International Journal of
Molecular Sciences

Special Issue Reprint

Advances in Research for Horticultural Crops Breeding and Genetics

Edited by
Ai-Sheng Xiong and Mengyao Li

mdpi.com/journal/ijms

MDPI

Advances in Research for Horticultural Crops Breeding and Genetics

Advances in Research for Horticultural Crops Breeding and Genetics

Guest Editors

Ai-Sheng Xiong
Mengyao Li

MDPI

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

Guest Editors

Ai-Sheng Xiong
State Key Laboratory of Crop Genetics and Germplasm Enhancement and Utilization
Nanjing Agricultural University
Nanjing
China

Mengyao Li
College of Horticulture
Sichuan Agricultural University
Chengdu
China

Editorial Office

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *International Journal of Molecular Sciences* (ISSN 1422-0067), freely accessible at: https://www.mdpi.com/journal/ijms/special_issues/0401486299.

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 978-3-7258-5855-2 (Hbk)
ISBN 978-3-7258-5856-9 (PDF)
https://doi.org/10.3390/books978-3-7258-5856-9

Contents

Shuangjuan Yang, Hao Tang, Xiaochun Wei, Yanyan Zhao, Zhiyong Wang, Henan Su, et al.
BrWAX3, Encoding a β-ketoacyl-CoA Synthase, Plays an Essential Role in Cuticular Wax Biosynthesis in Chinese Cabbage
Reprinted from: *Int. J. Mol. Sci.* **2022**, *23*, 10938, https://doi.org/10.3390/ijms231810938 . . . . . **1**

Ao-Qi Duan, Shan-Shan Tan, Yuan-Jie Deng, Zhi-Sheng Xu and Ai-Sheng Xiong
Genome-Wide Identification and Evolution Analysis of R2R3-MYB Gene Family Reveals S6 Subfamily R2R3-MYB Transcription Factors Involved in Anthocyanin Biosynthesis in Carrot
Reprinted from: *Int. J. Mol. Sci.* **2022**, *23*, 11859, https://doi.org/10.3390/ijms231911859 . . . . . **17**

Yan Wang, Zhiyi Wang, Jing Zhang, Zhenshan Liu, Hao Wang, Hongxia Tu, et al.
Integrated Transcriptome and Metabolome Analyses Provide Insights into the Coloring Mechanism of Dark-red and Yellow Fruits in Chinese Cherry [*Cerasus pseudocerasus* (Lindl.) G. Don]
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 3471, https://doi.org/10.3390/ijms24043471 . . . . . **34**

Hao Zhou, Wei Zhou, Xinzhuan Yao, Qi Zhao and Litang Lu
Genome-Wide Investigation and Functional Analysis Reveal That *CsGeBP4* Is Required for Tea Plant Trichome Formation
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 5207, https://doi.org/10.3390/ijms24065207 . . . . . **54**

Weijuan Han, Yiru Wang, Huawei Li, Songfeng Diao, Yujing Suo, Taishan Li, et al.
Transcriptome and Metabolome Reveal Distinct Sugar Accumulation Pattern between PCNA and PCA Mature Persimmon Fruit
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 8599, https://doi.org/10.3390/ijms24108599 . . . . . **74**

Lianxi Zhang, Yiping Wang, Maolan Yue, Leiyu Jiang, Nating Zhang, Ya Luo, et al.
FaMYB5 Interacts with FaBBX24 to Regulate Anthocyanin and Proanthocyanidin Biosynthesis in Strawberry (*Fragaria* × *ananassa*)
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 12185, https://doi.org/10.3390/ijms241512185 . . . . . **87**

Zheng Jiang, Lihui Du, Lei Shen, Jie He, Xin Xia, Longhao Zhang and Xu Yang
Genome-Wide Exploration and Expression Analysis of the CNGC Gene Family in Eggplant (*Solanum melongena* L.) under Cold Stress, with Functional Characterization of *SmCNGC1a*
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 13049, https://doi.org/10.3390/ijms241713049 . . . . . **101**

Yuanxiu Lin, Hao He, Yanling Wen, Shuaipeng Cao, Zisen Wang, Ziqing Sun, et al.
Comprehensive Analysis of the Pectate Lyase Gene Family and the Role of *FaPL1* in Strawberry Softening
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 13217, https://doi.org/10.3390/ijms241713217 . . . . . **122**

Lei Shen, Xin Xia, Longhao Zhang, Shixin Yang and Xu Yang
Genome-Wide Identification of Catalase Gene Family and the Function of *SmCAT4* in Eggplant Response to Salt Stress
Reprinted from: *Int. J. Mol. Sci.* **2023**, *24*, 16979, https://doi.org/10.3390/ijms242316979 . . . . . **136**

Marzena Mazurek, Aleksandra Siekierzyńska, Tomasz Piechowiak, Anna Spinardi and Wojciech Litwińczuk
Comprehensive Analysis of Highbush Blueberry Plants Propagated In Vitro and Conventionally
Reprinted from: *Int. J. Mol. Sci.* **2024**, *25*, 544, https://doi.org/10.3390/ijms25010544 . . . . . **154**

Zhihang Hu, Nan Zhang, Zhiyuan Qin, Jinwen Li, Ni Yang, Yi Chen, et al.
Differential Response of MYB Transcription Factor Gene Transcripts to Circadian Rhythm in Tea Plants (*Camellia sinensis*)
Reprinted from: *Int. J. Mol. Sci.* **2024**, *25*, 657, https://doi.org/10.3390/ijms25010657 . . . . . **172**

Lu Wei, Cheng Yang, Jerome Jeyakumar John Martin, Rui Li, Lixia Zhou, Shuanghong Cheng, et al.
Metabonomics and Transcriptomic Analysis of Free Fatty Acid Synthesis in Seedless and Tenera Oil Palm
Reprinted from: *Int. J. Mol. Sci.* **2024**, *25*, 1686, https://doi.org/10.3390/ijms25031686 . . . . . **193**

Kexin Wang, Xinyu Wang, Lijing Zhang, Yichen Chi, Yusong Luo, Wenlong Xu, et al.
Morphological Analyses and QTL Mapping of Mottled Leaf in Zucchini (*Cucurbita pepo* L.)
Reprinted from: *Int. J. Mol. Sci.* **2024**, *25*, 2491, https://doi.org/10.3390/ijms25052491 . . . . . **213**

International Journal of
Molecular Sciences

MDPI

Article

BrWAX3, Encoding a β-ketoacyl-CoA Synthase, Plays an Essential Role in Cuticular Wax Biosynthesis in Chinese Cabbage

Shuangjuan Yang [1], Hao Tang [1,2], Xiaochun Wei [1], Yanyan Zhao [1], Zhiyong Wang [1], Henan Su [1], Liujing Niu [1], Yuxiang Yuan [1,*] and Xiaowei Zhang [1,*]

[1] Institute of Horticulture, Henan Academy of Agricultural Sciences, Zhengzhou 450002, China
[2] College of Horticulture and Landscape, Henan Institute of Science and Technology, Xinxiang 453003, China
* Correspondence: yuxiangyuan126@126.com (Y.Y.); xiaowei5737@163.com (X.Z.)

Abstract: In this study, we identified a novel glossy mutant from Chinese cabbage, named SD369, and all wax monomers longer than 26 carbons were significantly decreased. Inheritance analysis revealed that the glossy trait of SD369 was controlled by a single recessive locus, *BrWAX3*. We fine-mapped the *BrWAX3* locus to an interval of 161.82 kb on chromosome A09. According to the annotated genome of *Brassica rapa*, *Bra024749* (*BrCER60.A09*), encoding a β-ketoacyl-CoA synthase, was identified as the candidate gene. Expression analysis showed that *BrCER60.A09* was significantly downregulated in all aerial organs of glossy plants. Subcellular localization indicated that the BrCER60.A09 protein functions in the endoplasmic reticulum. A 5567-bp insertion was identified in exon 1 of *BrCER60.A09* in SD369, which lead to a premature stop codon, thus causing a loss of function of the BrCER60.A09 enzyme. Moreover, comparative transcriptome analysis revealed that the 'cutin, suberine, and wax biosynthesis' pathway was significantly enriched, and genes involved in this pathway were almost upregulated in glossy plants. Further, two functional markers, BrWAX3-InDel and BrWAX3-KASP1, were developed and validated. Overall, these results provide a new information for the cuticular wax biosynthesis and provide applicable markers for marker-assisted selection (MAS)-based breeding of *Brassica rapa*.

Keywords: *Brassica rapa*; β-ketoacyl-CoA synthase; wax synthesis; map-based cloning

1. Introduction

The lipidic cuticle exists on the aerial surface of many land plants, working as a physical barrier to prevent nonstomatal water loss [1,2]. The main components of cuticles are cutin and cuticular wax. Cutin is a cross-linked polymer of modified long-chain fatty acids (C16 and C18 hydroxy and epoxy fatty acids) and glycerol, which provides mechanical strength to the surface layer [3,4]. Cuticular wax is a mixture of very-long-chain fatty acids (VLCFAs) and their derivatives [5,6]. Cuticles also protect plants from various biotic and abiotic stresses [7,8], profoundly affect plant-insect interactions [9], affect the pollen-stigma signaling [10], and prevent epidermal fusions [11].

Wax biosynthesis begins with de novo C16 and C18 fatty acid biosynthesis in the plastids of epidermal cells, and further elongates into VLCFAs in the endoplasmic reticulum (ER). The VLCFAs are then modified via two pathways: the alkane-forming pathway and the alcohol-forming pathway. The former generates aldehydes, alkanes, secondary alcohols, and ketones, while the latter produces primary alcohols and wax esters [2,5,6]. As the essential precursors of wax, VLCFAs are synthesized under the consecutive catalyzes of four different enzymes, namely, β-ketoacyl-CoA synthase (KCS), β-ketoacyl-CoA reductase (KCR), β-hydroxyacyl-CoA dehydratase (PAS2/HCD), and enoyl-CoA reductase (ECR/CER10), which formed a fatty acid elongase (FAE) complex [5,6,12]. Among the enzymes, KCSs determine substrate specificity, and are key rate-limiting condensing

enzymes. In *Arabidopsis*, a large family of 21 *KCS* genes has been annotated [13]. To date, several *KCSs* have been reported. *FAE1/KCS18* encodes a seed-specific condensing enzyme that catalyzes the elongation of C18 to C20 and C22 [14]. *KCS2* and *KCS20* are responsible for the elongation of C20 to C22 and are functionally redundant [15]. *KCS9* is involved in the elongation of C22 to C24 fatty acids, which play multiple roles in the production of suberins, cuticular waxes, and membrane lipids [16]. *KCS1* is required for the elongation of C24 VLCFAs [17]. *KCS5/CER60* and *KCS6/CER6* play redundant roles in the elongation of C26 to C28 during wax biosynthesis, among which *KCS6* (*CER6*) plays a major role [10,18–20]. Together with *KCS5/CER60* or *KCS6/CER6*, three BAHD acyltransferases members, CER2, CER2-like1, and CER2-like2, participate in the synthesis of VLCFAs longer than C28 [9,21,22].

For economically important leafy vegetables in *Brassica* species, such as *Brassica rapa* (A genome) and *Brassica oleracea* (C genome), the waxless and glossy phenotype is an important commodity trait in breeding [23,24]. Therefore, identifying genes responsible for the glossy phenotype facilitates the breeding of new varieties with a glossy green phenotype. In *Brassica oleracea*, *Bol018504* (*CER1*) is responsible for the glossy trait of four materials, in which the alkane-forming pathway of wax biosynthesis is significantly affected [4,25,26]. *Bol013612* (*CER4*) is the candidate gene in two glossy materials, in which the primary alcohols and wax esters from the alcohol-forming pathway are severely reduced [27,28]. Recently, Ji et al. (2021) identified a new glossy cabbage genotype and found that *Bo1g039030* (*BoCER2*) was the candidate gene, in which wax monomers longer than C28 were significantly decreased [23]. In *Brassica rapa*, only two wax biosynthesis-related genes have been identified. *BrWax1* was first reported and *Bra013809* (*CER2*) was the candidate gene [29]. *Bra032670* (*CER1*), with different sequence variations, was responsible for the glossy trait in three *B.rapa* materials [24,30,31].

In this study, we characterized a novel glossy green Chinese cabbage mutant, SD369, which showed a significant reduction of wax monomers longer than C26. Genetic analysis suggested that the glossy trait was controlled by a single recessive gene. Map-based cloning revealed that the *Bra024749* gene, which is homologous to *CER60* (*KCS5*) in *Arabidopsis*, was the candidate gene, which has never been reported in *Brassica* species. Furthermore, sequence analysis and expression analysis showed that a 5567-bp insertion blocked VLCFA elongation from C26 to C28 in SD369, thus causing a wax deficiency. Additionally, we developed and validated two functional markers. These findings will provide new insight into the plant cuticular wax metabolic networks and will promote molecular marker-assisted breeding in *B.rapa*.

2. Results

2.1. Phenotypic Characterization and Genetic Analysis of Glossy Trait in SD369

SD369 is a spontaneous wax-deficient mutant found in the Chinese cabbage field. In contrast to the typical waxy appearance of R16-11 (P_2), all aerial organs of SD369 (P_1), such as the leaves (Figure 1a), stems (Figure 1b), flower buds (Figure 1c), and seedpods, were glossy green. Cryo-SEM analysis revealed many more wax crystals on R16-11 (Figure 1h) than on the SD369 (Figure 1d). The wax crystals on R16-11 were mainly flaky and columnar. However, the leaves of SD369 were covered with only a few wax crystals (Figure 1d), and the wax crystal shape was granular.

F_1, F_2, BC_1P_1, and BC_1P_2 populations were generated to investigate the SD369 glossy trait inheritance. The F_1 plants were all waxy, indicating that the waxy phenotype was dominant in the glossy phenotype. In a small F_2 population, 102 plants were waxy and 40 were glossy, corresponding to a segregation ratio of 3:1 by the chi-square test (Table 1). A larger F_2 population showed similar results (3026 waxy: 954 glossy, $\chi^2 = 2.25 < \chi^2_{0.05} = 3.84$). A ratio of 1:1 (540 waxy:494 glossy, $\chi^2 = 2.05 < \chi^2_{0.05} = 3.84$) was obtained in the BC_1 P_1 population, while in the BC_1 P_2 population, all 200 individuals were waxy (Table 1). These results indicated that the glossy phenotype of SD369 is controlled by a single recessive gene (Table 1). We tentatively named this locus *BrWAX3*.

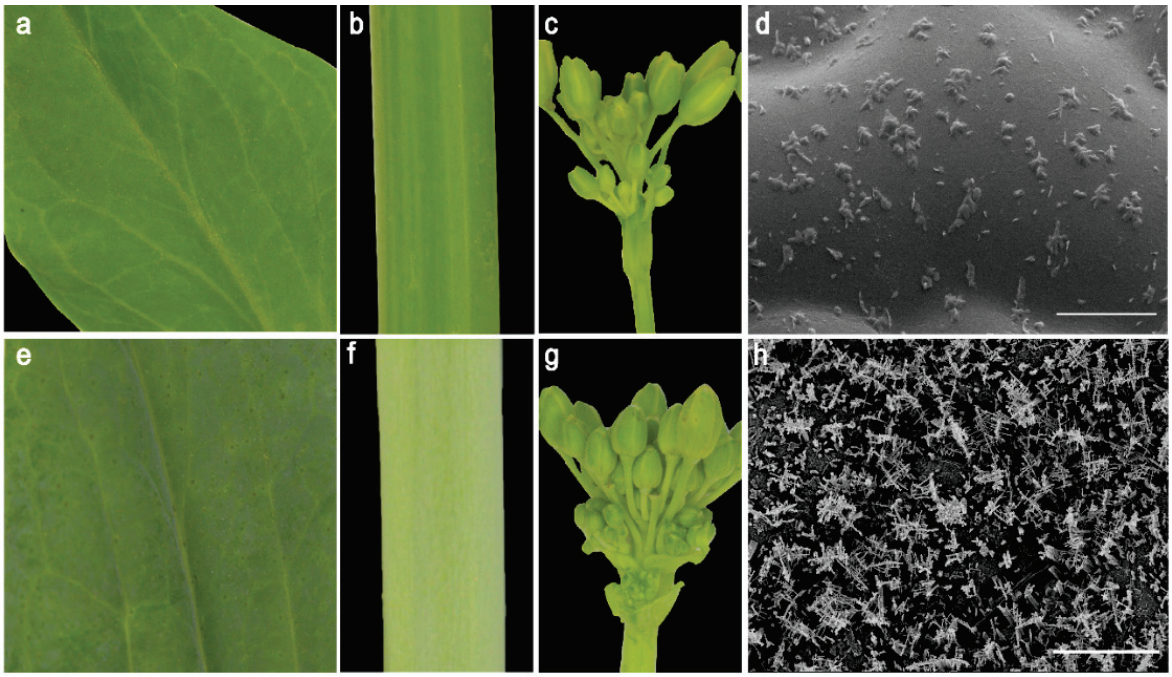


Figure 1. Phenotypic characterization of cuticular waxes in the two parent lines (SD369 and R16-11). The stem (**a**), leaf (**b**), and flower buds (**c**) of SD369 showed glossy phenotype at the bolting stage, as compared to the waxy appearance of R16-11 stem (**e**), leaf (**f**), and flower buds (**g**). Cryo-scanning electron microscopy images of leaves from SD369 (**d**) and R16-11 (**h**). Bar = 10 μm in (**d**,**h**).

Table 1. Genetic analysis of the glossy trait in crosses between SD369 and R16-11.

Population	Total	Waxy	Glossy	Expected ratio	χ^2	$\chi^2_{0.05}$
P_1 (SD369)	10	10	0	-	-	-
P_2 (R16-11)	10	0	10	-	-	-
F_1	15	15	0	-	-	-
F_2-small	142	102	40	3:1	0.76	3.84
F_2-large	3980	3026	954	3:1	2.25	3.84
BC_1P_1 (F_1 × SD369)	1020	540	494	1:1	2.05	3.84
BC_1P_2 (F_1 × R16-11)	300	300	0	-	-	-

2.2. Cuticular Wax Analysis via GC-MS

To investigate the reason of wax depletion in SD369, cuticular wax from W-bulk and G-bulk was examined by GC-MS. The wax load on waxy leaves reached, on average, 149.67 μg per g fresh weight, whereas wax loads on glossy leaves were severely reduced, with an average of 48.56 μg/g fresh weight, which reduced 68% of the wax load when compared to W-bulk (Figure 2b, Table S1). Wax composition analysis revealed that most products from alkane-forming pathways decreased severely in G-bulk. For instance, the C29 alkane, C30 aldehyde, and C29 ketones in G-bulk reached only 2.5, 0, and 4.4%, respectively, of the levels found on the leaves of W-bulk (Figure 2a, Table S1). However, the C25 alkane and C26 aldehyde were significantly increased in G-bulk than in W-bulk (Figure 2a, Table S1). Considering the products from the alcohol-forming pathway, the amount of C28 primary alcohol decreased by 62% in G-bulk, whereas the amount of C26 primary alcohol increased approximately 2.8-fold in G-bulk. Similarly, the amounts of C28 and C30 fatty acids decreased significantly, whereas the amount of C26 fatty acids increased significantly in G-bulk (Figure 2a, Table S1). Overall, wax components with chain lengths beyond C26 decreased severely in glossy plants, while shorter chains increased several-fold compared with those in the waxy plants. These findings suggested that the

glossy phenotype of SD369 might be caused by the interruption in VLCFA carbon-chain elongation from C26 to C28 during cuticular wax biosynthesis.

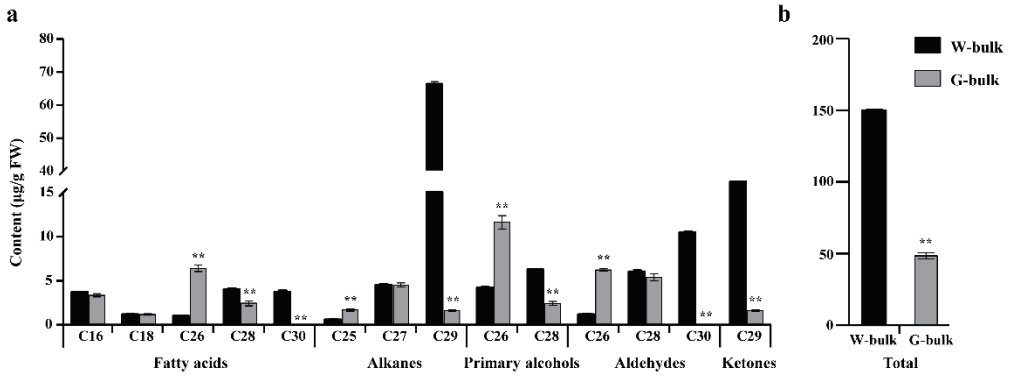


Figure 2. Cuticular wax composition in leaves from W-bulk and G-bulk. (**a**) Wax composition of W-bulk and G-bulk. (**b**) Total wax load of W-bulk and G-bulk. Error bars indicate SD (n = 3). ** *p*-value < 0.01.

2.3. Fine Mapping of the BrWAX3 Gene

To identify candidates of the *BrWAX3* gene, 50 waxy (W-pool) and 50 glossy (G-pool) individuals were selected from the F_2 population and used to construct two extreme pools for Bulked Segregant Analysis (BSA). In total, we obtained 168 and 173 million raw reads for the W-pool and G-pool (Table S2), and 1418,060 SNPs and 201,519 InDels were identified between the two DNA pools. Through sliding window analysis with the absolute value of Δ(SNP-index), a 6.5-Mb candidate region from 19.65 to 26.15 Mb on chromosome A09 was identified at a 0.01 confidence level (Figure 3a).

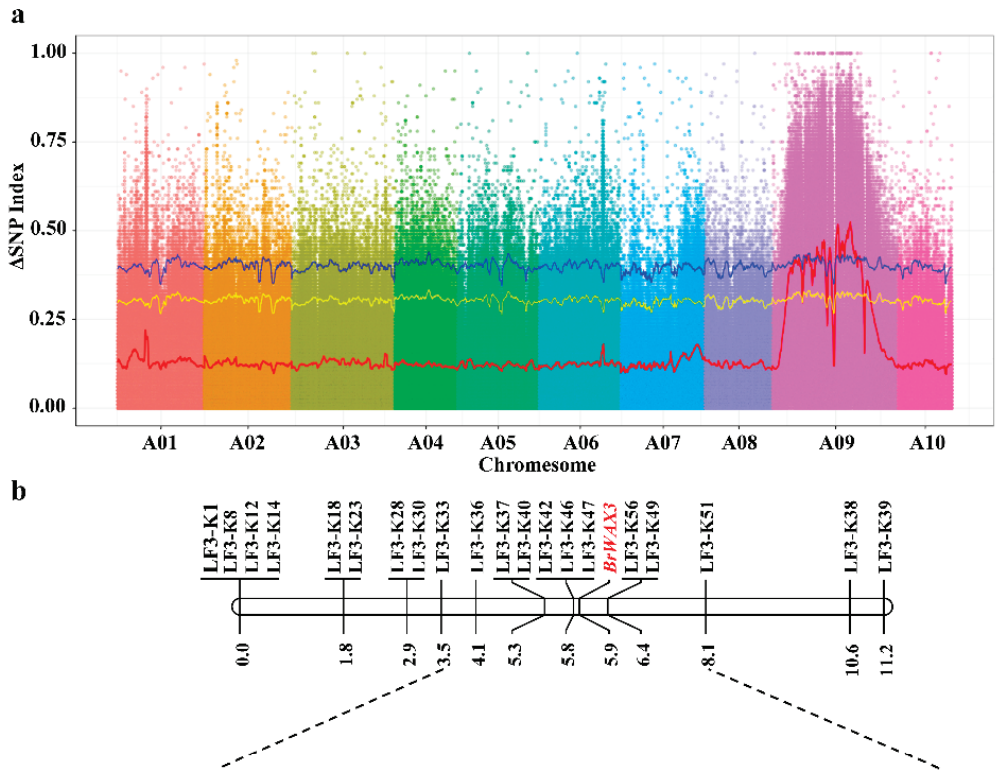


Figure 3. *Cont.*

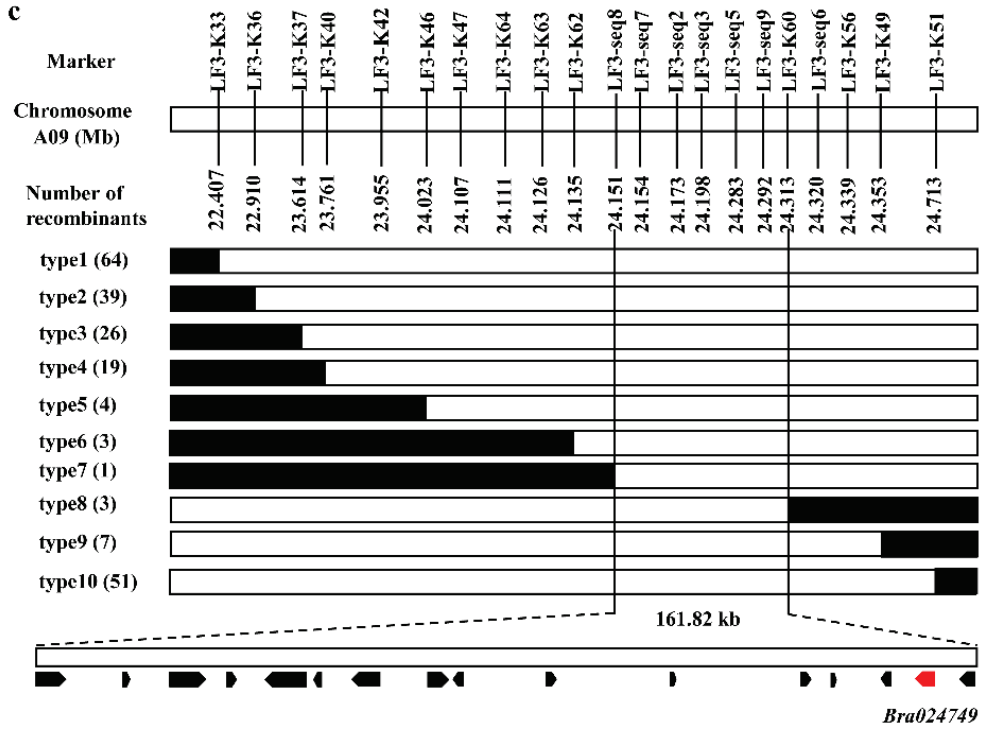


Figure 3. Map-based cloning of *BrWAX3* gene in Chinese cabbage. (**a**) BSA-seq analysis for *BrWAX3*. The Δ(SNP-index) was calculated at 1-Mb intervals with a 50-kb sliding window. One candidate region was identified on chromosome A09. (**b**) Initial mapping of *BrWAX3*. The genetic map of *BrWAX3* was showed with cM as the unit. (**c**) Fine mapping of *BrWAX3*. The *BrWAX3* gene was delimited to an interval between LF3-seq8 and LF3-K60 on chromosome A09, with an estimated physical length of 161.82 kb, and 16 genes were annotated in this region based on the reference genome sequence. The genetic structure of each recombinant type is depicted as white for homozygous glossy phenotype, black for heterozygous alleles, respectively. The number of each recombinant type is indicated in the brackets.

Based on BSA-seq analysis results, 56 KASP markers were developed, and 20 markers (Table S3) were polymorphic between the parents. Using these 20 KASP markers, 93 F_2 plants were genotyped for linkage analysis (Table S4). As shown in Figure 3b, the *BrWAX3* locus was initially mapped to a region on chromosome A09 between KASP markers LF3-K47 and LF3-K56, with a physical interval of 232 kb. The genetic distances between the *BrWAX3* locus and LF3-K47 and LF3-K56 were 0.1 and 0.5 cM, respectively (Figure 3b).

To further finely map the *BrWAX3* locus, 954 glossy F_2 plants were screened using flanking markers, LF3-K33 and LF3-K51, and a total of 115 recombinants (type 1 and type 10) were identified (Figure 3c). All the 115 recombinants were further genotyped using LF3-K36, LF3-K37, LF3-K40, LF3-K42, LF3-K46, LF3-K47, LF3-K56, and LF3-4K9, based on which 6 recombinants (type 6 and type 8) were identified (Figure 3c). Then, 4 KASP markers and 7 sequencing markers were further developed to genotype the 6 recombinants. The results delimited the *BrWAX3* gene to a 161.82-kb interval between markers LF3-seq8 and LF3-K60 (type 7 and type 8) (Figure 3c), with 1 recombinant between *BrWAX3* and LF3-seq8 and 3 recombinants between *BrWAX3* and LF3-K60. Five markers, namely, LF3-seq7, LF3-seq2, LF3-seq3, LF3-seq5, and LF3-seq9, co-segregated with the *BrWAX3* gene in the fine-mapping population (Figure 3c).

2.4. Candidate Gene Analysis

According to the annotation of *B.rapa* reference genome (V1.5), a total of 16 annotated or predicted genes were found within the 161.82-kb candidate interval (Table 2). Among these 16 genes, only the *Bra024749* gene, which is homologous to *CER60* in *Arabidopsis*,

could be the candidate gene (Table 2). *CER60* encodes a β-ketoacyl-CoA synthase that is involved in the biosynthesis of very long-chain fatty acids (VLCFAs) during cuticular wax biosynthesis.

Table 2. Annotated genes in the candidate interval of the *BrWAX3* locus.

Gene Name	Gene Position on A09	*Arabidopsis* Homolog	Gene Function
Bra024763	24151883-24153849	*AT1G25280*	Member of TLP family
Bra024762	24164515-24164946	*AT5G05020*	Pollen Ole e 1 allergen and extension family protein
Bra024761	24171584-24174217	*AT1G25320*	Leucine-rich repeat protein kinase family protein
Bra024760	24179389-24180542	*AT1G25340*	putative transcription factor (MYB116)
Bra024759	24185650-24190615	*AT1G25350*	glutamine-tRNA ligase, putative/glutaminyl-tRNA synthetase
Bra024758	24192301-24193325	*AT1G25370*	hypothetical protein (DUF1639)
Bra024757	24197064-24199869	*AT1G25375*	Metallo-hydrolase/oxidoreductase superfamily protein
Bra024756	24207077-24209333	*AT1G25380*	Encodes a mitochondrial-localized NAD+ transporter that transports NAD+ in a counter exchange mode with ADP and AMP in vitro
Bra024755	24210118-24212198	*AT1G25390*	Protein kinase superfamily protein
Bra024754	24226168-24227349	*AT2G05970*	F-box protein (DUF295)
Bra024753	24253251-24253559	*AT1G25422*	hypothetical protein
Bra024752	24279278-24280968	*AT1G29470*	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Bra024751	24285640-24285933	*AT1G25425*	CLAVATA3/ESR-RELATED 43
Bra024750	24299524-24301114	*AT1G25440*	B-box type zinc finger protein with CCT domain-containing protein
Bra024749	24304800-24306911	*AT1G25450*	Encodes KCS5, a member of the β-ketoacyl-CoA synthase family involved in the biosynthesis of VLCFA (very long chain fatty acids); CER60
Bra024748	24311235-24312823	*AT1G69710*	Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain-containing protein

The genomic sequence (gDNA) and coding sequence (CDS) of *Bra024749* in the parental lines were amplified and sequenced using the primer pairs BrWAX3-Ful-F and BrWAX3-Ful-R1 (Table S5). The results showed that the *Bra024749* gene in the waxy parent R16-11 was 2112 bp in length and contained 2 exons and 1 intron (Figure 4a). The CDS of the *Bra024749* gene in R16-11 was 1494 bp in length, which shares 87% identity with *CER60* (*KCS5*) in *Arabidopsis*. Therefore, *Bra024749* was also designated *BrCER60.A09* in this study. However, in the glossy parent SD369, the *Bra024749* gene was 7679 bp in length, which was caused by a 5567-bp insertion at 590 bp in the first exon (Figure 4a,c). The large fragment insertion disrupted the normal transcription and translation of *Bra024749* in SD369. As shown in Figure 4a, no CDS products were detected in glossy SD369 using full-length primer pairs BrWAX3-Ful-F and BrWAX3-Ful-R1. Five primer pairs (P2-P6) (Figures 4b and S1, Table S5) spanning the full length of *Bra024749* in SD369 were further designed. Two of the five cDNA products, amplified using primers BrWAX3-P3 and BrWAX3-P5 (Figure 4b), could not be detected, which explained why the full-length CDS of *Bra024749* in SD369 could not be detected when the full-length primer pairs were used. We also mapped the RNA sequencing (RNA-seq) reads from G-bulk to the *Bra024749* gDNA sequence of SD369, and the results showed that very few reads were mapped to the 3200–3450 bp region and the 5500–5650 bp region (Figures S2 and S3), among the amplified regions of primers BrWAX3-P3 and BrWAX3-P5, respectively, which supported that transcription of *Bra024749* in SD369 was interrupted. Most importantly, the 5567-bp insertion caused premature translation termination at the 205 amino acid position, which caused the loss of function of BrWAX3 (Figure 4c,d).

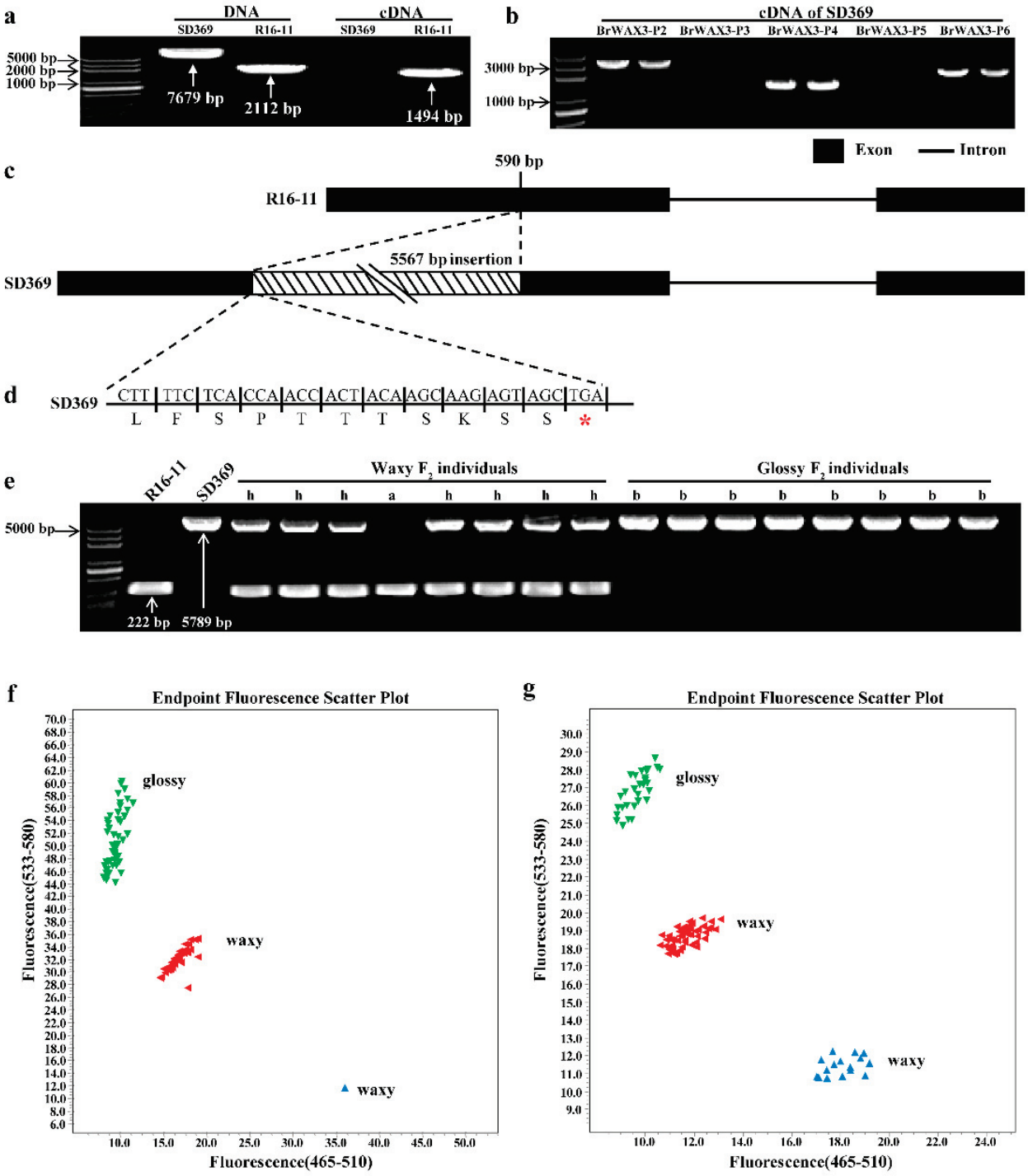


Figure 4. Candidate gene analysis of *BrWAX3*. (**a**) Amplification products of the full length primer using DNA and cDNA from SD369 and R16-11. The full length CDS of glossy SD369 could not be detected. (**b**) Amplification products of each fragmental primer using cDNA from SD369. (**c**) *BrWAX3* includes 2 exons and 1 intron. (**d**) A 5567-bp insertion in glossy SD369 caused a premature stop codon. (**e**) Validation of the functional marker BrWAX3-InDel in F_2 individuals. (**f**,**g**) Validation of the functional marker BrWAX3-KASP1 in BC_1P_1 population (**f**) and in (SD369 × SD2135)-F_2 population (**g**).

