) [13]. It was a hypothesis that perhaps *SlbHLH22* regulates abiotic stress-related genes via the trihelix family.

Aside from regulation by the transcription level, trihelix factor functions are often affected by post-transcription modification. Calcium/calmodulin kinase II (CaMKII) can phosphorylate GT-1 at T133 [14]. ShCIGT (SlGT-24) regulated abiotic tolerance by interacting with Snf1-related kinase 1 (SnRK1) [9]. NMR titration experiments suggested the phosphorylation site of GT-1 is located at the N-terminus of the third helix [15]. The N-terminal of PTL, a GT-2 factor, can be phosphorylated by SnRK1α1(AKIN10), an α-subunit of SnRK1 [16]. Meanwhile, ASR3 can be phosphorylated by MAMP-activated MPK4 [6]. Therefore, we speculated that trihelix factors might fulfill the necessary functions via phosphorylation.

In our experiment, transcriptome and metabolome analysis was used to reveal the molecular mechanism of a developmental malformation in OE, suggesting that the susceptibility of OE plants to *Xanthomonas* was related to increasing salt or oxidant tolerance. Extensive analysis indicated that *GTgamma* was suppressed downstream of SlbHLH22 protein, which was similar to that inhibited expression in salt stress. Deep analysis forecasted that the GTgamma protein might be phosphorylated at the post-transcription level. Therefore, our research provided a good foundation for studying the pathogenic mechanism of hypertrophy development and GTgamma's role in biotic and abiotic stress.

2. Materials and Methods

2.1. Plant Materials and Growth Conditions

Solanum lycopersicum Mill. var. Ailsa Craig (AC^{++}, WT) and *SlbHLH22* (Solyc03g097820, also called *SlUPA-like*) OE lines [3] were grown in a glasshouse under controlled conditions with 16-h-light/8-h-dark cycles, 25 °C-day/18 °C-night temperatures, 80% relative humidity, and 250 μmol m^{-2} s^{-1} luminous intensity. Flowers were tagged at the anthesis stage, immature green fruit was defined as 20 DPA (days past anthesis), mature green fruit as 35 DPA, and breaker fruit as 38 DPA with the color starting to generate a slight yellow shade. Other fruits from the 4th (B+4) and 7th (B+7) days after the breaker were harvested. Fruits at different ripening stages were collected, frozen immediately in liquid nitrogen, and stored at −80 °C until use [17].

2.2. Transcriptome and Metabolome Analysis

Total RNA was extracted from OE and WT leaves by using Trizol reagent (Invitrogen, Carlsbad, CA, USA), and the concentration and purity of RNA were measured by Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). The RNA integrity was measured by Agient 2100, LabChip GX (Santa Clara, CA, USA). Three biological replicates were sampled for each group (WT, OE). RNA and then transcriptomic experiments were conducted by BMKcloud, Beijing, China (http://www.biomarker.com.cn, accessed on 19 May 2023). Clean reads were obtained by removing adapters. Reads were then mapped to the *Solanaceae*

genome (https://solgenomics.net/, accessed on 19 May 2023) using HISAT2 and gene expression levels were quantified with HTseq (BMKcloud, Beijing, China) [18].

Samples were ground to powder using a grinder (MM 400, Retsch, Shanghai, China) and dissolved into an extraction solution to remove by ultrasonic extraction. The extracted metabolites were analyzed by LC-MS/MS with Waters Xevo G2-XS QTOF (Milford, CT, USA). The metabolomics experiments and conjoint analyses of transcriptome and metabolome sequencing were conducted by BMKcloud, Beijing, China (http://www.biomarker.com.cn/, accessed on 19 May 2023) [18].

2.3. Hormonal and Salt Treatments

A 35-day-old tomato seedling of AC^{++} planted in green house of Jiujiang University (Jiujiang, China) was used for hormonal and abiotic treatments with three biological replicates.

For hormonal treatment, all the potted tomato seedlings were sprayed with different hormonal (50 μM 3-Indoleacetic Acid IAA, 50 μM Gibberellin GA, 100 μM 1-Aminocyclopropane-1-Carboxylicacid, ACC, 100 μM Abscisic Acid ABA, 50 μM Methyl Jasmonic Acid MeJA; 50 μM Epibrassinolide EBR; 50 μM Uniconazole NA) (Coolaber, Beijing, Chia) and distilled water (the control). Plants were enclosed in plastic immediately and left for 0, 1, 4, 8, 12, 24 h; the leaves of the tomato seedlings were taken and stored at −80 °C until use [19–21].

Salinity treatments were operated by submerging the roots of the tomato seedlings in distilled water with 200 mM NaCl for 0, 1, 4, 8, 12, 24, 48 and 72 h; Roots and leaves from the treated seedlings were collected and stored at −80 °C until use [22].

2.4. RT-PCR

The total RNA was reverse-transcribed to cDNA. RT-PCR was performed using SYBR® Premix Ex Taq TM (TaKaRa, Dalian, China). RT-PCR primers were designed with Primer 5 (Supplementary Table S1). The tomato *SlCAC* and *SlEF1a* gene were used as an internal control of expression patterns and treatments. All the selected genes were calculated with three technical replicates.

2.5. Statistic Analysis

All data are means ± standard deviation of at least three independent experiments. Significance in a difference between the two groups was assessed by a Student's *t*-test (*, $p < 0.05$ or **, $p < 0.01$). The different letters above the column in the figures indicate that significant differences of $p < 0.05$ were assessed by ANOVA. These statistical programs were performed using DPS v2.1.3 software (Ruifeng, Hangzhou, China).

2.6. Computational Modeling

The structure of the peptides was drawn using SWISS-MODEL. The peptide was sent to the GRAMMX protein–protein docking server (Version 12.0). Conformation models were obtained. These docking conformations were sent to the Rosetta FlexPepDock 4.0 server to be refined from a complex between a protein receptor and an estimated conformation for a peptide, allowing full flexibility to the peptide and sidechain of the receptor. FlexPepDock 4.0 gave an output of predicted energies for the complex. Peptides were added to the CHARMM36 force field to correct any resulting mischarges [23].

3. Results

3.1. Metabolome Analysis of OE vs. WT

After metabolome analysis, different expressed genes (DEGs) encoding metabolic processes in OE were primarily clustered in "alanine, aspartate and glutamate metabolism", "carbon metabolism", "monoterpenoid biosynthesis", "taurine and hypotaurine metabolism", "tyrosine metabolism" and "zeatin biosynthesis" compared to those in WT. The different metabolic processes were most enriched in "ascorbate and aldarate metabolism", "carbon

metabolism", "pentose and glucuronate interconversions", and "vitamin B6 metabolism". They were further enriched in "arginine biosynthesis", "unsaturated fatty acids biosynthesis", "monoterpenoid biosynthesis", "phosphatidylinositol signaling system", "sulfur metabolism", "taurine and hypotaurine metabolism", "terpenoid backbone biosynthesis", and "zeatin biosynthesis" (Figure 1). The consistent results between metabolic processes and their DEGs were "carbon metabolism", "monoterpenoid biosynthesis", "taurine and hypotaurine metabolism", and "zeatin biosynthesis". Within these processes, six key metabolites were abundant, including Acetylserine/O-Acetyl-L-serine (OAS), Glucono-1,5-lactone, Gluconate, 2-Oxoglutarate (2-OG) and Loganate (Figure S1). Previous studies confirmed that these metabolites were helpful to salt or oxidant resistance [24–28].

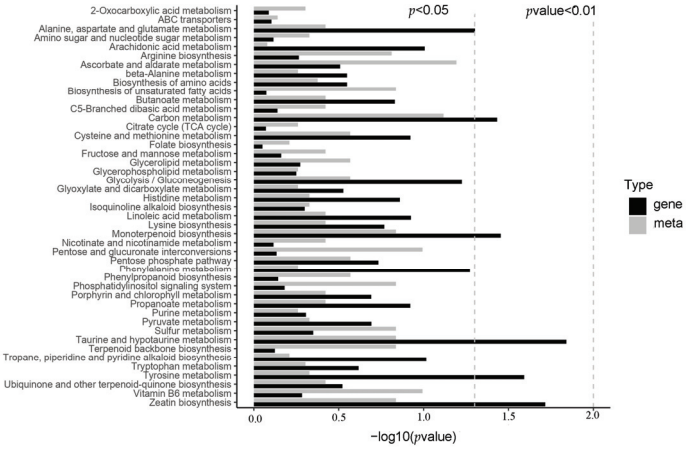


Figure 1. Pathways annotated with differential metabolic process and genes by KEGG analysis. Gene: related to metabolic process; Meta: metabolic process.

3.2. Transcriptome Analysis of OE vs. WT

To better understand the molecular mechanism of malformation developments in OE leaves, we performed transcriptome analysis in the mature leaves of OE vs. WT. Through RNA-seq analysis, we obtained 6 RNA-seq libraries and 24 to 27 million clean reads. After alignment with reference sequences, the alignment efficiency of clean reads ranged from 94.47% to 96.22% (Supplementary Table S2). Clearly, 2815 DEGs were identified, including 1299 upregulated and 1516 downregulated DEGs (Figure 2).

Gene ontology (GO) analysis clarified that upregulated DEGs remarkably converged on "amino acid" and the "sulfate transmembrane transport process" in the biological process (Figure 3A). In cellular component ontology, "integral component of membrane" and "plasma membrane" were the most abundant categories (Figure 3B). Genes involved in "amino acid transmembrane transporter activity", "sequence-specific DNA binding", "transcription factor activity" and "secondary sulfate transmembrane transporter activity" were enriched in the molecular function category (Figure 3C). Downregulated DEGs markedly gathered in "photosynthesis", "light harvesting in PSI", "protein-chromophore linkage", "responses to light stimulus", "flavonoid glucuronidation", "flavonoid synthesis", "DNA replication initiation" and "cell wall biogenesis" in biological processes (Figure 3D). In cellular component ontology, "photosystem", "plastoglobule", "MCM complex", "chloroplast", "cell wall", "nucleosome", "intracellular membrane-bounded organelle", and "THO complex" were the most abundant categories (Figure 3E). Genes involved in "chlorophyll" and "pigments binding" were enriched in the molecular function category (Figure 3F). These data suggested that the strongly repressed photosynthesis increased the substance transmembrane transport and transcription factor activities in OE.

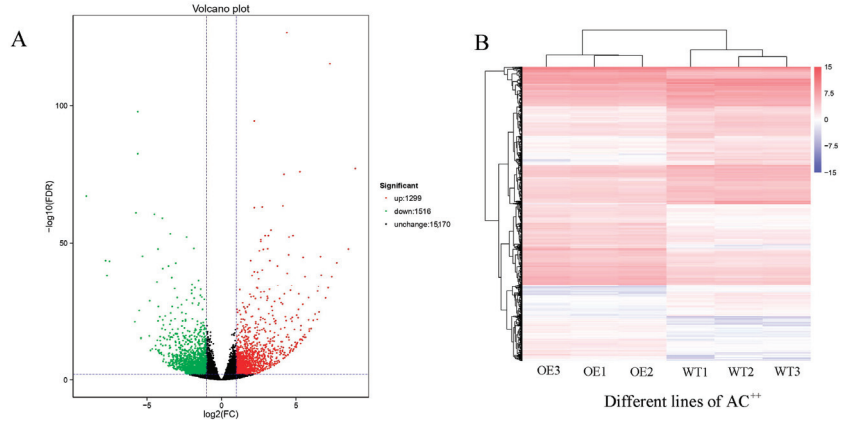


Figure 2. Comparing DEGs by volcano (**A**) and heatmap (**B**) pictures.

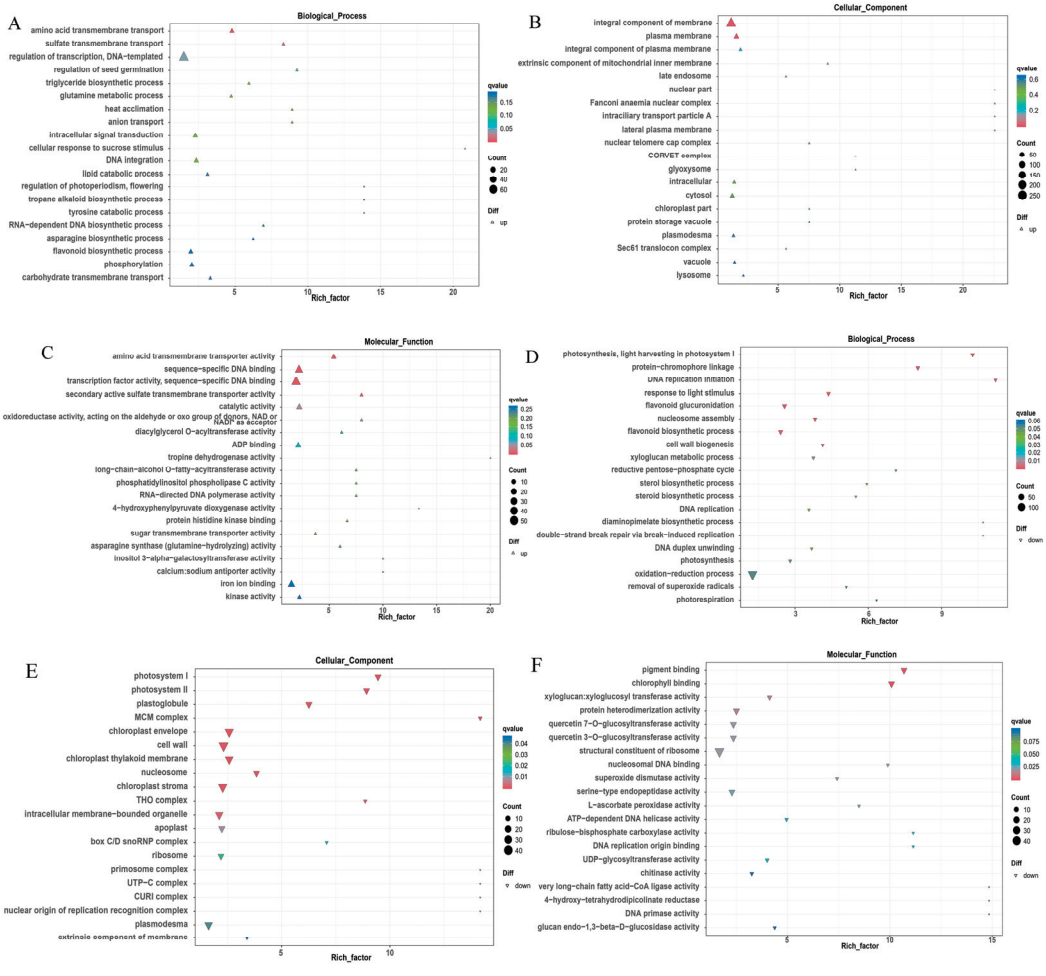


Figure 3. Go enrichment analysis of up—(**A**–**C**) and down—(**D**–**F**) regulated DEGs.