Based on the 5567-bp insertion, a functional marker BrWAX3-InDel (primer pairs BrWAX3-InDel-F and BrWAX3-InDel-R, Table S5), which could amplify a 222-bp and 5789-bp product from lines R16-11 and SD369, respectively, was developed. When Phanta® High-Fidelity DNA Polymerase (5 s/kb amplification rate) (Vazyme, Nanjing, China) with

30 s PCR extension time was used, all glossy F_2 individuals showed 5789-bp products, and waxy F_2 individuals presented either a homozygous 222-bp band or both bands (Figure 4e), which revealed that BrWAX3-InDel co-segregated with the cuticular wax phenotype in the F_2 population. We also assayed the BrWAX3-InDel marker in the BC_1P_1 population and another F_2 population (SD369 × SD2135)-F_2 via EasyTaq DNA Polymerase (1 min/kb amplification rate) (Trans, Beijing, China) in conjunction with a 30 s PCR extension time. The results also showed 100% consistency between the cuticular wax phenotype and genotype (Figures S4 and S5) with no band in glossy individuals and a 222-bp band in waxy individuals, as the 30 s PCR extension time (1 min/kb amplification rate) was not enough for the 5789-bp product in glossy plants but was sufficient for the 222-bp product in waxy plants. We also developed and validated a KASP marker BrWAX3-KASP1 based on the 5567-bp insertion (Figure 4f,g), which could be used for high-throughput genotyping systems.

Taken together, the above findings suggest that the *Bra024749* gene is the candidate gene for the cuticular wax gene *BrWAX3*.

2.5. Expression Pattern Analysis and Subcellular Localization of BrWAX3

The expression levels of *Bra024749* (*BrCER60.A09*) in different tissues of the two parent lines were examined by qRT-PCR analysis using primer pairs BrWAX3-qF and BrWAX3-qR (Table S5). The results suggested that the *Bra024749* transcript was found in various tissues, including the stems, leaves, sepals, petals, stamens, and pistils, but not in the roots, with the highest level in leaves (Figure 5a). The expression of *Bra024749* was much lower in SD369 than in the waxy parent R16-11 in any of the tissues we examined (Figure 5a).

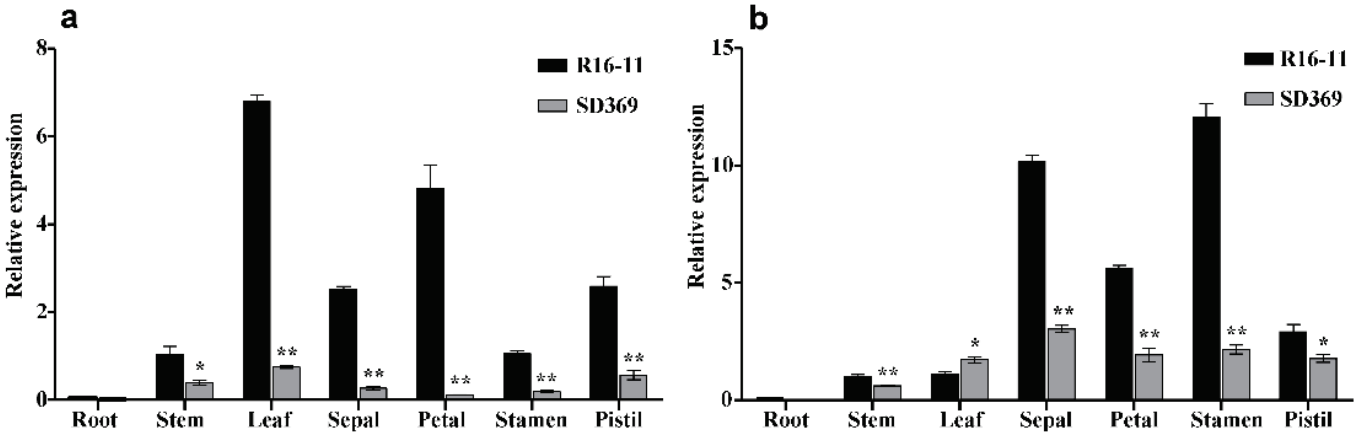


Figure 5. Gene expression data analysis. Quantitative RT-PCR of *BrWAX3* (*BrCER60.A09*) (**a**) and *Bra004034* (*BrCER6.A07*) (**b**) in different tissues of the two parents. The *BrGAPDH* was used as an internal control. Error bars indicate SE (n = 3). * $p < 0.05$. ** $p < 0.01$.

To evaluate the subcellular localization, a fusion protein of BrCER60-GFP under the drive of 35 S CaMV promoter was transiently expressed in tobacco leaf epidermal cells. The results showed that the green fluorescent signals from BrCER60-GFP were found in the ER (Figure 6a), exactly overlapping with the red fluorescent signals from the ER marker (Figure 6b–d), indicating that *Bra024749* (*BrCER60.A09*) was localized to the ER.

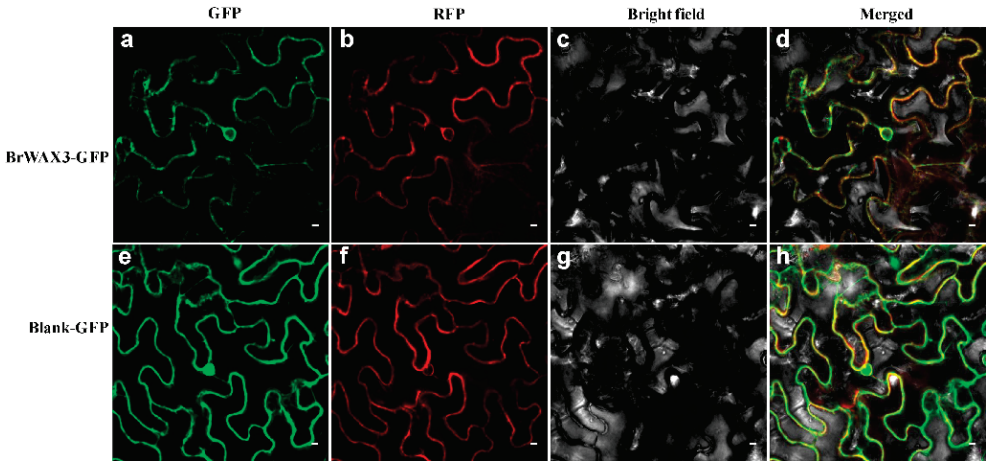


Figure 6. Subcellular localization analysis of BrWAX3 (BrCER60.A09). (**a**,**e**) The fusion construct BrWAX3-GFP (**a**) and the construct control (**e**) were transiently introduced into the tobacco leaf epidermal cells. (**b**,**f**) Fluorescence signals of an ER marker protein AtPIN5 fused with RFP. (**c**,**g**) Bright-field image. (**d**,**h**) Merged image of bright field and fluorescence.

2.6. Sequence and Expression Pattern Analysis of BrCER6.A07

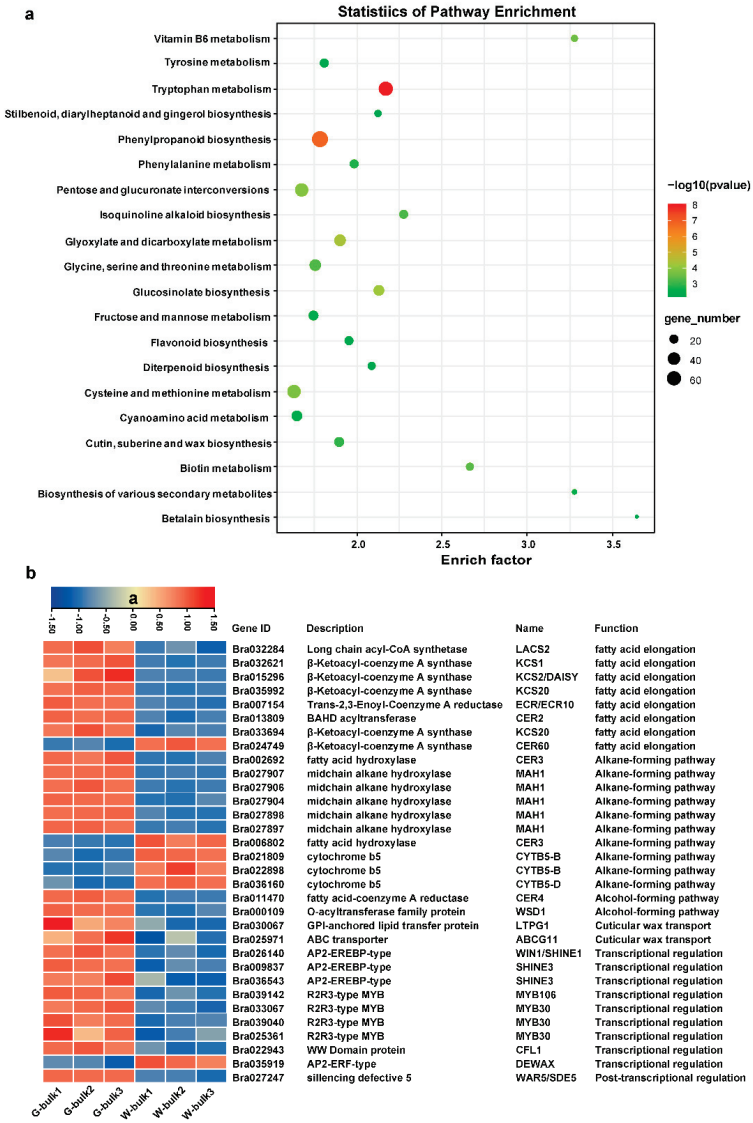
In *Arabidopsis*, *CER6* (*AT1G68530*) is a paralog of *CER60* (*AT1G25450*) [18]. Therefore, we blasted the coding sequence of *CER6* (*AT1G68530*) against the *B.rapa* genome, and found that the best-hit gene was *Bra004034* (*BrCER6.A07*), which shared 82.2% identity with the candidate *Bra024749* (*BrCER60.A09*) at the coding sequence level. We designed a gene-specific primer pair, BrCER6.A07-ful-F and BrCER6.A07-ful-R (Table S5), to amplify the full-length CDS of *BrCER6.A07* in SD369 and R16-11. The CDSs of *BrCER6.A07* from SD369 and R16-11 were submitted to GeneBank under accession numbers OPO46432 and OPO46431. Eight SNPs were identified between the coding sequences of R16-11 and SD369 (Figure S6, Table S6). Even though two SNPs caused nonsynonymous mutations, they did not affect protein function (Figure S7, Table S6). We also compared the expression level of *BrCER6.A07* in the parental lines using primer pairs BrCER6.A07-qF and BrCER6.A07-qR (Table S5). As shown in Figure 5b, the stamen showed the highest expression levels, whereas much lower levels were found in stems and leaves. Furthermore, *BrCER6.A07* showed a comparable level in stems between SD369 and R16-11 and was lower in leaves of waxy R16-11, which suggested that the *BrCER6.A07* gene was not responsible for glossy phenotype.

2.7. Transcriptome Analysis in Waxy and Glossy Stems

We performed comparative transcriptome analysis between the W-bulk and G-bulk to identify the gene regulatory networks involved in cuticular wax biosynthesis. We obtained approximately 261 million raw reads for the six cDNA libraries, ranging from 42.3 to 44.4 million reads per library (Table S7). The raw data were submitted to SRA database under accession number PRJNA860219 (accessed on 19 July 2022). Among the clean reads, 75.1–79.7% were uniquely mapped to the reference genome (Table S7). In total, we identified 5314 differentially expressed genes (DEGs) between the W-bulk and G-bulk, among which 2513 genes were upregulated and 2801 were downregulated in the G-bulk compared with the W-bulk.

KEGG pathway enrichment analysis revealed that 'cutin, suberine, and wax biosynthesis' was significantly enriched (Figure 7a). In accordance with the reduced amount of cuticular wax in glossy plants, the candidate gene *Bra024749* (*BrCER60.A09*) was significantly downregulated in G-bulk, but most of the other genes involved in wax biosynthesis and its regulation were upregulated (Figure 7b, Table S8). For example, the genes *LACS2* (*Bra032284*), *ECR* (*Bra007154*), *CER2* (*Bra013809*), *KCS2* (*Bra015296*), and *KCS20* (*Bra033694*), which participate in fatty acid elongation, were upregulated in G-bulk (Figure 7b, Table S8). Most

of the genes in alkane-forming pathways, such as *CER3* (*Bra002692*), *MAH1* (*Bra027907*, *Bra027906*, *Bra027904*, *Bra027898* and *Bra027897*), and genes in alcohol-forming pathway, such as *CER4* (*Bra011470*) and *WSD1* (*Bra000019*), were all upregulated in glossy plants (Figure 7b, Table S8). Additionally, the *SHINE1* (*Bra026140*), *SHINE3* (*Bra009837* and *Bra036543*), *MYB30* (*Bra033067*, *Bra039040* and *Bra025361*) and *MYB106* (*Bra039140*) genes, which participate in transcriptional regulation of cutin and wax biosynthesis, were also upregulated in G-bulk (Figure 7b, Table S8). Our qRT-PCR analysis of wax metabolism-related genes was consistent with the RNA-seq results (Figure S8). Cuticular waxes contain not only VLCFAs and their derivatives, but also other secondary metabolites, such as terpenoids, sterols, and flavonoids [6,24,32]. As expected, pathways of 'flavonoid biosynthesis', 'phenylalanine metabolism', 'glucosinolate biosynthesis', 'stilbenoid, diarylheptanoid and gingerol biosynthesis' were significantly enriched (Figure 7a), and most genes in these pathways were downregulated in G-bulk (Table S8), which was consistent with the reduced number of waxes in glossy plants.



Gene ID	Description	Name	Function
Bra032284	Long chain acyl-CoA synthetase	LACS2	fatty acid elongation
Bra032621	β-Ketoacyl-coenzyme A synthase	KCS1	fatty acid elongation
Bra015296	β-Ketoacyl-coenzyme A synthase	KCS2/DAISY	fatty acid elongation
Bra035992	β-Ketoacyl-coenzyme A synthase	KCS20	fatty acid elongation
Bra007154	Trans-2,3-Enoyl-Coenzyme A reductase	ECR/ECR10	fatty acid elongation
Bra013809	BAHD acyltransferase	CER2	fatty acid elongation
Bra033694	β-Ketoacyl-coenzyme A synthase	KCS20	fatty acid elongation
Bra024749	β-Ketoacyl-coenzyme A synthase	CER60	fatty acid elongation
Bra002692	fatty acid hydroxylase	CER3	Alkane-forming pathway
Bra027907	midchain alkane hydroxylase	MAH1	Alkane-forming pathway
Bra027906	midchain alkane hydroxylase	MAH1	Alkane-forming pathway
Bra027904	midchain alkane hydroxylase	MAH1	Alkane-forming pathway
Bra027898	midchain alkane hydroxylase	MAH1	Alkane-forming pathway
Bra027897	midchain alkane hydroxylase	MAH1	Alkane-forming pathway
Bra006802	fatty acid hydroxylase	CER3	Alkane-forming pathway
Bra021809	cytochrome b5	CYTB5-B	Alkane-forming pathway
Bra022898	cytochrome b5	CYTB5-B	Alkane-forming pathway
Bra036160	cytochrome b5	CYTB5-D	Alkane-forming pathway
Bra011470	fatty acid-coenzyme A reductase	CER4	Alcohol-forming pathway
Bra000109	O-acyltransferase family protein	WSD1	Alcohol-forming pathway
Bra030067	GPI-anchored lipid transfer protein	LTPG1	Cuticular wax transport
Bra025971	ABC transporter	ABCG11	Cuticular wax transport
Bra026140	AP2-EREBP-type	WIN1/SHINE1	Transcriptional regulation
Bra009837	AP2-EREBP-type	SHINE3	Transcriptional regulation
Bra036543	AP2-EREBP-type	SHINE3	Transcriptional regulation
Bra039142	R2R3-type MYB	MYB106	Transcriptional regulation
Bra033067	R2R3-type MYB	MYB30	Transcriptional regulation
Bra039040	R2R3-type MYB	MYB30	Transcriptional regulation
Bra025361	R2R3-type MYB	MYB30	Transcriptional regulation
Bra022943	WW Domain protein	CFL1	Transcriptional regulation
Bra035919	AP2-ERF-type	DEWAX	Transcriptional regulation
Bra027247	silencing defective 5	WAR5/SDE5	Post-transcriptional regulation

Figure 7. Transcriptome analysis in waxy and glossy stems. (**a**) Scatter plot of top 20 enriched KEGG pathways. Rich factor is the ratio of the DEG number to the background number in a certain pathway. The size of the dots represents the number of genes, and the color of the dots represents the range of the *p*-value. (**b**) Differentially expressed genes involved in cuticular wax metabolism. The heatmap colors are shown in log2(FPKM + 1). Three biological replicates of the W-bulk and G-bulk are shown.

3. Discussion

In the present study, the *BrWAX3* gene, which confers wax biosynthesis, was successfully and finely mapped to a physical interval of 161.82 kb. Several lines of evidence indicate that *Bra024749* (*BrCER60.A09*) is the candidate gene for *BrWAX3*. (1) Among the 16 annotated genes within the 161.82 kb interval, only one gene, *Bra024749*, which is homologous to *CER60/KCS5* in *Arabidopsis*, might be involved in cuticular wax biosynthesis. (2) The expression level of *BrCER60.A09* in all aerial organs was much lower in glossy plants than in waxy plants. (3) A 5567-bp insertion was found in glossy plants, which resulted in a premature stop codon and loss of function of the CER60 enzyme. (4) Two functional markers for *BrCER60.A09* co-segregated with the wax phenotype. (5) Subcellular localization analysis showed that the *BrCER60.A09* was localized to the ER, which is the site of wax biosynthesis. (6) Cuticular wax composition analysis showed a reduction of wax monomers with chain lengths beyond C26 and an increased proportion of shorter chains, which was in agreement with the CER60 function in *Arabidopsis* [19]. Overall, the 5567-bp insertion of *BrCER60.A09* in SD369 was the main cause of the glossy phenotype.

Studies in *Arabidopsis* showed that several *cer* mutants (*cer1, cer2, cer3,* and *cer6*) are male sterile due to defective pollen recognition or failed pollen hydration [18,33,34]. *CER6/KCS6* is involved in VLCFA elongation of C26 to C28 during cuticular wax biosynthesis. The *cer6* mutant showed a substantial reduction of derivatives beyond C26, nearly abolished stem wax accumulation, and exhibited conditional male sterility [10,18–20]. *CER60/KCS5*, a paralog of *CER6/KCS6*, plays a redundant role with *CER6/KCS6* in wax biosynthesis, but *CER6/KCS6* plays a major role [19]. The mutation in *CER60/KCS5* caused only a slight reduction in total wax amounts in leaves and flowers, and the wax amounts in stems barely changed [19]. In our study, the mutation of *BrWAX3* (*BrCER60.A09*) in SD369 caused a significant reduction of wax monomers with chain lengths beyond C26 and an increased number of shorter chains, which was the same as *CER60/KCS5* in *Arabidopsis* [19]. However, *BrWAX3* (*BrCER60.A09*) showed a higher expression level in stems and leaves than *Bra004034* (*BrCER6.A07*) did (Figures 5a and S9), and the mutation of *BrWAX3* (*BrCER60.A09*) in SD369 caused an obvious glossy appearance of its stems, leaves, and flower buds, reflecting the predominant role of *BrWAX3* (*BrCER60.A09*) in wax biosynthesis on aerial organs in Chinese cabbage. These results were different from those for *CER60/KCS5* in *Arabidopsis*, which is transcribed at a low level in all the mature shoot tissues [20] and plays a minor role in wax biosynthesis [19]. However, another study revealed that GUS expression driven by the *AtKCS5* promoter was much higher than that driven by *AtKCS6* [35], which is consistent with the results of our study. Additionally, we obtained many seeds after self-pollination of SD369, suggesting that the mutation of *BrWAX3* (*BrCER60.A09*) in glossy SD369 did not cause male sterility, which was different from the male sterility that occurred for the *cer6* mutant [10,18]. We speculated that the higher expression of *BrCER6.A07* in stamens might compensate for the loss of function of *BrWAX3* (*BrCER60.A09*) in stamens of SD369 and restore the fertility. Further studies, such as those involving the generation of *BrWAX3* (*BrCER60.A09*) and *BrCER6.A07* single and double mutants, the identification of their possible roles in wax biosynthesis, and functional analysis of these two genes, are needed to provide more evidence for the involvement of these genes in wax formation and pollen development in *B.rapa*.

The SD369 mutant and the cloned *BrWAX3* gene are important for breeding *B.rapa* varieties. *B.rapa* includes a variety of vegetables, such as Chinese cabbage, Pakchoi, Cai-xin, and purple cai-tai. The edible parts of Cai-xin and purple cai-tai are tender stems, for which the glossy appearance is preferred by customers [30]. Therefore, breeders can introduce the locus from SD369 to Cai-xin and purple cai-tai, which would create new varieties with glossy phenotype but does not influence male fertility (unlike the *cer1* mutant in *B.rapa*). We also developed two functional markers, BrWAX3-InDel and BrWAX3-KASP1, for *BrWAX3* (*BrCER60.A09*) based on the 5567-bp insertion, which could be used for molecular marker-assisted breeding either through agarose gel electrophoresis or by high-throughput genotyping platforms.

The global effect of *BrWAX3* (*BrCER60.A09*) mutation on gene expression in Chinese cabbage was also investigated via comparative transcriptome analysis. In contrast to the lower expression level of *BrWAX3* (*BrCER60.A09*) in glossy plants, most genes in the fatty acid elongation, alkane-forming, and alcohol-forming pathways of wax biosynthesis, and in the transcriptional and posttranscriptional regulation of wax biosynthesis, were upregulated in glossy plants, indicating that a feedback mechanism occurred in glossy plants, which was consistent with the feedback observed in *nwgl* glossy cabbage [25]. However, we could not confirm it was the downregulation of *BrWAX3* (*BrCER60.A09*) or the reduced amount of wax that caused the feedback. Additionally, genes involved in cutin biosynthesis, such as *CYP77A6* (*Bra029852*) and *CYP86A4* (*Bra032642*), were also upregulated in glossy plants (Table S8), which are similar to those in our previous studies [24].

4. Materials and Methods

4.1. Plant Materials

SD369 (P_1), a doubled haploid (DH) line of Chinese cabbage with glossy phenotype, and R16-11 (P_2), a DH line with wax phenotype, were used as the parents to construct the F_1, F_2, and BC_1 populations for inheritance analysis and map-based cloning. Additionally, another F_2 population, (SD369 $\times$ SD2135)-F_2, was generated for marker validation by crossing the glossy line SD369 with the waxy line SD2135. All generations were grown in open fields at Henan Academy of Agricultural Sciences. At the bolting stage, the glossy and waxy phenotypes were assessed visually. Chi-square test (χ^2) was used to examine the phenotype segregation ratios.

4.2. Cryo-Scanning Electron Microscopy (cryo-SEM) and Gas Chromatography-Mass Mpectrometry (GC-MS)

A Hitachi SU3500 (Japan) scanning electron microscope was used to observe the abundance and morphology of wax of leaves from SD369 and R16-11. The method for cryo-SEM was followed as in our previous study [24].

W-bulk and G-bulk were constructed by mixing equal amounts of leaves from 10 waxy or 10 glossy F_2 plants, respectively. In total, three biological replicates of W-bulk and G-bulk were constructed. The cuticular wax composition and components in W-bulk and G-bulk were assessed via GC-MS at Shanghai Jiao Tong University, which was performed as described in our previous study [24].

4.3. Identification of Candidate Genes via Bulked-Segregant Analysis Sequencing (BSA-Seq) and Kompetitive Allele-Specific PCR (KASP) Assays

Candidate genes were identified according to the BSA-seq method [36]. Two DNA pools were constructed by mixing equal amounts of DNA from 50 waxy F_2 individuals (W-pool) and 50 glossy F_2 individuals (G-pool). The two DNA pools were resequenced by Anoroad Biotech Co., Ltd. (Beijing, China) using 150-bp PE strategy. The raw data were deposited in the Sequence Read Archive (SRA) in NCBI as PRJNA859942. The Burrows-Wheeler Aligner (BWA) [37] was used to map the clean reads to the *B.rapa* reference genome V1.5 [38]. The SAMtools software (V1.3.1) [39] was used to detect the single-nucleotide polymorphism (SNP) and insertion/deletion (InDel) variants between the W-pool and G-pool. Then, we calculated the SNP-index and Δ(SNP-index) for all genomic positions in the W-pool and G-pool, which was performed as in our previous studies [24,40]. Finally, the absolute value of Δ(SNP-index) was calculated for sliding window analysis, with a 1-Mb window width and a 50-kb sliding window step.

We used KASP assay to conduct the initial linkage analysis of the *BrWAX3* gene, which was performed as our previous studies [24,40,41]. First, SNPs showing polymorphism between the two DNA pools and nearing the candidate BSA-seq region were selected for KASP marker development [40]. Then, KASP markers (Table S3), showing polymorphism between the two parents, were employed to genotype the F_2 population containing 93 individuals. The genetic linkage map was constructed using JoinMap 4.0 software [42], and followed as in our previous studies [24,40].

954 individuals with glossy phenotypes were used for fine mapping of the candidate gene.

4.4. Gene Cloning and Sequence Analysis

BrCER60.A09, the candidate gene of *BrWAX3*, and its homologue *BrCER6.A07*, were cloned using Phanta® High-Fidelity enzyme Mix (Vazyme, Nanjing, China) in a total volume of 50 μL reaction: 3 μL DNA template, 3 μL of both forward and reverse primers, 25 μL enzyme mix, and 16 μL ddH_2O. The PCR conditions were performed as in the manuals. The PCR products were sequenced by Sunya Biotech Co., Ltd. (Zhengzhou, China). The sequences of SD369 and R16 were aligned using DNAMAN.

4.5. RNA Extraction and Expression Analysis

Various tissue samples (root, stem, leaf, sepal, petal, stamen, and pistil) of SD369 and R16-11 were collected. The total RNA of each sample was extracted using RNAiso Plus reagent (TaKaRa, Japan), and the first-strand cDNA was synthesized using the TransScript One-Step gDNA Removal and cDNA Synthesis Kit (Trans, Beijing, China). Quantitative real-time PCR (qRT-PCR) was performed with SYBR Premix Ex TaqTM II (TaKaRa, Japan). The analysis of gene relative expression data was performed using the $2^{-\Delta\Delta Ct}$ method [43]. *BrGAPDH* was employed as the reference gene [24,40]. The primers are listed in Table S5.

4.6. Subcellular Localization

The coding sequences of *BrCER60.A09* without the stop codon were amplified from waxy R16-11 using the primer pairs BrWAX3-fulF and BrWAX3-fulR2 (Table S5). The amplified cDNA fragments were subsequently inserted into the modified pBWA(V)HS vector under the control of the 35 S promoter. The resultant binary plasmid was referred to as pBrCER60-GFP. A pAtPIN5-RFP construct was used as an ER marker [44]. Vector of pBrCER60-GFP, pAtPIN5-RFP, and the blank control vector were infiltrated into epidermal cells of tobacco leaves through Agrobacterium-mediated transformation. Fluorescence signals were observed with a confocal laser scanning microscope (Nikon C1, Japan).

4.7. Transcriptome Analysis

The W-bulk and G-bulk each with three replicates were subjected to mRNA sequencing by BioMarker Tech Co., Ltd. (Beijing, China). The clean reads of each sample were aligned to the *B.rapa* reference genome (V1.5) using HISAT2 software (V2.1.0) [45]. Then, the fragments per kilobase of transcript per million mapped reads (FPKM) value of each gene were calculated to estimate gene expression levels. Genes with a q-value ≤ 0.05 and $|\log2(\text{fold change})| \geq 1$ identified by DESeq2 (V1.6.3) [46] were recognized as differentially expressed genes (DEGs). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was implemented using TBtools [47].

5. Conclusions

The present study showed the molecular mechanism of wax deficiency in SD369. The *BrWAX3* locus was fine-mapped to an interval of 161.82 kb, and *Bra024749* (*BrCER60.A09*), which encodes a β-ketoacyl-CoA synthase, was the most likely candidate gene for *BrWAX3*. A 5567-bp insertion of *BrWAX3* (*BrCER60.A09*) in glossy SD369 caused a loss of protein function, thus blocking the VLCFA elongation of C26 to C28, and ultimately resulting in the glossy phenotype. The loss of function of *BrWAX3* (*BrCER60.A09*) in glossy plants also caused feedback of genes involved in cutin and wax biosynthesis pathways. Besides, two functional markers for *BrWAX3* were developed and validated. Our research will promote molecular research on wax synthesis in *Brassica rapa*.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms231810938/s1.

Author Contributions: Conceptualization, Y.Y. and X.Z.; methodology, S.Y.; software, H.T.; formal analysis, X.W. and H.S.; investigation, Y.Z., Z.W. and L.N.; resources, X.Z.; data curation, S.Y. and H.T.;

writing—original draft preparation, S.Y.; writing—review and editing, Y.Y. and X.Z.; supervision, X.Z.; funding acquisition, S.Y. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Programs for Excellent Youth of Henan Academy of Agricultural Science (grant number 2022 YQ10), the Natural Science Foundation of Henan Province (grant number 222300420197), the National Science Foundation of China (grant number 32202485), Sci-Tech Innovation Team of Henan Academy Agricultural Sciences (grant number 2022 TD06), Joint Research Project of Improved Agricultural Varieties in Henan Province (grant number 2022010504).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data from BSA-seq analysis have been deposited into the SRA database (https://www.ncbi.nlm.nih.gov/sra/) with accession number PRJNA859942 (accessed on 18 July 2022). The raw data from transcriptome analysis were deposited in the SRA database under accession number PRJNA860219 (accessed on 19 July 2022). The CDS sequence of *BrWAX3* from waxy R16-11 was deposited in Genbank under accession number OPO46430 (accessed on 28 July 2022).

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

References

1. Riederer, M.; Schreiber, L. Protecting against water loss: Analysis of the barrier properties of plant cuticles. *J. Exp. Bot.* **2001**, *52*, 2023–2032. [CrossRef] [PubMed]
2. Haslam, T.M.; Kunst, L. *Arabidopsis* ECERIFERUM2-LIKEs are mediators of condensing enzyme function. *Plant Cell Physiol.* **2021**, *61*, 2126–2138. [CrossRef] [PubMed]
3. Beisson, F.; Li-Beisson, Y.; Pollard, M. Solving the puzzles of cutin and suberin polymer biosynthesis. *Curr. Opin. Plant Biol.* **2012**, *15*, 329–337. [CrossRef] [PubMed]
4. Liu, Z.; Fang, Z.; Zhuang, M.; Zhang, Y.; Lv, H.; Liu, Y.; Li, Z.; Sun, P.; Tang, J.; Liu, D.; et al. Fine-Mapping and analysis of *Cgl1*, a gene conferring glossy trait in Cabbage (*Brassica oleracea* L. var. capitata). *Front. Plant Sci.* **2017**, *8*, 239. [CrossRef] [PubMed]
5. Lee, S.B.; Suh, M.C. Advances in the understanding of cuticular waxes in *Arabidopsis thaliana* and crop species. *Plant Cell Rep.* **2015**, *34*, 557–572. [CrossRef] [PubMed]
6. Bernard, A.; Joubès, J. *Arabidopsis* cuticular waxes: Advances in synthesis, export and regulation. *Prog. Lipid Res.* **2013**, *52*, 110–129. [CrossRef]
7. Bourdenx, B.; Bernard, A.; Domergue, F.; Pascal, S.; Léger, A.; Roby, D.; Pervent, M.; Vile, D.; Haslam, R.P.; Napier, J.A.; et al. Overexpression of *Arabidopsis ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiol.* **2011**, *156*, 29–45. [CrossRef]
8. Yang, M.; Yang, Q.; Fu, T.; Zhou, Y. Overexpression of the *Brassica napus BnLAS* gene in *Arabidopsis* affects plant development and increases drought tolerance. *Plant Cell Rep.* **2011**, *30*, 373–388. [CrossRef]
9. Haslam, T.M.; Mañas-Fernández, A.; Zhao, L.; Kunst, L. *Arabidopsis* ECERIFERUM2 is a component of the fatty acid elongation machinery required for fatty acid extension to exceptional lengths. *Plant Physiol.* **2012**, *160*, 1164–1174. [CrossRef]
10. Millar, A.A.; Clemens, S.; Zachgo, S.; Giblin, E.M.; Taylor, D.C.; Kunst, L. *CUT1*, an *Arabidopsis* gene required for cuticular wax biosynthesis and pollen fertility, encodes a very-long-chain fatty acid condensing enzyme. *Plant Cell* **1999**, *11*, 825–838. [CrossRef]
11. Sieber, P.; Schorderet, M.; Ryser, U.; Buchala, A.; Kolattukudy, P.; Métraux, J.P.; Nawrath, C. Transgenic *Arabidopsis* plants expressing a fungal cutinase show alterations in the structure and properties of the cuticle and postgenital organ fusions. *Plant Cell* **2000**, *12*, 721–738. [CrossRef] [PubMed]
12. Haslam, T.M.; Kunst, L. Extending the story of very-long-chain fatty acid elongation. *Plant Sci.* **2013**, *210*, 93–107. [CrossRef] [PubMed]
13. Beaudoin, F.; Wu, X.; Li, F.; Haslam, R.P.; Markham, J.E.; Zheng, H.; Napier, J.A.; Kunst, L. Functional characterization of the *Arabidopsis* beta-ketoacyl-coenzyme A reductase candidates of the fatty acid elongase. *Plant Physiol.* **2009**, *150*, 1174–1191. [CrossRef] [PubMed]
14. Millar, A.A.; Kunst, L. Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *Plant J.* **1997**, *12*, 121–131. [CrossRef]
15. Lee, S.B.; Jung, S.J.; Go, Y.S.; Kim, H.U.; Kim, J.K.; Cho, H.J.; Park, O.K.; Suh, M.C. Two *Arabidopsis* 3-ketoacyl CoA synthase genes, *KCS20* and *KCS2/DAISY*, are functionally redundant in cuticular wax and root suberin biosynthesis, but differentially controlled by osmotic stress. *Plant J.* **2009**, *60*, 462–475. [CrossRef]
16. Kim, J.; Jung, J.H.; Lee, S.B.; Go, Y.S.; Kim, H.J.; Cahoon, R.; Markham, J.E.; Cahoon, E.B.; Suh, M.C. *Arabidopsis* 3-ketoacyl-coenzyme a synthase9 is involved in the synthesis of tetracosanoic acids as precursors of cuticular waxes, suberins, sphingolipids, and phospholipids. *Plant Physiol.* **2013**, *162*, 567–580. [CrossRef]

17. Todd, J.; Post-Beittenmiller, D.; Jaworski, J.G. *KCS1* encodes a fatty acid elongase 3-ketoacyl-CoA synthase affecting wax biosynthesis in *Arabidopsis thaliana*. *Plant J.* **1999**, *17*, 119–130. [CrossRef]
18. Fiebig, A.; Mayfield, J.A.; Miley, N.L.; Chau, S.; Fischer, R.L.; Preuss, D. Alterations in *CER6*, a gene identical to *CUT1*, differentially affect long-chain lipid content on the surface of pollen and stems. *Plant Cell* **2000**, *12*, 2001–2008. [CrossRef]
19. Huang, H.; Ayaz, A.; Zheng, M.; Yang, X.; Zaman, W.; Zhao, H.; Lü, S. *Arabidopsis KCS5* and *KCS6* play redundant roles in wax synthesis. *Int. J. Mol. Sci.* **2022**, *23*, 4450. [CrossRef]
20. Hooker, T.S.; Millar, A.A.; Kunst, L. Significance of the expression of the CER6 condensing enzyme for cuticular wax production in *Arabidopsis*. *Plant Physiol.* **2002**, *129*, 1568–1580. [CrossRef]
21. Pascal, S.; Bernard, A.; Sorel, M.; Pervent, M.; Vile, D.; Haslam, R.P.; Napier, J.A.; Lessire, R.; Domergue, F.; Joubès, J. The *Arabidopsis cer26* mutant, like the *cer2* mutant, is specifically affected in the very long chain fatty acid elongation process. *Plant J.* **2013**, *73*, 733–746. [CrossRef] [PubMed]
22. Kim, J.; Kim, R.J.; Lee, S.B.; Chung Suh, M. Protein-protein interactions in fatty acid elongase complexes are important for very-long-chain fatty acid synthesis. *J. Exp. Bot.* **2022**, *73*, 3004–3017. [CrossRef] [PubMed]
23. Ji, J.; Cao, W.; Tong, L.; Fang, Z.; Zhang, Y.; Zhuang, M.; Wang, Y.; Yang, L.; Lv, H. Identification and validation of an *ECERIFERUM2-LIKE* gene controlling cuticular wax biosynthesis in cabbage (*Brassica oleracea* L. var. capitata L.). *Theor. Appl. Genet.* **2021**, *134*, 4055–4066. [CrossRef] [PubMed]
24. Yang, S.; Liu, H.; Wei, X.; Zhao, Y.; Wang, Z.; Su, H.; Zhao, X.; Tian, B.; Zhang, X.W.; Yuan, Y. *BrWAX2* plays an essential role in cuticular wax biosynthesis in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *Theor. Appl. Genet.* **2022**, *135*, 693–707. [CrossRef]
25. Zhu, X.; Tai, X.; Ren, Y.; Chen, J.; Bo, T. Genome-wide analysis of coding and long non-coding RNAs involved in cuticular wax biosynthesis in cabbage (*Brassica oleracea* L. var. capitata). *Int. J. Mol. Sci.* **2019**, *20*, 2820. [CrossRef]
26. Dong, X.; Ji, J.; Yang, L.; Fang, Z.; Zhuang, M.; Zhang, Y.; Lv, H.; Wang, Y.; Sun, P.; Tang, J.; et al. Fine-mapping and transcriptome analysis of *BoGL-3*, a wax-less gene in cabbage (*Brassica oleracea* L. var. capitata). *Mol. Genet. Genom.* **2019**, *294*, 1231–1239. [CrossRef] [PubMed]
27. Liu, D.; Tang, J.; Liu, Z.; Dong, X.; Zhuang, M.; Zhang, Y.; Lv, H.; Sun, P.; Liu, Y.; Li, Z.; et al. *Cgl2* plays an essential role in cuticular wax biosynthesis in cabbage (*Brassica oleracea* L. var. capitata). *BMC Plant Biol.* **2017**, *17*, 223. [CrossRef]
28. Liu, D.; Dong, X.; Liu, Z.; Tang, J.; Zhuang, M.; Zhang, Y.; Lv, H.; Liu, Y.; Li, Z.; Fang, Z.; et al. Fine mapping and candidate gene identification for wax biosynthesis locus, *BoWax1* in *Brassica oleracea* L. var. capitata. *Front. Plant Sci.* **2018**, *9*, 309. [CrossRef]
29. Zhang, X.; Liu, Z.; Wang, P.; Wang, Q.; Yang, S.; Feng, H. Fine mapping of *BrWax1*, a gene controlling cuticular wax biosynthesis in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *Mol. Breeding* **2013**, *32*, 867–874. [CrossRef]
30. Wang, C.; Li, Y.; Xie, F.; Kuang, H.; Wan, Z. Cloning of the *Brcer1* gene involved in cuticular wax production in a glossy mutant of non-heading Chinese cabbage (*Brassica rapa* L. var. *communis*). *Mol. Breeding* **2017**, *37*, 142. [CrossRef]
31. Liu, C.; Song, G.; Wang, N.; Huang, S.; Feng, H. A single SNP in *Brcer1* results in wax deficiency in Chinese cabbage (*Brassica campestris* L. ssp. pekinensis). *Sci. Hortic.* **2021**, *282*, 110019. [CrossRef]
32. Samuels, L.; Kunst, L.; Jetter, R. Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* **2008**, *59*, 683–707. [CrossRef] [PubMed]
33. Preuss, D.; Lemieux, B.; Yen, G.; Davis, R.W. A conditional sterile mutation eliminates surface components from *Arabidopsis* pollen and disrupts cell signaling during fertilization. *Genes Dev.* **1993**, *7*, 974–985. [CrossRef]
34. Hülskamp, M.; Kopczak, S.D.; Horejsi, T.F.; Kihl, B.K.; Pruitt, R.E. Identification of genes required for pollen-stigma recognition in *Arabidopsis thaliana*. *Plant J.* **1995**, *8*, 703–714. [CrossRef] [PubMed]
35. Singh, S.; Geeta, R.; Das, S. Comparative sequence analysis across Brassicaceae, regulatory diversity in *KCS5* and *KCS6* homologs from *Arabidopsis thaliana* and *Brassica juncea*, and intronic fragment as a negative transcriptional regulator. *Gene Expr. Patterns* **2020**, *38*, 119146. [CrossRef]
36. Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S.; et al. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **2013**, *74*, 174–183. [CrossRef]
37. Li, H.; Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **2010**, *26*, 589–595. [CrossRef]
38. Wang, X.; Wang, H.; Wang, J.; Sun, R.; Wu, J.; Liu, S.; Bai, Y.; Mun, J.H.; Bancroft, I.; Cheng, F.; et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **2011**, *43*, 1035–1039. [CrossRef]
39. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
40. Yang, S.; Liu, H.; Zhao, Y.; Su, H.; Wei, X.; Wang, Z.; Zhao, X.; Zhang, X.W.; Yuan, Y. Map-based cloning and characterization of *Br-dyp1*, a gene conferring dark yellow petal color trait in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *Front. Plant Sci.* **2022**, *13*, 841328. [CrossRef]
41. Yang, S.; Yu, W.; Wei, X.; Wang, Z.; Zhao, Y.; Zhao, X.; Tian, B.; Yuan, Y.; Zhang, X. An extended KASP-SNP resource for molecular breeding in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *PLoS ONE* **2020**, *15*, e0240042. [CrossRef] [PubMed]
42. Ooijen, J.; Van, J.W. *JoinMap 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations*, version 4; Kyazma B.V.: Wageningen, The Netherlands, 2006.

43. 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] [PubMed]
44. Mravec, J.; Skůpa, P.; Bailly, A.; Hoyerová, K.; Krecek, P.; Bielach, A.; Petrásek, J.; Zhang, J.; Gaykova, V.; Stierhof, Y.D.; et al. Subcellular homeostasis of phytohormone auxin is mediated by the ER-localized PIN5 transporter. *Nature* **2009**, *459*, 1136–1140. [CrossRef] [PubMed]
45. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef] [PubMed]
46. 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]
47. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef] [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Genome-Wide Identification and Evolution Analysis of R2R3-MYB Gene Family Reveals S6 Subfamily R2R3-MYB Transcription Factors Involved in Anthocyanin Biosynthesis in Carrot

Ao-Qi Duan, Shan-Shan Tan, Yuan-Jie Deng, Zhi-Sheng Xu * and Ai-Sheng Xiong *

State Key Laboratory of Crop Genetics and Germplasm Enhancement, Ministry of Agriculture and Rural Affairs Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in East China, College of Horticulture, Nanjing Agricultural University, 1 Weigang, Nanjing 210095, China
* Correspondence: xuzhisheng@njau.edu.cn (Z.-S.X.); xiongaisheng@njau.edu.cn (A.-S.X.)

Abstract: The taproot of purple carrot accumulated rich anthocyanin, but non-purple carrot did not. MYB transcription factors (TFs) condition anthocyanin biosynthesis in many plants. Currently, genome-wide identification and evolution analysis of R2R3-MYB gene family and their roles involved in conditioning anthocyanin biosynthesis in carrot is still limited. In this study, a total of 146 carrot R2R3-MYB TFs were identified based on the carrot transcriptome and genome database and were classified into 19 subfamilies on the basis of R2R3-MYB domain. These R2R3-MYB genes were unevenly distributed among nine chromosomes, and *Ka*/*Ks* analysis suggested that they evolved under a purified selection. The anthocyanin-related S6 subfamily, which contains 7 MYB TFs, was isolated from R2R3-MYB TFs. The anthocyanin content of rhizodermis, cortex, and secondary phloem in 'Black nebula' cultivar reached the highest among the 3 solid purple carrot cultivars at 110 days after sowing, which was approximately 4.20- and 3.72-fold higher than that in the 'Deep purple' and 'Ziwei' cultivars, respectively. The expression level of 7 MYB genes in purple carrot was higher than that in non-purple carrot. Among them, *DcMYB113* (*DCAR_008994*) was specifically expressed in rhizodermis, cortex, and secondary phloem tissues of 'Purple haze' cultivar, with the highest expression level of 10,223.77 compared with the control 'DPP' cultivar at 70 days after sowing. *DcMYB7* (*DCAR_010745*) was detected in purple root tissue of 'DPP' cultivar and its expression level in rhizodermis, cortex, and secondary phloem was 3.23-fold higher than that of secondary xylem at 110 days after sowing. Our results should be useful for determining the precise role of S6 subfamily R2R3-MYB TFs participating in anthocyanin biosynthesis in carrot.

Keywords: carrot; R2R3-MYB family; transcription factor; anthocyanin; evolution; expression profile

1. Introduction

Carrot (*Daucus carota* L.), an annual or biennial herb vegetable crops, is one of the 10 vegetable crops grown worldwide, and has a long history of cultivation. According to its taproot color, they can be divided into five types: white, yellow, orange, red and purple carrot [1–4]. Non-purple carrot is produced by a mutation of purple carrot. Non-purple carrots are well known for their orange pigmentation and carotenoid accumulation [5]. Purple carrot can be divided into solid purple carrot type and tissue-specific purple carrot types [6–9]. Nowadays, purple carrots are popular among consumers because they are an excellent source of anthocyanins. As the anthocyanin group, they accumulate in fleshy roots, as well as in flowers, petioles and fruits. Carrots can be used to produce anthocyanin-rich concentrate for the pigment industry [10,11].

MYB transcription factors, as one of the largest transcription factor families in plants, are characterized at most four repetitive sequences (R), each of which is about 53 amino

acids, forming H1, H2 and H3 α-helices [12]. In each repeat, H2 and H3 helices form helix-turn-helix structure respond for the DNA interaction of MYB domain [13]. According to the repeat number, MYB TFs were divided into four types: 1R-MYB, R2R3-MYB, 3R-MYB and 4R-MYB [14]. Among the four types, R2R3-MYB transcription factor is the most common one, including two tandem repeats [15]. R2R3-MYB usually forms a complex with bHLH and WD40 proteins, which directly acts on structural genes of anthocyanin synthesis pathway in plants, thus affecting anthocyanin synthesis and accumulation [16–18]. MYB and MYB-related sequences are available in the Plant Transcription Factor Database [19]. The classification of R2R3-MYB TF members and the anthocyanins accumulation have been reported in many species. The whole genome sequences of R2R3-MYB TFs in different species have also been identified, such as *Arabidopsis thaliana* (126 R2R3-MYB proteins) [20], *Gossypium raimondii* (205 R2R3-MYB proteins) [21], soybean (244 R2R3-MYB proteins) [22], *Oryza sativa* (88 R2R3-MYB proteins) [23], apple (222 apple typical R2R3 MYB proteins) [24], Chinese Pear (105 R2R3-MYB proteins) [25] and tea plants (122 CsR2R3-MYB proteins) [26].

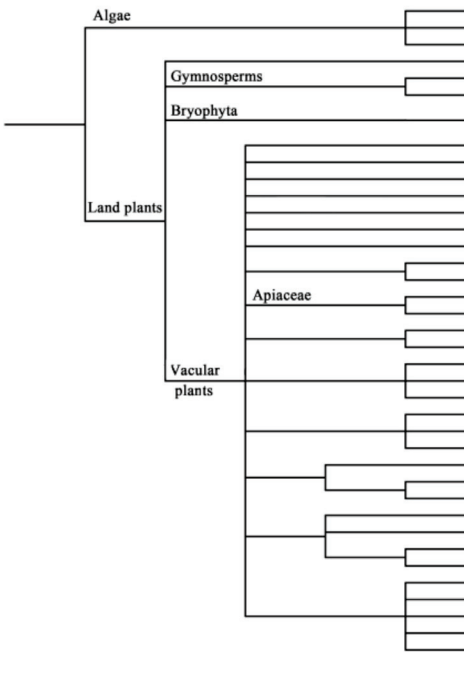
Anthocyanins are water-soluble natural pigments widely distributed in fruits and vegetables, with pH-dependent color changing from red to purple and then to blue. They are important secondary metabolites of plants, are synthesized in the cytosol and stored in vacuoles. Anthocyanins mainly contain six common aglycones and various types of glycosylations and acylations [27]. Over 600 kinds of anthocyanins have been identified, and the edible anthocyanins in nature include colored fruits and vegetables, such as apples, cherries, grapes, carrots, eggplant, celery, water dropwort and purple cabbage [28–34]. Anthocyanins are usually used as food colorants, which are beneficial to human health and have anti-cancer and anti-oxidation properties [29,35]. As a natural colorant, it is enjoying increasing loved by producers and consumers [30]. Among plants, pigments play an important role in attracting insects for seed transmission and pollination and improving the ability to resist abiotic and biological stresses.

In this study, to obtain further understanding in differences of anthocyanin synthesis between purple and non-purple carrot, we obtained the R2R3-MYB TFs members based on the carrot transcriptome and genome database, which are related to anthocyanin synthesis in carrot. Then, phylogenetic relationships, physicochemical properties, and anthocyanin contents of different color carrots were detected and analyzed. Expression patterns of seven selected *DcMYB* genes of the S6 subfamily in the taproots of six purple and three non-purple carrot cultivars were determined. This study will help establish the regulatory molecular mechanism of R2R3-MYB gene and pave the way for future functional research on anthocyanin in carrot.

2. Results

2.1. Evolution of MYB TFs of Carrot in Different Species

As shown in the Figure 1, the number of MYB TFs varies from several to several hundred in different species. The number of MYB family transcription factors in higher plants is generally higher than that in lower algae, which may be due to the replication and expansion of MYB TFs during the evolution of plants. The number of MYB family transcription factors in *Musa acuminata*, *Gossypium raimondii*, *Glycine max*, *Brassica rapa* and *Populus trichocarpa* was greater than 200. Compared with celery (3.33 Gb) [34,36], carrot has a smaller genome (~0.5 Gb) [37], but a similar number of MYB transcription factors.



species	Total MYB family factors	Genomic size (Mb)	Average number of MYB family factor per Mb
Chlamydomonas reinhardtii	16	130	0.123
Chlorella vulgaris	9	46	0.196
Ostreococcus sp. lucimarinus	11	13.25	0.830
Selaginella moellendorffii	24	212	0.113
Picea glauca	46	20800	0.002
Pinus taeda	35	20150	0.002
Physcomitrella patens	180	480	0.375
Citrus sinensis	107	367	0.292
Cucumis sativus	134	350	0.383
Nelumbo nucifera	134	929	0.144
Musa acuminata	303	523	0.579
Amborella trichopoda	60	748	0.080
Vitis vinifera	138	490	0.282
Phoenix dactylifera	129	685	0.188
Solanum lycopersicum	140	900	0.156
Solanum tuberosum	126	840	0.150
Daucus carota	157	480	0.327
Apium graveolens	154	3330	0.046
Theobroma cacao	154	430	0.358
Gossypium raimondii	299	775.2	0.386
Cicer arietinum	166	738	0.225
Glycine max	430	1100	0.391
Medicago truncatula	185	500	0.370
Prunus persica	163	265	0.615
Pyrus	184	527	0.349
Fragaria vesca	110	240	0.458
Carica papaya	98	370	0.265
Arabidopsis thaliana	168	125	1.344
Brassica rapa	293	485	0.604
Populus trichocarpa	266	480	0.554
Linum usitatissimum	206	373	0.552
Manihot esculenta	218	760	0.287
Ricinus communis	62	350	0.177
Hordeum vulgare	99	5000	0.020
Setaria italica	155	490	0.316
Zea mays	203	2300	0.088
Brachypodium distachyon	163	260	0.627
Oryza sativa Indica Group	121	430	0.281

Figure 1. Summary of MYB TFs among plants. *Daucus carota* was highlighted with purple box.

2.2. Classification and Phylogenetic Analysis of R2R3-MYB TFs in D. carota and A. thaliana and Analyses of Chromosomal Locations

Results showed that 146 R2R3-MYB TFs were identified based on carrot genome. The amino acid sequences of R2R3-MYB TFs in carrot and *A. thaliana* were aligned to draw a phylogenetic tree (Figure 2 and Supplementary Table S1). According to the evolutionary relationship of R2R3-MYB TFs between carrot and *A. thaliana*, MYB TFs of carrot were divided into 19 subfamilies (Figure 3). The number of MYB TFs in S14 and S20 subfamilies was the largest, with 9 in each subfamily. The second is the S6 subfamily, which has 7 TFs. In *A. thaliana*, the functions of R2R3-MYB in different subfamilies are widely different than in plants. *A. thaliana* S6 subfamily members include AtMYB75, AtMYB90, AtMYB113 and AtMYB114, which plays a regulatory role in anthocyanin biosynthesis [13,38,39]. Similarly, R2R3-MYB TFs in the S6 subfamily of carrot related to anthocyanin biosynthesis, including DcMYB6 (DCAR_000385), DcMYB7 (DCAR_010745), DcMYB8 (DCAR_010746), DcMYB9 (DCAR_010747), DcMYB11 (DCAR_010751), DcMYB113 (DCAR_008994) and DcMYB016 (DCAR_016451). According to the starting position of the R2R3-MYB gene in chromosomes, it was found that *DcMYB* genes were unevenly distributed among 9 carrot chromosomes (Figure 2B). Some *DcMYB* genes cannot eventually be mapped on any chromosome. Chromosome 3 contained the most *DcMYB* genes (20), followed by chromosome 1 (17), and the lowest numbers were found on chromosome 7 (10). More genes were observed at the bottom of chromosomes 1, 2, 3 and 6, and the genes on chromosomes 7 and 8 were evenly distributed.

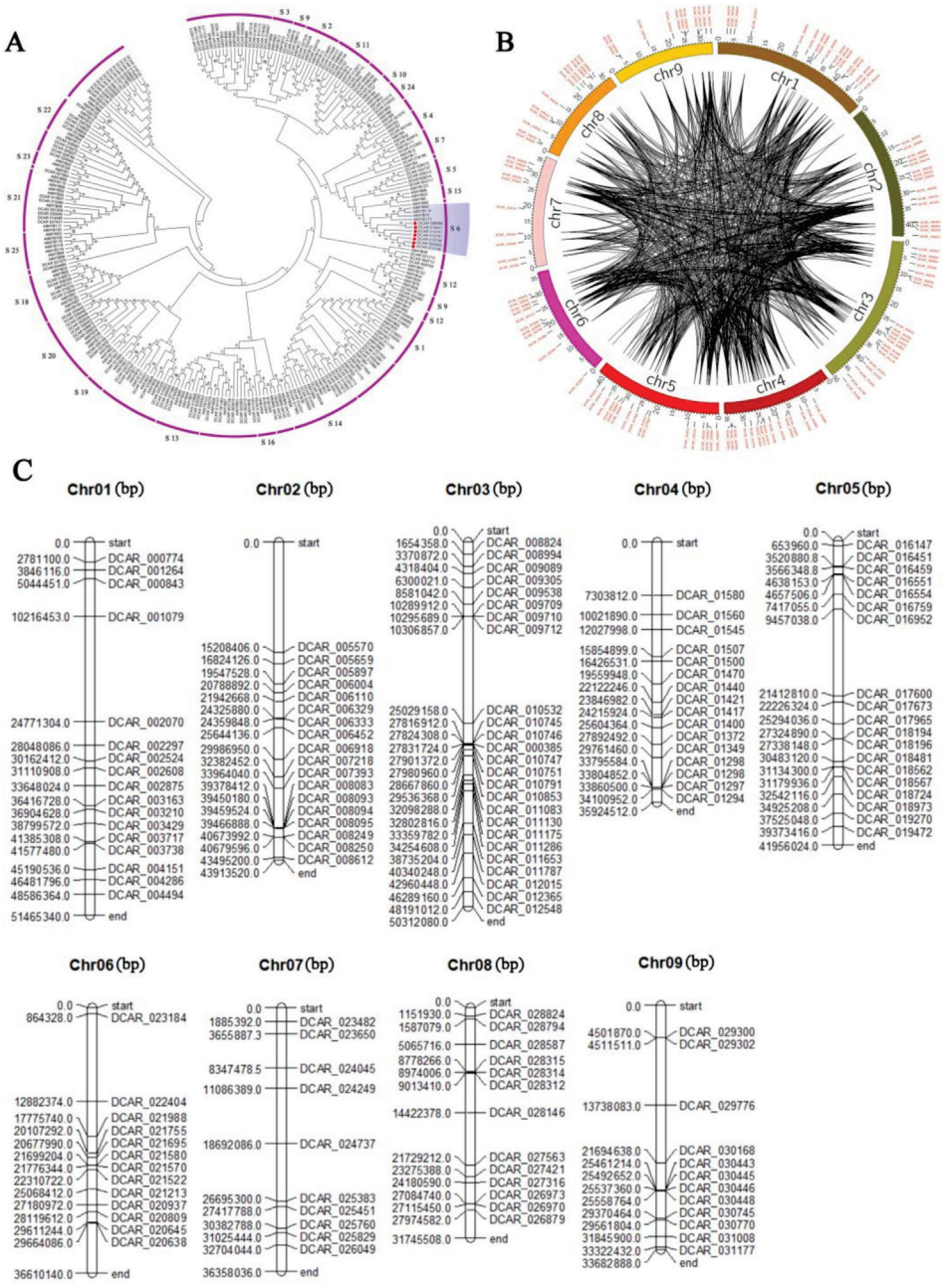


Figure 2. Phylogenetic tree of R2R3-MYB TFs of *Arabidopsis* and carrot and chromosomal determination and collinearity analysis of R2R3-MYB genes in carrot. Phylogenetic trees were constructed from the R2R3-MYB protein sequence of carrot and *Arabidopsis* (**A**). The phylogenetic trees were established by the neighbor-joining (NJ) method. (**B**,**C**) represent the carrot chromosomes. The black lines represent collinear gene pairs. The genes were mapped onto different chromosomes and the numbers along the chromosome boxes represent sequence lengths in megabases.

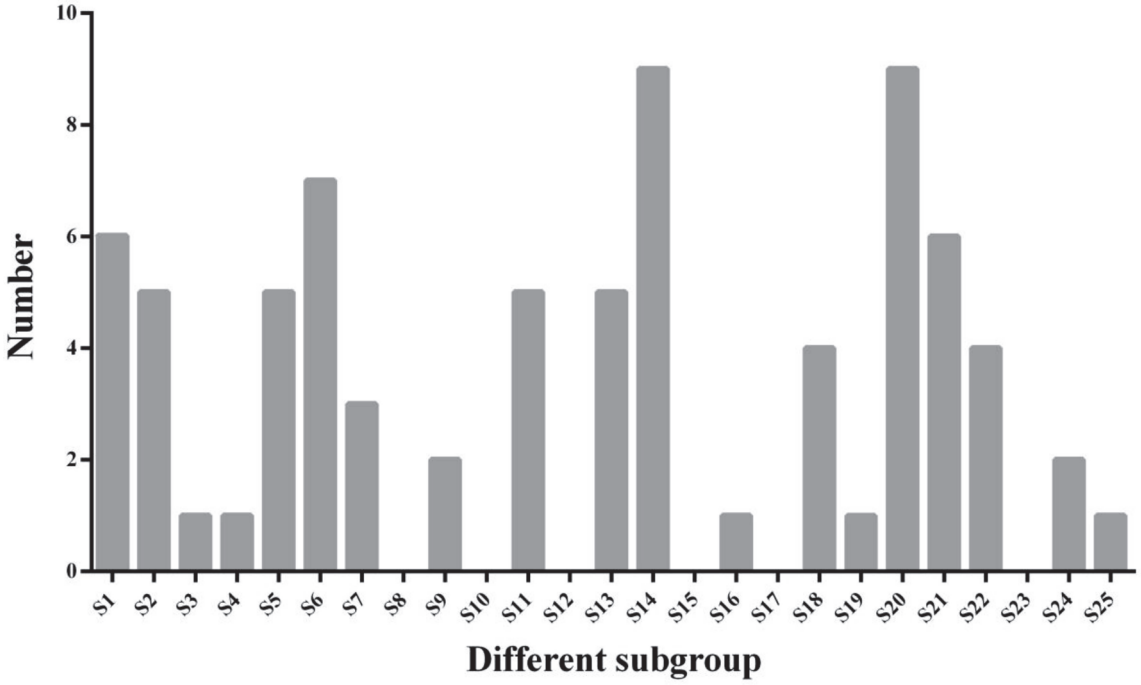


Figure 3. Distribution of carrot R2R3-MYB TFs subfamilies. R2R3-MYB TFs subfamilies represented on the Abscissa. Ordinate represents the number of each subfamily member.

2.3. Conserved Motifs of R2R3-MYB Proteins Based on the Analysis of MEME

According to the evolutionary relationship of R2R3-MYB TFs in carrot and *A. thaliana*, R2R3-MYB TFs in carrot were divided into different subfamilies. The conserved domains of all class members were analyzed by online software MEME. Through online analysis, a total of 10 conserved motifs were obtained, and the distribution of conserved motifs in different subfamilies is shown in the Figure 4. Closely related genes had the same motif compositions, suggesting that there were functional similarities between MYB proteins. Motifs 7 only exists in the S6 subfamily. Motif 9 only exists in the S20 subfamily, and the S20 subfamily contained the most motifs. Reports on the regulatory mechanism of *DcMYBs* in different subfamilies of carrot are still limited. Based on the homology with *Arabidopsis thaliana*, we infer the possible functions of *DcMYBs* subfamilies. Subfamilies S1 and S4 may participate in the response to biotic and abiotic stresses [40,41]. Subfamily S2 may control the biosynthesis of proanthocyanidins (PAs) [42,43]. Subfamily S3 and subfamily S21 may be involved in cell wall biosynthesis [13,44]. Subfamily S6 controls anthocyanin biosynthesis in plant tissues [13]. Subfamily S7 may control flavonol biosynthesis [45]. Subfamily S13 may play a role in influencing lignin deposition and stomatal aperture [46]. Subfamily S22 was proposed to regulate lateral root formation [13] and subfamily S25 was proposed to play roles in embryogenesis [47].

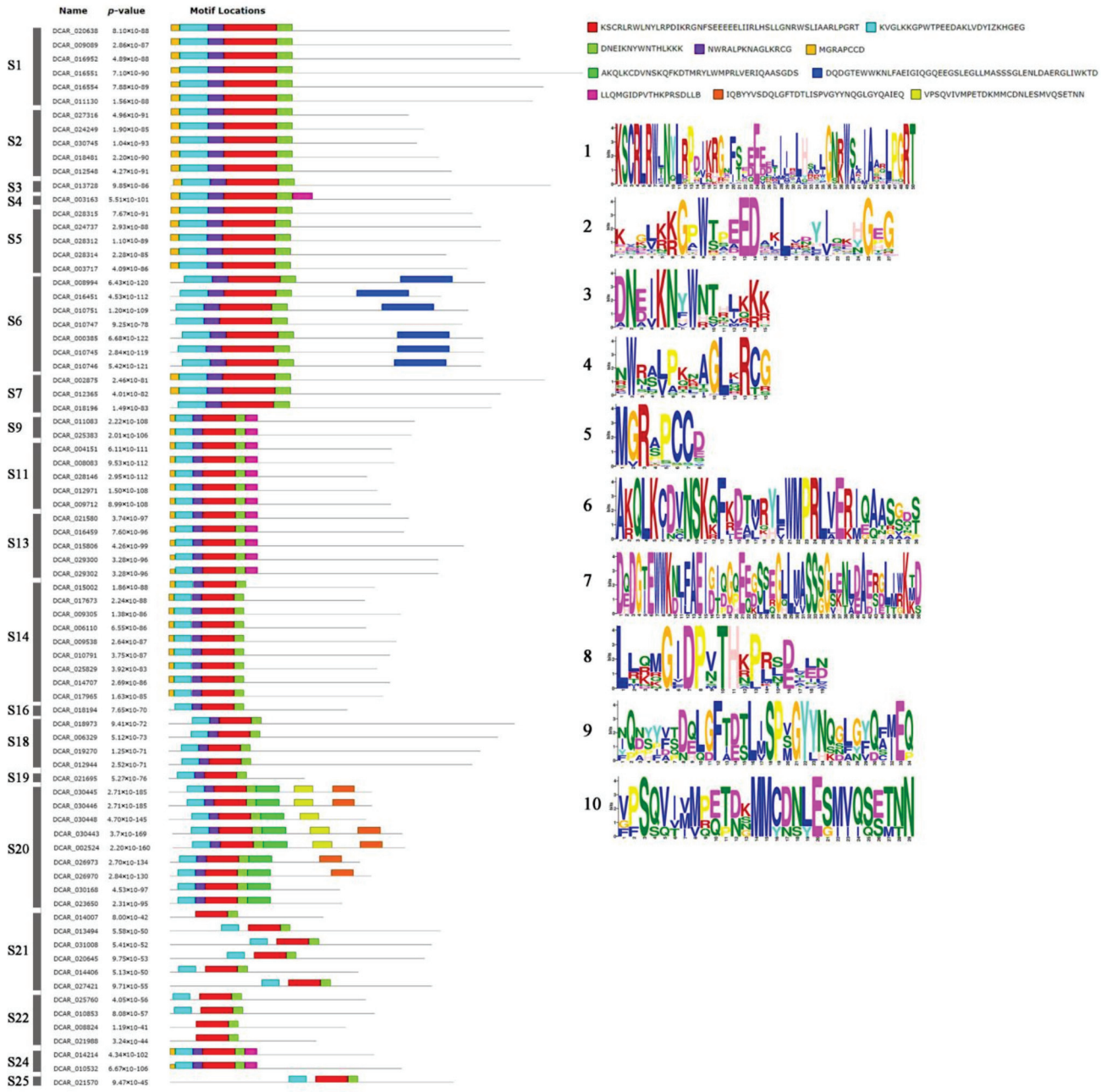


Figure 4. Sequence logos of conserved domains of R2R3-MYB proteins in carrot.

2.4. Selection Pressure in Carrot

The collinearity relationships of the *MYB* genes in carrot chromosomes are presented in Supplementary Tables S2 and S3. Positive selection promotes the evolution of animals and plants by accumulating favorable mutations, while purified selection promotes the evolution of animals and plants by eliminating harmful mutations. To determine which selection pressure promoted the evolution of R2R3-MYB gene in carrot, we used the CDS sequence of R2R3-MYB gene to calculate *Ka*, *Ks* value and *Ka*/*Ks* ratio (Figure 5 and Supplementary Table S2). The *Ks* value for a clear majority of R2R3-MYB paralog gene pairs in carrot is higher than *Ka*, and the *Ka*/*Ks* value of paralog gene pairs is mostly less than 1, indicating that the evolution of the R2R3-MYB gene in-species is mainly performed via purified selection. The *Ka*/*Ks* ratio is mostly concentrated in the range of 0.1–0.4, indicating strong purified selection.

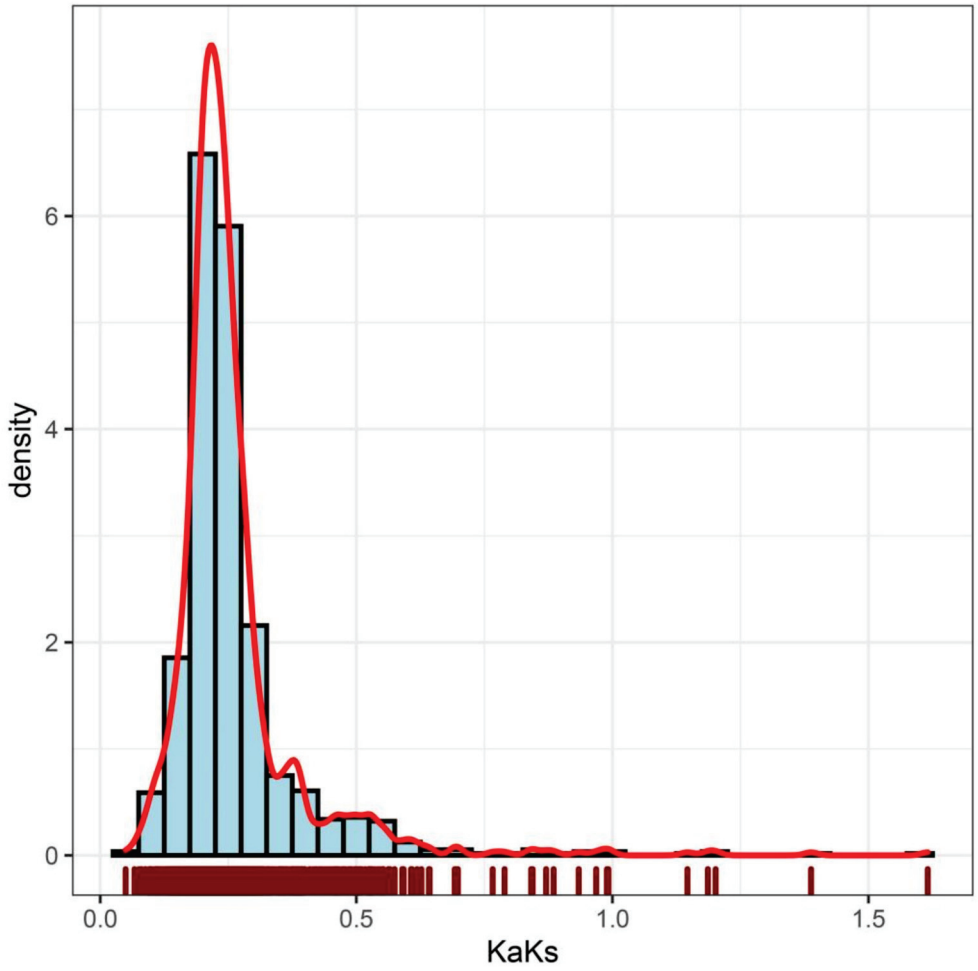


Figure 5. *Ka*/*Ks* density distribution in carrot.

2.5. Anthocyanin Content in Different Root Tissues of Nine Carrot Cultivars at Different Development Stages

Anthocyanins were distributed in different parts of carrot taproots, and it made the taproot appear purple and dark purple color (Figure 6). The anthocyanins of PPHZ taproots only accumulated in the rhizodermis, cortex, and secondary phloem, and the purple area increased with the development stage. Anthocyanins accumulated in the rhizodermis and cortex of CPP and ZL but did not exist in secondary phloem and secondary xylem (Figure 6). No purple or dark purple coloring was detected in KRD, MGH, SHBC taproots. Anthocyanins accumulated in rhizodermis, cortex, and secondary phloem and secondary xylem of DPP, BN and ZW carrot taproots in 70 and 110 days after sowing. Total anthocyanin content in different parts of six purple carrot cultivars taproots (DPP, BN, ZW, PPHZ, CPP, ZL) increased significantly with development process, and the anthocyanin-increased efficiency of taproot in solid purple carrots was higher than others (Figure 7). Among these six purple carrot cultivars, the anthocyanin accumulation of BN in the taproot was the highest at the two periods. At the 70 and 110 days after sowing, total anthocyanin accumulation in rhizodermis, cortex, and secondary phloem reached 71.00 and 128.47 mg/100 g fresh weight (fw) respectively, while that in the secondary xylem reached 25.78 and 38.04 mg/100 g fw respectively (Figure 7). No anthocyanin accumulation was detected in MGH, KRD and SHBC taproots. There was no significant difference in the anthocyanin accumulation between the rhizodermis and cortex of CPP and ZL, and the increase efficiency in the two development stages was lower. At the 70 days after sowing, the anthocyanin contents in the rhizodermis and cortex of CPP and ZL were 4.11 and 3.23

mg/100 g fw respectively, while 110-day-old rhizodermis and cortex' anthocyanin values were 4.50 and 5.04 mg/100 g fw, respectively.

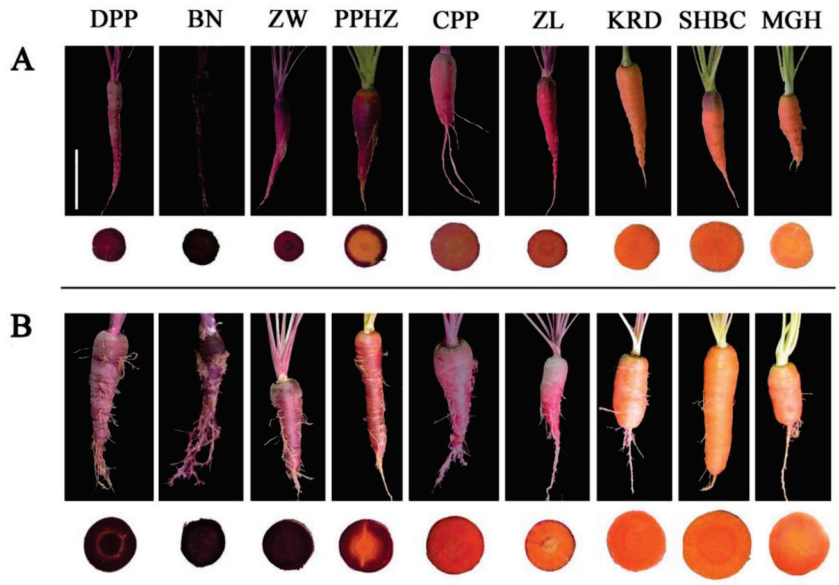


Figure 6. The color of cross-sections and the taproots of nine purple and non-purple carrots cultivars at two different stages. (**A**) represents the purple and nonpurple carrots grown for 70 days. (**B**) represents the purple and nonpurple carrots grown for 110 days. Cultivar abbreviations: DPP, Deep purple; BN, Black nebula; ZW, Ziwei; PPHZ, Purple haze; CPP, Cosmic purple; ZL, Zilong; KRD, Kurodagosun; SHBC, Sanhongbacun; MGH, Meiguihong. White bar represents 5 cm.

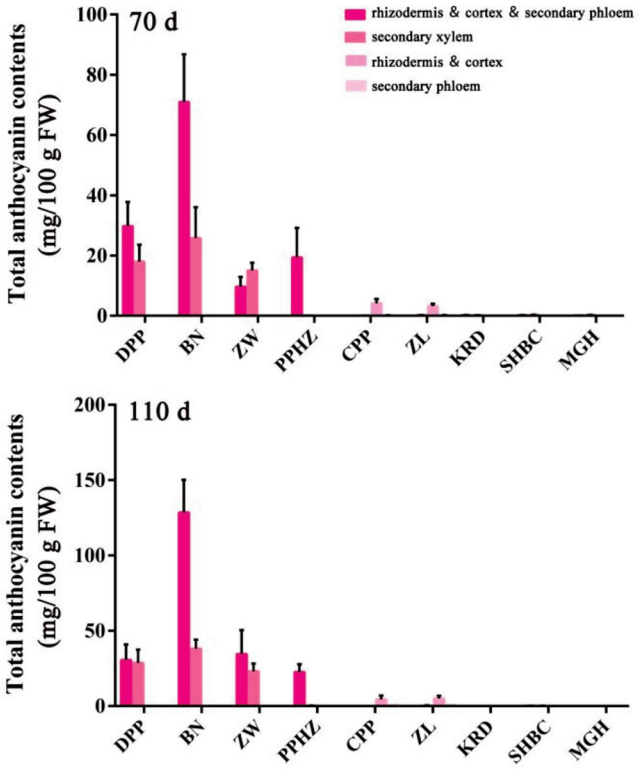


Figure 7. Total anthocyanin contents in the taproot of nine carrot cultivars at two. different stages. Error bars represent standard deviation (SD) of three replicates and are computed as cyanidin 3-O-galactoside equivalents.

2.6. Expression Profiles of DcMYB TFs in S6 Subfamily of Different Carrot Cultivars and Position of DcMYB Genes in S6 Subfamily on Chromosome

We selected 7 *DcMYB* genes in S6 subfamily and mapped their location onto chromosomes (Figure 8A). *DcMYB016* (*DCAR_016451*) was located on chromosome 5 and the rest on chromosome 3. *DcMYB113* (*DCAR_008994*) on chromosome 3 was far away from the other five transcription factors. Based on the carrot transcriptome data with different colors, transcripts of R2R3-MYB TFs in S6 subfamily were analyzed. The heat map relying on log2 (RPKM + 1) transformation value was constructed (Figure 8B). Yellow represents high expression level and blue represents low expression level. The results of transcriptome analysis showed that these *DcMYB* genes of S6 subfamily expression levels were different in carrot cultivars. *DcMYB7* (*DCAR_010745*) was highly expressed in purple tissues, but hardly expressed in non-purple carrot. *DcMYB6* (*DCAR_000385*) had the higher taproot mRNA levels in purple carrot cultivars than non-purple carrot. *DcMYB113* (*DCAR_008994*) was highly expressed in rhizodermis, cortex, and secondary phloem of PPHZ, but less expressed in other cultivars.