Our KEGG enrichment analysis is shown in Figure 4A. The pathway "galactose metabolism", "fatty acid degradation", "amino acids (valine, leucine and isoleucine) degradation", "tyrosine metabolism" and "α-linolenic acid metabolism" were primarily clustered. From a wider range of KEGG enrichment results, "protein processing in endoplasmic reticulum", "ubiquitin mediated proteolysis", "plant hormone signal transduction", and the "phosphatidylinositol signaling process" were also enriched (Figure S2). Downregulated DEGs clustered in "antenna proteins", "DNA replication", "ribosome", "glutathione metabolism", "steroid biosynthesis" and "ribosome biogenesis" (Figure 4B). These results point to accelerated energy consumption, decreased growth, and development processes in OE.

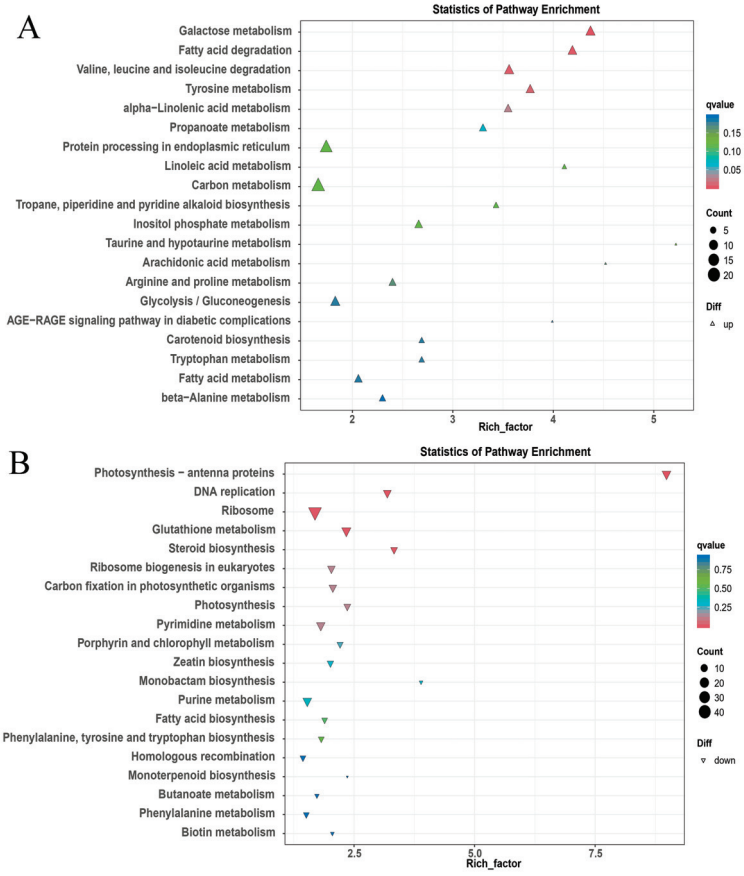


Figure 4. KEGG analysis of up—(**A**) and down—(**B**) regulated DEGs.

3.3. Analysis of the Transcription Factor among DEGs

GO analysis indicated that DEGs encoding transcription factors were significantly enriched in downstream genes. Through an amino acid blast in the NCBI and SGN databases (plantTFDB), 206 DEGs and 46 TF (transcription factors) families were obtained in OE (Table 1). Trihelix factors always take part in plant photosynthesis, growth, and development [29,30]. Four genes of the trihelix family in OE were clearly regulated, including upregulated *SlGT-31* (GT-2) and *SlGT-32* (SIP1) and downregulated *SlGT-34* (GT-2) and *SlGT-36* (GTgamma) (Figure 5). Recently, the role of the GTgamma subfamily in salt stress has been emphasized [12], but *GTgamma* gene responses in hypertrophy development have rarely been reported.

Table 1. Statistical analysis of all differentially expressed transcription factor genes.

Serial Number	TF Family	DEGs Numbers	Serial Number	TF Family	DEGs Numbers	Serial Number	TF Family	DEGs Numbers
1	AP2/ERF-AP2	2	17	E2F-DP	1	33	MYB-related	4
2	AP2/ERF-ERF	19	18	EIL	1	34	NAC	19
3	B3	5	19	GARP-ARR-B	1	35	NF-YA	4
4	B3-ARF	2	20	GARP-G2-like	2	36	NF-YB	1
5	BBR-BPC	1	21	GeBP	1	37	NF-YC	1
6	bHLH	15	22	GRAS	4	38	OFP	1
7	bZIP	10	23	HB-BELL	2	39	PLATZ	2
8	C2C2-CO-like	2	24	HB-HD-ZIP	15	40	RWP-RK	1
9	C2C2-Dof	5	25	HB-KNOX	2	41	SRS	1
10	C2C2-GATA	3	26	HB-other	4	42	TCP	6
11	C2C2-YABBY	2	27	HMG	2	43	Tify	2
12	C2H2	11	28	HSF	6	44	Trihelix	4
13	C3H	2	29	LOB	1	45	WRKY	8
14	CPP	1	30	MADS-MIKC	8	46	zf-HD	1
15	DBB	1	31	MADS-M-type	3			
16	DBP	1	32	MYB	16			

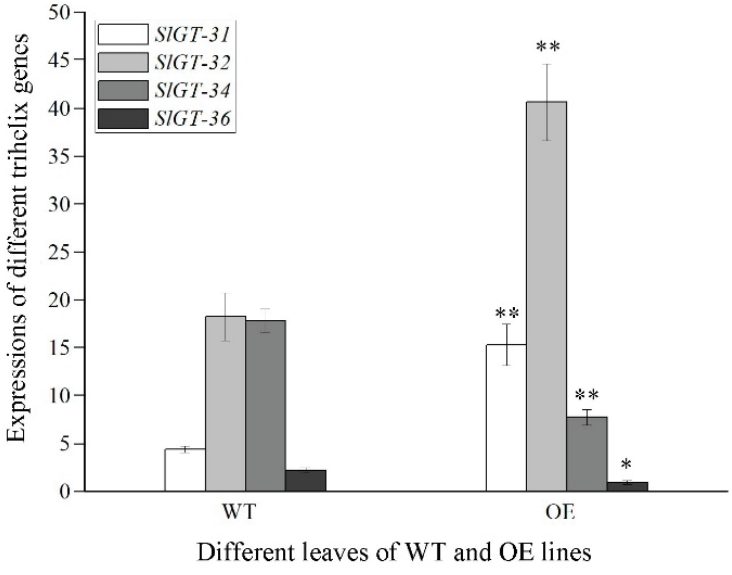


Figure 5. qRT-PCR validation of four differentially expressed trihelix genes in WT vs OE. All data are means ± standard deviation of at least three independent experiments. Significance in difference between the two groups was assessed by a Student's *t*-test using DPS software (*, $p < 0.05$; **, $p < 0.01$).

3.4. Expression Patterns of GTgamma Genes in AC^{++} and Their Responses to External Stimuli

Given that *GTgamma* is a downstream gene of SlbHLH22 protein and has a positive function in salt tolerance in rice [12], *GTgamma* responses to salt treatments and expression patterns were investigated in tomatoes. We tested the expression profiles of 11 different organs of the tomato cultivar AC^{++}. Two *GTgamma* genes (*SlGT-7* and *SlGT-36*) were expressed in the leaves of AC^{++}, especially *SlGT-36*. *SlGT-7* displayed significantly higher expressions in B+4 and B+7 (Figure 6A). *SlGT-36* transcripts accumulated the lowest in the B stage (Figure 6B). Thus, the expression patterns of two *GTgamma* genes exhibited tissue specificity.

To examine the endogenous response of *GTgamma* genes to salinity, 35-day-old tomato seedlings were watered with salinity (Figure 7). Both *SlGT-7* and *SlGT-36* were gradually induced to 2~2.5 fold at 12 h and then suddenly suppressed to less than 50% at 24 h in leaves. In the next two days, they remained at a low level (Figure 7A,B). In seedling roots, *SlGT-7* was gradually upregulated to about 4.5-fold within 48 h and then downregulated

(Figure 7C,D). The experimental results suggested that both *GTgamma* genes were repressed in leaves due to salinity stress.

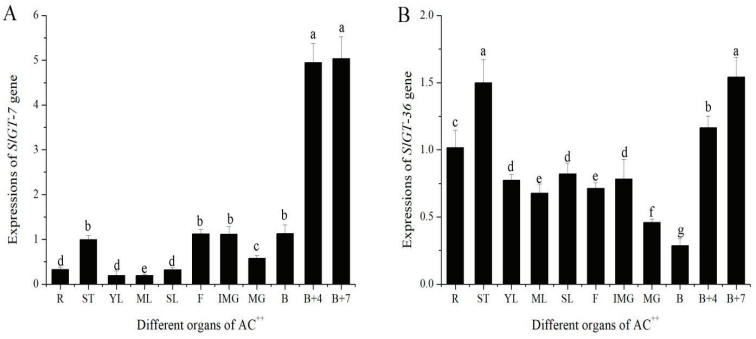


Figure 6. Expressions patterns of *GTgamma* genes, *SlGT-7* (**A**) and *SlGT-36* (**B**) in AC^{++}. R: roots; ST: stem; YL: young leaves; ML: mature leaves; SL: senescent leaves; F: flowers; IMG: immature green fruit; mature green fruit; breaker fruit; B+4: 4 days after breaker fruit; B+7: 7 days after breaker fruit; All data are means ± standard deviation of at least three independent experiments. The different letters above the column indicated that significant expressions of *GTgamma* genes among diverse organs were assessed by ANOVA ($p < 0.05$) using DPS software.

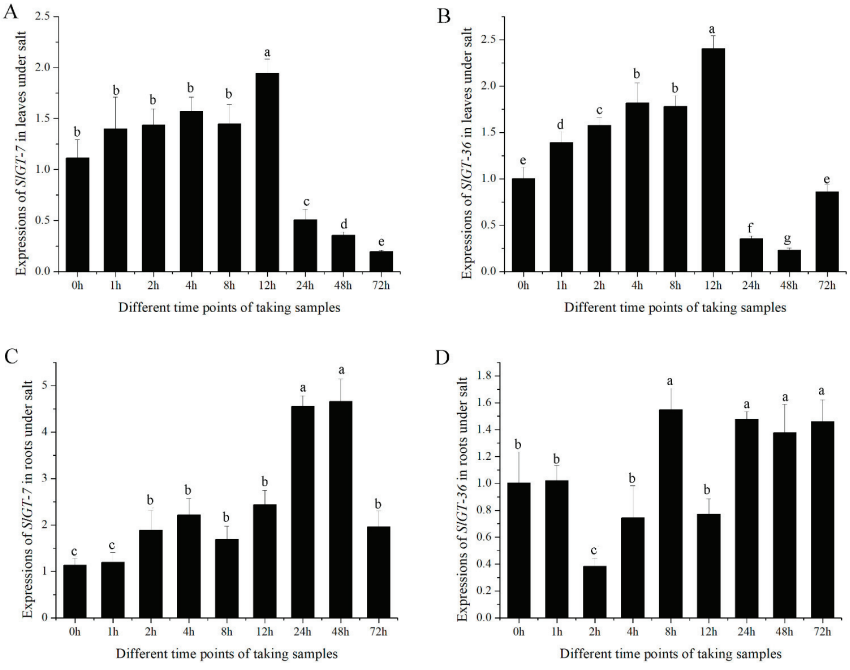


Figure 7. Expressions of *GTgamma* genes, *SlGT-7* (**A**,**C**) and *SlGT-36* (**B**,**D**) in salt stress. The leaves and roots of a 35-day-old AC^{++} seedling were used. All data are means ± standard deviation of at least three independent experiments. The different letters above the column indicate that significant expressions of *GTgamma* genes among diverse time points were assessed by ANOVA ($p < 0.05$) using DPS software.

To find the putative signaling pathway, *SlGT-7* and *SlGT-36* were treated with seven hormones. The expression levels of both *GTgamma* genes were higher in all hormonal treat-

ments than in water spraying after 8 h (Figure 8A−D). Within 24 h, *SlGT-7* and *SlGT-36* were maintaining higher levels than controls under ABA treatments (Figure 8A,C). In addition, both *GTgamma* genes showed sensitivity to other hormonal stimuli (Figure 8B,D). These results suggested that *GTgamma* genes might participate in the ABA signaling pathway.

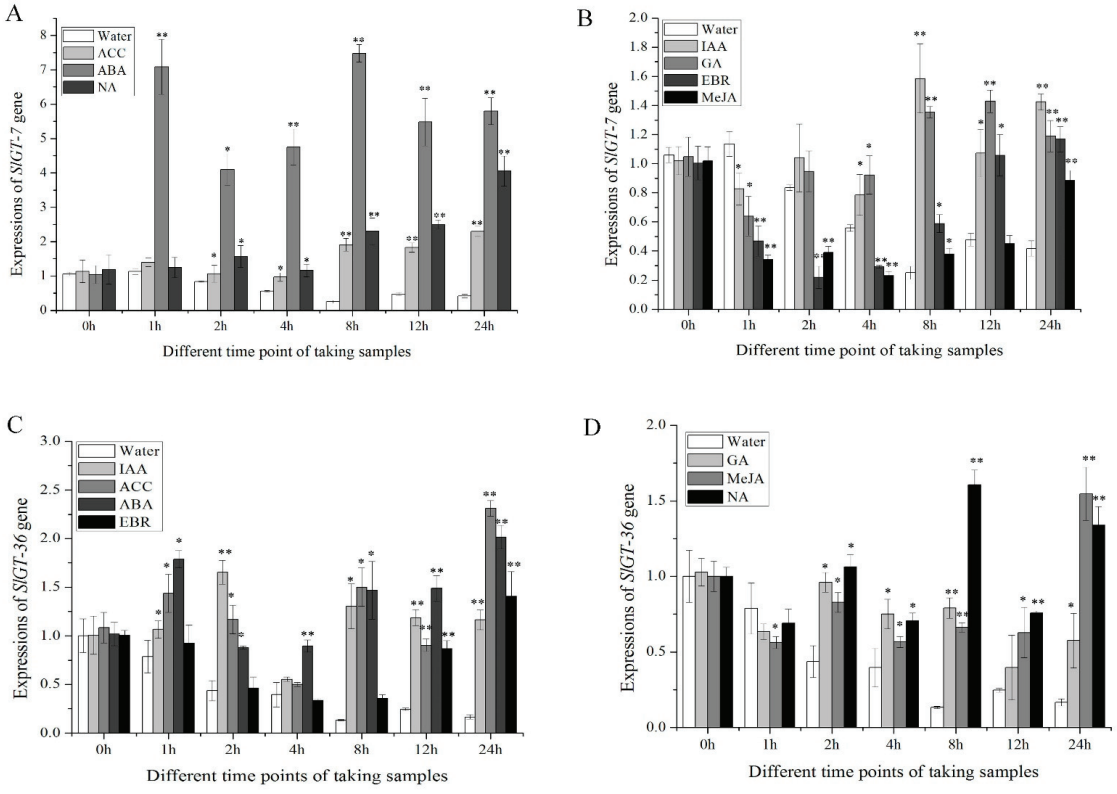


Figure 8. Expressions of two *GTgamma* genes, *SlGT-7* (**A**,**B**) and *SlGT-36* (**C**,**D**), in hormonal treatments. IAA: 3-Indoleacetic Acid; GA: Gibberellin; ACC: 1-Aminocyclopropane-1-Carboxylicacid; ABA: Abscisic Acid; MeJA: Methyl Jasmonic Acid; EBR: Epibrassinolide; NA: Uniconazole. The leaves of 35-day-old AC^{++} seedlings were used. All data are means ± standard deviation of at least three independent experiments. Significance in different expressions of *GTgamma* genes between hormonal treatments and control were assessed by a Student's *t*-test using DPS software (*, $p < 0.05$; **, $p < 0.01$).