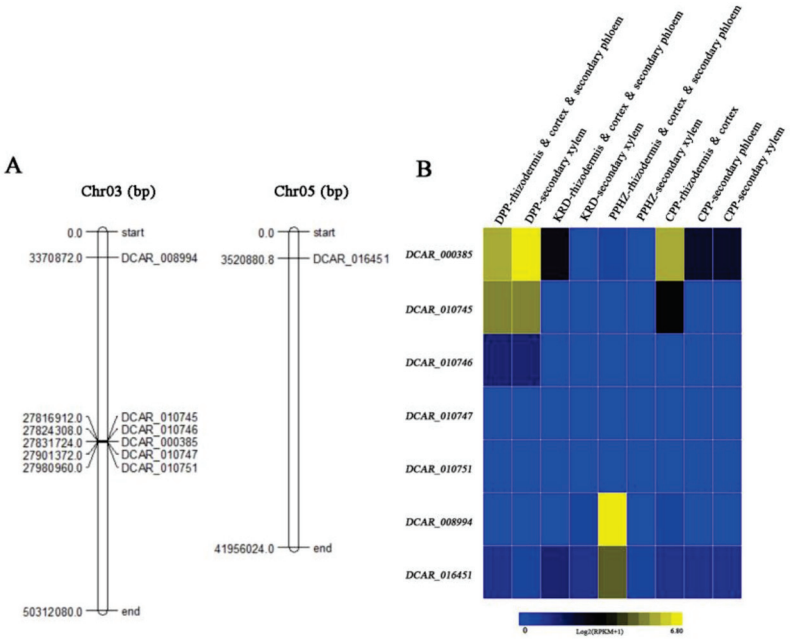


Figure 8. Site of MYB TFs in S6 subfamily on chromosome and transcript abundances of *DcMYB* genes in four different tissue parts of *D. carota* cvs. DPP, PPHZ, CPP and KRD. (**A**) represents the site of MYB TFs. Heat map was constructed based on the log2 (RPKM + 1) transformation values (**B**), and bar notes represent different expression levels. Yellow represents high expression level, and blue represents little expression level.

2.7. Relative Transcript Levels of S6 Subfamily DcMYBs in Different Root Tissues of Nine Carrot Cultivars

To better analyze the response of S6 subfamily members in root tissue of carrot cultivars, we sampled different parts in taproots of 9 purple and non-purple carrot cultivars (Figure 9). For RT-qPCR assay, 7 *DcMYB* genes (*DcMYB6* (*DCAR_000385*), *DcMYB7* (*DCAR_010745*), *DcMYB8* (*DCAR_010746*), *DcMYB9* (*DCAR_010747*), *DcMYB11* (*DCAR_010751*), *DcMYB113* (*DCAR_008994*) and *DcMYB016* (*DCAR_016451*)) from S6 subfamily were selected to determine the expression profiles of different root tissue during two development stages. At 70 days after sowing, the expression levels of *DcMYB6* (*DCAR_000385*) and *DcMYB7* (*DCAR_010745*) were higher in all purple carrot root tissue, and the transcripts of secondary xylem were lower than those of rhizodermis, cortex, and secondary phloem tissues. In ZL and CPP, the two gene transcripts, *DcMYB6 (DCAR_000385)* and *DcMYB7 (DCAR_010745)*, were detected in the rhizodermis and cortex.

DcMYB113 (DCAR_008994) and *DcMYB016 (DCAR_016451)* had the highest expression in the rhizodermis, cortex, and secondary phloem of PPHZ, which was consistent with the gene expression at 110 days after sowing.

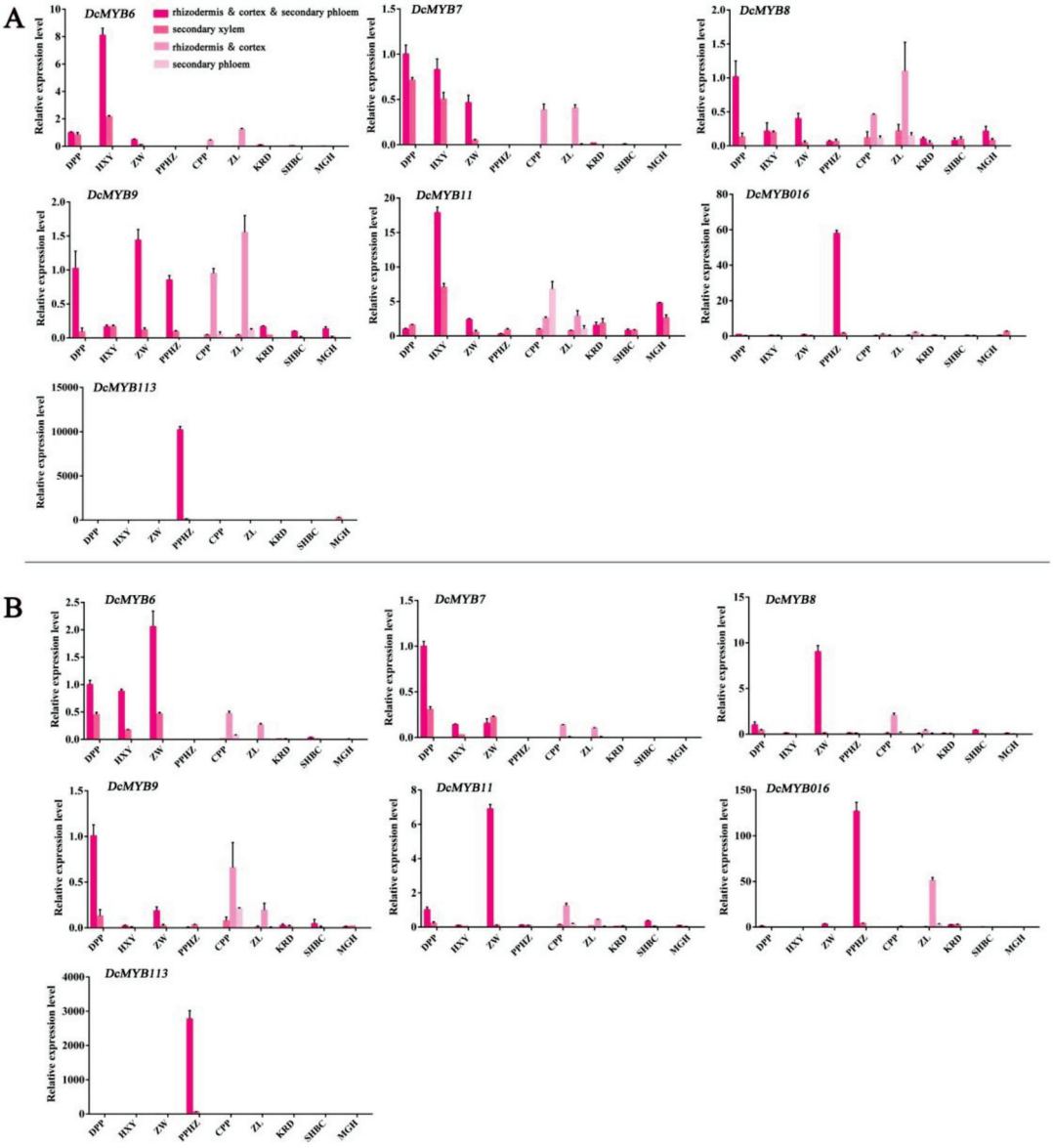


Figure 9. Expression profiles of S6 subfamily DcMYB genes (DcMYB6 (DCAR_000385), DcMYB7 (DCAR_010745), DcMYB8 (DCAR_010746), DcMYB9 (DCAR_010747), DcMYB11 (DCAR_010751), DcMYB016 (DCAR_016451) and DcMYB113 (DCAR_008994)) in different root tissues of carrot at two growth stages. (**A**) represents expression profiles of *DcMYB* genes in purple and nonpurple carrots at 70 days after sowing. (**B**) represents expression profiles of *DcMYB* genes in purple and nonpurple carrots at 110 days after sowing. The relative gene expression was calculated with the $2^{-\Delta\Delta CT}$ method. Error bars represent standard deviation (SD) of three replicates.

3. Discussion

Carrot is an important vegetable crop of *Apiaceae* family. In addition, celery and water dropwort and coriander also belong to common *Apiaceae* plants [4,48,49]. In addition to having rich anthocyanins [7,11,50], carrot has many active compounds, including carotenoids, volatile oil, vitamins, dietary fiber [4,48] and other nutrients.

Anthocyanins are very important secondary metabolites in many plants with functions on providing antioxidant protection and scavenging free radical. Anthocyanins are one of

the most health-protective ingredients in the human diet. Eating fruits and vegetables rich in anthocyanins is conducive to improving human health, and also can effectively prevent cardiovascular diseases and improve blood sugar balance [30]. The high expression of *MYB* gene in fruits and vegetables plays an important role in anthocyanin biosynthesis of plants. There are few studies on the genome-wide classification of R2R3-MYB gene family and its transcription in different root tissues of carrot. RT-qPCR was used to analyze the gene expression of S6 subfamily in purple and non-purple carrot at different developmental stages, which is helpful to further understand the molecular regulation mechanism of DcMYB transcription factor on carrot root development. These R2R3-MYB gene transcript levels were normalized to *DcActin1* [51].

MYB gene is widely distributed in plants and constitutes one of the largest transcription factor families. It acts a significant role in color accumulation, growth metabolism and stress response of plant organs. Since the first MYB transcription factor C1 was found in maize, with the development of genome sequencing and bioinformatics, more and more *MYB*-functional genes have been studied in different species [20,52–55]. In this study, we identified 146 R2R3-MYB factors, analyzed their phylogeny with 126 R2R3-MYB genes from *A. thaliana*, and divided them into 19 subfamilies. According to homology alignment, DcMYB TFs with similar motifs were divided into the same branch. The sequence structure of R2R3-MYB protein in different subfamilies is usually connected to biological function, and R2R3-MYB protein in S4-S7 subfamilies is related to phenylalanine metabolism regulating plant metabolite synthesis [38,39,42,56]. Consistent with sequence alignment and phylogenetic tree analysis, MEME analysis showed that the conserved domains of R2R3-MYB transcription factors in the same subfamilies were similar, indicating that these conserved motifs were associated with the regulatory functions of different subfamilies [39]. The number of R2R3-MYB transcription factors in higher plants is higher than that in lower plants, which may be due to the replication of plant genome during evolution. Besides, some *DcMYBs* could not be mapped to any chromosome, which might be due to the quality of the carrot genome sequence or a high level of heterozygosity [25]. High levels of heterozygosity within an individual might be an indication of low sample quality. The *Ka*/*Ks* ratios of clear R2R3-MYB genes in carrot indicated that the R2R3-MYB gene family in *Apiaceae* has generally undergone purifying selection and highly conserved evolution. Parts of the CDS of *R2R3-MYBs* have undergone positive selection, indicating that new gene functions might have been acquired. Such genes are rapidly evolving genes recently, which may be of great significance to the evolution of species.

Tandem and segmental duplication events had been hypothesized as the leading driving mechanism for the expansion and chromosomal organization into clusters of MYB gene families [57]. R2R3-MYB family members had been commonly found in gene clusters in genomes of many plant species. The S6 cluster is related to anthocyanin accumulation in *A. thaliana* and is the main subfamily of anthocyanin accumulation in *A. thaliana*, which primarily includes AtMYB75, AtMYB90, AtMYB113 and AtMYB114 [16,58]. Similar to *Arabidopsis*, MYB members classified as S6-MYB type in Petunia and tomato regulated anthocyanin biosynthesis [38,39]. Phylogenetic tree analysis of R2R3-MYB transcription factors in carrot and *A. thaliana* showed that 7 MYB TFs and *A. thaliana* S6 subfamily clustered together. DcMYB proteins in the same branch are highly correlated and may have similar functions. In all subfamilies, there were 9 MYBs in the subfamilies S14 and S20 that contained the most MYBs. The S14 subfamily is related to the formation of meristem in *A. thaliana*, among them, three R2R3-MYBs (*RAX* genes) were designated as regulators of axillary meristem. The *Rax* gene controls the early stage in the initiation of axillary meristem [59]. S20 subfamily *AtMYB62* was involved in response to phosphate starvation [60]. There were no R2R3-MYB members found in the S8, S10, S12, S15, S17 and S23 R2R3-MYB subfamilies. These results may imply the special characteristics R2R3-MYB genes in carrot species.

The expression pattern of R2R3-MYB genes in different developmental stages of carrot was studied by transcriptome analysis [61]. A total of 7 R2R3-MYB genes related to

anthocyanin accumulation were tested. Significant changes in the expression levels of all of the members in S6 subfamily were detected. Similarly, these data agree with the previous analysis described above, suggesting that several genes in S6 subfamily had been confirmed to be participants in the anthocyanin biosynthesis of carrot. Research showed that *DcMYB6* (*DCAR_000385*) had high transcript levels in the anthocyanin-pigmented carrot tissues with purple color, and it was detectable in the non-purple carrot cultivars. *DcMYB6* (*DCAR_000385*) can induce anthocyanin biosynthesis of *A. thaliana* [9]. *DcMYB7* (*DCAR_010745*), designated as the *DcMYB113-like* gene in the carrot genome, was always associated with purple root pigmentation of all purple carrots [6,7,11,17,58,62,63]. This corresponds to our transcriptional and heatmap values. In addition, *DcMYB016* (*DCAR_016451*) and *DcMYB113* (*DCAR_008994*) in the rhizodermis, cortex, and secondary phloem of PPHZ were significantly higher than that in other root tissues, suggesting that they may be a key gene regulating anthocyanin synthesis in rhizodermis, cortex, and secondary phloem. Previous studies on anthocyanin accumulation of carrot R2R3-MYB TFs have been reported, but there is no systematic classification of R2R3-MYB TF family. Systematic classification and transcriptional studies will provide us valuable tools to explore its potential application in carrot molecular breeding.

4. Materials and Methods

4.1. Plant Materials and Growth Conditions

Three solid purple root types ('Deep purple' (DPP), 'Black nebula' (BN) and 'Ziwei' (ZW)), three partial purple root types ('Purple haze' (PPHZ), 'Cosmic purple' (CPP) and 'Zilong' (ZL)) and three non-purple carrot cultivars ('Kurodagosun' (KRD), 'Sanhongbacun' (SHBC) and 'Meiguihong' (MGH)) were used as experimental materials and planted in the artificial climate room of Nanjing Agricultural University (32°04′ N, 118°85′ E, Nanjing, China). Organic soil, vermiculite and perlite (2:2:1, v/v) were used as substrates. The artificial climate chamber was set at 25 °C for 16 h in the daytime and 16 °C for 8 h in the evening, and the light intensity was 320 μmol m^{-2} s^{-1}, and the relative humidity was set at 75%. After 70 and 110 days after sowing plant growth, different tissues (rhizodermis, cortex, secondary phloem and secondary xylem) of carrot taproot were sampled. One part was used to extract RNA, and the other part was used to determine anthocyanin content. Three independent biological replicates of each carrot plant sample were prepared.

4.2. Identification of R2R3-MYB TFs in Carrot

The sequences of R2R3-MYB TFs in carrot were extracted based on the carrot genome database [64] and plant genome database (https://phytozome.jgi.doe.gov/pz/portal.html) (accessed on 1 May 2021) [52]. A total of 126 R2R3-MYB TFs sequences from A. thaliana were downloaded from the website TAIR (http://www.arabidopsis.org/) (accessed on 1 May 2021) [65]. The MYB TFs of other species are derived from Plant Transcription Factor Database [19]. The obtained candidate sequences are identified through Hmmer and Pfam number (PF00249), and the parameters were the default value.

4.3. Constructions of Phylogenetic Tree, Evolutionary Analysis and Sequence Features Analysis

Homologous genes were further analyzed by BLASTp search. Multiple sequences of R2R3-MYB protein from *A. thaliana* and carrot were aligned by Cluster W [66]. Then a phylogenetic tree was constructed by neighbor-joining method (bootstrap value of 1000) with MEGA 7 [67]. Position of MYB TFs in S6 subfamily on chromosome was anchored to a physical map with Mapchart v2.32. The conserved domains of R2R3-MYB TFs were analyzed by MEME search. Evolutionary relationship was obtained via online software (http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi) (accessed on 1 May 2021). The species evolution diagram was drawn by online software (http://itol.embl.de/upload.cgi) (accessed on 1 May 2021). The heat map of *DcMYBs* candidate gene expression in carrot was established by HemI 1.0 software (http://hemi.biocuckoo.org/faq.php) (accessed on 1 May 2021) [68].

4.4. Collinearity and Non-Synonymous Substitution Rate (Ka)/Synonymous Substitution Rate (Ks) Analyses

To further understand the structural information of R2R3-MYB gene family in carrot family, we analyzed the chromosomal localization and collinearity of R2R3-MYB gene family in carrot. MCScanX software was used to analyze collinearity of carrot R2R3-MYB gene family. In addition, the homologous gene pairs of R2R3-MYB gene were extracted and visualized with TBtools software.

In order to identify the selection pressure on R2R3-MYB genes in the process of evolution, we calculated the *Ka* and *Ks*. The ratio of *Ka* and *Ks* can be used to measure the selection pressure. Firstly, the coding sequence and nucleotide sequence of R2R3-MYB gene of carrot were compared with ParaAT2.0 software [69], and then *Ka*, *Ks* and *Ka*/*Ks* values were calculated by *KaKs*_calculator2.0 software [70,71]. When $Ka/Ks > 1$, it indicates positive selection. When $Ka/Ks < 1$, it indicates purified selection. When $Ka/Ks = 1$, it indicates neutral evolution.

4.5. Anthocyanin Measurement

Total anthocyanins were extracted from different tissues of carrot root by methanol–HCl method [3,9]. The fresh weight of different tissues of carrot root was determined after sampling, and then ground to powder in liquid nitrogen. The milled sample was transferred to the extract (50 % methanol, 49.9 % double distilled water, 0.1 % HCl, V/V) for the whole night. The 200 μL supernatant was absorbed to determine the absorbance for quantitative analysis at 530 nm, 620 nm and 650 nm. Optical density of anthocyanin $OD_\lambda = (OD_{530} - OD_{620}) - 0.1 \times (OD_{650} - OD_{620})$. The total anthocyanin content was expressed by the weight of cyanidin 3-O-galactoside per 100 g plant fresh weight [33]. Three biological replicates and three technical replicates were set for the extraction and determination of anthocyanin.

4.6. RT-qPCR Analysis

A total of 7 *DcMYB* genes of S6 subfamily were used for RT-qPCR analysis (Table 1). The specific primers were designed by Primer Premier 6.0 and synthesized by Genscript Inc. (Nanjing, China). Real-time quantitative PCR was performed using the instructions of the SYBR Premix *Ex Taq* kit (TaKaRa, Dalian, China). Using a 20 μL system, each PCR reaction system contained 2.0 μL of cDNA, 0.4 μL of forward and reverse quantification primers, 10 μL of SYBR Premix *Ex Taq*, and 7.2 μL of ddH_2O. The relative gene transcript levels were normalized to *DcActin1* [51] and calculated by the formula $2^{-\Delta\Delta CT}$ [72]. Each PCR reaction were performed with three independent biological replicates.

Table 1. The primer sequences used for RT-qPCR of S6 subfamily *DcMYB* genes.

Gene ID	Forward Primer (5′→3′)	Reverse Primer (5′→3′)
DCAR_000385	AGTGGCATCCTCAAGTGGTTCA	CCCAAATGTCACTCCAGCAACT
DCAR_010745	AGCGGCAACGACATTAACAACA	TTCATCTGGTAAGGCGGTGGTT
DCAR_010746	GCAGCAGCAACATCAACAACGA	TTCATCTGGTAAGGCGGTGGTT
DCAR_010747	TGCCACTACTTGTACCGCTACC	TCCTCCACCACTCGTTTCCATC
DCAR_010751	AAGCCTGTTCCGCAGACCTTAA	GGACGCCACTTGAGGACACAT
DCAR_008994	AGTGGCACCTTGTTCCTCAGAG	GCTGGCAATGATGGCTTCTTGT
DCAR_016451	GTGGCACCTTGTTCCTCAGAGA	TCGTCATTCAGTGGCAGTGTTG
DcActin1	CGGTATTGTGTTGGACTCTGGTGAT	CAGCAAGGTCAAGACGGAGTATGG

4.7. Statistical Analysis

The column charts were drawn with Graphpad Prism 6.0 software. Bars represent the mean values of three biological replicates ± standard deviation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms231911859/s1. Table S1. The sequence information of R2R3-

MYB TFs in carrot and *A. thaliana* identified in this study. Table S2. Information of Ka/Ks of the MYB genes in carrot chromosomes. Table S3. Information of collinear MYB gene pair in carrot chromosomes.

Author Contributions: A.-S.X., Z.-S.X. and A.-Q.D. initiated and designed the research, A.-Q.D. and S.-S.T. performed the experiments, Y.-J.D. and A.-Q.D. analyzed the data, A.-Q.D. wrote the paper. A.-S.X. revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by National Natural Science Foundation of China (32072563), Priority Academic Program Development of Jiangsu Higher Education Institutions Project (PAPD).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the other data sets supporting the conclusions of this article are included within the article and its additional files.

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

References

1. Arscott, S.A.; Tanumihardjo, S.A. Carrots of many colors provide basic nutrition and bioavailable phytochemicals acting as a functional food. *Compr. Rev. Food Sci. Food Saf.* **2010**, *9*, 223–239. [CrossRef]
2. Kammerer, D.; Carle, R.; Schieber, A. Detection of peonidin and pelargonidin glycosides in black carrots (*Daucus carota* ssp. *sativus* var. *atrorubens* Alef.) by high-performance liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **2010**, *17*, 2407–2412. [CrossRef]
3. Xu, Z.S.; Huang, Y.; Wang, F.; Song, X.; Wang, G.L.; Xiong, A.S. Transcript profiling of structural genes involved in cyanidin-based anthocyanin biosynthesis between purple and non-purple carrot (*Daucus carota* L.) cultivars reveals distinct patterns. *BMC Plant Biol.* **2014**, *14*, 262. [CrossRef] [PubMed]
4. Que, F.; Hou, X.L.; Wang, G.L.; Xu, Z.S.; Xiong, A.S. Advances in research on the carrot, an important root vegetable in the Apiaceae family. *Hortic. Res.* **2019**, *6*, 69. [CrossRef] [PubMed]
5. Li, T.; Deng, Y.J.; Liu, J.X.; Duan, A.Q.; Liu, H.; Xiong, A.S. DcCCD4 catalyzes the degradation of α-carotene and β-carotene to affect carotenoid accumulation and taproot color in carrot. *Plant J.* **2021**, *108*, 1116–1130. [CrossRef]
6. Bannoud, F.; Ellison, S.; Paolinelli, M.; Horejsi, T.; Senalik, D.; Fanzone, M.; Iorizzo, M.; Simon, P.W.; Cavagnaro, P.F. Dissecting the genetic control of root and leaf tissue-specific anthocyanin pigmentation in carrot (*Daucus carota* L.). *Theor. Appl. Genet.* **2019**, *132*, 2485–2507.
7. Bannoud, F.; Carvajal, S.; Ellison, S.; Senalik, D.; Talquenca, S.G.; Iorizzo, M.; Simon, P.W.; Cavagnaro, P.F. Genetic and transcription profile analysis of tissue-specific anthocyanin pigmentation in carrot root ploem. *Genes* **2021**, *12*, 1464–1483. [CrossRef]
8. Butler, N. Purple is the new orange: Anthocyanin regulation coming together in carrot. *Plant Physiol.* **2019**, *181*, 12–13.
9. Xu, Z.S.; Yang, Q.Q.; Feng, K.; Xiong, A.S. Changing carrot color: Insertions in DcMYB7 alter the regulation of anthocyanin biosynthesis and modification. *Plant Physiol.* **2019**, *181*, 195–207. [CrossRef]
10. Kumar, M.; Dahuja, A.; Sachdev, A.; Kaur, C.; Varghese, E.; Saha, S.; Sairam, K.V.S.S. Valorisation of black carrot pomace: Microwave assisted extraction of bioactive phytoceuticals and antioxidant activity using Box-Behnken design. *J. Food Sci. Technol.* **2019**, *56*, 995–1007.
11. Iorizzo, M.; Curaba, J.; Pottorff, M.; Simon, P.W.; Cavagnaro, P.F. Carrot anthocyanin genetics: Status and perspectives to optimize applications for the food colorant industry. *Genes* **2020**, *11*, 906. [CrossRef] [PubMed]
12. Ogata, K.; Hojo, H.; Aimoto, S.; Nakai, T.; Nakamura, H.; Sarai, A.; Ishii, S.; Nishimura, Y. Solution structure of a DNA-binding unit of Myb: A helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6428–6432.
13. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **2010**, *15*, 573–581. [CrossRef] [PubMed]
14. Jin, H.; Martin, C. Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol. Biol.* **1999**, *41*, 577–585. [PubMed]
15. Ng, W.K.; Abeysinghe, J.K.; Kamali, M. Regulating the regulators: The control of transcription factors in plant defense signaling. *Int. J. Mol. Sci.* **1999**, *19*, 3737.
16. 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.* **2010**, *53*, 814–827.
17. Kodama, M.; Brinch-Pedersen, H.; Sharma, S.; Holme, I.B.; Joernsgaard, B.; Dzhanfezova, T.; Amby, D.B.; Vieira, F.G.; Liu, S.; Gilbert, M.T.P. Identification of transcription factor genes involved in anthocyanin biosynthesis in carrot (*Daucus carota* L.) using RNA-Seq. *BMC Genom.* **2018**, *19*, 811.

18. Peng, Y.Y.; Wang, K.L.; Cooney, J.M.; Wang, T.C.; Espley, R.V.; Allan, A.C. Differential regulation of the anthocyanin profile in purple kiwifruit (*Actinidia species*). *Hortic. Res.* **2019**, *6*, 3.
19. Jin, J.P.; Tian, F.; Yang, D.C.; Meng, Y.Q.; Kong, L.; Luo, J.C.; Cao, G. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* **2017**, *45*, D1040–D1045. [CrossRef]
20. Stracke, R.; Werber, M.; Weisshaar, B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [CrossRef]
21. He, Q.; Jones, D.C.; Li, W.; Xie, F.; Ma, J.; Sun, R.; Wang, Q.; Zhu, S.; Zhang, B. Genome-wide identification of R2R3-MYB Genes and expression analyses during abiotic stress in *Gossypium raimondii*. *Sci. Rep.* **2016**, *6*, 22980. [CrossRef] [PubMed]
22. Du, H.; Yang, S.S.; Liang, Z.; Feng, B.R.; Lei, L.; Huang, Y.B.; Tang, Y.X. Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biol.* **2012**, *12*, 106. [CrossRef] [PubMed]
23. Katiyar, A.; Smita, S.; Lenka, S.; Rajwanshi, R.; Chinnusamy, V.; Bansal, K. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genom.* **2012**, *13*, 544. [CrossRef] [PubMed]
24. Cao, Z.H.; Zhang, S.Z.; Wang, R.K.; Zhang, R.F.; Hao, Y.J. Genome wide analysis of the apple MYB transcription factor family allows the identification of *MdoMYB121* gene confering abiotic stress tolerance in plants. *PLoS ONE* **2013**, *8*, e69955. [CrossRef]
25. Cao, Y.P.; Han, Y.H.; Li, D.H.; Lin, Y.; Cai, Y.P. MYB Transcription Factors in Chinese Pear (*Pyrus bretschneideri* Rehd.): Genome-wide identification, classification, and expression profiling during fruit development. *Front Plant Sci.* **2016**, *7*, 577.
26. Chen, X.J.; Wang, P.J.; Gu, M.Y.; Lin, X.Y.; Hou, B.H.; Zheng, Y.C.; Sun, Y.; Jin, S.; Ye, N.X. R2R3-MYB transcription factor family in tea plant (*Camellia sinensis*): Genome-wide characterization, phylogeny, chromosome location, structure and expression patterns. *Genomics* **2021**, *113*, 1565–1578. [CrossRef]
27. Kähkönen, M.; Heinonen, M. Antioxidant activity of anthocyanins and their aglycons. *J. Agric. Food Chem.* **2003**, *51*, 628–633.
28. Tang, B.Y.; Li, L.; Hu, Z.L.; Chen, Y.N.; Tan, T.T.; Jia, Y.H.; Xie, Q.L.; Chen, G.P. Anthocyanin accumulation and transcriptional regulation of anthocyanin biosynthesis in purple pepper. *J. Agric. Food Chem.* **2000**, *68*, 845–880.
29. Netzel, M.; Netzel, G.; Kammerer, D.R.; Schieber, A.; Carle, R.; Simons, L.; Bitsch, I.; Bitsch, R.; Konczak, I. Cancer cell antiproliferation activity and metabolism of black carrot anthocyanins. *Innov. Food Sci. Emerg. Technol.* **2007**, *8*, 365–372. [CrossRef]
30. He, J.; Giusti, M.M. Anthocyanins: Natural colorants with health-promoting properties. *Nnu. Rev. Food Sci. Technol.* **2010**, *1*, 163–187. [CrossRef]
31. Ahmad, T.; Cawood, M.; Iqbal, Q.; Ario, A.; Akhtar, S. Phytochemicals in Daucus carota and their health benefits-review article. *Foods* **2019**, *8*, 424. [CrossRef] [PubMed]
32. Ahmed, N.U.; Park, J.I.; Jung, H.J.; Hur, Y.K.; Nou, I.S. Anthocyanin biosynthesis for cold and freezing stress tolerance and desirable color in *Brassica rapa*. *Funct. Integr. Genom.* **2015**, *15*, 383–394. [CrossRef] [PubMed]
33. Feng, K.; Liu, J.X.; Duan, A.Q.; Li, T.; Yang, Q.Q.; Xu, Z.S.; Xiong, A.S. AgMYB2 transcription factor is involved in the regulation of anthocyanin biosynthesis in purple celery (*Apium graveolens* L.). *Planta* **2018**, *248*, 1249–1261. [CrossRef]
34. Li, M.Y.; Feng, K.; Hou, X.L.; Jiang, Q.; Xu, Z.S.; Wang, G.L.; Liu, J.X.; Wang, F.; Xiong, A.S. The genome sequence of celery (*Apium graveolens* L.), an important leaf vegetable crop rich in apigenin in the Apiaceae family. *Hortic. Res.* **2020**, *7*, 9. [CrossRef] [PubMed]
35. Petroni, K.; Tonelli, C. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* **2011**, *181*, 219–229. [CrossRef] [PubMed]
36. Song, X.M.; Sun, P.C.; Yuan, J.Q.; Gong, K.; Li, N.; Meng, F.B.; Zhang, Z.K.; Li, X.Y.; Hu, J.J.; Wang, J.P.; et al. The celery genome sequence reveals sequential paleo-polyploidizations, karyotype evolution and resistance gene reduction in apiales. *Plant Bio. Technol. J.* **2021**, *19*, 731–744. [CrossRef]
37. Iorizzo, M.; Ellison, S.; Senalik, D.; Zeng, P.; Satapoomin, P.; Huang, J.; Bowman, M.; Iovene, M.; Sanseverino, W.; Cavagnaro, P. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet.* **2016**, *48*, 657–666. [CrossRef]
38. Albert, N.W.; Lewis, D.H.; Zhang, H.B.; Schwinn, K.E.; Jameson, P.E.; Davies, K.M. Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant J.* **2011**, *65*, 771–784. [CrossRef]
39. Subban, P.; Prakash, S.; Mann, A.B.; Kutsher, Y.; Evenor, D.; Levin, I.; Reuveni, M. Functional analysis of MYB alleles from Solanum chilense and Solanum lycopersicum in controlling anthocyanin levels in heterologous tobacco plants. *Physiol. Plant* **2021**, *172*, 1630–1640. [CrossRef]
40. Seo, P.J.; Xiang, F.; Qiao, M.; Park, J.Y.; Lee, Y.N.; Kim, S.G.; Lee, Y.H.; Park, W.J.; Park, C.M. The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in *Arabidopsis*. *Plant Physiol.* **2009**, *151*, 275–289. [CrossRef]
41. Li, X.M.; Zhong, M.; Qu, L.; Yang, J.X.; Liu, X.Q.; Zhao, Q.; Liu, X.M.; Zhao, X.Y. *AtMYB32* regulates the ABA response by targeting ABI3, ABI4 and ABI5 and the drought response by targeting CBF4 in *Arabidopsis*. *Plant Sci.* **2021**, *310*, 110983. [CrossRef]
42. Lepiniec, L.; Debeaujon, I.; Routaboul, J.M.; Baudry, A.; Pourcel, L.; Nesi, N.; Caboche, M. Genetics and biochemistry of seed flavonoids. *Annu. Rev. Plant Biol.* **2006**, *57*, 405–430. [CrossRef] [PubMed]
43. Uematsu, C.; Katayama, H.; Makino, I.; Inagaki, A.; Arakawa, O.; Martin, C. Peace, a MYB-like transcription factor, regulates petal pigmentation in flowering peach 'Genpei' bearing variegated and fully pigmented flowers. *J. Exp. Bot.* **2014**, *65*, 1081–1094. [CrossRef] [PubMed]

44. Koshiba, T.; Yamamoto, N.; Tobimatsu, Y.; Yamamura, M.; Suzuki, S.; Hattori, T.; Mukai, M.; Noda, S.; Shibata, D.; Sakamoto, M.; et al. MYB-mediated upregulation of lignin biosynthesis in Oryza sativa towards biomass refinery. *Plant Biotechnol.* **2017**, *34*, 7–15. [CrossRef] [PubMed]
45. Pandey, A.; Misra, P.; Trivedi, P.K. Constitutive expression of *Arabidopsis* MYB transcription factor, *AtMYB11*, in tobacco modulates flavonoid biosynthesis in favor of flavonol accumulation. *Plant Cell Rep.* **2015**, *34*, 1515–1528. [CrossRef]
46. Kim, S.H.; Lam, P.Y.; Lee, M.H.; Jeon, H.S.; Tobimatsu, Y.; Park, O.K. The *Arabidopsis* R2R3 MYB transcription factor *MYB15* is a key regulator of lignin biosynthesis in effector-triggered immunity. *Front. Plant Sci.* **2020**, *11*, 583153. [CrossRef]
47. Zhang, Y.F.; Cao, G.Y.; Qu, L.J.; Gu, H.Y. Involvement of an R2R3-MYB transcription factor gene *AtMYB118* in embryogenesis in *Arabidopsis*. *Plant Cell Rep.* **2009**, *28*, 337–346. [CrossRef]
48. Li, M.Y.; Hou, X.L.; Wang, F.; Tan, G.F.; Xu, Z.S.; Xiong, A.S. Advances in the research of celery, an important Apiaceae vegetable crop. *Crit. Rev. Biotechnol.* **2018**, *38*, 172–183. [CrossRef]
49. Wang, X.J.; Luo, Q.; Li, T.; Meng, P.H.; Pu, Y.T.; Liu, J.X.; Zhang, J.; Liu, H.; Tan, G.F.; Xiong, A.S. Origin, evolution, breeding and omics of Apiaceae: A family of vegetables and medicinal plants. *Hortic. Res.* **2022**, *9*, uhac076. [CrossRef]
50. Iorizzo, M.; Cavagnaro, P.F.; Bostan, H.; Zhao, Y.; Zhang, J.; Yildiz, M.; Simon, P.W. A cluster of MYB transcription factors regulate anthocyanin biosynthesis in carrot (*Daucus carota* L.) root and leaf. *Front. Plant Sci.* **2019**, *9*, 1927. [CrossRef]
51. Wang, G.L.; Tian, C.; Jiang, Q.; Xu, Z.S.; Wang, F.; Xiong, A.S. Comparison of nine reference genes for real-time quantitative PCR in roots and leaves during five developmental stages in carrot (*Daucus carota* L.). *J. Hortic. Sci. Biotechnol.* **2016**, *91*, 264–270. [CrossRef]
52. Paz-Ares, J.; Ghosal, D.; Wienand, U.; Peterson, P.A.; Saedler, H. The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J.* **1987**, *6*, 3553–3558. [CrossRef] [PubMed]
53. Matus, J.T.; Aquea, F.; Arce-Johnson, P. Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and *Arabidopsis* genomes. *BMC Plant Biol.* **2008**, *8*, 83. [CrossRef] [PubMed]
54. Cavagnaro, P.F.; Chung, S.M.; Manin, S.; Yildiz, M.; Simon, P.W. Microsatellite isolation and marker development in carrot-genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. *BMC Genom.* **2011**, *12*, 386. [CrossRef]
55. Iorizzo, M.; Senalik, D.A.; Ellison, S.L.; Grzebelus, D.; Cavagnaro, P.F.; Allender, C.; Brunet, J.; Spooner, D.M.; Deynze, V.A.; Simon, P.W. Genetic structure and domestication of carrot (*Daucus carota* subsp. sativus) (Apiaceae). *Am. J. Bot.* **2013**, *100*, 930–938. [CrossRef] [PubMed]
56. Stracke, R.; Ishihara, H.; Huep, G.; Barsch, A.; Weisshaar, B. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* **2010**, *50*, 660–677. [CrossRef] [PubMed]
57. Feller, A.; Machemer, K.; Braun, E.L.; Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **2011**, *66*, 94–116. [CrossRef] [PubMed]
58. Borevitz, J.O.; Xia, Y.; Blount, J.; Dixon, R.A.; Lamb, C. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* **2000**, *12*, 2383–2393. [CrossRef]
59. Guo, D.S.; Zhang, J.Z.; Wang, X.L.; Han, X.; Wei, B.Y.; Wang, J.Q.; Li, B.X.; Yu, H.; Huang, Q.P.; Gu, H.Y. The WRKY transcription factor WRKY71/EXB1 controls shoot branching by transcriptionally regulating *RAX* Genes in *Arabidopsis*. *Plant Cell* **2015**, *27*, 3112. [CrossRef]
60. Ji, Q.; Wang, D.; Zhou, J.; Xu, Y.; Zhou, F. Genome-wide characterization and expression analyses of the MYB superfamily genes during developmental stages in Chinese jujube. *PeerJ* **2019**, *7*, e6353.
61. Xu, Z.S.; Feng, K.; Que, F.; Wang, F.; Xiong, A.S. A MYB transcription factor, *DcMYB6*, is involved in regulating anthocyanin biosynthesis in purple carrot taproots. *Sci. Rep.* **2017**, *7*, 45324. [CrossRef] [PubMed]
62. Cavagnaro, P.F.; Iorizzo, M.; Yildiz, M.; Senalik, D.; Parsons, J.; Ellison, S.; Simon, P.W. A gene-derived SNP-based high resolution linkage map of carrot including the location of QTL conditioning root and leaf anthocyanin pigmentation. *BMC Genom.* **2014**, *15*, 1118. [CrossRef] [PubMed]
63. Curaba, J.; Bostan, H.; Cavagnaro, P.F.; Senalik, D.; Mengist, M.F.; Zhao, Y.; Simon, P.W.; Iorizzo, M. Identification of an *SCPL* gene controlling anthocyanin acylation in carrot (*Daucus carota* L.) root. *Front. Plant Sci.* **2020**, *10*, 1770. [CrossRef] [PubMed]
64. Xu, Z.S.; Tan, H.W.; Wang, F.; Hou, X.L.; Xiong, A.S. CarrotDB: A genomic and transcriptomic database for carrot. *Database* **2014**, *2014*, bau096. [CrossRef] [PubMed]
65. Yon, R.S.; William, B.; Berardini, T.Z.; Chen, G.; David, D.; Aisling, D.; Margarita, G.H.; Eva, H.; Gabriel, L.; Mary, M. The *Arabidopsis* Information Resource (TAIR): A model organism database providing a centralized, curated gateway to *Arabidopsis* biology, research materials and community. *Nucleic Acids Res.* **2003**, *31*, 224–228.
66. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680. [CrossRef] [PubMed]

67. Sudhir, K.; Glen, S.; Koichiro, T. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874.
68. Deng, W.K.; Wang, Y.B.; Liu, Z.X.; Cheng, H.; Xue, Y. HemI: A toolkit for illustrating heatmaps. *PLoS ONE* **2014**, *9*, e111988. [CrossRef]
69. Zhang, Z.; Xiao, J.F.; Wu, J.Y.; Zhang, H.Y.; Liu, G.M.; Wang, X.M.; Lin, D. ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **2012**, *419*, 779–781. [CrossRef]
70. Nei, M.; Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **1986**, *3*, 418–426.
71. Wang, D.P.; Zhang, Y.B.; Zhang, Z.; Zhu, J.; Yu, J. *KaKs*_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [CrossRef]
72. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [CrossRef] [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Integrated Transcriptome and Metabolome Analyses Provide Insights into the Coloring Mechanism of Dark-red and Yellow Fruits in Chinese Cherry [*Cerasus pseudocerasus* (Lindl.) G. Don]

Yan Wang [1,2,†], Zhiyi Wang [1,2,†], Jing Zhang [1,2], Zhenshan Liu [1,2], Hao Wang [1,2], Hongxia Tu [1,2], Jingting Zhou [1,2], Xirui Luo [1], Qing Chen [1], Wen He [1,2], Shaofeng Yang [1,2], Mengyao Li [1], Yuanxiu Lin [1,2], Yunting Zhang [1,2], Yong Zhang [1], Ya Luo [1], Haoru Tang [1,2] and Xiaorong Wang [1,2,*]

[1] College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China
[2] Institute of Pomology and Olericulture, Sichuan Agricultural University, Chengdu 611130, China
* Correspondence: wangxr@sicau.edu.cn
† These authors contributed equally to this work.

Abstract: Chinese cherry [*Cerasus pseudocerasus* (Lindl.) G. Don] is an important fruit tree from China that has excellent ornamental, economic, and nutritional values with various colors. The dark-red or red coloration of fruit, an attractive trait for consumers, is determined by anthocyanin pigmentation. In this study, the coloring patterns during fruit development in dark-red and yellow Chinese cherry fruits were firstly illustrated by integrated transcriptome and widely-targeted metabolome analyses. Anthocyanin accumulation in dark-red fruits was significantly higher compared with yellow fruits from the color conversion period, being positively correlated to the color ratio. Based on transcriptome analysis, eight structural genes (*CpCHS*, *CpCHI*, *CpF3H*, *CpF3'H*, *CpDFR*, *CpANS*, *CpUFGT*, and *CpGST*) were significantly upregulated in dark-red fruits from the color conversion period, especially *CpANS*, *CpUFGT*, and *CpGST*. On contrary, the expression level of *CpLAR* were considerably higher in yellow fruits than in dark-red fruits, especially at the early stage. Eight regulatory genes (*CpMYB4*, *CpMYB10*, *CpMYB20*, *CpMYB306*, *bHLH1*, *CpNAC10*, *CpERF106*, and *CpbZIP4*) were also identified as determinants of fruit color in Chinese cherry. Liquid chromatography-tandem mass spectrometry identified 33 and 3 differential expressed metabolites related to anthocyanins and procyanidins between mature dark-red and yellow fruits. Cyanidin-3-*O*-rutinoside was the predominant anthocyanin compound in both fruits, while it was 6.23-fold higher in dark-red than in yellow fruits. More accumulated flavanol and procyanidin contents resulted in less anthocyanin content in flavonoid pathway in yellow fruits due to the higher expression level of *CpLAR*. These findings can help understand the coloring mechanism of dark-red and yellow fruits in Chinese cherry, and provide genetic basis for breeding new cultivars.

Keywords: *Cerasus pseudocerasus* (Lindl.) G. Don; transcriptome; metabolome; anthocyanin; candidate gene; transcription factor

1. Introduction

Anthocyanins and proanthocyanidins (PAs), known as flavonoids, belong to the group of the ubiquitous secondary metabolites. Anthocyanins are a group of important natural water-soluble pigments that commonly produce red/purple/blue colors to flowers and fruits of plants [1,2]. They have shown health-promoting properties, including antioxidant activity, cholesterol decomposition, visual acuity, and prevention of cardiovascular disease in humans [3]. In nature, anthocyanins existed as glycosides of polyhydroxy and polymethoxy derivatives mainly including cyanidin, pelargonidin, peonidin, delphinidin, malvidin, and petunidin [1,4]. The composition and proportion of anthocyanins determine

the coloration of plant tissues. PAs are essential taste factors affecting astringency and bitterness of fruits, which are also considered as important determinants of fruit quality [5].

Anthocyanins and PAs are synthesized by multiple enzyme-encoding structural genes via the flavonoid pathway. The phenylalanine forms anthocyanins, being catalyzed by phenylalanine ammonia-lyase (*PAL*), cinnamate 4-hydroxylase (*C4H*), 4-coumarate-CoA ligase (*4CL*), early biosynthesis genes (EBGs) (chalcone synthase (*CHS*), chalcone isomerase (*CHI*), flavanone 3-hydroxylase (*F3H*), flavonoid 3′-hydroxylase (*F3′H*), flavonoid 3′,5′-hydroxylase (*F3′5′H*)), and late biosynthesis genes (LBGs) (dihydroflavonol-4-reductase (*DFR*), anthocyanidin synthase/leucoanthocyanidin dioxygenase (*ANS*/*LDOX*), UDP-glucose: flavonoid-3-*O*-glucosyl-transferase (*UFGT*), and glutathione S-transferase (*GST*)) [1,6–9]. Leucoanthocyanidins and anthocyanins are further catalyzed by leucoanthocyanidin reductase (*LAR*) and anthocyanidin reductase (*ANR*) to form catechin and epicatechin, finally forming PAs [5]. In addition, all these structural genes are regulated by a MYB-bHLH-WD40 (MBW) complex at the transcriptional level [6,10–12].

Chinese cherry [*Cerasus pseudocerasus* (Lindl.) G.Don], belonging to the genus *Cerasus* of the family Rosaceae, is an economically important tetraploid fruiting cherry species [13,14]. It is native to China and has been widely cultivated across China as an important deciduous fruit with high economic and ornamental values [15]. Recently, cherry cultivation has been developing rapidly in China and has increasingly contributed to poverty alleviation and rural revitalization. Because of its diverse adaptability to various environments, Chinese cherry has not only been widely used as rootstock for cherry varieties, but it is also an excellent gene donor for intraspecific hybridization breeding program [16–18]. The fruits possess many valuable traits, including their various colors, unique taste, and abundant vitamins, fiber, minerals, and antioxidant compounds for healthy diets [15,19]. The majority of Chinese cherry germplasms have red, dark-red or orange-red fruit color, with a small number of accessions with black purple or yellow fruit color [20].

In cherry fruits, the difference between red and yellow color of the fruit peel and flesh is mainly dependent on the accumulation of anthocyanins. However, the component and content of specific anthocyanins vary with different cherry varieties [21]. Seven anthocyanins have been detected in four cherry species [22]. Cyanidin 3-rutinoside and cyanidin 3-glucosyl-rutinoside were the major components in both *C. pseudocerasus* and *C. vulgaris* [22], and cyanidin 3-rutinoside was predominant in *C. avium* [22–25]. Pelargonidin 3-*O*-glucoside, cyanidin 3-*O*-rutinoside, and pelargonidin 3-*O*-rutinoside were the chief compounds in the red *C. tomentosa* compared with the white fruit [26]. In *C. avium* and *C. tomentosa*, a large number of structural and regulatory genes involved in anthocyanin biosynthesis, transport and degradation pathway, have been identified by previous studies [26–29]. It is worth noting that the major anthocyanin component and the related genes/transcription factors (TFs) regulating anthocyanin biosynthesis varied from different cherry species. Therefore, it is necessary to explore the key candidate genes and potential molecular mechanism regulating anthocyanin levels of Chinese cherry fruit with different colors.

With the recent advancements in transcriptome and metabolome, the integrative analysis has provide an effective approach to illustrate the metabolic pathways and regulatory genes in fruit crops, such as sweet cherry [30], tomentosa cherry [26], strawberry [11], and longan [31]. In this study, integrative analysis of transcriptome and metabolome data were conducted to (i) analyze the color differences between dark-red and yellow Chinese cherry fruits, (ii) identify the differential expressed genes associated with anthocyanin biosynthesis during fruit development, and (iii) identify the key differential expressed metabolites related to anthocyanins and proanthocyanidins between mature dark-red and yellow fruits. These results provide insights for the identification of metabolites and candidate genes involved in anthocyanin formation and color change between dark-red and yellow Chinese cherry fruits, which lays a molecular foundation for color improvement and breeding program in the future.

2. Results

2.1. Color Phenotypic Characterization of Chinese Cherry during Fruit Development

Phenotypic observation at various developmental stages revealed significant differences in fruit peel color among the four Chinese cherry accessions. 'HP31', 'HF', and 'HP5' turned pink blush at S3 stage and the whole fruit (peel and flesh) color became dark-red, red, or light-red at S5 stage; whereas 'PZB' remained yellow throughout the corresponding developmental stage (Figure 1A). The a*, b*, and a*/b* ratio were also considered as indicators of fruit color (Figure 1B–D). The a*/b* ratio was significantly higher in dark-red fruits than in yellow fruits from the S3 stage, reaching the most significant differences at S5 stage compared with yellow fruits (Figure 1D). Consistent with the color change in the fruit, anthocyanin biosynthesis started at S2 stage, significant differences in anthocyanin content among the four accessions appeared at S3 stage, and anthocyanins accumulated in large quantities by S5 stage (Figure 1E). The total anthocyanin content was 23-fold higher in 'HP31' (75.96 mg/kg FW) than in 'PZB' (3.31 mg/kg FW) fruits at mature stage (S5), indicating that the anthocyanin accumulation was enhanced in the dark-red Chinese cherry fruits. Significant higher anthocyanin content was also detected in red fruits 'HF' and 'HP5' than that in 'PZB' (Figure 1E). The green fruits (S1) showed the highest total flavonoids content for the all four accessions, and decreased to at red stage (S4) for 'HP31' and increased to 6.08 mg/g at dark-red stage (S5) (Figure 1F). At mature stage, the highest and the lowest contents of flavonoids were detected in 'HP31' and 'HP5', respectively (Figure 1F).

2.2. Differentially Expressed Genes Analysis between Dark-red and Yellow Fruits

To identify the genes related to the fruit color formation between dark-red and yellow fruits, the fruits at S1-S5 stage for four accessions were subjected to RNA-seq. The raw transcriptome sequences from 60 samples have been submitted to the CNGB database under project number CNP0003682. After filtering the raw data, 2,800,020,582 clean reads were obtained, ranging from 41,582,094 to 48,104,774 per sample. The GC content was more than 46.54%, and the Q20 values ranged from 93.86% to 97.36%. The comparison rate exceeded 85.65% with the genome of Chinese cherry (unpublished data) as the reference genome (Supplementary Table S1).

We compared the transcriptome profiles of dark-red and yellow fruits to identify differentially expressed genes (DEGs) during fruit development. There were more DEGs between dark-red and yellow fruits than that during fruit development of the same accession (Supplementary Figure S1A). For each developmental stage between dark-red and yellow fruits, the down-regulated DEGs were more abundant than the up-regulated DEGs at S1, S2, and S4 stages, but the up-regulated DEGs were more abundant than the down-regulated DEGs at S3 and S5 stages (Supplementary Figure S1A). At total of 477, 591, 1146, 1682, and 1022 common DEGs were identified between dark-red and yellow fruits at S1-S5 stages, respectively (Supplementary Figure S1B). Within each accession, the gene expressions differences were the most significant from stage S2 to S4. A total of 102, 243, 578, and 115 common DEGs were identified in the S1 vs. S2, S2 vs. S3, S3 vs. S4, S4 vs. S5 comparison groups within dark-red fruits (Supplementary Figure S1C). A total of 1047, 2833, 2027, and 1936 DEGs were identified in the S1 vs. S2, S2 vs. S3, S3 vs. S4, S4 vs. S5 comparison groups in yellow fruits (Supplementary Figure S1A).

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis provided additional information about the enriched biological pathways, including "metabolism", "biosynthesis of other secondary metabolites", "starch and sucrose metabolism", "amino acid metabolism", and "flavonoid biosynthesis", and so on (Supplementary Figure S2). Based on the statistical significance criterion for multiple testing correlation (correlated p-value), "metabolism", "biosynthesis of other secondary metabolites", "flavonoid biosynthesis", and "transporters" were significantly enriched at S1-S5 stages between dark-red and yellow fruits (Supplementary Figure S3). Among them, structural genes such as *CHS*, *CHI*, *DFR*, *LAR*, and *ANR* were screened between dark-red and yellow fruits. During developmental stage S2-S4 within dark-red fruits, *C4H*, *CHS*, *CHI*, *F3'H*, *DFR*, and *UFGT* were enriched in

the flavonoid biosynthesis pathway (Supplementary Figure S4). In yellow fruits, *C4H*, *FLS*, *DFR*, *LAR*, *ANR*, and *UFGT* were enriched in S2 vs. S3 comparison, and *CHI*, *LAR*, and *UFGT* were identified in S3 vs. S4 comparison (Supplementary Figure S5). These results suggested obvious difference in anthocyanin biosynthesis between dark-red and yellow fruits, especially during color conversion period.

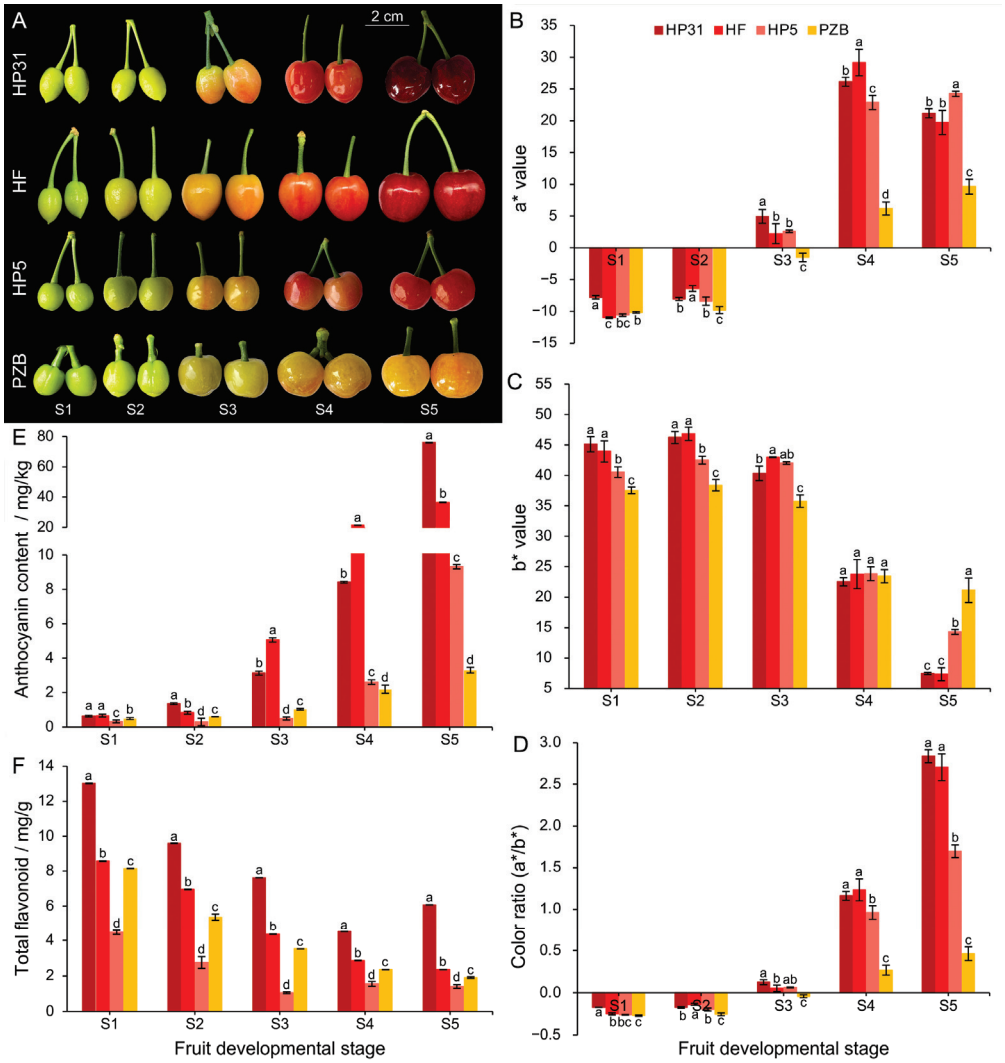


Figure 1. Color phenotypic characterization of Chinese cherry during fruit development. (**A**) Fruit phenotypes during development in dark-red and yellow fruits. (**B**) The a* values. (**C**) The b* values. (**D**) Color ratio (a*/b*). The a* coordinate represents the red (positive)-to-green (negative) scale, the b* coordinate represents the yellow (positive)-to-blue (negative) scale. Color ratio (a*/b*) represents the comprehensive color index [32]. (**E**) Total anthocyanin content. (**F**) Total flavonoids content. Error bars indicate ± standard deviation (SD) from three independent biological replicates. The lower case letters indicate significant difference at 0.05 level.

2.3. Weight Gene Co-Expression Network Association Analysis

The weighted co-expression network analysis (WGCNA) was conducted based on the normalized expression data for 46,058 genes from all 60 samples. After filtering, 11,515 genes were retained and classified into 12 distinct gene modules (Supplementary Figure S6). Module–trait relationship analysis revealed that six, two and six modules were significantly ($p < 0.01$) correlated with color ratio, anthocyanin content and flavonoid content, respectively (Figure 2A). Among these modules, the MEblack module of 80 genes were highly positively correlated with both color ratio ($r^2 = 0.88$, $p = 1 \times 10^{-20}$) and anthocyanin content

(r^2 = 0.70, $p = 6 \times 10^{-10}$) (Figure 2A). The KEGG pathway enrichment analysis revealed that the MEblack module genes were significantly enriched in "flavonoid biosynthesis" (ko00941) and "biosynthesis of other secondary metabolites" (ko09110) (Figure 2B). Gene ontology (GO) enrichment analysis showed that they were significantly enriched in molecular function and biological process category, mainly involved in "flavonoid metabolic process" (GO: 0009812) and "flavonoid biosynthetic process" (GO: 0009813) (Figure 2C).

The hierarchical clustering heatmap was constructed based on FPKM (fragments per kilobase per million fragments) values in each sample, which illustrated the expression patterns of 80 genes in MEblack module among four accessions at five developmental stages (Figure 2D). This module harbored most of structural genes involved in anthocyanin biosynthesis and transport pathway, such as *CHS*, *CHI*, *F3'H*, *DFR*, *ANS*, *UFGT*, and *GST*. Moreover, we annotated two reported MYB homologs involved in anthocyanin biosynthesis in other species. MSTRG.6844 and MSTRG.18742 encode TFs homologous to *MYB4*-like and *MYB10* from *Prunus avium*, and they are related to anthocyanin biosynthesis [29,33]. As an important part of the MBW transcription factor complex regulating the anthocyanin biosynthesis, a *bHLH1* gene (MSTRG.42352), homologous to bHLH from *Arabidopsis thaliana* [34] was also identified (Figure 2D). Hub genes were screened in the MEblack module (Figure 2E), including a protein detoxification 33 (MSTRG.7465), D-xylose-proton symporter-like 2 (MSTRG.15909), MSTRG.28943, ubiquitin-conjugating enzyme E2 22 (MSTRG.34925), probable E3 ubiquitin-protein ligase ARI2 (MSTRG.7464) (Supplementary Table S2).

2.4. Expression of Genes and TFs Related to Anthocyanins Biosynthesis

To further investigate the regulatory mechanism underlying anthocyanin accumulation in Chinese cherry fruit, we focused on the 80 MEblack module genes and other structural genes involved in the flavonoid biosynthesis pathway (Figure 3A, Supplementary Table S3). As for the expression levels, no significant differences were detected in the three genes (*PAL*, *C4H*, and *4CL*) involved in phenylpropanoid pathway between dark-red and yellow fruits. The EBGs (*CHS*, *CHI*, *F3H*, and *F3'H*) and LBGs (*DFR*, *ANS*, and *UFGT*) were significantly up-regulated in the dark-red fruits, especially from S3 to S5 stages (Figure 3A), consistent with the high anthocyanin content in dark-red fruits (Figure 1). GST genes was proposed to be involved in the anthocyanin transport, which showed 3.33~4.75-fold higher $\log_2$fold change values at mature stage in dark-red fruits, being the most significant DEG in the anthocyanin accumulation process (Supplementary Table S3). Interestingly, the expression of *LAR* was up-regulated in yellow fruits especially at the early stage (S1–S2), although their levels were relatively low compared with other genes (Supplementary Table S3). This might imply that more procyanidin content was accumulated in yellow fruits.

It is widely known that anthocyanin biosynthesis was primarily regulated by the MBW protein complex and other TFs. Two TFs, *MYB10* and *bHLH1*, selected from both DEGs and WGCNA, were significantly up-regulated in the dark-red fruits, especially at the later stages (Figure 3B). In addition to them, we also obtained 6 MYB, 5 bHLH, and 1 WD40 in Chinese cherry fruits (Figure 3B, Supplementary Table S3). According to transcriptional levels of candidate genes, 5 MYB and 3 bHLH were upregulated, while *MYB4*, *bHLH148*, *bHLH10*, and *WD40* were down-regulated in dark-red fruits. Consistent with previous studies [35–41], other TFs including 6 NAC, 5 MADS, 8 ERF, 9 WRKY, and 3 bZIP were also selected from the DEGs (Figure 3B). These TFs may also exert an effect on or participate in the regulation of structural and regulatory genes in anthocyanin biosynthesis.

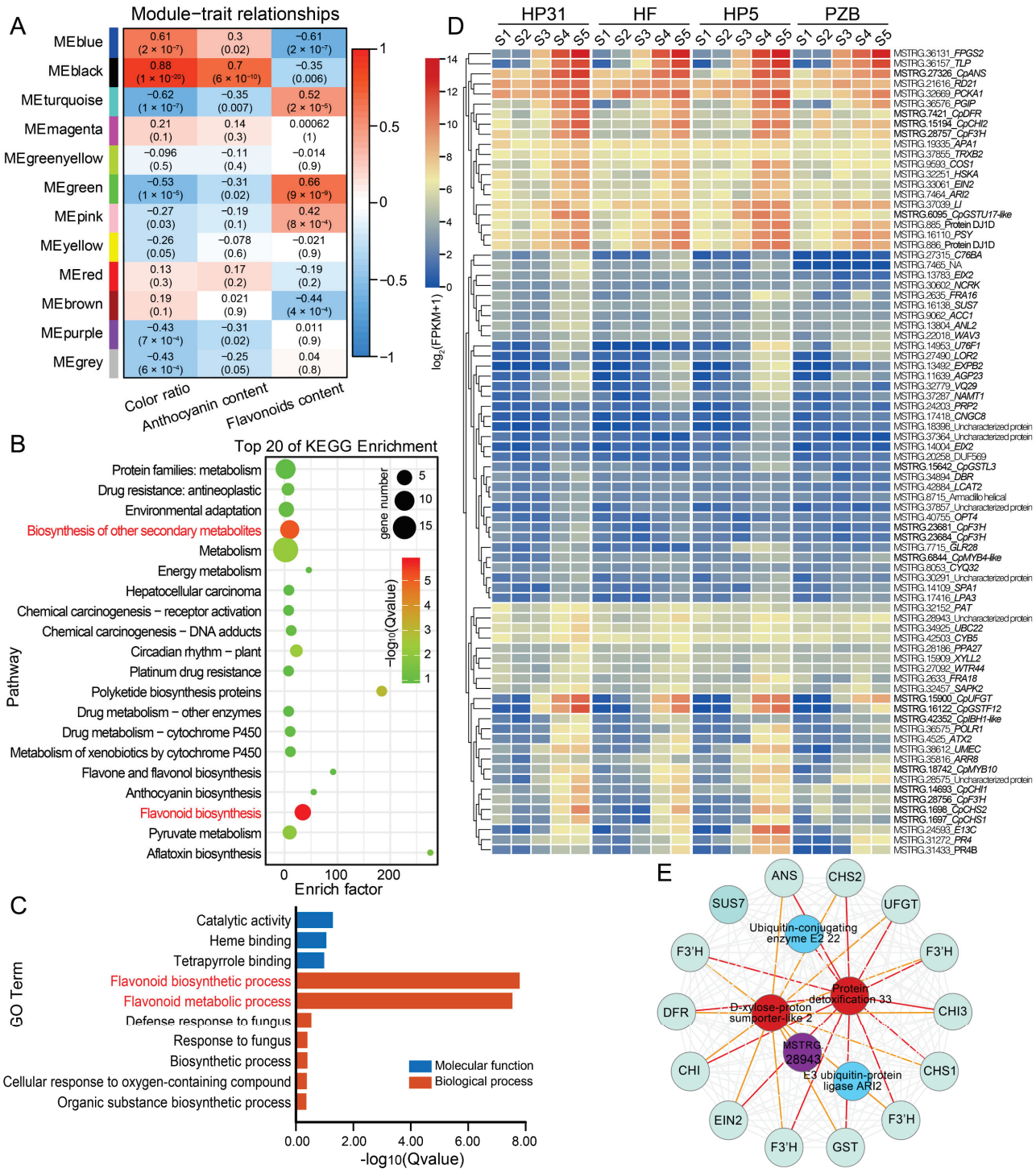


Figure 2. WGCNA recognized gene networks and key candidate genes associated with anthocyanin biosynthesis during fruit development of Chinese cherry. (**A**) Module–trait relationships based on Pearson correlations. The color key from blue to red represents r^2 values from −1 to 1. (**B**) KEGG enrichment analysis of with the top 20 KEGG pathways in the genes in the MEblack module. (**C**) GO enrichment analysis of genes in the MEblack module. (**D**) Heatmap of the expression level of genes in the MEblack module. The color scale of the heatmap represents expression levels as $\log_2$(FPKM + 1). (**E**) Genes whose expression was highly correlated in the MEblack module.

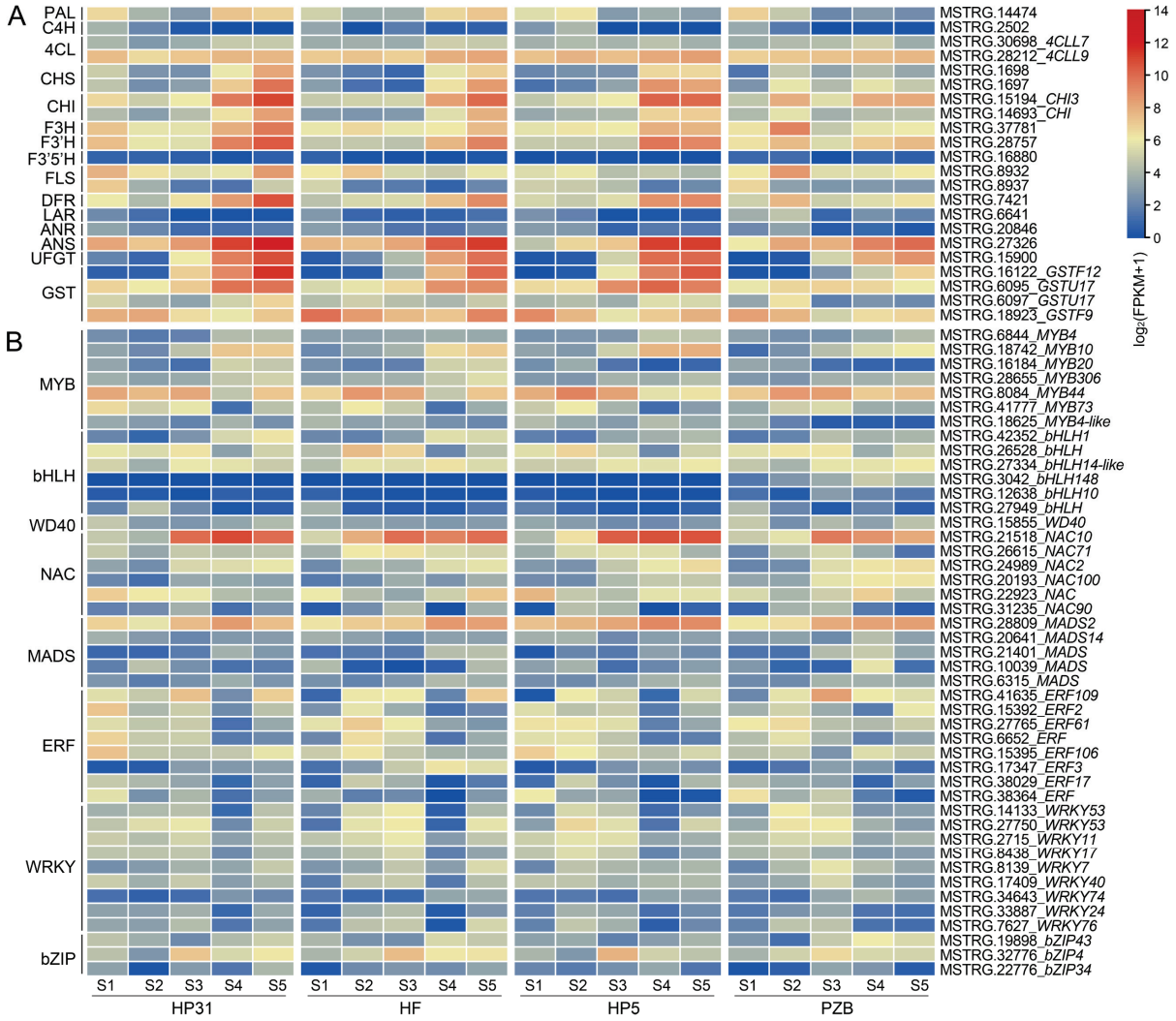


Figure 3. Expression heatmap of structural genes (**A**) and transcription factors (**B**) associated with anthocyanin biosynthesis. The heatmap represents the expression of corresponding genes in Chinese cherry fruits, and from blue to red in the heatmap indicates the expression levels of genes ranging from low to high. The color scale of the heatmap represents expression levels as $\log_2$(FPKM + 1).

2.5. RT-qPCR Validation

To verify the reliability of RNA-seq data, real-time PCR was performed to determine the expression levels of seventeen genes (Figure 4). During fruit development, their expression levels gradually upregulated from S1 to S2 stage and dramatically increased from S3 stage, then reached the highest at S5 stage in dark-red fruits. On the contrary, the expression levels slightly up-regulated from S1 to S2 stage, while significantly down-regulated at S3 stage and recovered slightly at S4 and S5 stages in yellow fruits. *CpGST* and *CpANS* showed the highest expression level at the later stages for dark-red fruits (Figure 4). At mature stage, the most significantly different genes were *CpGST*, *CpF3H*, and *CpUFGT*, with about 8.14-, 6.08-, and 5.50-fold $\log_2$fold change values in the 'PZB' vs. 'HP31' comparison, respectively. However, *CpLAR* revealed distinct expression profile, which was continuously decreasing during the fruit development in dark-red fruits, while it kept significant higher level in yellow fruits especially at the early stage (Figure 4). The TFs, *CpMYB10*, *CpMYB20*, *CpMYB306*, *CpbHLH1*, *CpNAC10*, and *CpERF106* revealed significantly higher expression levels in dark-red than in yellow fruits, especially at S4 and S5 stages (Figure 4). By contrast, the expression level of *CpMYB4* was significant lower in dark-red than in yellow fruits, suggesting its negative role in anthocyanin biosynthesis

(Figure 4). In addition, the expression levels of TFs in different tissues illustrated that these TFs revealed much higher expression levels in fruits and red flower bud than that in root, stem and leaf (Supplementary Figure S7). Finally, the correlation analyses exhibited significant correlation coefficients ranging from 0.7221 to 0.9964 (except for *CpF3H*, R^2 = 0.5378; *CpMYB4*, R^2 = 0.1187) between RNA-seq and RT-qPCR (Figure 4), supporting the accuracy of the transcriptome data.

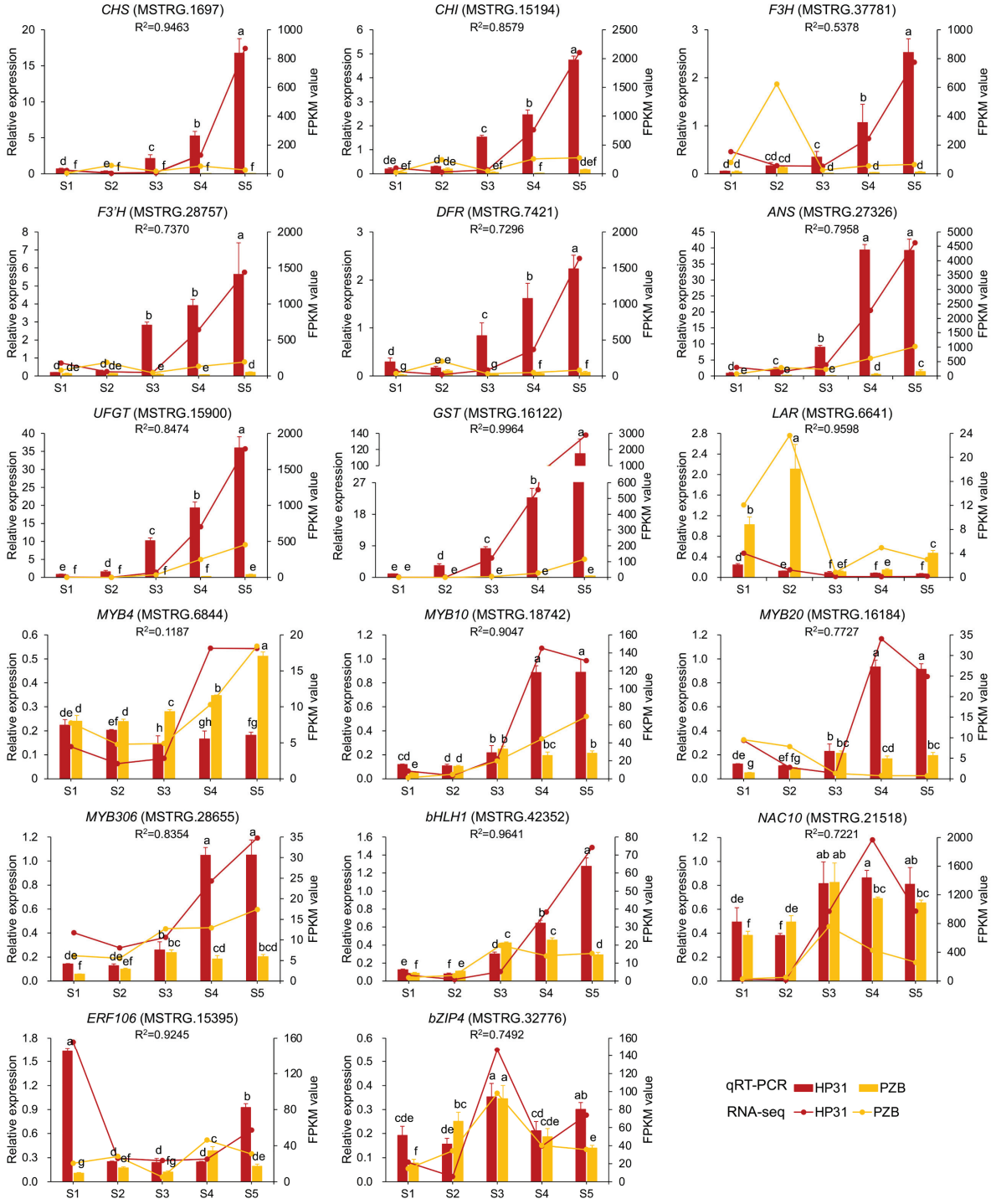


Figure 4. Expression patterns of candidate genes related to anthocyanin accumulation in Chinese cherry during fruit development. Error bars indicated ± SD from three independent replicates. The lower case letter indicated significant difference at 0.05 level for the relative expression level. The expression levels of these genes in 'HF' and 'HP5' were not shown in this figure.