3.5. Three-Dimensional Structures of SlGT-7 and Its Potential Phosphorylation Site

Transcription factors have a critical role in plant physiology and development, and most of these events are commonly mediated by protein phosphorylation [15,16]. To anticipate the posttranscriptional modification of GTgamma factors, a three-dimensional model of SlGT-7 was built. Using SWISS-MODEL, the lowest energy structure of SlGT-7 is shown as ribbon models in Figure 9A. In this model, two classical domains were found including triple-helix (Helix 1, Helix 2 and Helix 3) and the fourth helix at the C-terminal. SlGT-7 looked like an ellipse with a hole on one side (Figure 9B). ATP molecules putatively entered into the hole and interacted with SlGT-7 at the lowest energy (−6.35 kcal/mol) (Figure 9C–E). Further analysis showed that five amino acids (S^{96}, Y^{196}, Q^{199}, N^{200} and R^{201}) inside the hole interacted with the ATP molecules via hydrogen bonds (Figure 9F). The distance estimation of γ-phosphate to five amino acids implied that S^{96} in Helix 1 was the potential phosphorylation site.

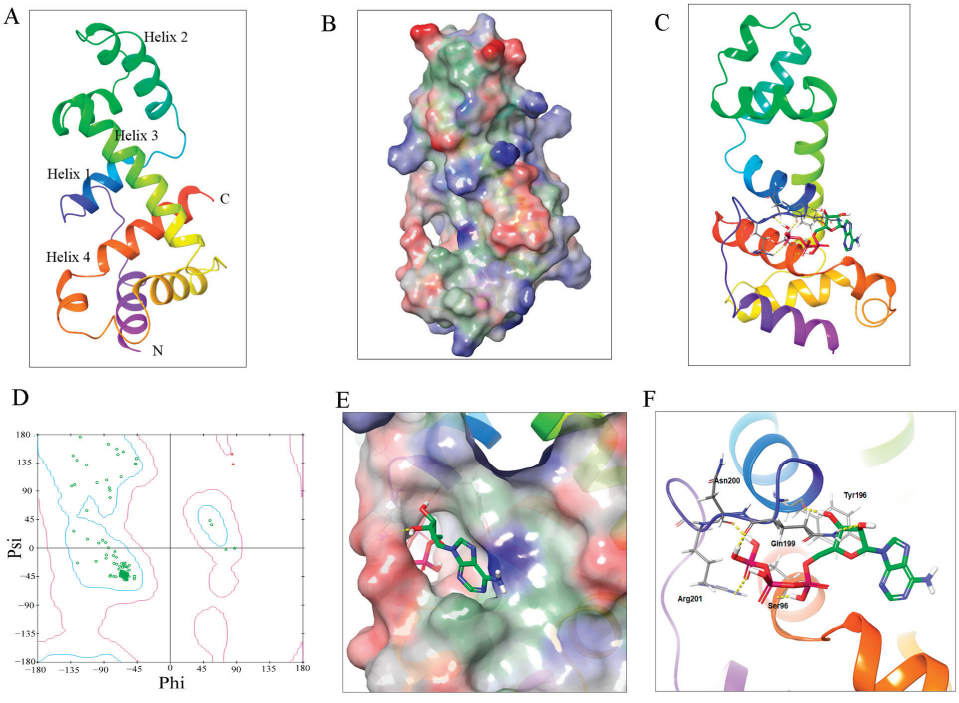


Figure 9. Construction of SlGT-7 model and interaction between SlGT-7 and ATP molecular by autodock. (**A**): The ribbon models of SlGT-7; (**B**): Three-dimensional model of SlGT-7 protein. Electrostatic potential: Positive (blue), negative (red) and hydrophobic (green); (**C**): Interactions of SlGT-7 and ATP molecular as ribbon models; (**D**): Ramachandran plot showing the lowest energy of all the amino acids interacting with the ATP molecular. Phi and Psi represent the rotation angle of the C-N and C-C bonds of α carbon in every peptide unit, respectively. Blue curves indicate the low energy and red the high energy. The dot represents amino acid; (**E**): The putative action site of the SlGT-7 model and ATP molecular; (**F**): The binding of SlGT-7 and ATP by hydrogen bonds.

4. Discussion

Xanthomonas delivers TALes into plant cells to overcome a plant's defense [1]. Like a transcription factor, AvrBS3, one TALe targets *UPA20* to induce hypertrophy development in pepper leaves, which promotes the infection of *Xanthomonas* [2,3] and *SlUPA-like* (*SlbHLH22*) functions in tomato leaves [2,4]. To reveal the malformation development of OE leaves in more depth, transcriptome and metabolome analyses were carried out in WT vs. OE. The metabolome results showed that the following metabolites were over-accumulated: Acetylserine, O-Acetyl-L-serine (OAS), Glucono-1,5-lactone, Gluconate, 2-Oxoglutarate (2-OG), and Loganate (Figure 1). OAS accumulations are related to resistance to salt stress [24,25], which was analogous to the biological function of the GTgamma factor in rice [12]. Gluconate induces increased abiotic stress resistance in plants [28]. 2-OG is linked to the metal toxicity alleviatory of tomato and hormonal synthesis in the sulfate-dependent or independent pathway [26,31], which was similar to our results in the GO analysis (Figure 3). Through RNA-seq analysis, 1299 and 1516 DEGs were, respectively, up- and downregulated (Figure 2). The transcriptome enrichment results indicated that weak photosynthesis, high-energy consumption, increased transcription factor activity, and sulfate transmembrane transport occurred in OE (Figures 3 and 4). Loganate has the capability of scavenging against superoxide radicals [25]. In addition, *SlbHLH22* (also called *SlUPA-like*) enhances plant salinity [13,32]. Therefore, both transcriptome and metabolome

analyses suggested that the hypertrophy phenotypes of OE lines might be connected with promoting salt or oxidative resistance.

Further research showed that the *GTgamma* gene was not only suppressed in hypertrophy leaves, but also inhibited by salt stress. The GO analysis showed that these biological processes, e.g., "light harvesting", "photosynthesis", "responses to light stimulus", "flavonoid synthesis", etc., were prominently restrained in OE, which always took place in the trihelix factor [30,33]. Fortunately, four trihelix genes exhibited remarkable regulation: increased *SlGT-31*(GT-2) and *SlGT-32* (SIP1) and decreased *SlGT-34* (GT-2) and *SlGT-36* (GTgamma) (Figure 5). Furthermore, six metabolites (Acetylserine, OAS, Glucono-1,5-lactone, Gluconate, 2-OG and Loganate) had a possible role in promoting salt or oxidant tolerance [24–28]. It was reported that GTgamma played the role of a positive regulator in salt stress in rice and that *SlbHLH22* boosted salt resistance in tomatoes [11–13]. These results implied that *GTgamma*, as downstream genes of SlbHLH22 protein, might perform a salt-resistant function in tomatoes. Figure 7 shows that both *GTgamma* genes were prominently inhibited by salt stress, implying a consistent role in malformation development of the OE line and salt stress.

Through an extensive analysis of the *GTgamma* genes, we found that two *GTgamma* genes were expressed in AC^{++} leaves, especially *SlGT-36*, indicating the reason why only one *GTgamma* gene was repressed by *SlbHLH22* in hypertrophy. Tissue-specific expression patterns were present when *SlGT-7* transcripts were specifically expressed in B+4 and B+7 stages fruit and *SlGT-36* in all tissues except B stage fruit (Figure 6), which was slightly different from Yu et al. [34], indicating the following different varieties: AC^{++} and LA1777. In addition, *SlGT-7* was remarkably upregulated by ABA, which was very similar to *OsGTgamma-1* [11]. Both *SlGT-7* and *SlGT-36* responded to all selected phytohormone, indicating their versatile role in plant growth and development (Figure 8). Moreover, we also found that water inhibited *SlGT-7* and *SlGT-36* expressions by over 60% in the leaves of AC^{++} seedlings (Figure 8). Whether SlGT-7 was involved in the regulation of water stress needs more evidence.

Protein posttranslational modification is a fine-tuned mechanism in abiotic or biotic resistance [6,9,15–17]. Therefore, we hypothesized that GTgamma performed this function via phosphorylation but required further experimental evidence support. We constructed a three-dimensional model of SlGT-7 as a candidate. We discovered the interactions between ATP and SlGT-7 in a putative hole (Figure 9). We also predicted that S^{96} was the most likely phosphorylation site. It was commonly believed that protein kinases transfer γ-phosphate from ATP to Ser (S), Thr (T), or Tyr (Y) during protein modification [35]. Our model implied that S^{96} got closer to the γ-phosphate of ATP than others, suggesting the phosphorylation site of S^{96} (Figure 9F). In short, our present findings about the posttranslational modification model of the GTgamma protein provide the foundation for an in-depth study of the hypertrophy development of OE lines and the regulatory role of downstream genes in tomatoes.

5. Conclusions

Xanthomonas injects TALes into the host cells to suppress plant immune defense. One TALe, AvrBS3, activates the plant target gene: pepper *upa20*. The overexpression of *SlbHLH22* (also called *SlUPA-like*), i.e., the orthology of *upa20*, causes the hypertrophy and susceptibility of *Xanthomonas* in tomatoes. The metabolome analysis showed that specific metabolites were over-accumulated in OE with a potential role in promoting salt resistance. The transcriptome analysis verified that OE plants suffered from high energy consumption, weak photosynthesis, and increased transcription factors activity. *GTgamma* gene expression was suppressed by *SlbHLH22*. Furthermore, it was simultaneously inhibited by salt stress, indicating GTgamma's role in the formation of hypertrophy development via the salt stress response. Extensive analysis proved that both *GTgamma* genes expressed in leaves were induced by ABA. Moreover, the GTgamma protein had a putative phosphorylation site at

S[96]. Our results provide the basis for disclosing the pathogenic mechanism of hypertrophy development medicated by the GTgamma subfamily.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo13121195/s1, Table S1: The primers used for qRT-PCR analysis; Table S2: Quality control of RNA-seq reads in different samples; Figure S1: Key metabolites in four metabolic processes; Figure S2: Comprehensive analysis of up—(A) and down—(B) regulated DEGs by KEGG.

Author Contributions: B.C.: design, funding acquisition, laboratory experiments, data analysis, and manuscript drafting. M.Y.: laboratory experiments. J.B.: performed laboratory experiments. Z.Z.: manuscript drafting. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the National Natural Science Foundation of China (No. 31960605 and No. 32160711).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available within the article and the Supplementary Materials.

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

References

1. Romer, P.; Recht, S.; Strauss, T.; Elsaesser, J.; Schornack, S.; Boch, J.; Wang, S.; Lahaye, T. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol.* **2010**, *187*, 1048–1057. [CrossRef] [PubMed]
2. Kay, S.; Hahn, S.; Marois, E.; Hause, G.; Bonas, U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **2007**, *318*, 648–651. [CrossRef] [PubMed]
3. Lukhovitskaya, N.I.; Solovieva, A.D.; Boddeti, S.K.; Thaduri, S.; Solovyev, A.G.; Savenkov, E.I. An RNA virus-encoded zinc-finger protein acts as a plant transcription factor and induces a regulator of cell size and proliferation in two tobacco species. *Plant Cell* **2013**, *25*, 960–973. [CrossRef]
4. Cui, B.; Hu, Z.; Hu, J.; Zhang, Y.; Yin, W.; Zhu, Z.; Feng, Y.; Chen, G. Overexpression of *SlUPA-like* induces cell enlargement, aberrant development and low stress tolerance through phytohormonal pathway in tomato. *Sci. Rep.* **2016**, *6*, 23818. [CrossRef] [PubMed]
5. Mao, H.; Zhang, W.; Lv, J.; Yang, J.; Yang, S.; Jia, B.; Song, J.; Wu, M.; Pei, W.; Ma, J.; et al. Overexpression of cotton Trihelix transcription factor GhGT-3b_A04 enhances resistance to *Verticillium dahliae* and affects plant growth in *Arabidopsis thaliana*. *J. Plant Physiol.* **2023**, *283*, 153947. [CrossRef]
6. Li, B.; Jiang, S.; Yu, X.; Cheng, C.; Chen, S.; Cheng, Y.; Yuan, J.S.; Jiang, D.; He, P.; Shan, L. Phosphorylation of trihelix transcriptional repressor ASR3 by MAP KINASE4 negatively regulates *Arabidopsis* immunity. *Plant Cell* **2015**, *27*, 839–856. [CrossRef]
7. Wang, Y.; Tang, M.; Zhang, Y.; Huang, M.; Wei, L.; Lin, Y.; Xie, J.; Cheng, J.; Fu, Y.; Jiang, D.; et al. Coordinated regulation of plant defense and autoimmunity by paired trihelix transcription factors ASR3/AITF1 in *Arabidopsis*. *New Phytol.* **2023**, *237*, 914–929. [CrossRef]
8. Zhang, Q.; Zhong, T.; E, L.; Xu, M.; Dai, W.; Sun, S.; Ye, J. GT Factor ZmGT-3b Is Associated with Regulation of Photosynthesis and Defense Response to *Fusarium graminearum* Infection in Maize Seedling. *Front. Plant Sci.* **2021**, *12*, 724133. [CrossRef]
9. Yu, C.; Song, L.; Song, J.; Ouyang, B.; Guo, L.; Shang, L.; Wang, T.; Li, H.; Zhang, J.; Ye, Z. *ShCIGT*, a Trihelix family gene, mediates cold and drought tolerance by interacting with SnRK1 in tomato. *Plant Sci.* **2018**, *270*, 140–149. [CrossRef]
10. David, S.; Levin, E.; Fallik, E.; Alkalai-Tuvia, S.; Foolad, M.R.; Lers, A. Physiological genetic variation in tomato fruit chilling tolerance during postharvest storage. *Front. Plant Sci.* **2022**, *13*, 991983. [CrossRef]
11. Fang, Y.; Xie, K.; Hou, X.; Hu, H.; Xiong, L. Systematic analysis of GT factor family of rice reveals a novel subfamily involved in stress responses. *Mol. Genet. Genom. MGG* **2010**, *283*, 157–169. [CrossRef]
12. Liu, X.; Wu, D.; Shan, T.; Xu, S.; Qin, R.; Li, H.; Negm, M.; Wu, D.; Li, J. The trihelix transcription factor OsGTgamma-2 is involved adaption to salt stress in rice. *Plant Mol. Biol.* **2020**, *103*, 545–560. [CrossRef] [PubMed]
13. Waseem, M.; Rong, X.; Li, Z. Dissecting the Role of a Basic Helix-Loop-Helix Transcription Factor, SlbHLH22, under Salt and Drought Stresses in Transgenic *Solanum lycopersicum* L. *Front. Plant Sci.* **2019**, *10*, 734. [CrossRef] [PubMed]
14. Marechal, E.; Hiratsuka, K.; Delgado, J.; Nairn, A.; Qin, J.; Chait, B.T.; Chua, N.H. Modulation of GT-1 DNA-binding activity by calcium-dependent phosphorylation. *Plant Mol. Biol.* **1999**, *40*, 373–386. [CrossRef] [PubMed]

15. Nagata, T.; Niyada, E.; Fujimoto, N.; Nagasaki, Y.; Noto, K.; Miyanoiri, Y.; Murata, J.; Hiratsuka, K.; Katahira, M. Solution structures of the trihelix DNA-binding domains of the wild-type and a phosphomimetic mutant of *Arabidopsis* GT-1: Mechanism for an increase in DNA-binding affinity through phosphorylation. *Proteins* **2010**, *78*, 3033–3047. [CrossRef] [PubMed]
16. O'Brien, M.; Kaplan-Levy, R.N.; Quon, T.; Sappl, P.G.; Smyth, D.R. PETAL LOSS, a trihelix transcription factor that represses growth in *Arabidopsis thaliana*, binds the energy-sensing SnRK1 kinase AKIN10. *J. Exp. Bot.* **2015**, *66*, 2475–2485. [CrossRef]
17. Fu, M.; Li, F.; Zhou, S.; Guo, P.; Chen, Y.; Xie, Q.; Chen, G.; Hu, Z. Trihelix transcription factor SlGT31 regulates fruit ripening mediated by ethylene in tomato. *J. Exp. Bot.* **2023**, *74*, 5709–5721. [CrossRef]
18. Fang, X.; Wang, Y.; Cui, J.; Yue, L.; Jiang, A.; Liu, J.; Wu, Y.; He, X.; Li, C.; Zhang, J.; et al. Transcriptome and metabolome analyses reveal the key genes related to grain size of big grain mutant in Tartary Buckwheat (*Fagopyrum tartaricum*). *Front. Plant Sci.* **2022**, *13*, 1079212. [CrossRef]
19. Liu, W.; Zhang, Y.; Li, W.; Lin, Y.; Wang, C.; Xu, R.; Zhang, L. Genome-wide characterization and expression analysis of soybean trihelix gene family. *PeerJ* **2020**, *8*, e8753. [CrossRef]
20. Zhu, M.; Chen, G.; Zhang, J.; Zhang, Y.; Xie, Q.; Zhao, Z.; Pan, Y.; Hu, Z. The abiotic stress-responsive NAC-type transcription factor SlNAC4 regulates salt and drought tolerance and stress-related genes in tomato (*Solanum lycopersicum*). *Plant Cell Rep.* **2014**, *33*, 1851–1863. [CrossRef] [PubMed]
21. Yang, Y.; Mao, L.; Guan, W.; Wei, X.; Shao, Y.; Luo, Z.; Lin, X.; Li, L. Exogenous 24-epibrassinolide activates detoxification enzymes to promote degradation of boscalid in cherry tomatoes. *J. Sci. Food Agric.* **2021**, *101*, 2210–2217. [CrossRef] [PubMed]
22. Zhu, M.; Chen, G.; Zhou, S.; Tu, Y.; Wang, Y.; Dong, T.; Hu, Z. A new tomato NAC (NAM/ATAF1/2/CUC2) transcription factor, SlNAC4, functions as a positive regulator of fruit ripening and carotenoid accumulation. *Plant Cell Physiol.* **2014**, *55*, 119–135. [CrossRef] [PubMed]
23. Abdelhedi, O.; Nasri, R.; Mora, L.; Jridi, M.; Toldra, F.; Nasri, M. In silico analysis and molecular docking study of angiotensin I-converting enzyme inhibitory peptides from smooth-hound viscera protein hydrolysates fractionated by ultrafiltration. *Food Chem.* **2018**, *239*, 453–463. [CrossRef] [PubMed]
24. Apodiakou, A.; Hoefgen, R. New insights into the regulation of plant metabolism by O-acetylserine: Sulfate and beyond. *J. Exp. Bot.* **2023**, *74*, 3361–3378. [CrossRef] [PubMed]
25. Kusakari, K.; Fukuhara, T.; Motoyama, A.; Ochiai, N.; Watanabe, T.; Sugimoto, Y. The corrected structure of depressoside, an antioxidative iridoid glucoside extracted from the flowers of Gentiana urnula Harry Sm. *Nat. Prod. Res.* **2016**, *30*, 954–959. [CrossRef]
26. Araujo, W.L.; Martins, A.O.; Fernie, A.R.; Tohge, T. 2-Oxoglutarate: Linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Front. Plant Sci.* **2014**, *5*, 552. [CrossRef]
27. Hubberten, H.M.; Klie, S.; Caldana, C.; Degenkolbe, T.; Willmitzer, L.; Hoefgen, R. Additional role of O-acetylserine as a sulfur status-independent regulator during plant growth. *Plant J.* **2012**, *70*, 666–677. [CrossRef]
28. Ding, F.; Wang, R.; Chen, B. Effect of exogenous ammonium gluconate on growth, ion flux and antioxidant enzymes of maize (*Zea Mays* L.) seedlings under NaCl stress. *Plant Biol.* **2019**, *21*, 643–651. [CrossRef]
29. Green, P.J.; Kay, S.A.; Chua, N.H. Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the *rbcS-3A* gene. *EMBO J.* **1987**, *6*, 2543–2549. [CrossRef]
30. Kaplan-Levy, R.N.; Brewer, P.B.; Quon, T.; Smyth, D.R. The trihelix family of transcription factors—Light, stress and development. *Trends Plant Sci.* **2012**, *17*, 163–171. [CrossRef] [PubMed]
31. Alamri, S.; Alsubaie, Q.D.; Al-Amri, A.A.; Al-Munqedi, B.; Ali, H.M.; Kushwaha, B.K.; Singh, V.P.; Siddiqui, M.H. Priming of tomato seedlings with 2-oxoglutarate induces arsenic toxicity alleviatory responses by involving endogenous nitric oxide. *Physiol. Plant* **2021**, *173*, 45–57. [CrossRef] [PubMed]
32. Waseem, M.; Li, N.; Su, D.; Chen, J.; Li, Z. Overexpression of a basic helix-loop-helix transcription factor gene, *SlbHLH22*, promotes early flowering and accelerates fruit ripening in tomato (*Solanum lycopersicum* L.). *Planta* **2019**, *250*, 173–185. [CrossRef] [PubMed]
33. Le Gourrierec, J.; Li, Y.F.; Zhou, D.X. Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. *Plant J.* **1999**, *18*, 663–668. [CrossRef]
34. Yu, C.; Cai, X.; Ye, Z.; Li, H. Genome-wide identification and expression profiling analysis of trihelix gene family in tomato. *Biochem. Biophys. Res. Commun.* **2015**, *468*, 653–659. [CrossRef]
35. Minguez, P.; Parca, L.; Diella, F.; Mende, D.R.; Kumar, R.; Helmer-Citterich, M.; Gavin, A.C.; van Noort, V.; Bork, P. Deciphering a global network of functionally associated post-translational modifications. *Mol. Syst. Biol.* **2012**, *8*, 599. [CrossRef]

metabolites

MDPI

Article

Isolation and Identification of Alkaloid Genes from the Biomass of *Fritillaria taipaiensis* P.Y. Li

Nong Zhou [1,2], Chun-Mei Mei [1], Fu-Gui Chen [1], Yu-Wei Zhao [1], Ming-Guo Ma [3,*] and Wei-Dong Li [1,*]

[1] College of Pharmacy, Nanjing University of Chinese Medicine, Nanjing 210023, China; 20120038@sanxiau.edu.cn (N.Z.); 20210695@njucm.edu.cn (C.-M.M.); 20210690@njucm.edu.cn (F.-G.C.); 20221018@njucm.edu.cn (Y.-W.Z.)
[2] College of Food and Biological Engineering, Chongqing Three Gorges University, Chongqing 404120, China
[3] College of Materials and Science and Technology, Beijing Forestry University, Beijing 100083, China
* Correspondence: mg_ma@bjfu.edu.cn (M.-G.M.); liweidong0801@njucm.edu.cn (W.-D.L.)

Abstract: Background/Objectives: *Fritillaria taipaiensis* P.Y. Li is a valuable traditional Chinese medicinal herb that utilizes bulbs as medicine, which contain multiple alkaloids. Biomass, as a sustainable resource, has promising applications in energy, environmental, and biomedical fields. Recently, the biosynthesis and regulatory mechanisms of the main biomass components of biomass have become a prominent research topic. Methods: In this article, we explored the differences in the heterosteroidal alkaloid components of *F. taipaiensis* biomass using liquid chromatography–mass spectrometry and high-throughput transcriptome sequencing. Results: The experimental results demonstrated significant differences in the eight types of heterosteroidal alkaloid components among the biomass of *F. taipaiensis*, including peimisine, imperialine, peimine, peiminine, ebeinone, ebeiedine, ebeiedinone, and forticine. Transcriptomic analysis revealed substantial significant differences in gene expression patterns in the various samples. Three catalytic enzyme-coding genes, 3-hydroxy-3-methylglutaryl coenzyme A synthase (*HMGS*), 3-hydroxy-3-methylglutaryl coenzyme A reductase (*HMGR*), and terpene synthase (*TPS*), were speculated to contribute to the regulation of the differential accumulation of alkaloid synthesis in *F. taipaiensis* bulbs. A strong positive correlation was observed between the transcriptional level of the *TPS* gene and the alkaloid content of *F. taipaiensis* biomass, suggesting that *TPS* may be a key gene in the biosynthesis pathway of alkaloids. This finding can be used for subsequent gene function verification and molecular regulatory network analysis. Conclusions: This work provides fundamental data and novel insights for the subsequent research on alkaloid biosynthesis in *F. taipaiensis*.

Keywords: biomass; alkaloid genes; identification; metabolic group; transcriptome

Citation: Zhou, N.; Mei, C.-M.; Chen, F.-G.; Zhao, Y.-W.; Ma, M.-G.; Li, W.-D. Isolation and Identification of Alkaloid Genes from the Biomass of *Fritillaria taipaiensis* P.Y. Li. *Metabolites* **2024**, *14*, 590. https://doi.org/10.3390/metabo14110590

Academic Editors: Yan Li and Yanjie Zhang

Received: 28 August 2024
Revised: 20 October 2024
Accepted: 24 October 2024
Published: 31 October 2024

1. Introduction

Biomass is a sustainable resource, that has wide applications in the energy, environment, and biomedical fields [1–3]. Currently, research has focused on the separation, component analysis, and active alkaloid ingredients of biomass [4]. In particular, the biosynthesis and regulatory mechanisms of the main active components of biomass have garnered increasing attention [5,6]. Recently, transcriptome sequencing has emerged as a promising strategy for exploring gene expression regulation [7,8]. Transcriptomes can reflect the specific expression of biomass genes at a specific time and space, effectively demonstrating the differential expression of genes after the introduction of biomass to different regions, resulting in inconsistent secondary metabolites and variations in effective ingredients [9]. For instance, Zhao et al. reported transcriptome analysis and plotted the gene expression profile of *Fritillaria cirrhosa* [10], which is a valuable traditional Chinese medicinal herb that uses bulbs as medicine and is listed in the Pharmacopoeia of the People's Republic of China under the name of "ChuanBeiMu" [11].

Heterosteroidal alkaloids are a class of alkaloids characterized by a C_{27} skeleton system, comprising five to six carbon rings in their basic skeleton structure. Their distinctive ring structure is the C-nor-D-homo-[14(13→12)-abeo] ring system, which includes the unit of the peritendinous ring [4]. Heterosteroidal alkaloids are mainly found in plants of the Liliaceae family, specifically within the genera *Fritillaria* and *Veratrum*, and are the primary components responsible for the pharmacological activity in these two plant genera [12]. Li et al. reported a transcriptome analysis of the cytochrome P450 oxidase (FcCYP) gene family in *F. cirrhosa*, indicating a positive correlation between FcCYP transcriptional expression and the accumulation of steroid alkaloids [13]. However, limited studies focus on the components and biosynthesis mechanisms of alkaloids in biomass [14,15]. In particular, the gene function of alkaloid biosynthesis and the regulatory mechanism of transcription factors in biomass remain unclear [16]. *Fritillaria taipaiensis* P.Y. Li is a perennial herb in the Liliaceae family, primarily distributed in southwestern China [17]. Its bulbs contain multiple alkaloids that can cool down body temperature, relieve coughing and asthma, and resolve phlegm; it is also listed in the Pharmacopoeia of the People's Republic of China as "ChuanBeiMu" [18].

Herein, in order to explore the alkaloid biosynthesis key gene and its regulation mechanism in *F. taipaiensis*, the alkaloid and differentially accumulated metabolite content in the *F. taipaiensis* bulb from different regions was determined and the differentially expressed genes (DEGs) in the biomass of *F. taipaiensis* were identified using integrated transcriptomics technology. This study will contribute to improving the quality of biomass by transcriptomics regulation.

2. Materials and Methods

2.1. Samples

Based on a preliminary investigation by the research group, the geographical distribution and ecological differences in specimens led to the determination of representative sampling points for biomass *F. taipaiensis* (including wild species). In July 2022, >300 3-year-old biomass *F. taipaiensis* plants were randomly excavated using the snake-shaped method and each collection point was mixed at multiple points. Bulbs of biomass of *F. taipaiensis* plants were collected as samples. Samples from each production area were biologically replicated three times, washed with water, immediately frozen in liquid nitrogen, and divided into two parts. One part was stored at −80 °C for total RNA extraction while the other part was dried by hot air at a temperature of 50–55 °C for metabolite extraction. Information and codes regarding the origins of biomass *F. taipaiensis* from different regions are shown in Table 1. The soil property information is presented in Table S1. The experimental samples were subjected to high-throughput sequencing analysis by Beijing Novogene Technology Co., Ltd. (Beijing, China). Three independent biological replicates were used for each sample.

Table 1. The information and codes of the different biomass samples.

Sample Group	Collection Site	Latitude and Longitude	Elevation/m
WXY	Yinchangping, Hongchiba, Wuxi County, Chongqing	31°37′46.30″ N/108°56′00.27″ E	2105
WX	Narrow Neck, Hongchiba Community, Wuxi County, Chongqing	31°37′32.05″ N/108°57′20.71″ E	1769
FJ	Meng Yuan Medicine Valley, Xicao Village, Xinglong Town, Fengjie County, Chongqing	30°45′47.47″ N/109°36′38.96″ E	1915
WS	Renziping, Lihe Village, Dangyang Township, Wushan County, Chongqing	30°46′07.21″ N/109°36′33.99″ E	1538
CK	Xiaolongtan, Sihe Village, Mingzhong Township, Chengkou County, Chongqing	31°41′14.08″ N/108°56′52.59″ E	1562

Table 1. *Cont.*

Sample Group	Collection Site	Latitude and Longitude	Elevation/m
NL	Yangjiancao, Cuiyu Township, Ninglang County, Lijiang City, Yunnan Province	27°31′04.91″ N/100°35′39.14″ E	3372
TB	Tangkou Village, Zuitou Township, Taibai County, Baoji City, Shaanxi Province	34°03′25.82″ N/107°23′40.79″ E	1625
GY	Qinglin Village, Lijia Town, Chaotian District, Guangyuan City, Sichuan Province	32°35′10.06″ N/106°13′17.34″ E	1760

2.2. Metabolomics Study

Approximately 1.0 g of biomass of *F. taipaiensis* powder (sieved through a 200-mesh sieve) was selected. This powder was added to 25 mL of 60% ethanol, soaked for 20 min, and refluxed at 90 °C for 1 h for extraction. These steps were repeated twice. The continuous filtrate (1 mL) was accurately measured and passed through a SPF column, which was then diluted and eluted with different ratios of methanol in a gradient ranging from low to high concentrations. The diluted sample was then centrifuged at 12,000 r.min^{-1} for 10 min, after which the supernatant was pipetted, and it was passed through a 0.22 μm filter membrane [19].