2.6. Comparison of Metabolites between Dark-red and Yellow Fruits

To further confirm the key differential expressed metabolites (DEMs) in anthocyanin biosynthesis pathway between dark-red and yellow fruits, we obtained the metabolic profiling of 'HP31' and 'PZB' fruits at mature stage (S5) using liquid chromatography tandem mass spectrometry (LC-MS/MS). The raw data has been deposited to the MetaboLights database under accession number MTBLS6752. The principal component analysis (PCA) results revealed significant differences between them (Supplementary Figure S8). The two principal components, PC1 and PC2, were 44.35% and 11.50% in the fruits, respectively. Based on orthogonal partial least squares discriminant analysis (OPLS-DA), the Q^2Y value was 1, supporting the reliability of the metabolomic data (Figure 5A).

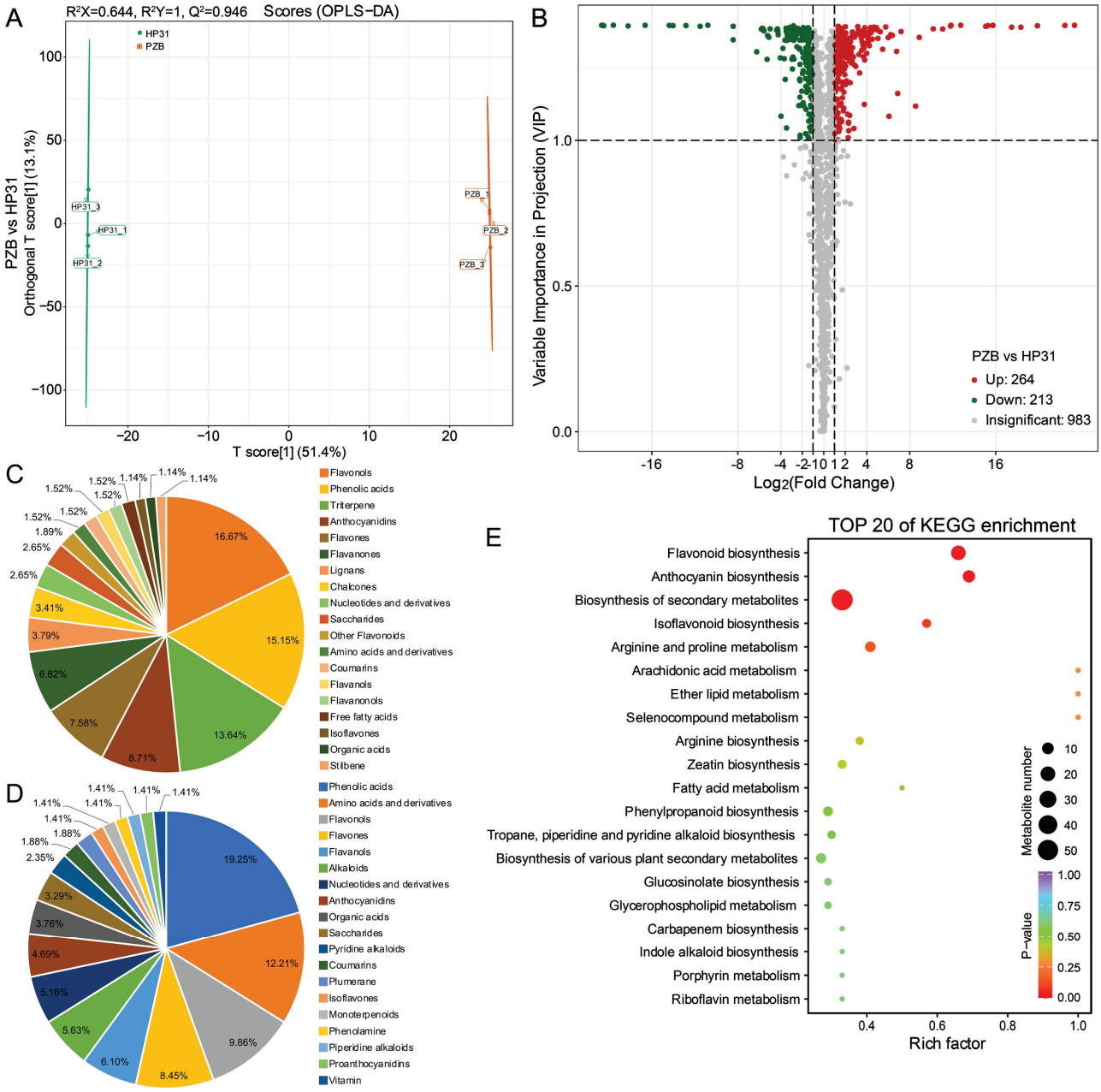


Figure 5. Overall analysis of the metabolomics data between dark-red and yellow fruits of Chinese cherry. (**A**) OPLS-DA. (**B**) Volcano plots of the metabolic profile in dark-red and yellow fruits. (**C**,**D**) Type and number of up-regulated and down-regulated metabolites. The TOP 19 types of DEMs were shown here. (**E**) TOP 20 of KEGG enrichment in the differential metabolites.

A total of 1460 metabolites in the fruits were identified by using UPLC-MS/MS (Supplementary Figure S8). The most abundant metabolites were flavonoids (380, 26.03%), followed by phenolic acids (243, 16.64%), others (142, 9.73%), amino acids and their derivatives (123, 8.42%), and terpenoids (113, 7.74%) (Supplementary Figure S9). Five anthocyanins and three flavonols existed only in dark-red fruits, and three anthocyanins and five flavonols were specifically detected in yellow fruits (Supplementary Table S4). Based on thresholds ($|\log_2$fold change$| \geq 1$, variable importance in projection (VIP) ≥ 1), we obtained 477 DEMs between dark-red and yellow fruits (Figure 5B). Among them, 264 metabolites were up-regulated and 213 metabolites were down-regulated in dark-red fruits. Notably, anthocyanidins, flavones, and flavanones were up-regulated, whereas flavanols were down-regulated in dark-red fruits (Figure 5C,D). The KEGG enrichment analysis revealed that "flavonoid biosynthesis" (19 metabolites, ko00941), "anthocyanin biosynthesis" (11 metabolites, ko00942), and "biosynthesis of secondary metabolites" (54 metabolites, ko01110) were significant (Figure 5E).

2.7. Identification and Comparison of Anthocyanin and Procyanidin Compounds

A total of 52 anthocyanin compounds and 11 procyanidin compounds were identified from dark-red and yellow fruits (Supplementary Table S5). Among them, 36 DEMs were identified, including 33 anthocyanins (15 cyanidins, 6 delphinidins, 5 pelargonidins, 3 peonidins, 2 malvidins, and 2 petunidins), and 3 procyanidins (Figure 6A). The dark-red fruits accumulated more anthocyanin than yellow fruits, being 6.50-fold greater accumulation (Figure 6B). On the contrary, yellow fruits showed 1.84 times of total procyanidin in comparison to dark-red fruits (Figure 6B), among which procyanidin B2 (46.9928% in dark-red, 58.0467% in yellow) and B3 (35.5550% in dark-red, 25.2852% in yellow) were the major procyanidin compounds in both of them (Figure 6C).

In both dark-red and yellow Chinese cherry, cyanidin-3-*O*-rutinoside was the highest accumulated anthocyanin, accounting for 81.7369% and 85.3393% of the total anthocyanins, respectively, while it was 6.23-fold higher in dark-red than in yellow fruit (Figure 6C, Supplementary Table S5). Cyanidin-3-*O*-(2″-*O*-glucosyl)glucoside, pelargonidin-3-*O*-rutinoside, and peonidin-3-*O*-rutinoside were also the major anthocyanin compounds (2.4323%~3.6232%) in dark-red fruits, while delphinidin-3-*O*-rutinoside-7-*O*-glucoside (4.8844%) and delphinidin-3-*O*-(6″-*O*-*p*-coumaroyl)glucoside (3.2979%) were the second and third largest compounds in yellow fruit (Figure 6C). In the comparison, the greatest different metabolite was pelargonidin-3-*O*-rutinoside, being 1.05×10^7-fold higher in dark-red than in yellow fruits (Figure 6C, Supplementary Table S5). The heatmap based on relative expression levels of the 36 DEMs showed that 23 anthocyanins were up-regulated, while 10 anthocyanins and 3 procyanidins were down-regulated in dark-red compared with yellow fruits (Figure 6D). Procyanidin B2 was 2.27 times in yellow fruits than that in dark-red fruits (Figure 6C). Therefore, it was obvious that the fruit color differences between dark-red and yellow fruits were not only determined by their total anthocyanin contents, but also by their anthocyanin components and percentages, especially cyanidin, pelargonidin, and peonidin derivatives, as well as the total procyanidin contents.

2.8. Comparison of Other Flavonoid Compounds in Flavonoid Pathway

To better understand the difference in flavonoid pathway between dark-red and yellow fruits, we further compared the type and number of other flavonoid compounds. The largest three sub-classes of flavonoids were flavonols (132, 34.74%), flavones (96, 25.26%), and flavanones (34, 8.95%) (Supplementary Table S5). In addition to anthocyanins, a total of 163 DEMs belonging to flavonoids were detected in the yellow vs. dark-red comparison, with 107 up-regulated and 56 down-regulated metabolites (Supplementary Figure S10). Among them, flavanones, flavanonols, and other flavonoids were all up-regulated, and the majority of DEMs from flavonols, flavones, and chalcones were up-regulated in dark-red fruits. Interestingly, 13 flavanol compounds were down-regulated among 17 differential

expressed flavanols in dark-red fruits, suggesting that the accumulation of flavanol content was much higher in yellow fruits than in dark-red fruits (Supplementary Figure S10).

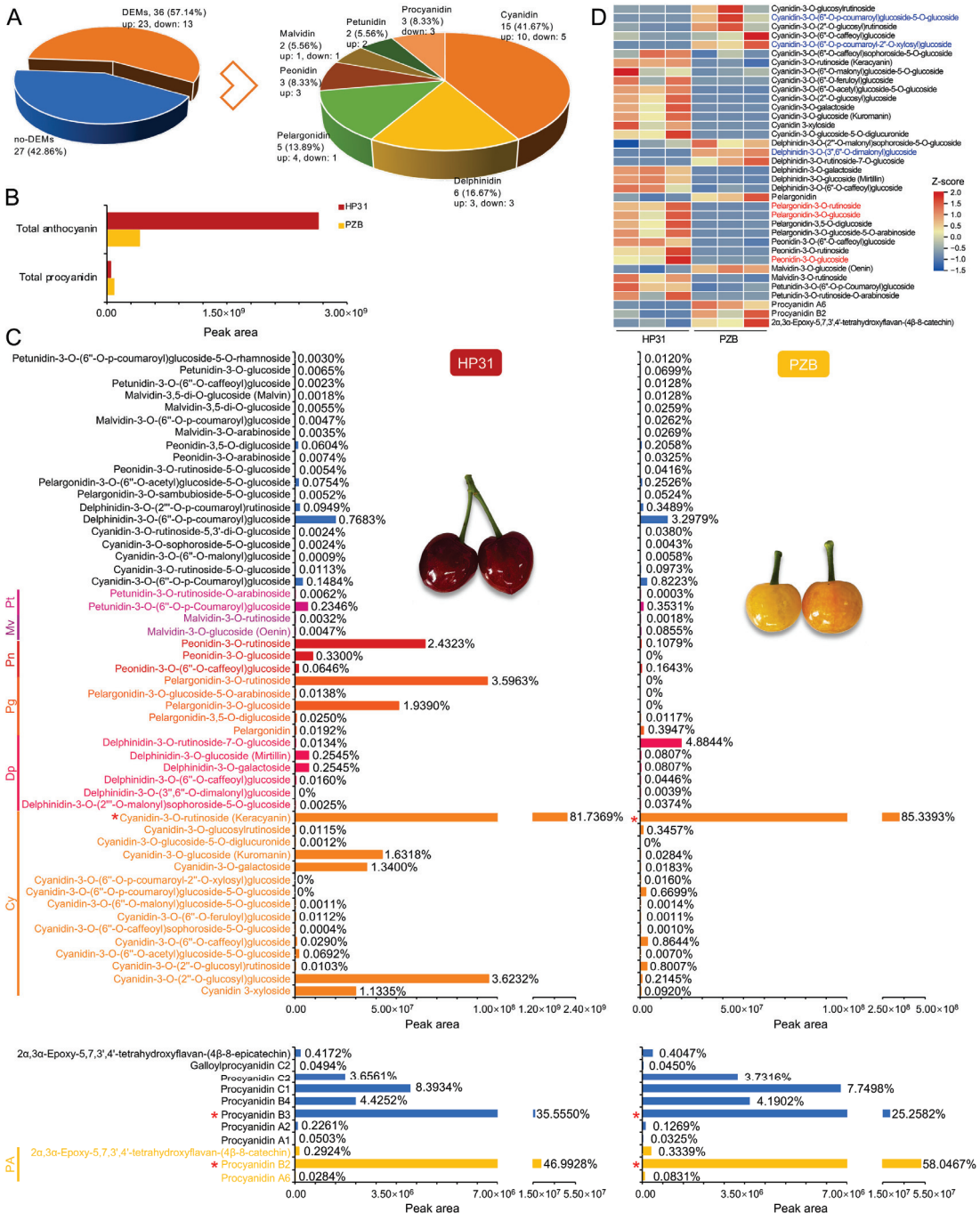


Figure 6. Comparison of anthocyanin and procyanidin compounds between dark-red and yellow fruits of Chinese cherry. (**A**) Type and number of differentially expressed metabolites. (**B**) The relative contents (peak area) of total anthocyanin and total procyanidin. (**C**) The relative contents and percentage of anthocyanin (upper) and procyanidin (lower) components in dark-red and yellow fruits of Chinese cherry. The blue font indicates the no-DEMs of procyanidins, and other colorful font indicates the 36 DEMs. Red asterisk indicates the predominant anthocyanin and procyanidin compounds. Abbreviations: Cy, cyanidin; Dp, delphinidin; Pg, pelargonidin; Pn, peonidin; Mv, malvidin; Pt, petunidin; PA, procyanidin. (**D**) Heatmap of the 36 DEMs. The color key from blue to red represents the relative content from −1.5 to 2.0. Red and blue font color indicates the top three up-regulated and down-regulated anthocyanins in the yellow vs. dark-red comparison.

3. Discussion

3.1. Comparison of Anthocyanin and Procyanidin Compounds in Dark-red and Yellow Chinese Cherry Fruits

The fruit color is largely dependent on anthocyanins classes and their concentrations. Generally, pelargonidin is reportedly as indicating an orange-red color, while cyanidin and peonidin indicate a deep red or purplish-red color [42]. Cyanidin and its glycoside derivatives have been reported as the primary anthocyanins in red-colored cherries [21,26,43]. Cyanidin 3-*O*-rutinoside and cyanidin 3-*O*-glucoside are the major anthocyanin components in sweet cherry [24,44,45] and sour cherry fruits [46,47]. Pelargonidin 3-*O*-glucoside, cyanidin 3-*O*-rutinoside, and pelargonidin 3-*O*-rutinoside were the three dominant anthocyanin compounds in red tomentosa cherry [22,26]. Cyanidin-3-*O*-glucoside was the most abundant anthocyanin in Chinese dwarf cherry, followed by pelargonidin-3-*O*-glucoside [48,49]. In Chinese cherry, four cyanidin-based anthocyanins were detected, and cyanidin 3-rutinoside and cyanidin 3-glucosyl-rutinoside were the two major compounds for red fruits [22]. In the present study, a total of 52 anthocyanins belonging to 6 types were firstly isolated in both dark-red and yellow fruits, although their contents of major types were much lower in yellow fruits (Figure 6B,C). Cyanidin-3-*O*-rutinoside was the largest anthocyanin compound in both of them (>80%), generally similar to previous report [22], but it was 6.23-fold higher in dark-red than in yellow fruits. Meanwhile, two pelargonidins, one peonidin, and one cyanidin were also up-regulated in dark-red fruits, although their proportions were relatively low within total anthocyanin content (Figure 6C). Two delphinidin derivatives were also the important compounds in yellow fruits, accounting for 4.88% and 3.30% (Figure 6C). These results indicated that the color difference between dark-red and yellow Chinese cherry was not only dependent on the total anthocyanin content, but also on the anthocyanin components and proportion. In addition, our results further supported that the predominant anthocyanin compounds were similar among Chinese cherry, sweet cherry and sour cherry [22–25,44,46], while it was obvious different from tomentosa cherry [26] and Chinese dwarf cherry [48,49]. This was generally consistent with their traditional taxonomy classifications, being assigned into subgenus *Cerasus* and *Microcerasus* of genus *Cerasus*, respectively [15].

Procyanidin B type was predominant in cherry fruits as described by previous reports [25,26,50,51]. Procyanidin B2 and B4 accumulated from large green stage, decreasing their accumulation as the fruits ripened in sweet cherry cultivar 'Lapins' [52]. Procyanidin B2 was the major procyanidin (about 95%) in both red and white tomentosa cherry, while no significant difference was detected in total procyanidins between them [26]. The most abundant compounds in Chinese dwarf cherry genotypes were procyanidin B1 (24.54~48.79%) and B2 (4.90~20.35%) [50]. In this study, procyanidin B2 was the largest compound among 11 detected procyanidins in Chinese cherry, accounting for about half, followed by procyanidin B3 and C1 (Figure 6C). All the relative contents of procyanidin compounds were higher in yellow than dark-red fruits, with three up-regulated compounds (Figure 6B,C). Therefore, the greater procyanidins accumulation was also responsible for the light color of yellow fruits, strongly supported by the higher content of flavanol (the precursor of procyanidin) in yellow fruits (Supplementary Figure S10).

3.2. Key Candidate Genes Involved in Anthocyanin Biosynthesis of Chinese Cherry

It has been widely reported that a series of structural genes, including *PAL*, *C4H*, *4CL*, *CHS*, *CHI*, *F3H*, *F3'H*, *DFR*, *ANS*/*LDOX*, and *UFGT*, co-regulated anthocyanin biosynthesis in many fruit crops. In the present study, a regulatory network of gene expression regulating anthocyanin biosynthesis and transport and the key differential expressed metabolites in Chinese cherry fruit was summarized, as shown in Figure 7. Combining the transcriptome data and RT-qPCR results, the up-regulation of structural genes (*CpCHS*, *CpCHI*, *CpF3H*, *CpF3'H*, *CpDFR*, *CpANS*, and *CpUFGT*) in (dark-red fruits enhanced flux in anthocyanin pathways. However, the low expression levels of these genes in yellow fruits generated a lack of stable anthocyanin synthesis. The DFR enzyme can selectively catalyze three kinds

of substrates to synthesize three specific products: leucodelphinidin, leucopelargonidin, and leucocyanidin. The expression of F3′H promoted the synthesis of dihydroquercetin through DFR and UFGT, further forming cyanidin-3-*O*-rutinoside. Both the cyanidin and pelargonidin contents of dark-red fruits are considerably higher than that of yellow fruits (Figure 6C). Thus, the high expression of *CpF3′H* and *CpDFR* determined the synthesis of specific anthocyanin component, which is consistent with the anthocyanin biosynthesis in grape [53]. Among these genes, *CpANS* showed the highest expression level, and *CpF3H* and *CpUFGT* exhibited the biggest difference (Figure 4). This suggested that the above seven genes are potential key genes regulating anthocyanin biosynthesis in Chinese cherry, especially *CpF3H*, *CpANS*, and *CpUFGT*. This was largely consistent with the results in many other Rosaceae fruit crops such as sweet cherry [27,28], tomentosa cherry [26], and peach [54].

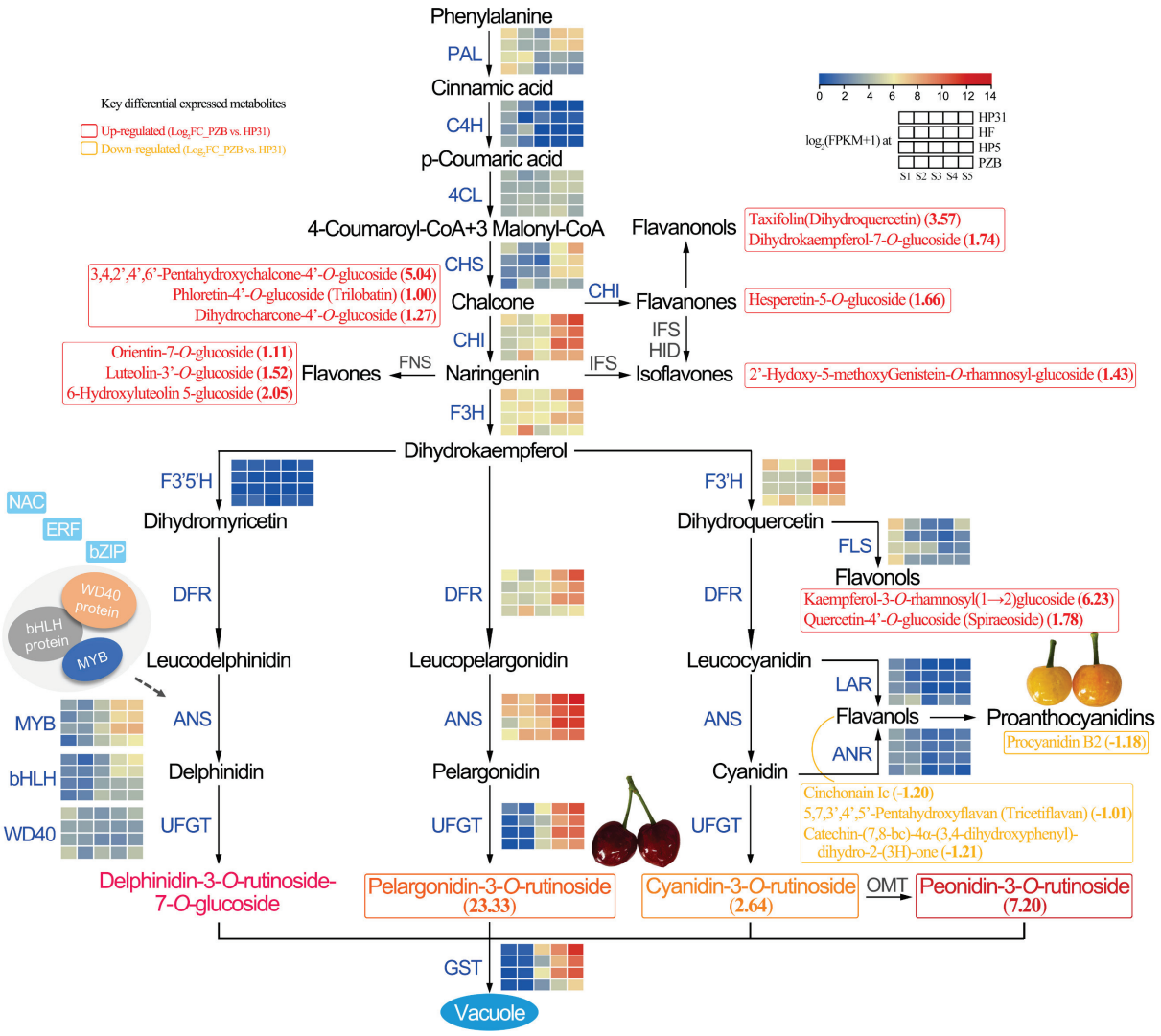


Figure 7. Regulatory network of anthocyanin biosynthesis in dark-red and yellow fruits of Chinese cherry. Grids represent the gene expression levels (log_2(FPKM + 1) values): S1, S2, S3, S4, and S5, left to right. The red and yellow font represents the key up-regulated and down-regulated metabolites (log_2fold change (PZB vs. HP31)). The black dotted line represents that the direct correlation was uncertain.

GSTs are known to participate in the anthocyanin transport and accumulation, while the absence of GSTs often results in an anthocyanin-less phenotype with reduced pigmen-

tation [7,8]. Here, four GST genes significantly increased in dark-red fruits compared to those in yellow fruits (Figure 3). Therefore, the upregulation of *CpGST* played a vital role in anthocyanin transport and resulted in the accumulation in vacuoles. Interestingly, the expression level of *CpGST* was about 8.91-fold greater at S5 than S4 stage in yellow fruits (Figure 4), which generated yellow with blush at mature stage (Figure 1A). This suggested that the anthocyanin biosynthesis pathway was fluent in yellow fruits, but the reduction in precursors finally resulted in the less accumulation of anthocyanins due to the lower expression levels of EBGs (*CpCHS*, *CpCHI*, *CpF3H*, and *CpF3'H*).

LAR and ANR are regarded as two key enzymes in procyanidin biosynthesis. As reported in *Medicago truncatula*, The loss of function of LAR in seed coats decreased the levels of procyanidins [55]. The higher expression level of *CpLAR* in yellow fruits (Figure 4) probably contribute to the more accumulated flavanols from the early stage, finally forming more procyanidins at mature stage (Figure 6B,C). It has been reported that competition existed between anthocyanins and flavonols biosynthesis in red apple fruits during fruit ripening [56], and apple flowers [57]. Higher amount flavonols were detected for all stages of white flower development than red flower in peach [58]. However, this study exhibited a different branch to flavanols and procyanidins rather than flavonols in bicolored (yellow) fruits in Rosaceae family.

3.3. Transcription Factors Involved in Anthocyanin Biosynthesis of Chinese Cherry

Anthocyanin metabolism is also regulated by a series of transcription factors, such as MBW (MYB-bHLH-WD40) protein complex, NAC, WRKY, ERF, and bZIP families. MYB TFs are reportedly associated with the regulation of anthocyanin biosynthesis and accumulation in sweet cherry [29], apple [7], and blackberry [59]. bHLH commonly interacts with MYB and WD40 to regulate anthocyanin biosynthesis jointly [60]. Based on to the expression levels, we identified seven MYB and six bHLH TFs (Figure 3B). Anthocyanin biosynthesis has been proven to be positively regulated by MYB TFs, such as *MYB10* [29,61,62], *MYB20* [63], and *MYB306* [64], by bounding to the promoter of structural genes or interacting with bHLH genes. *MYB4*, a repressor activated by *bHLH3*, prevents the formation of the MBW through competitive binding with *bHLH3* to inhibit the accumulation of anthocyanins by down-regulation of CHS, ANS, and DFR in mulberry [65], and bananas [66]. In addition, *bHLH148* and *bHLH10* were specifically expressed in yellow Chinese cherry (Figure 3B), suggesting their possible negative role in regulating anthocyanin biosynthesis. A WD40 protein, homologous to Arabidopsis TTG1 [67], was identified, which has also been characterized from apple [68] and strawberry [69]. The expression level of WD40 was significantly higher in yellow than dark-red fruits (Figure 3B), indicating its negative regulating role. Therefore, the anthocyanin biosynthesis in Chinese cherry fruit is regulated by the relevant MBW protein complex (Figure 7).

We also obtained two NAC, two WRKY, one ERF, and two bZIP that were significantly upregulated in dark-red fruits. These TFs have been proven to directly or indirectly regulate anthocyanin biosynthesis by binding to an MYB promoter or through protein-protein interactions in other fruits [35,37–39]. For example, *MdNAC52* can bind to the promoters of *MdMYB9* and *MdMYB11* to increase the anthocyanin content by regulating *MdLAR* in apple [70]. The interaction of *PyWRKY26* and *PybHLH3* could co-target the *PyMYB114* promoter, which resulted in anthocyanin accumulation in red-skinned pear [41]. *MdbZIP44* enhances *MdMYB1* binding to downstream target gene promoters to promote anthocyanin biosynthesis in apple [39]. Overall, eight regulatory genes were confirmed as determinants of fruit color in Chinese cherry. Our findings can enrich the key candidate genes and metabolites involved in the anthocyanins biosynthesis in Chinese cherry fruits, which are of great importance for molecular marker-assisted breeding.

4. Materials and Methods

4.1. Plant Materials

A total of four Chinese cherry accessions including 'HP31' (dark-red), 'HF' (red), 'HP5' (light-red) and 'PZB' (yellow) were grown under field conditions at the cherry germplasm resources of Sichuan Province (Chengdu City), China. Full flowering was set at 0 days after full bloom (DAFB) when 50% of flowers were open in the trees. Fruit samples were collected based on fruit phenology (Table 1). For 'HP31', the sampling was: green (S1), 37 DAFB; light green (S2), 41 DAFB; pink blush (S3), 46 DAFB; red (S4), 51 DAFB; dark red (S5), 55 DAFB. For 'HF', the sampling was: green (S1), 37 DAFB; light green (S2), 43 DAFB; pink blush (S3), 46 DAFB; light red (S4), 49 D AFB; red (S5), 53 DAFB. For 'HP5', the sampling was: green (S1), 36 DAFB; light green (S2), 40 DAFB; pink blush (S3), 45 DAFB; orange red (S4), 50 DAFB; light red (S5), 54 DAFB. For 'PZB', the sampling was: green (S1), 37 DAFB; light green (S2), 43 DAFB; straw yellow (S3), 48 DAFB; yellow (S4), 52 DAFB; yellow with blush (S5), 58 DAFB (Supplementary Table S6). Fruits were collected for their uniform size, same appearance, and no defects. Ten cherries were analyzed to measure their color parameter, and the other thirty fruits were immediately frozen in liquid nitrogen and stored at −80°C for subsequent analysis.

Table 1. Division of fruit phenology of Chinese cherry with different fruit color.

Fruit Phenology	Stage	Color Difference	Fruit Color		
			Dark-Red	Red/Light-Red	Yellow
Green ripening period	S1	a*	−11.34~−7.74	−10.56~−6.77	−10.00
		b*	43.13~44.35	40.11~45.75	37.63
		a*/b*	−0.26~−0.17	−0.26~−0.14	−0.27
Color conversion period	S2	a*	−7.98~−5.85	−8.21~1.08	−9.62
		b*	46.24~46.85	42.32~47.66	38.32
		a*/b*	−0.17~−0.12	−0.19~0.15	−0.25
	S3	a*	5.27~16.74	2.25~10.29	−1.66
		b*	33.80~40.28	31.02~43.00	35.72
		a*/b*	0.13~0.52	0.05~0.26	−0.05
Fruit ripening period	S4	a*	26.54~28.82	24.43~27.09	6.31
		b*	21.91~23.48	12.27~26.70	24.07
		a*/b*	1.22~1.23	0.93~2.20	0.26
	S5	a*	20.43~21.13	22.16~23.22	8.95
		b*	7.27~8.24	8.82~16.58	20.57
		a*/b*	2.55~2.98	1.68~2.53	0.43

Note: The standard was summarized based on eight representative Chinese cherry accessions with different fruit color.

4.2. Fruit Color Assessment

The fruit peel color was measured using a HunterLab chromameter (Konica Minolta, Inc., Tokyo, Japan) according to the CIE system. Positive a* values indicated red and purple, and negative values indicated green and blue. Positive b* values represented yellow, and negative values represented blue. Three sets of a* and b* values were measured for the equatorial part of each fruit and used to calculate the color ratio (a*/b*) [32]. Three biological replicates per sample point were analyzed, with ten cherries for each replicate.

4.3. Total Anthocyanin and Flavonoid Content Measurement

The extraction and measurement of total anthocyanin content was conducted using a pH differential method [71]. About 1.5 g fruit was extracted with 15 mL of extraction solution (acetone:methanol:water:acetic acid = 2:2:1:0.5), after a water bath at 40 °C, the mixture was centrifuged at 8,000 g for 25 min, and the supernatant was used for determination. Two buffer systems were employed, with 0.4 M potassium chloride (pH 1.0) and 0.4 M dibasic sodium (pH 4.5). Total anthocyanin content was calculated according to the

equation: A = $[(A_{510} - A_{700})_{pH\ 1.0} - (A_{510} - A_{700})_{pH\ 4.5}]$, which was converted into mg cyanidin 3-glucoside per 1,000 g fresh weight (FW). Three independent biological replicates per sample point were analyzed.

Aluminum chloride method was utilized for measurement of the total flavonoid content [72]. Stock solution was prepared by dissolving 50 mg quercetin in 50 mL of methanol. About 5 mg fruit extracts were mixed in 5 mL of distilled water and 0.3 mL of 5% $NaNO_2$. 0.6 mL of 10% $AlCl_3$ and 2 mL of 1.0 M NaOH were combined in the above solution after 5 min. The absorbance of the reaction mixture was measured at 510 nm using a spectrophotometer. Total flavonoid content was calculated as quercetin equivalents (mg/g FW), and performed in triplicates.

4.4. Transcriptome Analysis

Three biological replicates with mixed three fruits each, including fruit peel and flesh, were used for RNA-seq. A total of 1 μg RNA per sample was used for RNA preparations. The mRNA molecules were purified using oligo(dT)-attached magnetic beads and then were fragmented into small pieces using fragmentation reagent. The first-strand and second-strand cDNAs were synthesized using random hexamer-primed reverse transcription. End repair, polyadenylation, adapter ligation, PCR amplification, and library quality control were carried out following the DNABSEQ RNA-Seq library preparation protocol. A total of 60 libraries were sequences on an MGI2000 platform to generate raw 150 bp paired-end reads. For downstream analyses, high-quality clean reads were obtained by filtering low-quality reads and those containing adapters or poly-N in SOAPnuke software (−n 0.01–1.20, −q 0.4–A 0.25-cutAdaptor).

The reference genome database and gene annotation files were extracted from Chinese cherry (unpublished data). Clean reads were mapped onto the reference genome using Hisat2 v.2.1.0. The read counts and FPKM values for each gene were calculated in StringTie v.1.3.5. The DEGs were identified using EdegR package with the screening conditions $|\log_2 \text{fold change}| \geq 2$ and DFR ≤ 0.01. Structural and regulatory genes involved in anthocyanin biosynthesis pathway were screened from DEGs.

WGCNA was conducted using the WGCNA R package with default settings (v1.4.1717). All genes were imported into WGCNA to construct co-expression modules using the automatic network construction function block-wise Modules. Correlations between modules and color ratio, and anthocyanin and flavonoid content at each developmental stage were analyzed with respect to all genes in each module. Significant trait-related modules were identified based on high correlation values. Using default settings, genes from the MEblack module were exported for Cytoscape software (v.3.9.1) [73].

4.5. Real-Time PCR Analysis

The expression levels of nine structural genes and eight transcriptional factors in the anthocyanin biosynthesis pathway were determined by RT-qPCR. The gene-specific primers were designed by Primer 5.0 and shown in Supplementary Table S7. Total RNA was extracted from the fruits at different developmental stages using the Plant Total RNA Isolation Kit (SK8631; Sangon Biotech, Shanghai, China). The cDNA was synthesized from RNA using the PrimeScriptTM RT-PCR Kit (RR047A; TaKaRa Bio, Kusatsu, Japan). The RT-qPCR was performed in a 20 μL reaction volume using the TransStart® Green qPCR SuperMix (TransGen Biotech Co., Ltd., Beijing, China) on a CFX96 TouchTM Real-Time PCR detection system (Bio-Rad, Hercules, CA, USA). The reaction procedure is as follows: 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s, 60 °C for 30 s, and 72 °C for 30 s. The $2^{-\Delta\Delta CT}$ method was used to calculate the gene expression levels with the geometric mean of the two housekeeping genes (cherry actin and ubiquitin). Three independent biological replicates were analyzed for each sample point.

4.6. Metabolome Analysis

The freeze-dried matured cherry samples with mixed peel and flesh were crushed using a mill (MM400, Retsch, Germany) with a zirconia bread at 30 Hz for 1.5 min. Approximately 50 mg of samples was extracted overnight at 4 °C with 1.2 mL of 70% methanol before performing centrifugation at 12,000 rpm for 3 min. The supernatants were pooled and filtered with a microporous membrane (0.22 μm). The relative quantification of widely targeted metabolites in Chinese cherry fruit using an UPLC-ESI-MS/MS system (UPLC, ExionLC^{TM} AD; MS, Applied Biosystems 6500 Q TRAP, https://sciex.com.cn/, accessed on 21 October 2022). Quantification of metabolites was carried out using a scheduled multiple reaction monitoring method. Metabolite profiling and metabolomics data analyses were conducted by Metware Biotechnology Co., Ltd. (Wuhan, China).

PCA and OPLS-DA were conducted to verify the differences and reliability of metabolites. DEMs were determined by a VIP ≥ 1 and absolute $\log_2$fold change ($\geq$1). Then the DEMs were mapped to the KEGG database and their significance was determined by hypergeometric test's *p*-values.

4.7. Statistical Analysis

The data was analyzed using IBM SPSS Statistics software (v25.0). The results were expressed as mean ± standard deviation (SD). A $p \leq 0.05$ was considered a statistically significant difference (Tukey's test).

5. Conclusions

In summary, this is the first study to investigate the coloring patterns and the corresponding accumulation of anthocyanin between dark-red and yellow Chinese cherry fruits. Based on LC-MS/MS, we identified 33 and 3 differential expressed metabolites related to anthocyanins and proanthocyanidins between mature dark-red and yellow fruits. The anthocyanins were mainly up-regulated, while the proanthocyanidins were all down-regulated in dark-red fruits. By transcriptome analysis, eight biosynthesis genes (*CpCHS*, *CpCHI*, *CpF3H*, *CpF3′H*, *CpDFR*, *CpANS*, *CpUFGT*, and *CpGST*) were significantly more highly expressed in dark-red fruits, especially *CpANS*, *CpUFGT*, and *CpGST*. *CpLAR* was higher in yellow fruits than dark-red fruits, especially at the early stage. Eight regulatory genes (*CpMYB4*, *CpMYB10*, *CpMYB20*, *CpMYB306*, *bHLH1*, *CpNAC10*, *CpERF106*, and *CpbZIP4*) were also identified as determinants of fruit color in Chinese cherry. These findings can enrich the key genes and metabolites involved in the anthocyanins biosynthesis in Chinese cherry, which are of great importance for molecular marker-assisted breeding.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24043471/s1.