Mass spectrometry (MS) of the samples was performed in both positive and negative ion modes using a dynamic background subtraction function with a scanning time of 840 ms, equipped with an automatic calibration delivery system. The chromatographic conditions of the samples were determined by using an Agilent Zorbax Extended C18 UHPLC Column (Agilent, Beijing, China).

2.3. Transcriptomics Study

RNA integrity and total RNA were accurately assessed using an Agilent 2100 Bioanalyzer (Agilent, Beijing, China). Total RNA that met the standards was utilized for library construction. After passing the quality test, the library was subjected to high-throughput sequencing using an Illumina NovaSeq 6000 (Illumina, San Diego, CA, USA) by Beijing Novogene Technology Co., Ltd (Beijing, China).

New gene predictions were made using StringTie (version 1.3.3b). This software utilized the official National Center for Biotechnology Information Protein Sequence database (RefSeq non-redundant proteins (NR)), Nucleoside Sequence (NT) database, Kyoto Encyclopedia of Genes and Genomes (KEGG), SwissProt Protein Quality database (manually annotated and reviewed protein sequence database), Protein Family (PFAM) database, Gene Ontology (GO) database, and Eukaryotic Protein Homologous Groups (KOG) database to annotate the gene functions of the obtained transcripts. DESeq2 (1.20.0) was employed for intergroup differential expression analyses.

To verify the DEGs identified, three candidate genes (3-hydroxy-3-methylglutaryl coenzyme A synthase (*HMGS*), 3-hydroxy-3-methylglutaryl coenzyme A reductase (*HMGR*), and terpene synthase (*TPS*) were selected for quantitative reverse transcription PCR (qRT-PCR). First, 2 μg of total RNA was converted to complementary DNA (cDNA) using a one-step gDNA removal and cDNA synthesis kit (TransGene, Beijing, China). After diluting the samples five times, real-time fluorescence quantitative PCR was performed using an Eppendorf MasterCycle PCR instrument (Eppendorf, Hamburg, Germany). The *rpl16* gene (F: TTCGTGCTACATTCGTAGGGGGTC; R: GTTCCATTGCGGAGTTCGG) of the biomass of *F. taipaiensis* was used as an internal reference gene and each sample was analyzed in triplicate. Primers were designed using Primer Premier 6.0 and Beacon Designer 7.8 software (Table 2) and the relative expression of the gene was calculated using the $2^{-\Delta\Delta Ct}$ method. The PCR products were analyzed and purified by agarose gel electrophoresis and sequenced by Nanjing Aoqing Biotechnology Co., Ltd. [20].

Table 2. qRT-PCR primer information.

Gene Name	KO Number	Forward Primer (5′→3′)	Reverse Primer (5′→3′)	Temperature (°C)
HMGS	K01641	CCAACCTTGCGAGCGAATA	TTGTAGGGAGAATGGAACACGA	59.60
HMGR	K00021	TGCGAGGCTCCTGGCTACT	CTTTGCTGGACCTGTTATACTTCA	60.70
TPS	K03527	AGGATCAAATCACGGGAGGC	ATCAATCCTCCAAGTCGCCC	59.82

2.4. Data Analysis

The data are presented as mean ± standard deviation, the data were analyzed by a single factor (one-way ANOVA) for significance analysis via SPSS 25.0 and Excel 2003 ($\alpha = 0.05$), and correlation analysis was conducted on SPSS 25.0 using the Pearson method ($\alpha = 0.05$). The QIIME software (version 1.9.1) was used to calculate the Unifrac distance. Principal component analysis (PCA) was performed by vegan bag on R-software (version 4.3.2). GraphPad Prism 8 and R-software (version 4.3.2) were used for graphics analysis and Adobe Illustrator (version 28.1.0) was used for graphic processing.

3. Results

3.1. Chemical Constituents and Differential Metabolites

Potential compounds were inferred by analyzing the tandem mass spectrometry (MS/MS) fragmentation patterns of the substances (Table S2). In total, 71 compounds were identified in the positive ion mode, 34 compounds were identified in the negative ion mode, and 83 compounds were identified in both positive and negative ion modes. This included thirty alkaloids (Compounds **1**–**2**, **4**–**5**, **8**–**18**, **22**–**36**), six alkaloid glycosides (Compounds **3**, **6**–**7**, **19**–**21**), two fatty amides (Compounds 37–38), nine nucleosides (Compounds **39**–**46**, **48**), one peptide (Compound **47**), six fatty acids (Compounds **49**–**53**, **59**), six aromatic acids (Compounds **54**–**58**, **60**), ten lipids (Compounds **61**, **63**, **65**, **69**, **71**–**73**, **75**–**77**), one acid anhydride (Compound **62**), three aromatic aldehydes (Compounds **64**, **66**, **74**), one aromatic alcohol (Compound **67**), one aromatic ketone (Compound **68**), one phenol (Compound **70**), one fatty alcohol (Compound **78**), and five sugar components (Compounds **79**–**83**).

To clarify the differences in chemical compositions among different biomass samples, an orthogonal partial least squares discriminant analysis (OPLS-DA) model was developed (Figure 1). The OPLS-DA model established in both positive and negative ion modes yielded good prediction results. All simulated R^2 and Q^2 values on the left were less than the true values on the right, demonstrating reliable model validation (Figure 1A,B). The model identified a total of thirty-eight differential components in the positive ion mode and two differential components in the negative ion mode, totaling thirty-eight differential components (Figure 1C,D). These differential metabolites are mainly steroid alkaloids, including peimine, songbeisine, songbeinine, imperialine, delafrinone, peimenine, ebeinone, isodelavine, isoforticine, petilidine, puqiedine, ebeinine, delavine chuanbeinone, edebedi-none, puqiedinone, songbeinone, peimisine-3-O-β-D-glucopyranoside, isoverticine, delafrine, forticine, sipeimine-3-O-β-D-glucoside, isoverticine-β-N-oxide, yibeinoside B, yibeinoside A, puqiedinone-3-O-β-D-glucopyranoside, imperialine-β-N-oxide, cirrhosinine B, solanidine, cis-13 docosenoamide, 1,6-caprolactam, glyceryl monostearate, 1-monolinolein, linolenic acid, linoleic acid, and palmitic acid.

The relative content of the discrepant components (components sharing the same mass-to-charge ratio/retention time might be isomers) was analyzed. As depicted in Figure 2A, disparities exist in the relative content of each component among different origins and the content variations are conspicuously distinct. A clustering tendency emerges among the different components across different origins. Eight types of heterosteroidal alkaloids with significant differences among the different biomass samples were selected: peimisine, imperialine, peimine, peiminine, ebeinone, ebeiedine, ebeiedinone, and forticine. The response values of these eight significantly different indicator components from different biomass samples were statistically analyzed. As presented in Figure 2B, the contents of peimisine, imperialine, peiminine, ebeinone, and ebeiedinone were relatively high in the

samples of WXY, WX, WS, NL, and FJ and were low in the other samples. The peimine was highest in the sample of WX, followed by the samples of CK and NL, and was relatively low in the other samples. Ebeiedine content was highest in WX and lower in the other samples. Forticine was relatively high in two samples of WX and GY, whereas it was relatively low in the other samples. These results further indicate that the differential components represent significant differences in abundance, demonstrating the chemical diversity of different biomass samples.

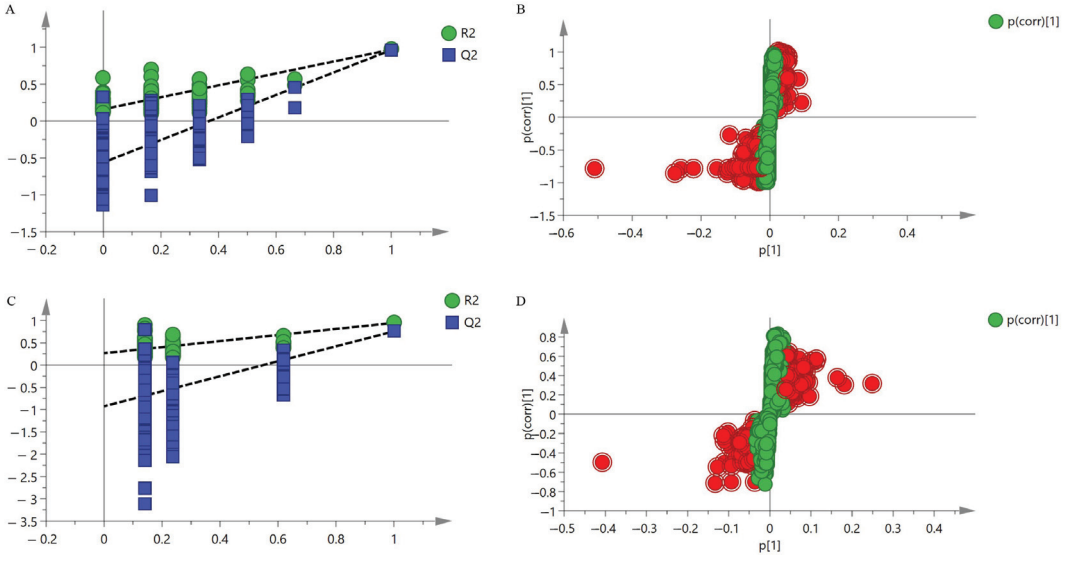


Figure 1. Plot of the OPLS-DA permutation test and S-plot scores for different biomass samples. (**A**) Permutation test in positive ion mode; (**B**) S-plot diagram in positive ion mode; (**C**) Permutation test in negative ion mode; (**D**) S-plot diagram in negative ion mode.

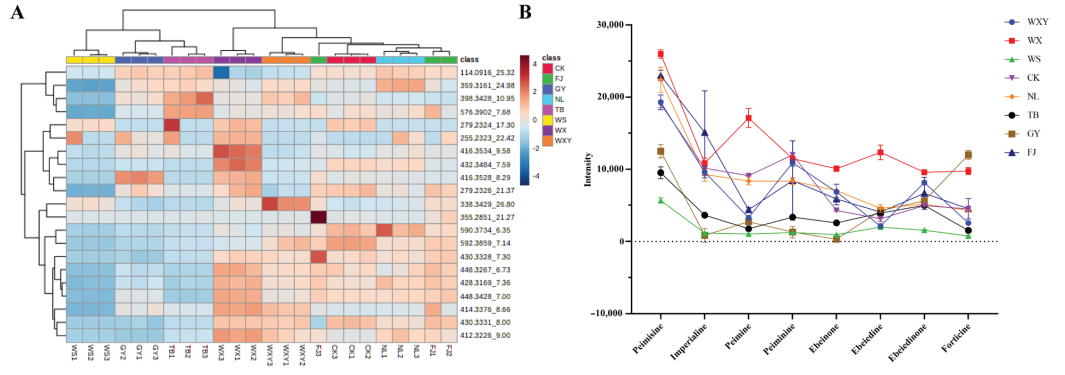


Figure 2. Results of semi-quantitative analysis of differential metabolites in different biomass samples. (**A**) Differential metabolite heat map with the classification tree of metabolites on the left and the classification tree of samples above; (**B**) Differential metabolite line graph.

3.2. Transcriptome High-Throughput Sequencing Analysis

To further investigate the key genes affecting the accumulation of alkaloids in different biomass samples, cultivated species from seven habitats were collected (with three biological replicates in each group) and sequencing libraries were constructed for transcriptome

sequencing (Table S3). Using Trinity software (version 2.15.1), high-quality sequences obtained after quality control were assembled, resulting in 359,529 transcripts and 174,827 unigenes (Figure S1). These results indicate high transcriptome sequencing and assembly integrity, which can be used for subsequent annotation analysis.

We compared all the unigenes obtained through transcriptional assembly using different protein databases to predict functional genes. The results showed that the highest number of unigenes (63,506) was annotated in the NR database, whereas the lowest number of unigenes was annotated in the KOG database, with 13,508 (Figure 3A). The numbers of unigenes annotated in the other databases were NT (35,586), KO (22,200), SwissProt (43,241), PFAM (45,206), and GO (45,198). The proportions of annotated unigenes in the seven databases were consistent with the numbers of annotated unigenes. Except for the comparison between the KOG and KO databases, where the content of annotated unigenes was <20%, the percentages of annotated unigenes in the remaining five databases were NR (36.32%), NT (20.35%), SwissProt (24.73%), PFAM (25.85%), and GO (25.85%; Figure 3B). The annotated unigenes were further enriched to clarify their biological functions. GO functional enrichment analysis demonstrated that the mainly annotated unigenes pertained to biological processes, such as cellular processes, metabolic processes, biological regulation, localization, and response to stimuli. Unigenes are mainly members of cellular components, such as cellular anatomical entities and intercellular and protein-containing complexes. From a molecular functional perspective, the main functions of the unigenes include binding, catalytic, and transporter activities (Figure 3C). The transcriptome sequencing results showed that the length of the transcripts varied between 100 and 3000 bp, with an average transcript length of approximately 500 bp, which could be used for subsequent analyses (Figure 3D).

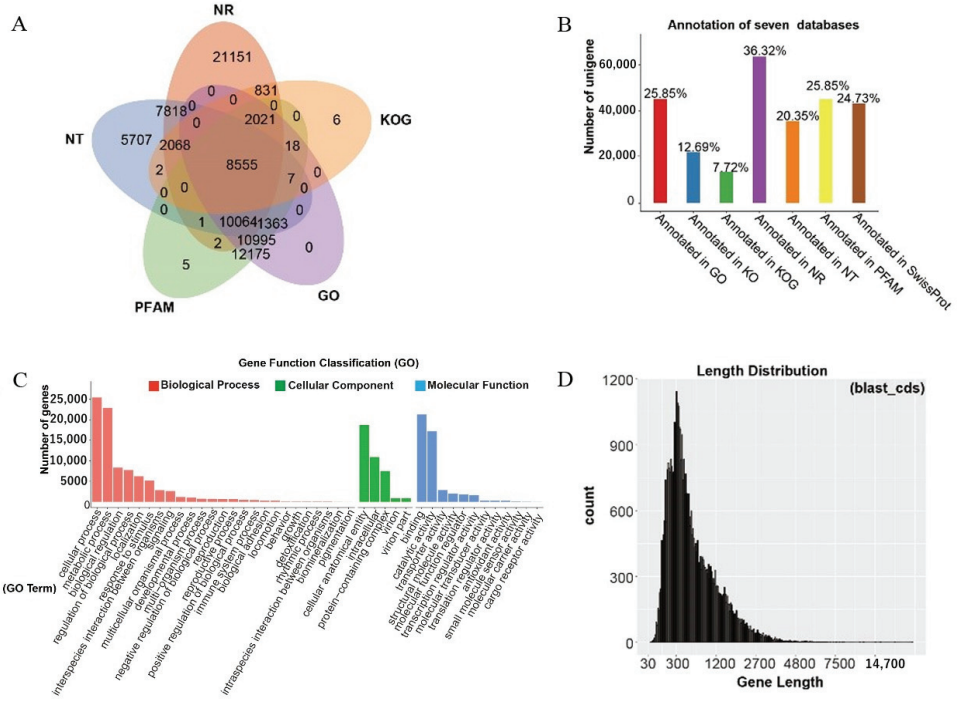


Figure 3. Transcript coding gene annotation. (**A**) Venn diagrams annotated in five databases (NR, NT, KOG, PAMF, GO). (**B**) Histogram of the proportion of annotatable unigenes in the seven databases. (**C**) Annotatable unigene GO clustering analysis. (**D**) Annotatable unigene length distribution map.

The square of the Pearson correlation coefficient (R^2) between samples was developed to represent the correlation between RNA samples from different biomass regions. The results showed that the correlation between each group of samples ranged from 0.685 to 0.878, with the maximum difference observed being between samples from WS and TB (0.685), whereas samples from WX and NL exhibited the highest correlation (0.878; Figure 4A). We further analyzed the number of DEGs in each sample, which was consistent with Pearson's correlation coefficient. The WS sample had the highest number of DEGs compared with the TB sample, totaling 18,407 DEGs, of which 9710 genes were upregulated and 8697 genes were downregulated. Notably, samples from the TB showed the most significant differences compared to samples from other groups, with a total of >10,000 DEGs compared to each group. The number of DEGs between samples from the GY and samples from other groups displayed the second-highest difference, totaling over 8000 (Figure 4B). The FJ sample displayed a lower change in its DEG number in comparison to the other samples. The DEG number of the WXY group, was higher than that of the groups WX and NL, whereas the samples of groups TB and GY showed greater differences.

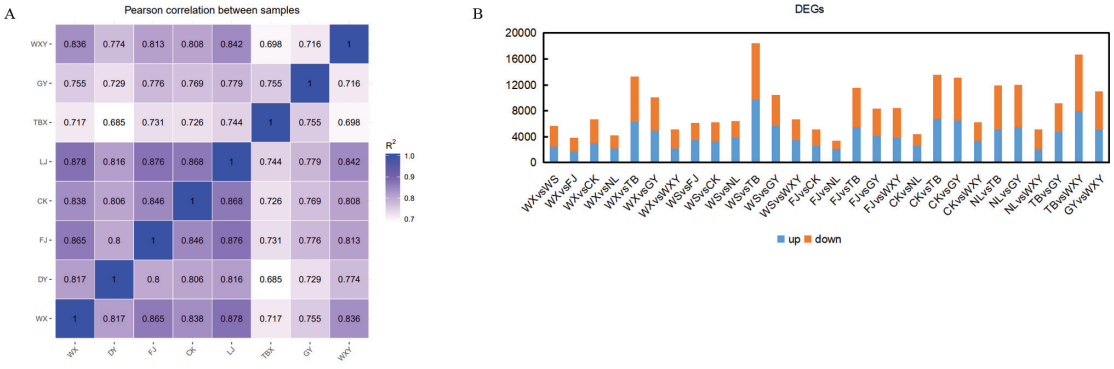


Figure 4. Differences in the number of transcripts in the biomass from different biomass samples. (**A**) Squared Pearson's correlation coefficients of transcripts from different biomass samples. (**B**) Number of differentially expressed genes in different biomass samples.

To further analyze the differences in gene expression between samples, an integrated analysis of differentially expressed genes (DEGs) across various groups and principal component analysis (PCA) were conducted. The number of DEGs was consistent with the observed differences in gene expression. The heat map of DEG clustering indicated that samples of groups NL, FJ, WX, CK, WS, and WXY were clustered into one branch, with the similarity between samples from different cultivation groups being greater than that between them and WXY. In WXY, the expression levels of various genes were more balanced compared with the other samples in the same branch and the expression levels of genes related to cell growth and death; biosynthesis of other secondary metabolites; and metabolisms of carbohydrate, lipid, cofactors, and vitamins were relatively high. In NL and FJ, the expression levels of most genes were relatively low compared with the other groups in the same branch (Figure 5A). The TB and GY groups were clustered separately into a single branch, with the higher expression levels of the genes related to the signal transduction, replication, and repair compared with the other groups, exhibiting different expression patterns compared to the WXY sample (Figure 5A). PCA revealed that samples from WXY, the four samples of WX, CK, FJ, and WS, and NL were clustered together, whereas those of the GY and TB samples were clustered separately; this is consistent with the findings of the DEG number and clustering analysis (Figure 5B). In summary, gene expression in different biomass samples is influenced by various conditions.

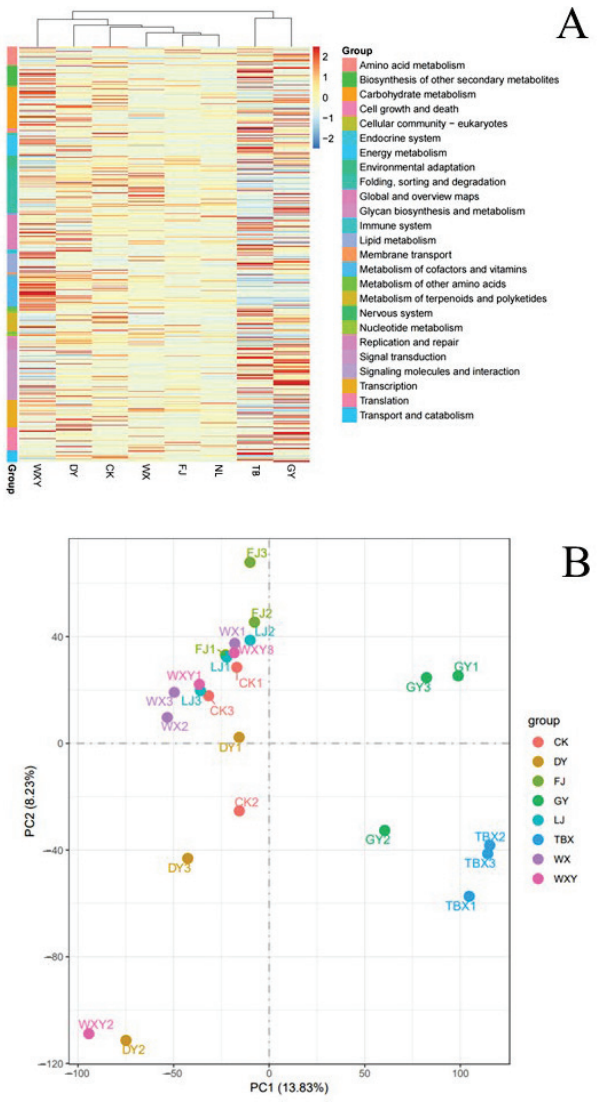


Figure 5. Differential gene clustering and principal component analysis. (**A**) Heat map of differential gene clustering. (**B**) Differential gene principal component analysis plot.

Differences in gene expression between cultivated and wild species in different biomass samples were analyzed using the WX sample as a representative of cultivated varieties. The results showed 240 DEGs (117 upregulated genes and 123 downregulated genes) in WXY and WX (Figure 6A), with the difference between these samples and the NL being primarily attributed to 163 DEGs (42 upregulated and 121 downregulated genes; Figure 6B). However, a higher DEG number was provided from the samples of WXY, GY, and TB. Specifically, the number of DEGs between the WXY and GY samples was 2631 (1417 upregulated genes and 1214 downregulated genes; Figure 6C), whereas the number of DEGs between the WXY and TB samples was 3215 (1689 upregulated and 1526 downregulated genes; Figure 6D), nearly 10 times the number of DEGs derived from the WXY, WX, and NL samples. The number of DEGs varied greatly among the different samples, indicating that the gene regulation of the WX and NL samples was more similar to that of

wild species, whereas the gene regulation of the TB and GY samples was similar to that of cultivated species.

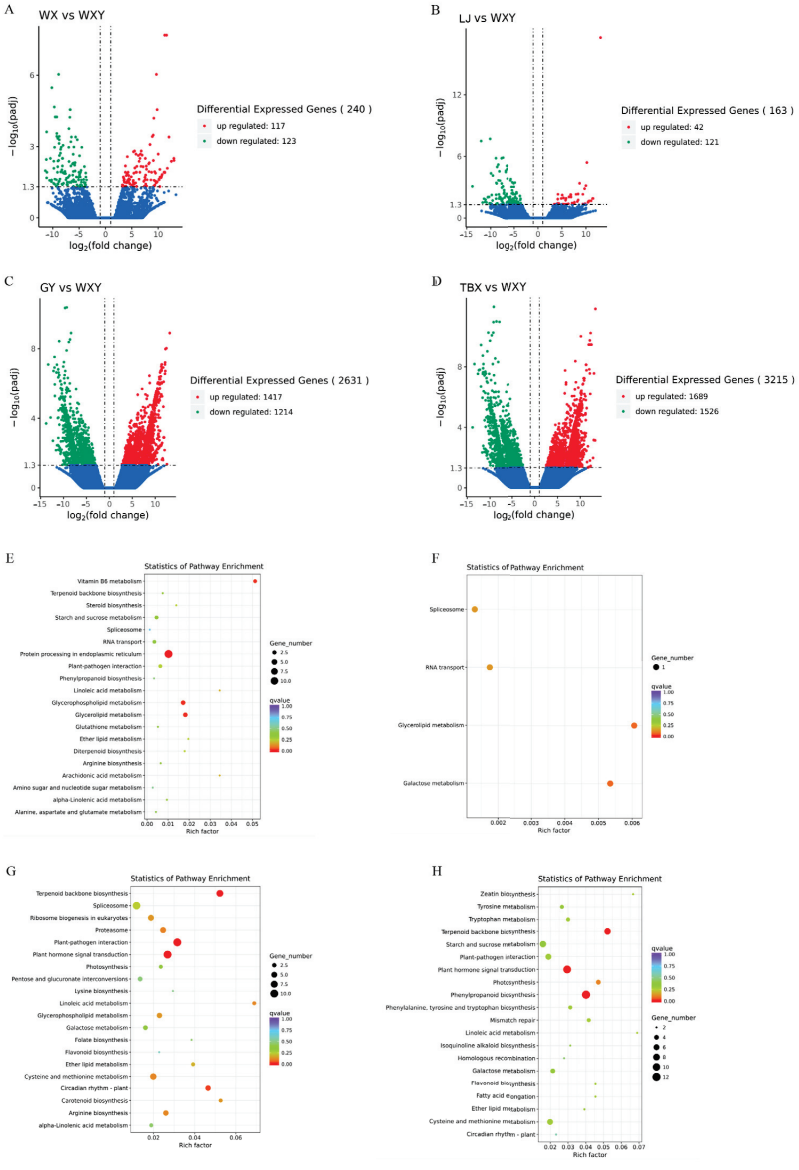


Figure 6. Differential gene expression and KEGG enrichment analysis between wild and cultivated cultivars. (**A**–**D**) Volcano map of differentially expressed genes between wild (WXY) and cultivated samples from (**A**) WX, (**B**) LJ, (**C**) GY, and (**D**) TB. (**E**–**H**) The differentially expressed gene KEEG-rich in the wild sample (WXY) and cultivated samples from (**E**) WX, (**F**) LJ, (**G**) GY, and (**H**) TB.

The KEGG pathway enrichment analysis of DEGs from the above four different samples showed that the DEGs in WXY and WX were concentrated in pathways such as protein processing in the endoplasmic reticulum, glycerol phospholipid metabolism, and glycerol lipid metabolism (Figure 6E). There were no significant differences in the metabolic pathways compared with NL (Figure 6F). In contrast, the DEGs of the WXY, GY, and TB

samples were enriched in the biosynthesis of the terpenoid backbone, a vital step in alkaloid synthesis, suggesting that the samples of WXY, GY, and TB have different genes regulating alkaloid biosynthesis than those of the WXY, GY, and TB samples. This finding may be directly related to the alkaloid content of the sample region. Differential genes associated with the plant hormone signal transduction pathway were enriched in the comparisons of WXY vs. TB and WXY vs. GY samples, suggesting that their differences in quality may be due to hormones. Additionally, genes involved in phenylpropanoid biosynthesis were differentially enriched in the samples of WXY and TB, whereas genes involved in the plant hormone signal transduction pathway were differentially enriched in the samples of WXY and GY. Genes related to plant–pathogen interactions and circadian rhythms were enriched in the WXY samples compared to their levels in the GY sample.

3.3. Genes Related to the Biosynthesis of Alkaloids in Biomass

Based on the GO enrichment analysis of DEGs, we found that the DEGs in the "terpenoid skeleton biosynthesis" and "phenylpropaneoid biosynthesis" pathways impacted the various biomass samples. To further clarify the formation mechanism and identify key genes, we screened genes in the "terpenoid skeleton biosynthesis" and "phenylpropanoid biosynthesis" pathways. The DEGs involved in the "terpenoid skeleton biosynthesis" pathway were clustered into two branches, with Branch 1 comprising the samples of WXY, FJ, WX, WS, TB, and NL and Branch 2 comprising the samples of GY and CK (Figure 7A). This division reflects the upregulation of key enzymes like geranylgeranyl diphosphate synthase and hydroxymethylglutaryl-CoA reductase in TB, GY, WS, and CK, which are indicative of an enhanced terpenoid biosynthesis capacity. Genes involved in the "phenylpropanoid biosynthesis pathway" were clustered into two branches based on their expression patterns, with samples from TB and GY having similar trends and being clustered into Branch 1 and the other samples with significant differences between their expression trends being clustered among different branches (Figure 7B). Comprising samples TB and GY, Branch 1 shows a coordinated expression of key genes, such as 'Cinnamyl-alcohol dehydrogenase' and 'Phenylalanine ammonia-lyase', essential for lignin synthesis and enhancing the plant structural integrity and stress response. Additionally, 'Caffeoyl-CoA O-methyltransferase', critical for flavonoid and lignan biosynthesis, also exhibited similar expression in these samples. Conversely, other samples displayed diverse expression profiles, with genes like '4-coumarate-CoA ligase' and 'Peroxidase'—involved in the initial steps of phenylpropanoid formation and defense against oxidative stress—forming distinct clusters. This expression landscape underscores the varied metabolic capabilities and adaptive responses across different biomass samples. The differential gene expression patterns in the terpenoid skeleton, "terpenoid skeleton biosynthesis," and "phenylpropanoid biosynthesis" pathways were similar to the total gene expression patterns in different samples, suggesting that the differential genes in these two pathways may determine the differences and uniqueness of region samples. Terpenes are natural bioactive compounds found in many medicinal biomass sources. It has been reported that >20,000 terpene compounds have been discovered, including many important steroid hormone drugs and ginsenosides. The characteristic components of biomass, including imperialine, peimisine, peimine, and peiminine, are steroid alkaloids. Despite the large number of steroidal compounds, their synthesis mechanisms are similar, primarily being the cytoplasmic methyl valerate (MVA) and cytoplasmic methyl erythritol 4-phosphate (MEP) pathways. *HMGR* is a crucial regulatory factor in the synthesis of terpenoid frameworks. It catalyzes the conversion of 3-hydroxy-3-methylglutaryl CoA to methoxyvalerate, which is an irreversible rate-limiting reaction, with *HMGR* expression directly affecting steroidal saponin content. *HMGS* catalyzes the synthesis of acetylacetyl-CoA to form HMG-CoA, which is used for subsequent reactions in the synthesis of 3-hydroxy-3-methylglutaryl CoA synthesis. *TPS* is speculated to be located in chloroplasts, owing to its potential participation in terpene synthesis by engaging in the MEP pathway. The fragments per kilobase of transcript per million mapped reads (FPKM) values of the *HMGS*, *HMGR*, and *TPS* genes in different biomass samples are

shown in Table S4. In each group, except for GY and TB, the FPKM values of the *HMGR* genes were higher than those of the WXY sample. The FPKM values of the *HMGS* genes in the TB and GY samples were 0, indicating significant differences compared to the other group samples. The FPKM value of the *TPS* gene in the TB sample was also 0, whereas the FPKM values of the *TPS* gene in the WS and GY samples were significantly lower than those in the other group samples.