Author Contributions: Conceptualization, X.W.; Methodology, Y.W. and Z.W.; Software, Y.W., Z.W. and J.Z. (Jing Zhang); Validation, J.Z. (Jing Zhang) and H.W.; Formal analysis, Y.W. and Z.W.; Investigation, Z.L., H.T. (Hongxia Tu), J.Z. (Jingting Zhou), X.L., Q.C. and W.H.; Resources, M.L., Y.L. (Yuanxiu Lin), and Y.Z. (Yunting Zhang); Data curation, Y.Z. (Yong Zhang), and Y.L. (Ya Luo); Writing—original draft, Y.W.; Writing—review and editing, Y.W. and X.W.; Supervision, H.T. (Haoru Tang); Project administration, S.Y.; Funding acquisition, Y.W. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by Cherry Resources Sharing and Service Platform of Sichuan Province, Chengdu Technological Innovation Research and Development Project (2022YF05-01017-SN), Tianfu Talent Project of Chengdu City (2021-CF02-0162396-RC-4096), The Project of Rural Revitalization Research Institute in Tianfu New Area of Sichuan Province (XZY1-04), Undergraduate Innovation Training Program in Sichuan Agricultural University (S202210626089), and Shuangzhi Project Innovation Team of Sichuan Agricultural University (P202107).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Jaakola, L. New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* **2013**, *18*, 477–483.
2. Cruz, L.; Basílio, N.; Mateus, N.; De Freitas, V.; Pina, F. Natural and synthetic flavylium-based dyes: The chemistry behind the color. *Chem. Rev.* **2021**, *122*, 1416–1481.
3. Zeng, Y.X.; Song, J.J.; Zhang, M.M.; Wang, H.W.; Zhang, Y.; Suo, H.Y. Comparison of in vitro and in vivo antioxidant activities of six flavonoids with similar structures. *Antioxidants* **2020**, *9*, 732. [CrossRef]
4. Tena, N.; Martín, J.; Asuero, A.G. State of the art of anthocyanins: Antioxidant activity, sources, bioavailability, and therapeutic effect in human health. *Antioxidants* **2020**, *9*, 451. [CrossRef]
5. Yu, D.; Huang, T.; Tian, B.; Zhan, J. Advances in biosynthesis and biological functions of proanthocyanidins in horticultural plants. *Foods* **2020**, *9*, 1774. [CrossRef]
6. Shen, N.; Wang, T.F.; Gan, Q.; Liu, S.A.; Wang, L.; Jin, B. Plant flavonoids: Classification, distribution, biosynthesis, and antioxidant activity. *Food Chem.* **2022**, *383*, 132531. [CrossRef]
7. Jiang, S.H.; Chen, M.; He, N.B.; Chen, H.L.; Wang, N.; Sun, Q.G.; Zhang, T.L.; Xu, H.F.; Fang, H.C.; Wang, Y.C. MdGSTF6, activated by MdMYB1, plays an essential role in anthocyanin accumulation in apple. *Hortic. Res.* **2019**, *6*, 40. [CrossRef]
8. Zhao, Y.; Dong, W.Q.; Zhu, Y.C.; Allan, A.C.; Lin-Wang, K.; Xu, C.J. *PpGST1*, an anthocyanin-related glutathione S-transferase gene, is essential for fruit coloration in peach. *Plant Biotechnol. J.* **2020**, *18*, 1284–1295.
9. Qi, X.L.; Liu, C.L.; Song, L.L.; Dong, Y.X.; Chen, L.; Li, M. A sweet cherry glutathione S-transferase gene, *PavGST1*, plays a central role in fruit skin coloration. *Cells* **2022**, *11*, 710.
10. Lafountain, A.M.; Yuan, Y.W. Repressors of anthocyanin biosynthesis. *New Phytol.* **2021**, *293*, 933–949. [CrossRef]
11. Lin, Y.X.; Hou, G.Y.; Jiang, Y.Y.; Liu, X.Y.; Yang, M.; Wang, L.X.; Long, Y.; Li, M.Y.; Zhang, Y.T.; Wang, Y.; et al. Joint transcriptomic and metabolomic analysis reveals differential flavonoid biosynthesis in a high-flavonoid strawberry mutant. *Front. Plant Sci.* **2022**, *13*, 919619.
12. Sun, L.P.; Huo, J.T.; Liu, J.Y.; Yu, J.Y.; Zhou, J.L.; Sun, C.D.; Wang, Y.; Leng, F. Anthocyanins distribution, transcriptional regulation, epigenetic and post-translational modification in fruits. *Food Chem.* **2023**, *411*, 135540. [CrossRef]
13. Yu, D.J. *Classification of Fruit Trees in China*; Agricultural Press: Beijing, China, 1979.
14. Wang, Y.; Du, H.M.; Zhang, J.; Chen, T.; Chen, Q.; Tang, H.R.; Wang, X.R. Ploidy level of Chinese cherry (*Cerasus pseudocerasus* Lindl.) and comparative study on karyotypes with four *Cerasus* species. *Sci. Hortic.* **2018**, *232*, 46–51.
15. Yü, D.J.; Lu, L.T.; Ku, T.C.; Li, C.L.; Chen, S.X. *Flora of China*; Science Press: Beijing, China, 1986; Volume 38.
16. Zhang, X.M.; Yan, G.H.; Zhou, Y.; Wang, J.; Duan, X.W.; Zhang, K.C. A new sweet cherry rootstock cultivar 'Jingchun 1'. *Acta Hortic. Sinica* **2021**, *48*, 2787–2788.
17. Zhang, X.M.; Yan, G.H.; Zhou, Y.; Wang, J.; Duan, X.W.; Zhang, K.C. A new sweet cherry rootstock cultivar 'Landing 3'. *Acta Hortic. Sinica* **2021**, *48*, 2789–2790.
18. Wu, Y.J.; Song, Q.Q.; Yuan, Y.; Guo, F.Q.; Wu, K.X.; Dong, M.M. In vitro efficiency of embryo rescue of intra- and interspecific hybrid crosses of sweet cherry and Chinese cherry cultivars. *Sci. Hortic.* **2021**, *275*, 109716. [CrossRef]
19. Huang, X.J.; Wang, X.R.; Chen, T.; Chen, J.; Tang, H.R. Research progress of genetic diversity in *Cerasus pseudocerasus* and their wild relative populations, and utilize progress of cultivation resources. *J. Fruit Sci.* **2013**, *30*, 470–479.
20. Wang, Y.; Hu, G.P.; Liu, Z.S.; Zhang, J.; Ma, L.; Tian, T.; Wang, H.; Chen, T.; Chen, Q.; He, W.; et al. Phenotyping in flower and main fruit traits of Chinese cherry [*Cerasus pseudocerasus* (Lindl). G.Don]. *Sci. Hortic.* **2022**, *296*, 110920.
21. Blando, F.; Oomah, B.D. Sweet and sour cherries: Origin, distribution, nutritional composition and health benefits. *Trends Food Sci. Tech.* **2019**, *86*, 517–529.
22. Cao, J.P.; Jiang, Q.; Lin, J.Y.; Li, X.; Sun, C.D.; Chen, K.S. Physicochemical characterisation of four cherry species (*Prunus* spp.) grown in China. *Food Chem.* **2015**, *173*, 855–863. [CrossRef]
23. Gonçalves, A.C.; Campos, G.; Alves, G.; Garcia-Viguera, C.; Moreno, D.A.; Silva, L.R. Physical and phytochemical composition of 23 Portuguese sweet cherries as conditioned by variety (or genotype). *Food Chem.* **2021**, *335*, 127637. [CrossRef] [PubMed]
24. Karagiannis, E.; Sarrou, E.; Michailidis, M.; Tanou, G.; Ganopoulos, I.; Bazakos, C.; Kazantzis, K.; Martens, S.; Xanthopoulou, A.; Molassiotis, A. Fruit quality trait discovery and metabolic profiling in sweet cherry genebank collection in Greece. *Food Chem.* **2021**, *342*, 128315. [CrossRef] [PubMed]
25. De Leo, M.; Iannuzzi, A.M.; Germano, M.P.; D'angelo, V.; Camangi, F.; Sevi, F.; Diretto, G.; De Tommasi, N.; Braca, A. Comparative chemical analysis of six ancient italian sweet cherry (*Prunus avium* L.) varieties showing antiangiogenic activity. *Food Chem.* **2021**, *360*, 129999. [PubMed]
26. Zhang, A.D.; Yang, H.Y.; Ji, S.; Tian, C.; Chen, N.; Gong, H.S.; Li, J.Z. Metabolome and transcriptome analyses of anthocyanin mechanisms reveal metabolite variations and key candidate genes involved in the pigmentation of *Prunus tomentosa* Thunb. cherry fruit. *Front. Plant Sci.* **2022**, *13*, 938098. [CrossRef]
27. Liu, Y.; Shen, X.J.; Zhao, K.; Ben, Y.; Guo, X.W.; Zhang, X.M.; Li, T.H. Expression analysis of anthocyanin biosynthetic genes in different colored sweet cherries (*Prunus avium* L.) during fruit development. *J. Plant Growth Regul.* **2013**, *32*, 901–907.

28. Shen, X.J.; Zhao, K.; Liu, L.L.; Zhang, K.C.; Yuan, H.Z.; Liao, X.; Wang, Q.; Guo, X.W.; Li, F.; Li, T.H. A role for PacMYBA in ABA-regulated anthocyanin biosynthesis in red-colored sweet cherry cv. Hong Deng (*Prunus avium* L.). *Plant Cell Physiol.* **2014**, *55*, 862–880. [CrossRef] [PubMed]
29. Jin, W.M.; Wang, H.; Li, M.F.; Wang, J.; Yang, Y.; Zhang, X.M.; Yan, G.H.; Zhang, H.; Liu, J.S.; Zhang, K.C. 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]
30. Yang, H.Y.; Tian, C.P.; Ji, S.J.; Ni, F.Z.; Fan, X.G.; Yang, Y.Q.; Sun, C.C.; Gong, H.S.; Zhang, A.D. Integrative analyses of metabolome and transcriptome reveals metabolomic variations and candidate genes involved in sweet cherry (*Prunus avium* L.) fruit quality during development and ripening. *PLoS One* **2021**, *16*, e0260004.
31. Yi, D.B.; Zhang, H.N.; Lai, B.; Liu, L.Q.; Pan, X.L.; Ma, Z.L.; Wang, Y.C.; Xie, J.H.; Shi, S.Y.; Wei, Y.Z. Integrative analysis of the coloring mechanism of red Longan pericarp through metabolome and transcriptome analyses. *J. Agr. Food Chem.* **2021**, *69*, 1806–1815. [CrossRef]
32. Arias, R.; Lee, T.C.; Logendra, L.; Janes, H. Correlation of lycopene measured by HPLC with the *L**, *a**, *b** color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content. *J. Agric. Food Chem.* **2000**, *48*, 1697–1702.
33. Sabir, I.A.; Manzoor, M.A.; Shah, I.H.; Liu, X.J.; Zahid, M.S.; Jiu, S.T.; Wang, J.Y.; Abdullah, M.; Zhang, C.X. MYB transcription factor family in sweet cherry (*Prunus avium* L.): Genome-wide investigation, evolution, structure, characterization and expression patterns. *BMC Plant Biol.* **2022**, *22*, 2. [CrossRef] [PubMed]
34. Hou, Q.C.; Zhao, W.; Lu, L.; Wang, L.L.; Zhang, T.Y.; Hu, B.B.; Yan, T.W.; Qi, Y.C.; Zhang, F.; Chao, N.; et al. Overexpression of HLH4 inhibits cell elongation and anthocyanin biosynthesis in *Arabidopsis thaliana*. *Cells* **2022**, *11*, 1087.
35. Zhang, S.Y.; Chen, Y.X.; Zhao, L.L.; Li, C.Q.; Yu, C.Q.; Yu, J.Y.; Li, T.T.; Yang, W.Y.; Zhang, S.N.; Su, H.Y.; et al. A novel NAC transcription factor, MdNAC42, regulates anthocyanin accumulation in red-fleshed apple by interacting with MdMYB10. *Tree Physiol.* **2020**, *40*, 413–423. [PubMed]
36. Jaakola, L.; Poole, M.; Jones, M.O.; KäMäRäInen-Karppinen, T.; KoskimäKi, J.J.; Hohtola, A.; HäGgman, H.; Fraser, P.D.; Manning, K.; King, G.J.; et al. A SQUAMOSA MADS box gene involved in the regulation of anthocyanin accumulation in bilberry fruits. *Plant Physiol.* **2010**, *153*, 1619–1629. [CrossRef] [PubMed]
37. Ma, H.Y.; Yang, T.; Li, Y.; Zhang, J.; Wu, T.; Song, T.T.; Yao, Y.C.; Tian, J. The long noncoding RNA MdLNC499 bridges MdWRKY1 and MdERF109 function to regulate early-stage light-induced anthocyanin accumulation in apple fruit. *Plant Cell* **2021**, *33*, 3309–3330.
38. Chen, Q.S.; Xu, X.Y.; Xu, D.; Zhang, H.S.; Zhang, C.K.; Li, G. WRKY18 and WRKY53 coordinate with HISTONE ACETYLTRANSFERASE1 to regulate rapid responses to sugar. *Plant Physiol.* **2019**, *180*, 2212–2226. [CrossRef] [PubMed]
39. An, J.P.; Yao, J.F.; Xu, R.R.; You, C.X.; Wang, X.F.; Hao, Y.J. Apple bZIP transcription factor MdbZIP44 regulates abscisic acid-promoted anthocyanin accumulation. *Plant Cell Environ.* **2018**, *41*, 2678–2692.
40. Fang, Z.Z.; Lin-Wang, K.; Jiang, C.C.; Zhou, D.R.; Lin, Y.J.; Pan, S.L.; Espley, R.V.; Ye, X.F. Postharvest temperature and light treatments induce anthocyanin accumulation in peel of 'Akihime' plum (*Prunus salicina* Lindl.) via transcription factor PsMYB10.1. *Postharvest Biol. Tec.* **2021**, *179*, 111592. [CrossRef]
41. Li, C.; Wu, J.; Hu, K.D.; Wei, S.W.; Sun, H.Y.; Hu, L.Y.; Han, Z.; Yao, G.F.; Zhang, H. *PyWRKY26* and *PybHLH3* cotargeted the *PyMYB114* promoter to regulate anthocyanin biosynthesis and transport in red-skinned pears. *Hortic. Res.* **2020**, *7*, 37.
42. Bueno, J.M.; Sáez-Plaza, P.; Ramos-Escudero, F.; Jiménez, A.M.; Fett, R.; Asuero, A.G. Analysis and antioxidant capacity of anthocyanin pigments. Part II: Chemical structure, color, and intake of anthocyanins. *Crit. Rev. Anal. Chem.* **2012**, *42*, 126–151. [CrossRef]
43. Brozdowski, J.; Waliszewska, B.; Loffler, J.; Hudina, M.; Veberic, R.; Mikulic-Petkovsek, M. Composition of phenolic compounds, cyanogenic glycosides, organic acids and sugars in fruits of black cherry (*Prunus serotina* Ehrh.). *Forests* **2021**, *12*, 762. [CrossRef]
44. Acero, N.; Gradillas, A.; Beltran, M.; García, A.; Mingarro, D.M. Comparison of phenolic compounds profile and antioxidant properties of different sweet cherry (*Prunus avium* L.) varieties. *Food Chem.* **2019**, *279*, 260–271. [CrossRef] [PubMed]
45. Turturică, M.; Stănciuc, N.; Bahrim, G.; Râpeanu, G. Investigations on sweet cherry phenolic degradation during thermal treatment based on fluorescence spectroscopy and inactivation kinetics. *Food Bioprocess Tech.* **2016**, *9*, 1706–1715. [CrossRef]
46. Homoki, J.R.; Nemes, A.; Fazekas, E.; Gyémánt, G.; Balogh, P.; Gál, F.; Al-Asri, J.; Mortier, J.; Wolber, G.; Babinszky, L.; et al. Anthocyanin composition, antioxidant efficiency, and a-amylase inhibitor activity of different Hungarian sour cherry varieties (*Prunus cerasus* L.). *Food Chem.* **2016**, *194*, 222–229. [PubMed]
47. Damar, I.; Eksi, A. Antioxidant capacity and anthocyanin profile of sour cherry (*Prunus cerasus* L.) juice. *Food Chem.* **2012**, *135*, 2910–2914. [CrossRef] [PubMed]
48. Wu, Q.; Yuan, R.Y.; Feng, C.Y.; Li, S.S.; Wang, L.S. Analysis of polyphenols composition and antioxidant activity assessment of Chinese dwarf Cherry (*Cerasus humilis* (Bge.) Sok.). *Nat. Prod. Commun.* **2019**, *14*, 1934578X19856509.
49. Liu, S.W.; Li, X.Y.; Guo, Z.Y.; Zhang, X.; Chang, X.D. Polyphenol content, physicochemical properties, enzymatic activity, anthocyanin profiles, and antioxidant capacity of *Cerasus humilis* (Bge.) Sok. genotypes. *J. Food Quality* **2018**, *1*, 5479565.
50. Wang, P.F.; Mu, X.P.; Du, J.J.; Gao, Y.G.; Bai, D.H.; Jia, L.T.; Zhang, J.C.; Ren, H.Y.; Xue, X.F. Flavonoid content and radical scavenging activity in fruits of Chinese dwarf cherry (*Cerasus humilis*) genotypes. *J. Forest Res.* **2018**, *29*, 55–63. [CrossRef]
51. Martini, S.; Conte, A.; Tagliazucchi, D. Phenolic compounds profile and antioxidant properties of six sweet cherry (*Prunus avium*) cultivars. *Food Res. Int.* **2017**, *97*, 15–26.

52. Ponce, C.; Kuhn, N.; Arellano, M.; Time, A.; Multari, S.; Martens, S.; Carrera, E.; Sagredo, B.; Donoso, J.M.; Meisel, L.A. Differential phenolic compounds and hormone accumulation patterns between early- and mid-maturing sweet cherry (*Prunus avium* L.) cultivars during fruit development and ripening. *J. Agr. Food Chem.* **2021**, *69*, 8850–8860. [CrossRef]
53. Jeong, S.T.; Goto-Yamamoto, N.; Hashizume, K.; Esaka, M. Expression of the flavonoid 3′-hydroxylase and flavonoid 3′,5′-hydroxylase genes and flavonoid composition in grape (*Vitis vinifera*). *Plant Sci.* **2006**, *170*, 61–69. [CrossRef]
54. Cao, K.; Ding, T.Y.; Mao, D.M.; Zhu, G.R.; Fang, W.C.; Chen, C.W.; Wang, X.W.; Wang, L.R. Transcriptome analysis reveals novel genes involved in anthocyanin biosynthesis in the flesh of peach. *Plant Physiol. Bioch.* **2018**, *123*, 94–102. [CrossRef] [PubMed]
55. Liu, C.; Wang, X.; Shulaev, V.; Dixon, R.A. A role for leucoanthocyanidin reductase in the extension of proanthocyanidins. *Nat. Plants* **2016**, *2*, 16182. [CrossRef]
56. Cheng, C.Z.; Guo, Z.W.; Li, H.; Mu, X.P.; Wang, P.F.; Zhang, S.; Yang, T.Z.; Cai, H.C.; Wang, Q.; Lu, P.T.; et al. Integrated metabolic, transcriptomic and chromatin accessibility analyses provide novel insights into the competition for anthocyanins and flavonols biosynthesis during fruit ripening in red apple. *Fronti. Plant Sci.* **2022**, *975356*, 1–15.
57. Chen, W.F.; Xiao, Z.C.; Wang, Y.L.; Wang, J.H.; Zhai, R.; Wang, K.L.; Espley, R.; Ma, F.W.; Li, P.M. Competition between anthocyanin and kaempferol glycosides biosynthesis affects pollen tube growth and seed set of *Malus*. *Hortic. Res.* **2021**, *8*, 173. [PubMed]
58. Ahmad Khan, I.; Cao, K.; Guo, J.; Li, Y.; Wang, Q.; Yang, X.W.; Wu, J.L.; Fang, W.C.; Wang, L.R. Identification of key gene networks controlling anthocyanin biosynthesis in peach flower. *Plant Sci.* **2022**, *316*, 111151.
59. Gutierrez, E.; García-Villaraco, A.; Lucas, J.A.; Gradillas, A.; Gutierrez-Mañero, F.J.; Ramos-Solano, B. Transcriptomics, targeted metabolomics and gene expression of blackberry leaves and fruits indicate flavonoid metabolic flux from leaf to red fruit. *Front. Plant Sci.* **2017**, *8*, 472. [CrossRef]
60. Wang, L.H.; Tang, W.; Hu, Y.W.; Zhang, Y.B.; Sun, J.Q.; Guo, X.H.; Lu, H.; Yang, Y.; Fang, C.B.; Niu, X.L.; et al. A MYB/bHLH complex regulates tissue-specific anthocyanin biosynthesis in the inner pericarp of red-centered kiwifruit *Actinidia chinensis* cv. Hongyang. *Plant J.* **2019**, *99*, 359–378. [CrossRef]
61. Espley, R.V.; Hellens, R.P.; Putterill, J.; Stevenson, D.E.; Kutty-Amma, S.; Allan, A.C. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* **2007**, *49*, 414–427.
62. Xi, W.P.; Feng, J.; Liu, Y.; Zhang, S.K.; Zhao, G.H. The R2R3-MYB transcription factor PaMYB10 is involved in anthocyanin biosynthesis in apricots and determines red blushed skin. *BMC Plant Biol.* **2019**, *19*, 287.
63. Geng, P.; Zhang, S.; Liu, J.Y.; Zhao, C.H.; Wu, J.; Cao, Y.P.; Fu, C.X.; Han, X.; He, H.; Zhao, Q. MYB20, MYB42, MYB43, and MYB85 regulate phenylalanine and lignin biosynthesis during secondary cell wall formation. *Plant Physio.* **2020**, *182*, 1272–1283. [CrossRef]
64. Wang, S.; Zhang, Z.; Li, L.X.; Wang, H.B.; Zhou, H.; Chen, X.S.; Feng, S.Q. Apple MdMYB306-like inhibits anthocyanin synthesis by directly interacting with MdMYB17 and MdbHLH33. *Plant J.* **2022**, *110*, 1021–1034. [PubMed]
65. Li, H.; Yang, Z.; Zeng, Q.W.; Wang, S.B.; Luo, Y.W.; Huang, Y.; Xin, Y.C.; He, N.J. Abnormal expression of *bHLH3* disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hortic. Res.* **2020**, *7*, 83.
66. Deng, G.M.; Zhang, S.; Yang, Q.S.; Gao, H.J.; Sheng, O.; Bi, F.C.; Li, C.Y.; Dong, T.; Yi, G.J.; He, W.D.; et al. *MaMYB4*, an R2R3-MYB repressor transcription factor, negatively regulates the biosynthesis of anthocyanin in Banana. *Front. Plant Sci.* **2021**, *11*, 600704. [PubMed]
67. Walker, A.R.; Davison, P.A.; Bolognesi-Winfield, A.C.; James, C.M.; Srinivasan, N.; Blundell, T.L.; Esch, J.J.; David Marks, M.; Gray, J.C. The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* **1999**, *11*, 1337–1349. [PubMed]
68. Brueggemann, J.; Weisshaar, B.; Sagasser, M. A WD40-repeat gene from *Malus* × *domestica* is a functional homologue of *Arabidopsis thaliana TRANSPARENT TESTA GLABRA1*. *Plant Cell Rep.* **2010**, *29*, 285–294. [CrossRef]
69. Schaart, J.G.; Dubos, C.; De La Fuente, I.R.; Van Houwelingen, A.M.M.L.; De Vos, R.C.H.; Jonker, H.H.; Xu, W.J.; Routaboul, J.M.; Lepiniec, L.; Bovy, A.G. Identification and characterization of MYBbHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* × *ananassa*) fruits. *New Phytol.* **2013**, *197*, 454–467. [CrossRef]
70. Sun, Q.G.; Jiang, S.H.; Zhang, T.L.; Xu, H.F.; Fang, H.C.; Zhang, J.; Su, M.Y.; Wang, Y.C.; Zhang, Z.Y.; Wang, N.; et al. Apple NAC transcription factor *MdNAC52* regulates biosynthesis of anthocyanin and proanthocyanidin through *MdMYB9* and *MdMYB11*. *Plant Sci.* **2019**, *289*, 110286.
71. Lee, J.; Durst, R.W.; Wrolstad, R.E. Determination of total monomeric anthocyanin pigment content of fruit juices, beverages, natural colorants, and wines by the pH differential method: Collaborative study. *J. AOAC Int.* **2005**, *88*, 1269–1278. [CrossRef]
72. Marinova, D.; Ribarova, F.; Atanassova, M. Total phenolics and total flavonoids in Bulgarian fruits and vegetables. *J. Uni. Chem. Tech. Meta.* **2005**, *40*, 255–260.
73. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowshi, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Genome-Wide Investigation and Functional Analysis Reveal That *CsGeBP4* Is Required for Tea Plant Trichome Formation

Hao Zhou [1], Wei Zhou [1], Xinzhuan Yao [1], Qi Zhao [2,*] and Litang Lu [1,2,*]

[1] College of Tea Sciences, Institute of Plant Health & Medicine, Guizhou University, Guiyang 550025, China
[2] The Key Laboratory of Plant Resources Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Institute of Agro-Bioengineering/College of Life Sciences, Guizhou University, Guiyang 550025, China
* Correspondence: zhaoqizq@yeah.net (Q.Z.); ltlv@gzu.edu.cn (L.L.)

Abstract: Tea plant trichomes not only contribute to the unique flavor and high quality of tea products but also provide physical and biochemical defenses for tea plants. Transcription factors play crucial roles in regulating plant trichome formation. However, limited information about the regulatory mechanism of transcription factors underlying tea plant trichome formation is available. Here, the investigation of trichome phenotypes among 108 cultivars of Yunwu Tribute Tea, integrated with a transcriptomics analysis of both hairy and hairless cultivars, revealed the potential involvement of *CsGeBPs* in tea trichome formation. In total, six *CsGeBPs* were identified from the tea plant genome, and their phylogenetic relationships, as well as the structural features of the genes and proteins, were analyzed to further understand their biological functions. The expression analysis of *CsGeBPs* in different tissues and in response to environmental stresses indicated their potential roles in regulating tea plant development and defense. Moreover, the expression level of *CsGeBP4* was closely associated with a high-density trichome phenotype. The silencing of *CsGeBP4* via the newly developed virus-induced gene silencing strategy in tea plants inhibited trichome formation, indicating that *CsGeBP4* was required for this process. Our results shed light on the molecular regulatory mechanisms of tea trichome formation and provide new candidate target genes for further research. This should lead to an improvement in tea flavor and quality and help in breeding stress-tolerant tea plant cultivars.

Keywords: *Camellia sinensis*; trichome formation; *CsGeBP*; transcriptional regulation

1. Introduction

The tea plant (*Camellia sinensis* (L.) O. Kuntze) is one of the most popular nonalcoholic beverage crops worldwide. Tea trichomes, also referred to as 'Cha Hao', greatly contribute to the unique flavor and health benefits of tea products since they are rich in secondary metabolites, such as catechins, theanine, caffeine, flavonols, and various volatiles [1]. Generally, the distribution of tea trichomes is associated with the degree of tissue tenderness and varies significantly among different tea plant varieties and cultivars [2–5]. Apical buds and young leaves are the main materials processed for tea products and are usually enriched with trichomes, which is an important feature associated with superior tea quality for many brands of commodity teas, such as white teas, high-quality black teas, and green teas [1,6]. Therefore, trichome density is considered one of the most important criteria for tea quality evaluation. Moreover, tea plant trichomes also act as protective barriers against abiotic and biotic stresses via divergent strategies, such as the reflection of ultraviolet and high lights, reduction in water loss under high temperatures, prevention of leaf freezing in response to cold stress, and resistance to pathogen and herbivore attacks [1,6–12]. Due to the importance of trichomes for both tea quality and defense, insights into the molecular bases of trichome initiation, growth, and development are needed to improve tea quality and the breeding of stress-resistant tea cultivars.

The formation of plant trichomes is regulated by numerous factors, including environmental cues, hormones, and regulatory genes [13]. Among them, a series of transcription factors playing vital regulatory roles in trichome initialization, growth, and development have been identified in a number of plant species, such as *Arabidopsis* [4,14–18], cucumber (*Cucumis sativus*) [19], tomato (*Solanum lycopersicum*) [20], *Brassica campestris* [21] and *B. napus* [22], and the functions of some genes have been explored in depth. Among these, positive regulatory factors have been proven to enhance the development of trichomes by forming an MYB-bHLH-WD40 protein (MBW) transcriptional complex and subsequently activating the downstream homeodomain–leucine zipper (HD-Zip) transcription factor *GLABRA2* (*GL2*) and other effectors [7,23–25]. The components of the MBW complex are functionally redundant and mainly include (i) R2R3 MYB-related transcription factors, GLABRA1 (GL1), MYB23, and MYB5; (ii) bHLH-like transcription factors, GLABRA3 (GL3), ENHANCER OF GLABRA3 (EGL3), TRANSPARENT TESTA 8 (TT8), and MYC-1; and (iii) a WD40-repeat protein, TRANSPARENT TESTA GLABRA1 (TTG1) [13,25–34]. Moreover, SAD2 was also identified as a positive regulator due to its function in maintaining the stability of the MBW complex [35]. In addition, CAPRICE (CPC), TRIPTY CHON (TRY), ENHANCER OF TRY AND CPC1 (ETC1), and ETC2 were identified as negative regulators of the formation of trichomes via interactions with GL3, EGL3, and TTG1 [13,36].

Tea trichomes are differentiated from epidermal cells and belong to the unicellular, unbranched, and non-glandular type of surface hairs, which are similar to those of *Arabidopsis* [1,37]. However, unlike the model plant *Arabidopsis*, the lack of a stable transgenic system in tea plants constrained the progress of studies on the transcription-factor-regulated mechanisms of tea trichome formation conducted in recent decades. Until recently, an R2R3 MYB transcription factor CsMYB1 was confirmed to regulate trichome formation in tea plants by forming an MBW complex with CsGL3 and CsWD40, thereby activating the trichome regulator genes *CsGL2* and *CsCPC*, and the galloylated *cis*-catechins biosynthesis genes *anthocyanidin reductase* and *serine carboxypeptidase-like 1A* [4,38]. In addition, comparative transcriptomics research was conducted between hairless 'Chuyeqi' (CYQ) and hairy 'Budiaomao' (BDM) tea plants, and a total of 208 transcription factors were differentially expressed [5]. Among them, *CsMYB75*, *CsNOK*, and *CsATML1* might enhance trichome development by up-regulating their expression levels, while *CsSPL6* and *CsSPL12* were regarded as negative regulators in trichome formation due to their higher expression levels in the hairless cultivar 'CYQ' than those in the hairy 'BDM' [5]. Another comparative transcriptomic study was carried out between the hairless 'Rongchunzao' and hairy 'Fudingdabai' tea plant cultivars, which provided several candidate regulatory transcription factors associated with trichome development, including gene members of the *HD-Zip*, *ZEP*, *SPL*, *MADS-box*, *TCP*, *and GRF* families [2]. Additionally, several transcription factor gene families, such as *bHLHs* and *CPC-like* genes, were identified from the tea plant genomes and were thought to play roles in tea trichome formation [3,39]. However, the specific mechanisms associated with tea trichome formation regulated by genes identified using comparative transcriptomics and genome-wide research are still elusive. Recently, virus-induced gene silencing (VIGS) technology has been successfully established in tea plants, enabling the efficient and accurate analysis of gene functions in tea plants [40]. Mutants can be easily obtained without genetic transformation in tea plants using VIGS technology, and the resulting silencing of targeted genes can be maintained for a relatively long time. This newly developed VIGS technology introduces a new method of conducting functional studies on regulatory genes involved in tea plant trichome formation.

The glabrous-enhancer-binding protein (GeBP) family members are plant-specific transcription factors whose members share a central DNA-binding domain. They were proven to play critical regulatory roles in cell expansion in *Arabidopsis* [41]. Another study revealed that AtGeBP interacted with the trichome-formation-related gene *AtGL1* in yeast and in vitro, and its expression was regulated by *KNAT1*, a meristem establishment gene [42,43]. These findings suggest that *AtGeBP* might participate in plant trichome formation. In addition, it was reported that GeBP family genes were involved in plant responses

to phytohormones such as gibberellin [18,42,44,45], cytokinin [45–48], and auxin [45]. Furthermore, their roles in plant responses to environmental stresses such as heavy metals and pathogens were also reported [38,45,49]. For example, *GeBP-like 4* (*GPL4*) was induced in response to cadmium, copper, and zinc stresses in *Arabidopsis* [49]. Thus far, the genome-wide identification and characterization of the *GeBP* gene family have been conducted in soybean [50], mango (*Mangifera indica*) [38], moso bamboo (*Phyllostachys edulis*) [51], tomato (*Solanum lycopersicum*) [52] and nine Gramineae crops (*Brachypodium distachyon*, *Hordeum vulgare*, *Oryza sativa* ssp. *Indica*, *Oryza sativa* ssp. *japonica*, *Oryza rufipogon*, *Sorghum bicolor*, *Setaria italica*, *Triticum aestivum*, *Zea mays*) [45]. However, the specific biological functions of most *GeBPs* have not been characterized. Furthermore, the function and regulatory mechanism of *CsGeBPs* in tea plant trichome formation are also elusive.

In this study, we integrated the investigations of trichome phenotypes and transcriptome profiling of different tea plant cultivars. The *CsGeBP* family members were screened out as candidate regulatory genes regulating tea plant trichome formation. The phylogenetic relationship, gene, protein structural features, and tissue-specific and environmental-responsive expression patterns of *CsGeBP* family members were characterized. The subsequent correlation analysis revealed a close relationship between the gene expressions of *CsGeBP4* and a high-density trichome phenotype, and the indispensable regulatory role of *CsGeBP4* in trichome formation was further verified by the VIGS strategy performed in tea plants, indicating that *GeBP4* is required for regulating tea plant trichome growth and development. This study provides reliable in vivo evidence for the involvement of the *GeBP* family member in plant trichome formation and proposes a new perspective on the potential transcriptional regulation mechanism involving *CsGeBP4* during tea plant trichome formation. This establishes a theoretical framework and valuable foundation for future research to improve tea flavor and quality, as well as to continue to breed stress-tolerant tea plant cultivars.

2. Results

2.1. Association of Trichome Phenotype with Transcription Factor CsGeBPs

To understand the trichome morphological variations in different tea plant cultivars and the underlying genetic factors regulating trichome formation, we investigated the trichome phenotypes on the apical buds of 108 cultivars of Yunwu Tribute Tea (*Camellia sinensis* (L.) Kuntze var. niaowangensis Q. H. Chen), whose tender leaves are used as raw materials for producing one of the internationally famous teas, Guiding Snow Bud [53]. The trichome density index (TDI), analyzed using ImageJ software (https://imagej.en.softonic.com/, accessed on 15 May 2022), was used to evaluate and quantify the trichome phenotypes in tea plants. TDIs were continuously distributed in 108 tea plant cultivars (Figure 1A). Six tea plant cultivars showing significant variations of TDIs were selected to present typical trichome phenotypes with TDIs ranging from high to low levels among the whole population (Figure 1B,C). The apical buds of cultivar No. 43 were covered with the densest trichomes among the six cultivars, and its TDI was 3.3-fold higher than that of cultivar No. 36, which showed the lowest trichome density compared with the other cultivars (Figure 1B,C).

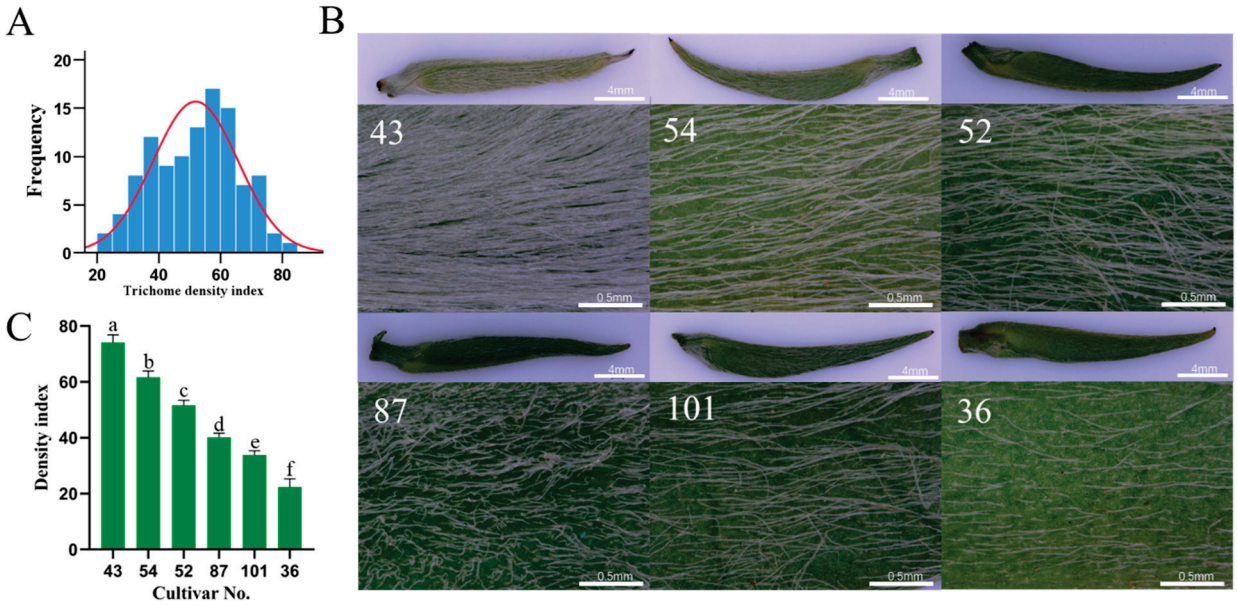


Figure 1. Trichome phenotypes among different tea plant cultivars. (**A**) Normal distribution test of the tea trichome density phenotype among the 108 cultivars of Yunwu Tribute Tea. The normal curves are shown in red. (**B**) Trichome phenotypes in the apical buds of selected tea plant cultivars (No. 43, No.54, No. 52, No. 87, No. 101, and No. 36). (**C**) Trichome density index of the apical buds of selected tea plant cultivars. Lowercase letters indicate significant differences among different samples.