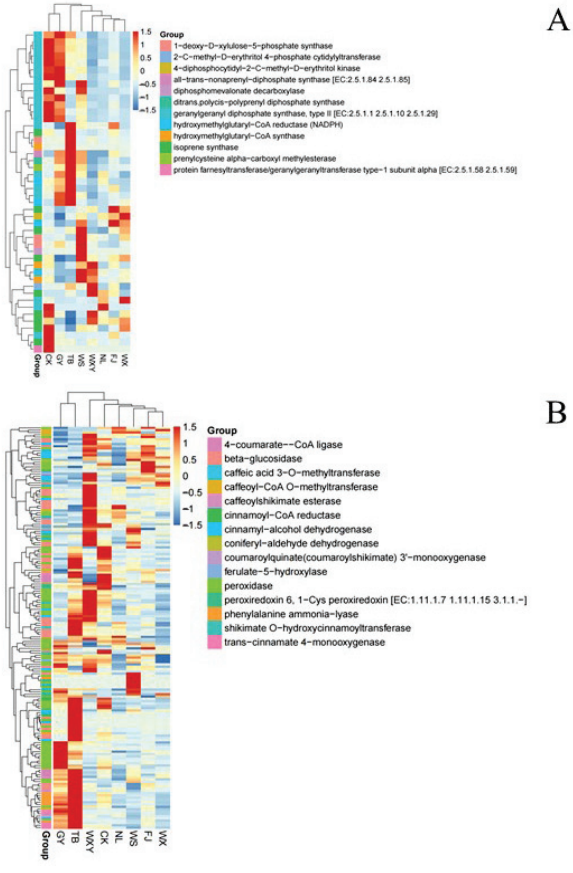


Figure 7. Expression analysis of key genes for the biosynthesis of active ingredients in biomass. (**A**) Heat map of differential gene clustering in terpene skeleton biosynthetic pathways. (**B**) Heat map of differential gene clustering in the phenylpropanoid biosynthetic pathway.

3.4. qRT-PCR Validation of Candidate Genes and Determination of Four Alkaloid Components

The relative expression levels of the functional genes in eight biomass samples were calculated as shown in Figure S2. Among the DEGs, a significant difference was observed in the expression level of *HMGR*. The expression level of the *TPS* gene was highest in the FJ sample and lowest in the NL sample. There were no significant differences in the *HMGS* expression level among the samples. The qRT-PCR results of DEGs were consistent with transcriptome data, verifying the reliability of the transcriptome sequencing results. The contents of imperialine, peimisine, peimine, and peiminine in different biomass samples were measured (Table S5). Overall, the alkaloid content in the WX, WS, FJ, and CK samples was significantly higher compared with that of the other samples, followed by the NL and GY samples, whereas the TB sample had the relatively lowest content.

4. Discussion

4.1. Relationship Between the Alkaloid Content and the Candidate Gene Transcription Level

Research has shown the main quality determinant factor, such as the alkaloid content in the bulbs of *Fritillaria. spp* plants, is closely related to their varieties, growth status, and planting environment. Lu et al. [21] found *F. cirrhosae* accumulated higher levels of major steroidal alkaloids than *F. thunbergii*. Transcription factors involved in the biosynthesis of steroidal alkaloids may contribute to the differences in alkaloid content in the two variants. Harvesting times, processing methods [22], and even lights [23] were reported to affect the contents of alkaloids, further determining the quality of the bulbs of *F. cirrhosa*. However, there is relatively little research on the quality of *F. taipaiensis*, the changes in the biomass and alkaloid content of *F. taipaiensis* from different regions are not clear, and the key regulatory genes are also unknown. To investigate the relationship between the alkaloid content of different types of biomass samples and transcription levels of the *TPS*, *HMGR,* and *HMGS* candidate genes (Figure S3), correlation analyses between the alkaloid content of different biomass types and relative expression levels of the candidate genes were performed. The trend of the relative expression levels of the *HMGS* and *HMGR* genes is similar to the trend of the content of berberine and berberine B content, whereas the trend of the relative expression levels of the *TPS* genes is similar to the trend of the content of berberine A content. However, certain differences exist between the transcription levels of the three candidate genes and the trend of changes in the content of berberine and total alkaloids. The results indicate that the expression levels of different genes significantly differ in relation to the alkaloid content in various biomass types.

Pearson bivariate correlation analysis was applied to explore the correlation between the expression levels of *HMGS*, *HMGR,* and *TPS* genes and alkaloids in the mature bulbs of different biomass samples (Table S6). Only the expression level of *TPS* was highly positively correlated with Peimine content ($r = 0.790$, $p = 0.02$).

4.2. Analysis of the Components of F. taipaiensis Using UPLC-Q-TOF-MS/MS Technology

An analytical method for determining the chemical composition of different biomass samples using UPLC-Q-TOF-MS/MS (Agilent, Beijing, China) technology was established. To the best of our knowledge, an accumulation of different metabolites in biomass samples was analyzed for the first time. In total, 83 compounds were achieved in the LC–MS positive and negative ion modes, including alkaloids, nucleosides, organic and fatty acids, carbonyl compounds, esters, and sugars (mainly alkaloids). Otherwise, findings determined that alkaloids are the main bioactive components underlying the pharmacological properties of biomass samples [24], affirming the results of previous studies on the *Fritillaria* genus, including *F. cirrhosa* and *F.unibracteata* [25–27]. Therefore, this study greatly enriched the components of the biomass of alkaloid-related compounds.

4.3. Differences in the Medicinal Quality of F. taipaiensis from Different Regions

Using the relevant metabolite information, hierarchical cluster analysis, PCA, and OPLS-DA were conducted on different biomass samples, determining that different biomass samples from different regions exhibit significant differences in their chemical compositions, illustrating that the chemical composition and types of *F. taipaiensis* are significantly influenced by regional environmental factors. This result is consistent with previous studies on *F. thunbergii*, *Paris polyphylla* var. *yunnanensis*, and *Panax quinquefolius* [28–30]. Eight heterosteroidal alkaloid differential compounds were identified by OPLS-DA. UPLC–MS was used to quantitatively analyze the four main alkaloid differential metabolites of the different biomass samples. Thus, these compounds exhibited significant differences in content between the biomass samples. The contents of the WX, CK, and FJ samples were much higher than those in the other samples. However, the content in the WS sample was slightly lower than that in the WX, CK, and FJ samples. Thus, the medicinal quality of the bulbs of different biomass samples is habitat-dependent. However, there is limited

in-depth research on the relationship between the changes in alkaloid content and quality of biomass samples.

4.4. DEGs in the Samples from Different Regions

The study of genetic resources has become a focus of recent research [21]. Recently, with the in-depth development of RNA-seq high-throughput sequencing technology, scholars have focused on the biosynthesis pathways of active ingredients and the functions of their related enzyme genes [31]. Transcriptome sequencing detected 359,529 transcripts and 174,827 unigenes, with 7035 unigenes co-annotated in seven databases (KOG, NT, KO, SwissProt, PFAM, and GO), accounting for 4.02% of the total unigenes. Overall, 468,036 DEGs were identified among different sample combinations, of which 242,857 were upregulated and 225,179 were downregulated, greatly enriching the transcriptomic and biological information of biomass samples. The DEGs between different samples in the KEGG database were mainly enriched in steroidal alkaloid synthesis, including terpenoid skeleton and phenylpropanoid biosyntheses, as well as stress response pathways, such as plant–pathogen interactions and plant hormone signal transduction. These findings provide important foundational data and research directions for revealing the biosynthesis mechanism of steroidal alkaloids and molecular mechanisms of biomass samples.

4.5. Discovery of Key Genes Involved in Differential Alkaloid Synthesis in F. taipaiensis from Different Regions

Previous studies have shown two main pathways for the synthesis of steroidal alkaloids: the MVA and MEP pathways [32]. Isopentenyl pyrophosphate is a common intermediate in both the MVA and MEP pathways. Isopentenyl pyrophosphate undergoes a series of enzymatic reactions that ultimately form steroid alkaloids [21]. Zhao et al. [10] found that in *F. cirrhosa*, the MEP pathway other than the MVA pathway was the main route for steroidal alkaloid biosynthesis using transcriptome analysis. We measured the alkaloid content of different biomass samples and found that the alkaloid content of the WXY samples was significantly higher than the GY and TB samples. Transcriptome KEGG annotation analysis revealed that in the terpenoid skeleton biosynthesis pathway, 21 DEGs were enriched between the GY and WXY samples while 28 DEGs were enriched between the TB and WXY samples. The key rate-limiting enzyme genes, such as *HMGS* and *HMGR*, in the terpenoid skeleton biosynthesis pathway were significantly downregulated in both the GY and TB samples, indicating that the downregulation of key genes in the terpenoid skeleton biosynthesis pathway is an important factor in alkaloid synthesis. *HMGS* and *HMGR* may be the key genes responsible for the differences in alkaloid synthesis across biomass samples, providing important data and insights for further exploration of the molecular mechanisms that affect the differences and uniqueness of various biomass samples.

4.6. Differences in the Gene Expression of F. taipaiensis from Different Regions

Transcriptome sequencing analysis of the bulbs from different biomass samples in this study showed that there were 240 DEGs (117 upregulated and 123 downregulated) between the WXY and WX samples and 163 DEGs (42 upregulated and 121 downregulated) between the WXY and NL samples. There were no significant differences in gene expression, which may be related to the introduction of NL-cultivated varieties originating from the WX sample. PCA showed that WXY grouped with the samples from WX, CK, FJ, WS, and NL. This indicates little difference in gene expression was observed between WXY and the other samples, possibly due to genetic basis factors. Zhou et al. reported similar results regarding the introduction and cultivation of Epimedium [33]. Therefore, further research is required to determine the magnitude of the effects of internal and external conditions on the metabolites.

4.7. Analysis of Systematic Signal Transduction Networks in F. taipaiensis from Different Regions

Biomass systems have complex and systematic signal transduction networks that actively respond to changes in internal factors and the external environment, thereby resisting environmental stress. In this network system, hormones are the main endogenous factors in plant stress responses, stress resistance regulation, and growth regulation [34]. KEGG enrichment analysis of WXY vs. GY samples in this study demonstrated that DEGs in WXY vs. TB samples were significantly enriched in the hormone signal transduction pathway. The plant–pathogen interaction pathway primarily acts on the complex, precise, and multilevel immune system formed during the long-term evolution of biomass [35]. The KEGG enrichment analysis results between the WXY and TB samples in this study showed significantly enriched differential genes in the plant–pathogen interaction pathway and differential genes, such as *CDPK*, *NOS*, and *WRKY25*, were upregulated in the WXY samples. It is speculated that there are certain differences in resistance to biological stress (pathogen infection) between the wild variety (WXY) and GY and TB samples.

4.8. Functional Speculation of Key Genes Involved in Alkaloid Synthesis in F. taipaiensis

With the continuous development of omics technology, multi-omics joint analysis is applied to analyze the complex biological mechanisms of biomass. In the joint analysis of the transcriptome and metabolome, the association analysis between DEGs in the transcriptome and differentially detected metabolites in the metabolome can be used to analyze changes at two levels, cause and effect; lock in key pathways related to metabolite changes; and construct a core regulatory network to reveal its internal laws [36]. Through transcriptome analysis, a large number of genes have been identified to play a regulatory role in alkaloid synthesis. Many transcription factors, including *bHLH*, *MYC*, *MYB*, and *WRKY*, regulate the alkaloid content by affecting the gene expression in the MVA pathway. *SQLE*, *CAS1*, *SMT1*, *SMO1*, *SMO2*, *SC5DL*, *DHCR7*, *DHCR24*, *CYP710A*, *3β-HSD*, *CYP90D2*, and *CYP374A6* genes were found to be differentially expressed in the bulb of *F. cirrhosa* [37].

Based on the differential metabolites of different biomass samples, the key genes involved in the synthesis pathways of the aforementioned substances were preliminarily validated, along with three key candidate genes: *HMGS*, *HMGR*, and *TPS*. The expression level of *TPS* was significantly positively correlated with the total alkaloid and peimine contents, whereas *HMGS* and *HMGR* had little effect on alkaloid content. The key genes play crucial roles in the upstream of the terpenoid biosynthesis pathway, including *HMGR*, farnesyl diphosphate synthase (*FPS*), squalene synthase (*SS*), and cycloartenol synthase (*CAS*), which have been well characterized, while the secondary metabolic pathways following the cycloartenol generation are rarely studied. Li et al. [13] revealed that *FcCYP450* of the predicted *CYP450* family, which was responsible for C-22, C-23, and C-26 hydroxylation in the steroidal alkaloid biosynthesis pathway, was related to the isosteroidal alkaloid biosynthesis. *TPS*s convert acyclic C5 to C20 cis- or trans-prenyl diphosphate intermediates into the C5-hemiterpenes, such as isoprene, C10-monoterpenoids, C15-sesquiterpenoids, or C20-diterpenoids [38]. Here, we found the expression level of *TPS* was highly correlated with alkaloid contents. It is speculated that *TPS* may be the main rate-limiting gene alkaloid biosynthesis, although further experimental evidence is needed. In the near future, using genetic engineering technology to enhance *TPS* gene expression may be a feasible approach to increase the amount of alkaloids in biomass samples.

In this study, three genes involved in alkaloid synthesis were selected for qRT-PCR. The upregulation and downregulation of expression trends in different biomass samples were consistent with the transcriptome FPKM values. This further proves the authenticity and reliability of the transcriptome sequencing results. Therefore, this study captured DEGs among different biomass samples through RNA high-throughput sequencing technology and systematically studied the genes involved in growth, development, and secondary metabolism as a whole, further elucidating the molecular mechanisms of active ingredient formation in biomass samples.

5. Conclusions

In summary, eight types of heterosteroidal alkaloids with significant differences among the *F. taipaiensis* biomass samples from different origins were selected, including peimisine, imperialine, peimine, peiminine, ebeinone, ebeiedine, ebeiedinone, and forticine. The *F. taipaiensis* samples can be clustered into two branches of DEGs: (1) NL, FJ, WX, CK, WS, and WXY; (2) TB and GY. The DEGs in the "terpenoid skeleton biosynthesis" and "phenylpropaneoid biosynthesis" pathways impacted the samples the most significantly. Eighty-three potential compounds were inferred in the biomass samples, including thirty alkaloids, six alkaloid glycosides, two fatty amides, nine nucleosides, one peptide, six fatty acids, six aromatic acids, ten lipids, one acid anhydride, three aromatic aldehydes, one aromatic alcohol, one aromatic ketone, one phenol, one fatty alcohol, and five sugar components. The alkaloid contents in samples WX, CK, and FJ were much higher than those in the other samples. The accumulation of active alkaloid components in different biomass samples was strongly correlated with the expression levels of catalytic enzyme genes involved in biosynthesis. Thus, *TPS* may be the main rate-limiting gene regulating alkaloid biosynthesis. This study reveals the biosynthesis and regulatory mechanisms of the main active ingredients and provides promising applications in the biomedical fields.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo14110590/s1, Figure S1: Unigenes-Transcript length distribution in the transcriptome; Figure S2: qRT-PCR validation of the DEGs in biomass from different regions; Figure S3: The relationship between the total alkaloid content of biomass and the expression of *TPS*, *HMGR*, and *HMGS* genes; Table S1: Soil properties of the sample collection areas; Table S2: Mass spectrometry information of the compounds of biomass; Table S3: Transcriptome sequencing data quality analysis; Table S4: The contents of Chinese and Imperialine, Peimisine, Peimine, and Peiminine in biomass from different origins; Table S5: The correlation analyze between the expression levels of *HMGS*, *HMGR* and *TPS* genes and alkaloid content in mature bulbs of biomass from different regions; Table S6: The correlation analyze between the expression levels of *HMGS*, *HMGR* and *TPS* genes and alkaloid content in mature bulbs of biomass from different regions.

Author Contributions: Methodology, F.-G.C., Y.-W.Z., and N.Z.; investigation, F.-G.C., Y.-W.Z., and C.-M.M.; writing-original draft preparation, N.Z. and C.-M.M.; writing-review and editing, M.-G.M. and W.-D.L.; supervision, W.-D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Major Project of Natural Science Foundation of Chongqing Education Commission (KJZD-M202301204).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw data of this study can be download at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1036838 (accessed on 25 October 2024), with password: 7xfu.

Acknowledgments: This work was supported by the Major Project of Natural Science Foundation of Chongqing Education Commission (KJZD-M202301204).

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Tang, X.; Zuo, M.; Li, Z.; Liu, H.; Xiong, C.X.; Zeng, X.H.; Sun, Y.; Hu, L.; Liu, S.J.; Lei, T.Z.; et al. Green processing of lignocellulosic biomass and its derivatives in deep eutectic solvents. *Chemsuschem* **2017**, *10*, 2696–2706. [CrossRef] [PubMed]
2. Hajiali, F.; Jin, T.; Yang, G.; Santos, M.; Lam, E.; Moores, A. Mechanochemical Transformations of Biomass into Functional Materials. *Chemsuschem* **2022**, *15*, e202102535. [CrossRef] [PubMed]
3. Antolini, E. Lignocellulose, cellulose and lignin as renewable alternative fuels for direct biomass fuel cells. *Chemsuschem* **2020**, *14*, 189–207. [CrossRef] [PubMed]
4. Wang, A.W.; Liu, Y.M. Progress in the study of isosteroidal alkaloids and pharmacological activities. *Nat. Prod. Res. Dev.* **2022**, *34*, 164–175.

5. Guerriero, G.; Hausman, J.F.; Strauss, J.; Ertan, H.; Siddiqui, K.S. Lignocellulosic biomass: Biosynthesis, degradation, and industrial utilization. *Eng. Life Sci.* **2016**, *16*, 1–6. [CrossRef]
6. Fang, Z.T.; Jin, J.; Ye, Y.; He, W.Z.; Shu, Z.F.; Shao, J.N.; Fu, Z.S.; Lu, J.L.; Ye, J.H. Effects of Different Shading Treatments on the Biomass and Transcriptome Profiles of Tea Leaves (*Camellia sinensis* L.) and the Regulatory Effect on Phytohormone Biosynthesis. *Front. Plant Sci.* **2022**, *13*, 909765. [CrossRef]
7. Maher, C.A.; Kumar-Sinha, C.; Cao, X.H.; Kalyana-Sundaram, S.; Han, B.; Jing, X.J.; Sam, L.; Barrette, T.; Palanisamy, N.; Chinnaiyan, A.M. Transcriptome sequencing to detect gene fusions in cancer. *Nature* **2009**, *458*, 97. [CrossRef]
8. Franks, A.E.; Glaven, R.H.; Lovley, D.R. Real-Time Spatial Gene Expression Analysis within Current-Producing Biofilms. *Chemsuschem* **2012**, *5*, 1092–1098. [CrossRef]
9. Kim, J.A.; Roy, N.S.; Lee, I.H.; Choi, A.Y.; Choi, B.S.; Yu, Y.S.; Park, N.I.; Park, K.C.; Kim, S.; Yang, H.S. Genome-wide transcriptome profiling of the medicinal plant *Zanthoxylum planispinum* using a single-molecule direct RNA sequencing approach. *Genomics* **2019**, *111*, 973–979. [CrossRef]
10. Zhao, Q.; Li, R.; Zhang, Y.; Huang, K.J.; Wang, W.G.; Li, J. Transcriptome analysis reveals in vitro-cultured regeneration bulbs as a promising source for targeted Fritillaria cirrhosa steroidal alkaloid biosynthesis. *3 Biotech* **2018**, *8*, 191. [CrossRef]
11. Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China*, 1st ed.; China Medical Science Press: Beijing, China, 2015; p. 36.
12. Hu, Z.; Zong, J.F.; Yili, M.; Yu, M.H.; Aisa, H.A.; Hou, A.J. Isosteroidal alkaloids from the bulbs of *Fritillaria tortifolia*. *Fitoterapia* **2018**, *131*, 112–118. [CrossRef] [PubMed]
13. Li, R.; Xiao, M.T.; Li, J.; Zhao, Q.; Wang, M.C.; Zhu, Z.W. Transcriptome analysis of CYP450 family members in *Fritillaria cirrhosa* D. Don and profiling of key *CYP450s* related to isosteroidal alkaloid biosynthesis. *Genes* **2023**, *14*, 219. [CrossRef]
14. Peng, W.; Li, Z.; Wang, S.; Wang, B.J. Unravelling the C-C and C-N coupling mechanism for the CYP96T1-catalyzed biosynthesis of Amaryllidaceae alkaloids. *Mol. Catal.* **2023**, *550*, 113609. [CrossRef]
15. Jiao, C.Y.; Wei, M.K.; Fan, H.H.; Song, C.; Wang, Z.J.; Cai, Y.P.; Jin, Q. Transcriptomic analysis of genes related to alkaloid biosynthesis and the regulation mechanism under precursor and methyl jasmonate treatment in *Dendrobium officinale*. *Front. Plant Sci.* **2022**, *13*, 941231. [CrossRef] [PubMed]
16. Zhang, H.T.; Hedhili, S.; Montiel, G.; Zhang, Y.X.; Chatel, G.; Pré, M.; Gantet, P.; Memelink, J. The basic helix-loop-helix transcription factor *CrMYC2* controls the jasmonate-responsive expression of the *ORCA* genes that regulate alkaloid biosynthesis in Catharanthus roseus. *Plant J.* **2011**, *67*, 61–71. [CrossRef] [PubMed]
17. Zhou, N.; Mu, M.J.; Yang, M.; Zhou, Y.; Ma, M.G. The effect of microbial fertilizer on the growth, rhizospheric environment and medicinal quality of *Fritillaria taipaiensis*. *Horticulturae* **2021**, *7*, 500. [CrossRef]
18. Sheng, L.; Zhou, N.; Fu, S.Z.; Yi, D.Y.; Jia, H.; Chen, H.Y. Pharmacognostical study on cultivated *Fritillaria taipaiensis*. *J. Chin. Med. Mater.* **2014**, *37*, 45–49.
19. Mu, M.J.; Zhang, G.D.; Zhang, H.; Yang, M.; Guo, D.Q.; Zhou, N. Correlation between rhizospheric microorganisms distribution and alkaloid content of *Fritillaria taipaiensis*. *China J. Chin. Mater. Med.* **2019**, *4*, 2331–2335.
20. Zhang, M.R.; Fu, S.B.; Xie, H.M.; Xie, J.H.; Gong, P.Z.; Wang, S.; Ye, B.G. Comparative study on the content of isosteroid alkaloids in *Fritillaria cirrhosa* from different producing areas. *W. China J. Pharm.* **2022**, *37*, 67–74.
21. Lu, Q.X.; Rui, L.; Liao, J.Q.; Hu, Y.Q.; Gao, Y.D.; Wang, M.C.; Li, J.; Zhao, Q. Integrative analysis of the steroidal alkaloids distribution and biosynthesis of bulbs Fritillariae Cirrhosae through metabolome and transcriptome analyse. *BMC Genom.* **2022**, *23*, 551. [CrossRef]
22. Ma, B.; Ma, J.; Li, B.; Tao, Q.; Gao, J.; Yan, Z. Effects of different harvesting times and processing methods on the quality of cultivated *Fritillaria cirrhosa* D. Don. *Food Sci. Nutr.* **2021**, *9*, 2853–2861. [CrossRef] [PubMed]
23. Chen, C.C.; Lee, M.R.; Wu, C.R.; Ke, H.J.; Xie, H.M.; Tsay, H.S.; Agrawal, D.C.; Chang, H.C. LED Lights affecting morphogenesis and isosteroidal alkaloid contents in *Fritillaria cirrhosa* D. Don-an important Chinese medicinal herb. *Plants* **2020**, *9*, 1351. [CrossRef] [PubMed]
24. Wang, A.W.; Liu, Y.M.; Zhu, M.M.; Ma, R.X. Isosteroidal alkaloids of *Fritillaria taipaiensis* and their implication to Alzheimer's disease: Isolation, structural elucidation and biological activity. *Phytochemistry* **2022**, *201*, 113279. [CrossRef] [PubMed]
25. Geng, Z.; Liu, Y.F.; Gou, Y.; Zhou, Q.M.; He, C.J.; Guo, L.; Zhou, J.; Xiong, L. Metabolomics study of cultivated bulbus *Fritillariae cirrhosae* at different growth stages using UHPLC-QTOF-MS coupled with multivariate data analysis. *Phytochem. Anal.* **2018**, *29*, 290–299. [CrossRef]
26. Ge, S. Analysis of Chemical Constituents Differences Between *Fritillaria cirrhosa* D. Don and *Fritillaria thunbergii* and Its Metabolomics Based on UPLC-Q-TOF-MS/MS Technology. Master's Thesis, Qinghai Normal University, Xining, China, 2022.
27. Li, Y.B.; Zhang, L.; Wu, H.Y.; Wu, X.; Ju, L.; Zhang, Y.J. Metabolomic study to discriminate the different *Bulbus fritillariae* species using rapid resolution liquid chromatography-quadrupole time-of-flight mass spectrometry coupled with multivariate statistical analysis. *Anal. Methods* **2014**, *6*, 2247–2259. [CrossRef]
28. Wu, J.L.; Zhang, X.B.; Hu, J.; Li, M.; Jing, Z.X.; Yuan, Q.; Li, S.Q.; Zhang, C.C. Quality of *Fritillaria thunbergii* Miq. regionalization based on compositional content and ecological factors. *J. Zhejiang Chin. Med. Univ.* **2021**, *45*, 835–841.
29. Wang, Q.; Ding, Y.; Yang, M.; Guo, D.Q.; Huang, Y.; Zhang, C.T.; Pang, Y.F.; Zhou, N. Correlation analysis of quality origin and phenotypic characters of *Paris polyphylla* var. *yunnanensis*. *China J. Chin. Mater. Med.* **2019**, *44*, 3203–3212.

30. Sun, X.; Qian, Q.Y.; Zheng, S.H.; Cheng, H.M.; Huang, L.F. Quality ecotype of *Panax quinquefolium* L. based on hereditychemistry-ecology characteristics. *Acta Pharm. Sin.* **2019**, *54*, 1695–1705.
31. Hao, D.C.; Chen, S.L.; Xiao, P.G.; Liu, M. Application of high-throughput sequencing in medicinal plant transcriptome studies. *Drug Develop. Res.* **2013**, *73*, 487–498. [CrossRef]
32. Liao, H.; Quan, H.G.; Huang, B.H.; Ji, H.Y.; Zhang, T.; Chen, J.; Zhou, J.Y. Integrated transcriptomic and metabolomic analysis reveals the molecular basis of tissue-specific accumulation of bioactive steroidal alkaloids in *Fritillaria unibracteata*. *Phytochemistry* **2023**, *214*, 113831. [CrossRef]
33. Zhou, M.; Zheng, W.; Guo, B.L.; Chen, A.J.; Ma, B.P. Study on the quality of cultivated *Epimedium brevicornu* based on UHPLC-PDA-Q-TOF/MSE technique. *J. Pharm.* **2020**, *55*, 995–1003.
34. Shuai, H.W.; Meng, Y.J.; Chen, F.; Zhou, W.G.; Luo, X.F.; Yang, W.Y. Hormonal signaling responses to plant shade stress. *J. Bot.* **2018**, *53*, 139–148.
35. Zhang, H.; Mao, R.; Wang, Y.Z.; Zhang, L.; Wang, C.Y.; Lv, S.K.; Liu, X.L.; Wang, Y.J.; Ji, W.Q. Transcriptome-wide alternative splicing modulation during plant-pathogen interactions in wheat. *Plant Sci.* **2019**, *288*, 110160. [CrossRef]
36. Liu, J.J.; Han, L.J.; Li, G.D.; Zhang, A.L.; Liu, X.L.; Zhao, M.Z. Transcriptome and metabolome profiling of the medicinal plant *Veratrum mengtzeanum* reveal key components of the alkaloid biosynthesis. *Front. Genet.* **2023**, *14*, 1023433. [CrossRef] [PubMed]
37. Kumar, P.; Ashrita; Acharya, V.; Warghat, A.R. Comparative transcriptome analysis infers bulb derived in vitro cultures as a promising source for sipeimine biosynthesis in *Fritillaria cirrhosa* D. Don (Liliaceae, syn. Fritillaria roylei Hook.)-High value Himalayan medicinal herb. *Phytochemistry* **2021**, *108*, 112631. [CrossRef]
38. Tholl, D. Biosynthesis and biological functions of terpenoids in plants. *Adv. Biochem. Eng. Biotechnol.* **2015**, *148*, 63–106.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Metabolites Editorial Office
E-mail: metabolites@mdpi.com
www.mdpi.com/journal/metabolites

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-3936-0