The RNA-seq transcriptome profiling on apical buds of tea plant cultivars No. 43 and No. 36 was further compared to search for the key regulators affecting trichome formation in tea plants. The top 30 up-regulated transcription factor genes with the highest fold changes in expression levels in the hairy cultivar No. 43 compared with the hairless cultivar No. 36 were extracted from the transcriptome data (Figure 2). These genes belong to 13 gene families, and members from *bHLH*, *bZIP*, *C2H2*, *ERF*, *GRAS*, *MYB*, *NAC*, *Trihelix*, and *WARK* families were identified as differentially expressed genes in previous comparative transcriptome studies conducted between hairy and hairless tea plant varieties [2,4,5]. Among the rest four gene families, only *GeBPs* were reported to be involved in plant trichome formation in previous studies [42,43,50], while their regulatory functions in tea plant trichome formation have not yet been studied. In the present study, a candidate *GeBP* gene (gene ID: CSS0010019.1; *CsGeBP4*) was expressed at higher levels in the hairy cultivar than in the hairless cultivar (Figure 2), indicating that members of the *GeBP* family might also be related to trichome formation in tea plants. Therefore, we further investigated the structural features of genes and proteins, phylogenetic relationships, and expression variations, as well as the functional patterns of *GeBP* family members, aiming to specify their regulatory roles in trichome growth and development in tea plants.

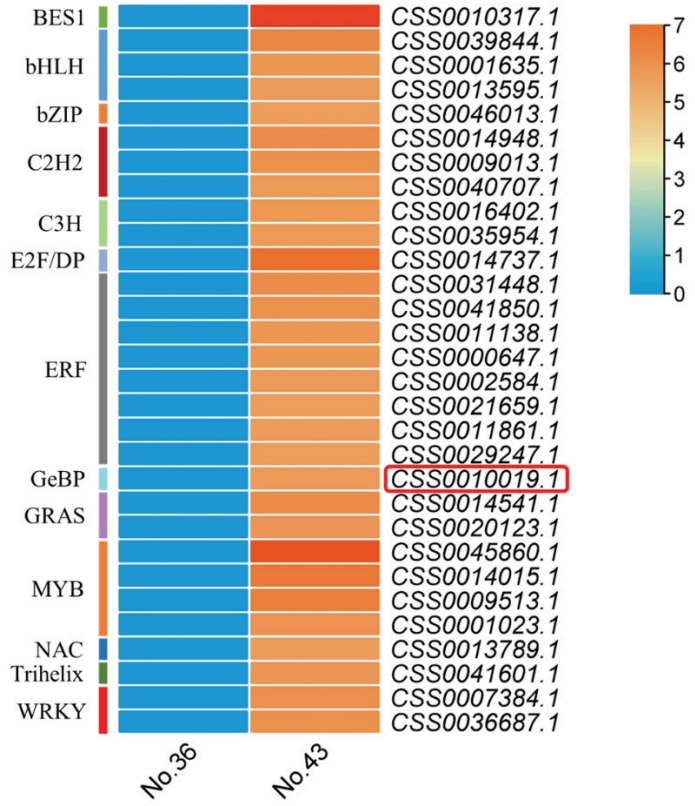


Figure 2. Heatmap of expression changes in putative transcription factors involved in trichome formation. Red indicates up-regulation, and blue indicates genes expressed at background levels. The scale bar represents the log2-based fold changes among two cultivars (hairless No. 36 and hairy No. 43) of Yunwu Tribute Tea. The gene ID of *CsGeBP4* is indicated with a red box.

2.2. Identification and Phylogenetic Analysis of GeBP Family in Tea Plants

A total of six *GeBP* genes were identified in the tea plant genome and named *CsGeBP1* to *CsGeBP6* based on their chromosomal localization order (Table 1). The biophysical and chemical properties of *GeBPs* are shown in Table 1. The genomic DNA size of *GeBPs* varied from 673 bp (CsGeBP2) to 2624 bp (CsGeBP4). The sequence lengths of GeBP proteins were between 221 (CsGeBP2) and 410 (CsGeBP4) amino acids. Their molecular weights and theoretical isoelectric points varied from 24.6 kDa (CsGeBP2) to 45.0 kDa (CsGeBP4) and 4.63 (CsGeBP3) to 9.85 (CsGeBP2), respectively. Furthermore, all of the CsGeBP members were localized to the nucleus based on the bioinformatics prediction.

Table 1. Physical and molecular properties of CsGeBPs identified in tea plant genome.

Gene ID	Gene Name	Chromosome No.	Length (bp)	Intron	Exon	Amino Acid (aa)	Molecular Weight (Da)	Isoelectric Point	Subcellular Localization
CSS0047893.1	CsGeBP1	Chr1	1363	0	1	370	40,951.80	8.28	Nucleus
CSS0008596.1	CsGeBP2	Chr1	673	1	2	221	24,638.69	9.85	Nucleus
CSS0012610.1	CsGeBP3	Chr6	1675	1	2	384	42,104.87	4.63	Nucleus
CSS0010019.1	CsGeBP4	Chr6	2624	1	2	410	45,031.43	4.65	Nucleus
CSS0022087.1	CsGeBP5	Chr7	881	1	2	283	31,931.65	9.40	Nucleus
CSS0046322.1	CsGeBP6	Chr14	1325	0	1	330	36,379.28	5.41	Nucleus

To explore the evolutionary relationships of GeBPs in tea plants with their homologs in other plant species, a multi-species phylogenetic tree of GeBPs from tea plants, *Arabidopsis*, rice, and soybean was constructed using the neighbor-joining method (Figure 3; Table S1). A total of 44 GeBPs were divided into four major groups. The six CsGeBPs only appeared in two groups, and close homology was found in three pairs of CsGeBPs (CsGeBP1 and

CsGeBP2; CsGeBP3 and CsGeBP4; CsGeBP5 and CsGeBP6). Group I contained the largest numbers of GeBPs, including six GmGeBP, four CsGeBPs, four OsGeBPs, and two AtGeBPs, accounting for 36% of the total GeBPs. Fifteen GeBPs were classified into Group II, with ten, three, and two GeBPs from *Arabidopsis*, soybean, and tea plants, respectively. None of the OsGeBPs were included in this group, indicating that GeBP family members in Group II might be specific to dicots. Group III and IV consisted of four AtGeBPs and nine OsGeBPs, respectively, suggesting that unequal loss and expansion of GeBPs might appear during species differentiation.

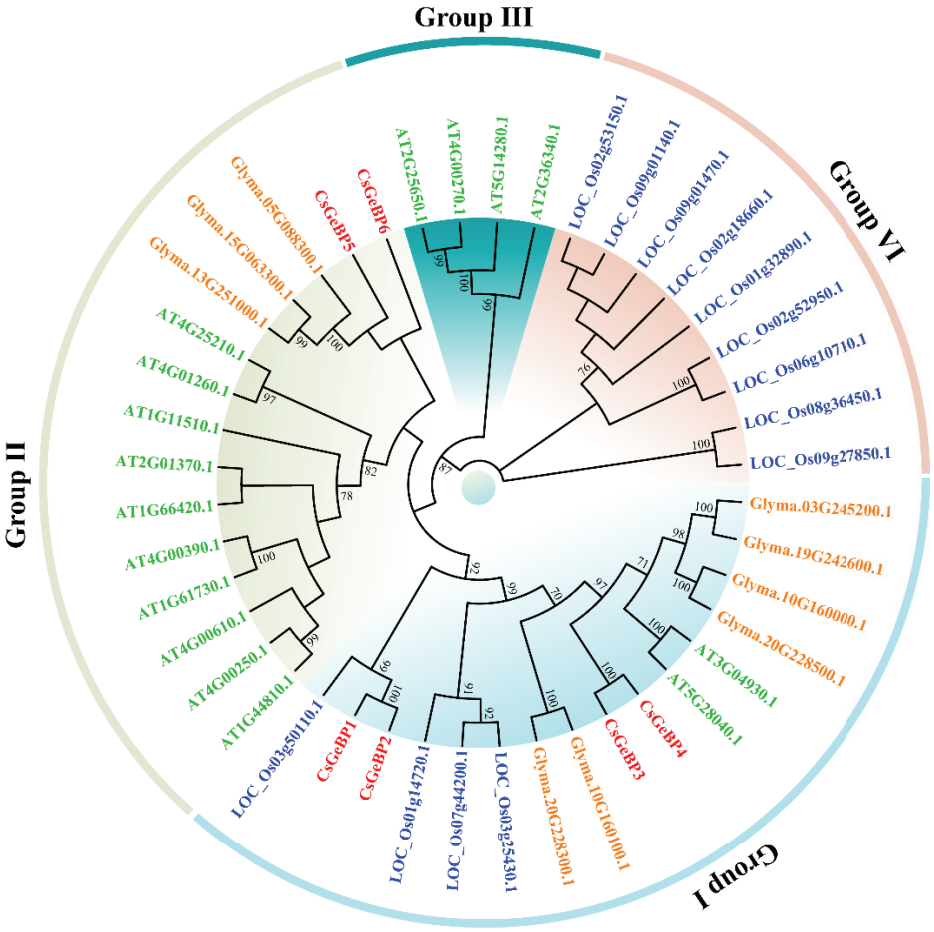


Figure 3. Phylogenetic relationships and classifications of GeBP proteins from tea plants (*Camellia sinensis*), *Arabidopsis thaliana*, rice (*Oryza sativa*), and soybean (*Glycine max*). The phylogenetic tree was constructed via MEGA X, using the neighbor-joining (NJ) method and 1000 bootstrap replicates. All of the GeBPs are divided into four groups. Gene IDs are indicated in different colors representing proteins from different plant species. Red represents tea plants, green represents *Arabidopsis*, blue indicates rice, and orange represents soybean.

2.3. Chromosomal Distribution and Gene Structure of CsGeBPs

The *CsGeBP* gene sequences were mapped onto the tea plant genome to investigate their chromosomal localization. The physical position of each *CsGeBP* gene on the tea plant chromosome is marked in Figure 4. The six *CsGeBP* genes were unevenly distributed on four out of fifteen chromosomes. In detail, *CsGeBP1* and *CsGeBP2* were localized on chromosome 1, while *CsGeBP3* and *CsGeBP4* were distributed on chromosome 6. In addition, *CsGeBP5* and *CsGeBP6* were mapped on chromosomes 7 and 14, respectively. These data provided extra evidence for the analysis of phylogenetic relationships among *CsGeBPs*.

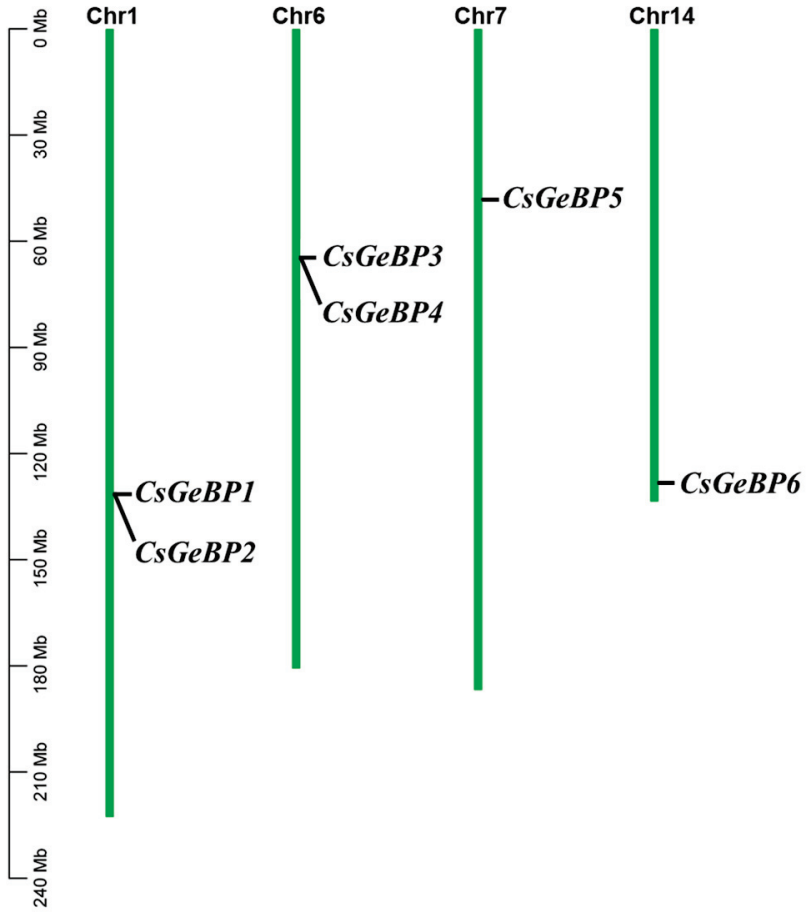


Figure 4. Distribution of *CsGeBP* genes on the tea plant chromosomes. All of the six *CsGeBP* genes were localized to tea chromosomal regions.

The analysis of exon–intron organization was performed to understand the gene structural features of the *CsGeBP* family genes (Figure 5). *CsGeBP1*, *CsGeBP4*, and *CsGeBP6* possessed one exon, while *CsGeBP2*, *CsGeBP3*, and *CsGeBP5* contained two exons. Furthermore, *CsGeBP1* and *CsGeBP6* were found to be without introns, and the other four *CsGeBPs* contained only one intron, indicating that the gene structure of *CsGeBPs* might be stable and not prone to being alternatively spliced in the process of gene replication.

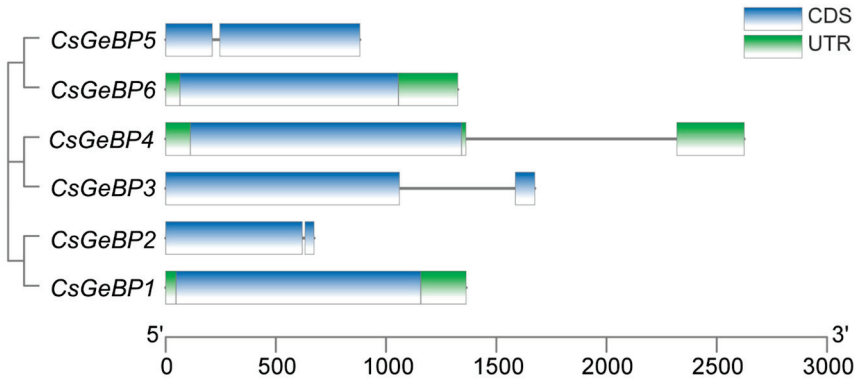


Figure 5. Exon–intron structure of *CsGeBPs*. Green boxes indicate untranslated 5′—and 3′—regions, blue boxes represent exons, and black lines indicate introns.

2.4. Conserved Motif Analysis of the CsGeBP Proteins

A total of fifteen distinct motifs were predicted in the CsGeBP family using the MEME program (Figure 6). The number of motifs contained in each CsGeBP ranged from 5 to 10. Both CsGeBP3 and CsGeBP4 contained ten conserved motifs; CsGeBP1 contained nine conserved motifs; and CsGeBP2, CsGeBP5, and CsGeBP6 contained five conserved motifs, respectively. Motif 1 was the most conservative motif due to its presence in all CsGeBPs, and Motifs 2, 3, 6, and 9 were widespread in at least four of the CsGeBPs, indicating these motifs were conserved during the evolution of gene families. CsGeBP3 and CsGeBP4 shared nine common motifs, and CsGeBP1 and CsGeBP2 contained five common motifs, which was highly consistent with their close phylogenetic relationships.

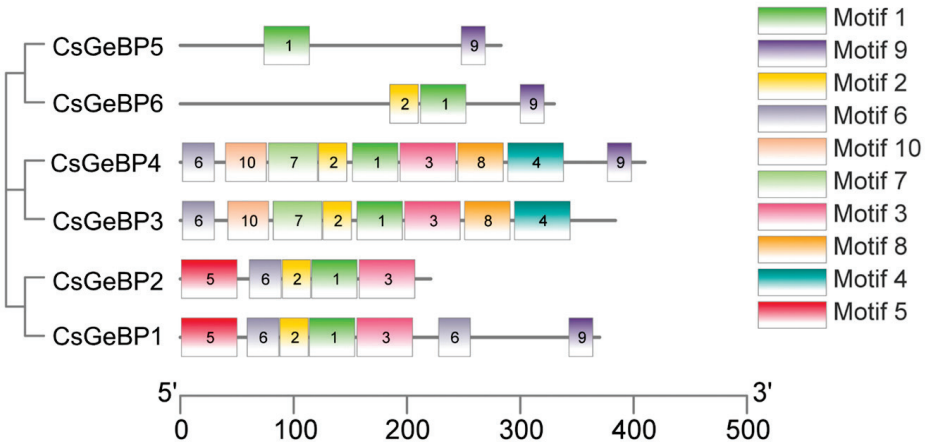


Figure 6. Conserved protein motif analysis of the CsGeBPs. Motifs 1 to 10 displayed by different colors represent different conserved protein motifs. The order of the motifs corresponds to their position within an individual protein sequence.

2.5. Cis-Regulatory Elements: Analysis of CsGeBP Gene Promoters

The modular composition of *cis*-regulatory elements in gene promoter regions plays key roles in regulating gene expression patterns in response to both internal signals and environmental factors. The promoter sequences of *CsGeBP* genes were analyzed to identify *cis*-regulatory elements (Figure 7; Table S2). A total of 161 *cis*-regulatory elements representing 29 non-redundant elements were found in the promoter regions of *CsGeBP* genes, and they were divided into four main groups: light response (39.8%), stress response (34.8%), hormone response (21.7%) and plant developmental regulation (3.7%) (Figure 7A). In the hormone-responsive group, elements associated with MeJA response were highly enriched in *CsGeBP* promoters, followed by ABA- and auxin-responsive elements (Figure 7B). Among all the environmental factors, light might be the most influential factor in the expression of *CsGeBPs*, followed by anoxic stress (Figure 7A,C). In addition, three elements involved in meristem expression, circadian control, and photosynthesis, respectively, were also found in *CsGeBP* promoters (Figure 7D). These results were consistent with the multiple roles of *GeBPs* in plant responses to phytohormones and environmental factors, as well as in plant growth reported by previous studies [38,41–43,45–49]. Furthermore, the composition modes of *cis*-regulatory elements were similar between *CsGeBP1* and *CsGeBP2*, as well as *CsGeBP3* and *CsGeBP4*, respectively, which was consistent with their corresponding phylogenetic relationships (Table S2).

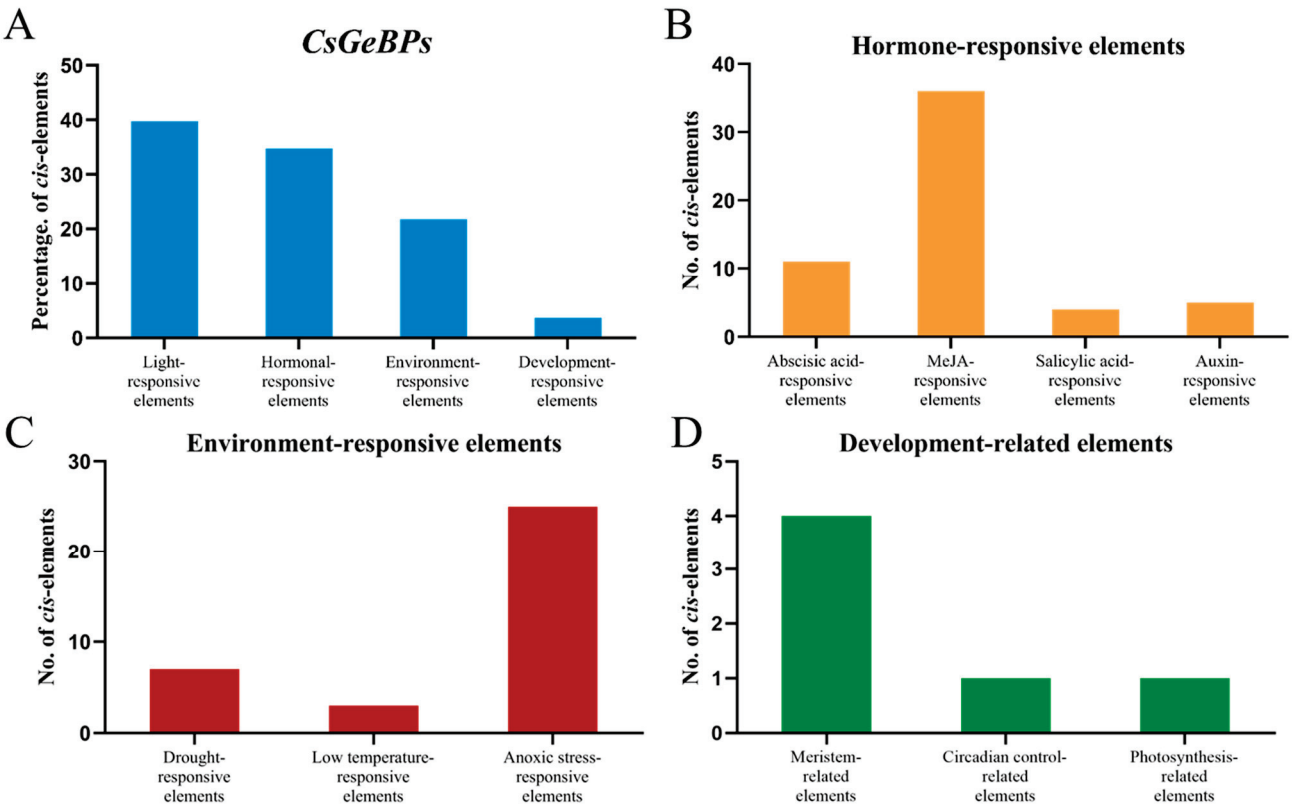


Figure 7. Analysis of the *cis*-regulatory elements in the promoter regions of *CsGeBP* genes. (**A**) The percentage of light-responsive elements, hormone-responsive elements, environment-responsive elements, and plant-growth-related elements in all *CsGeBP* family members. (**B**) Different hormone (ABA, MeJA, auxin, salicylic acid)-responsive elements in the *cis*-element regions of *CsGeBP* genes. (**C**) Different environmental stress (drought, low temperature, and anoxic stress)-responsive elements in *cis*-element regions of *CsGeBP* genes. (**D**) Different plant development-related elements in *cis*-element regions of *CsGeBP* genes.

2.6. Expression Pattern Analysis of CsGeBP Gene in Tea Plants

The expression patterns of the *CsGeBP* genes identified using publicly available RNA-Seq data from eight tissues (apical bud, young, mature, and old leaves, stem root, flower, and fruit) of the tea plant 'Shuchazao' were analyzed (Figure 8) [54,55]. *CsGeBP4* and *CsGeBP5* were expressed at higher levels in almost all tissues compared with other gene members (Figure 8). *CsGeBP3* was highly enriched in old leaf, root, flower, and fruit, while the expression levels of *CsGeBP1*, *CsGeBP2*, and *CsGeBP6* were lower in most tested tissues in comparison with other genes (Figure 8). These findings suggested that *CsGeBP3*, *CsGeBP4*, and *CsGeBP5* might play critical regulatory roles during tea plant growth and development. However, the expression patterns of these genes were not well-correlated to the tender degree of leaves, as well as the general distribution patterns of tea trichomes, indicating that the regulatory roles of *CsGeBP* genes in tea trichome formation might be distinct among different tea plant varieties.

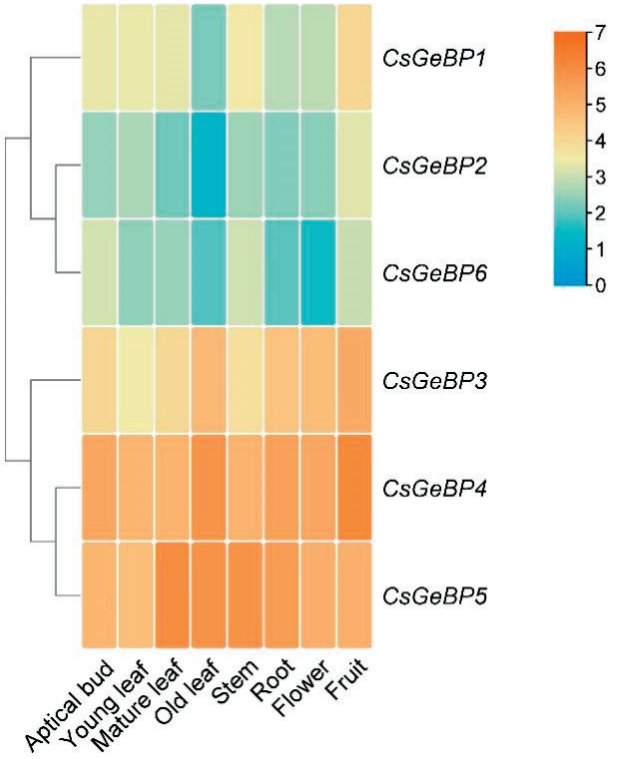


Figure 8. Expression profiles of *CeGeBPs* in different tissues of tea plants. Log2-based fold changes were used to create the heatmap based on the RNA-seq data downloaded from TPIA (http://tpia.teaplant.org, accessed on 13 June 2022). The gene expression level is displayed in different colors on the map, as shown in the bar at the upper right corner.

To elucidate the expression profiles of the *CsGeBP* genes that are responsive to environmental factors, their expression levels in response to drought and cold treatments were analyzed using the downloaded abiotic stress-responsive transcriptome data (Figure 9) [54,56–58]. *CsGeB2*, *CsGeB3*, and *CsGeBP4* showed similar change patterns, being slightly induced in response to drought stress for 72 h, although the expression level of *CsGeB2* was lower than the expression levels of *CsGeB3* and *CsGeBP4* under both control and treatment conditions (Figure 9A). *CsGeB5* displayed a slight decrease after treatment with drought stress for 72 h, while *CsGeB1* and *CsGeBP6* showed no significant changes in response to drought stress and were expressed at low levels in both control and stress-treated plants (Figure 9A). Furthermore, *CsGeB2*, *CsGeB3*, *CsGeB4*, and *CsGeBP5* were all induced by cold stress, while the expression level of *CsGeB2* was lower than the other three genes in both control and stressed plants (Figure 9B). Moreover, the expression levels of *CsGeB1* and *CsGeBP6* were not significantly affected by cold treatments (Figure 9B). These results suggest that *CsGeBP* family members might play critical roles in tea plant defense against environmental stresses.

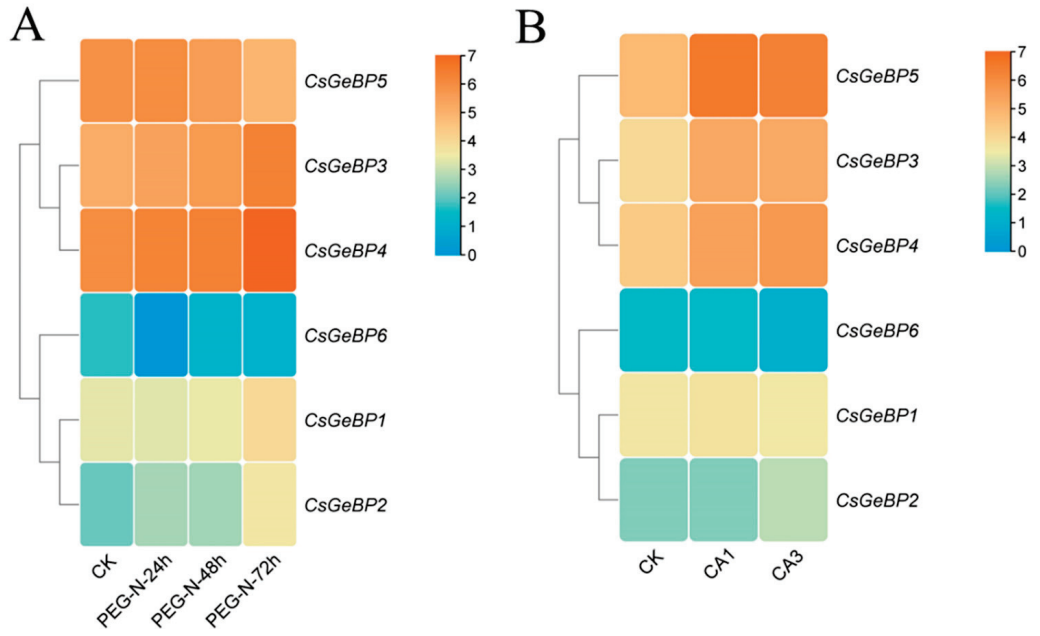


Figure 9. Expression profiles of *CsGeBPs* under drought (**A**) and cold (**B**) stresses. Log2-based fold changes were used to create the heatmap based on the RNA-seq data downloaded from TPIA (http://tpia.teaplant.org, accessed on 13 June 2022). The gene expression levels are displayed in different colors on the map, as shown in the bar at the upper right corner. Drought treatments were conducted by applying 25% polyethylene glycol (PEG) treatment for 0, 24, 48, and 72 h [56]. Cold treatments included non-acclimated (CK), fully acclimated (CA1), and de-acclimated stages (CA3) [57].

2.7. CsGeBP4 was Highly Related to the High-Dense Trichome Phenotype of Tea Plants

To confirm the regulatory roles of *CsGeBPs* in the trichome formation, the expression levels of *CsGeBPs* in the apical buds of different cultivars of Yunwu Tribute Tea were verified using qRT-PCR (Figure 10A). *CsGeBP1*, *CsGeBP3*, and *CsGeBP4* were highly expressed in tea plant cultivars with relatively higher TDIs (Figures 10A and 1B,C). Among them, the expression level of *CsGeBP4* showed the highest correlation with the high-density trichome phenotype ($r = 0.98$, $p = 0.001$) (Figure 10B). We further tested the tissue expression patterns of *CsGeBP4* by using apical buds, the first and the second leaves of three representative cultivars. The transcript level of *CsGeBP4* in each tea plant cultivar was higher on the apical buds, where trichomes are mostly present, than those on the first and second leaves, although the expression level of *CsGeBP4* was similar between the first and the second leaves (Figure 10C). These results strongly indicate that *CsGeBP4* could be a critical regulator involved in trichome formation in tea plants.

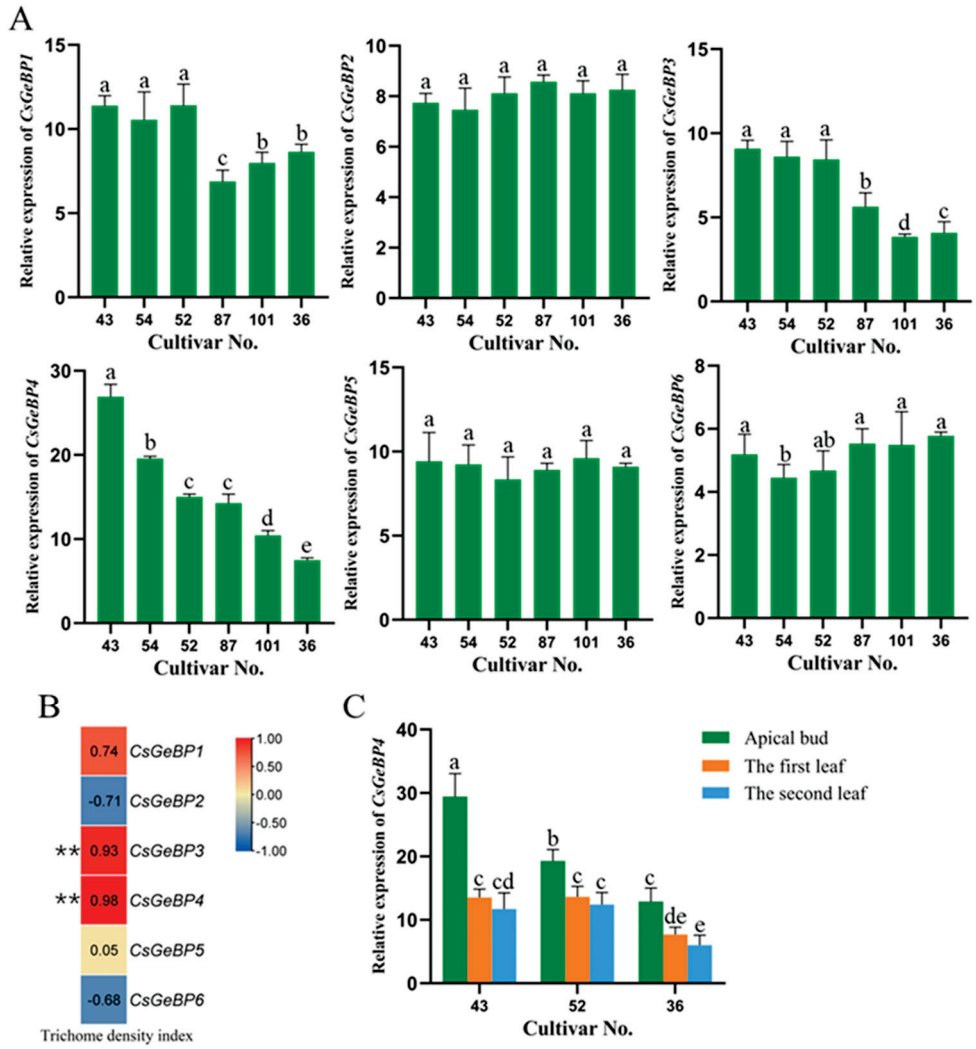


Figure 10. Expression profiles of *CsGeBPs* and their correlation with the tea plant trichome phenotypes. (**A**) Expression levels of *CsGeBPs* in the apical buds of different cultivars of Yunwu Tribute Tea. (**B**) Pearson correlation analysis of the expression of *CsGeBPs* and the trichome density index of tea plants. "**" represents a significantly high degree of correlation ($p < 0.01$). Red indicates a positive, and blue represents a negative correlation. (**C**) The expression level of *CsGeBP4* in the apical buds, the first and the second leaves of different cultivars of Yunwu Tribute Tea. Lowercase letters indicate significant differences among different samples ($p < 0.05$).

2.8. Silencing of CsGeBP4 Inhibited Trichome Development in Tea Plants

The in vivo function of *CsGeBP4* in trichome formation was further verified using the newly developed VIGS technology in the tea plants [40]. The apical buds of *CsGeBP4*-silenced tea plants displayed defective in trichome development, and their TDIs were significantly reduced by 2.3-fold compared with wild-type and control plants (Figure 11A,B). The transcript level of *CsGeBP4* was significantly repressed in *CsGeBP4*-silenced tea plants compared with wild-type and control plants, as verified using qRT-PCR (Figure 11C). These results strongly suggest that *CsGeBP4* was required for tea trichome formation.

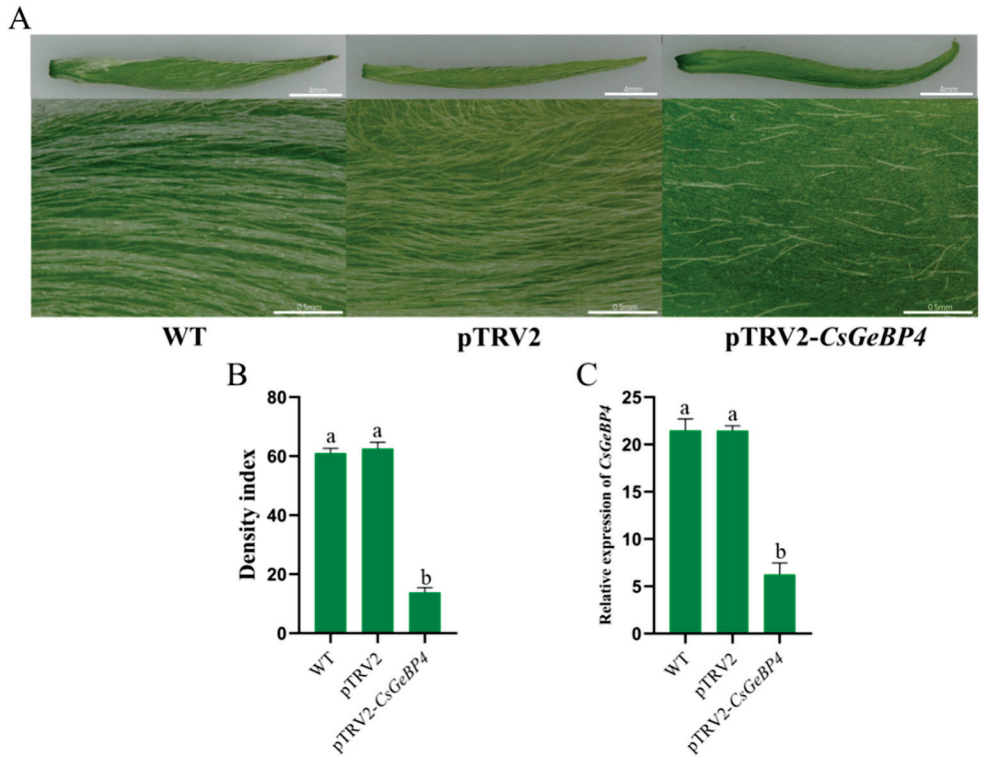


Figure 11. TRV-mediated gene silencing of the *CsGeBP4* in the apical buds of tea plants. Trichome phenotype (**A**), trichome density index (**B**), and the expression level of *CsGeBP4* in the tea cuttings from apical buds of tea plant variety, 'Fudingdabai' (**C**). WT, wild-type tea cuttings; pTRV2, tea cuttings infected by pTRV1 + pTRV2 *Agrobacterium*; pTRV2-*CsGeBP4*, tea cuttings infected by pTRV1 + pTRV2-*CsGeBP4 Agrobacterium*. Lowercase letters indicate significant differences among different samples. TRV, tobacco rattle virus; WT, wild type.

3. Discussion

Tea plant trichomes contribute to the unique flavor and nutritional quality of tea and protect tea plants from various environmental stresses by producing numerous metabolites, such as catechins, theanine, caffeine, flavonoids, and volatiles [1]. However, the regulatory mechanism underlying trichome initiation, growth, and development in tea plants is not fully understood. By integrating morphological, transcriptomic, bioinformatics, as well as qRT-PCR- and VIGS-mediated gene silencing strategies, we identified and characterized the *CsGeBP* gene family potentially involved in trichome formation and further specified the regulatory role of *CsGeBP4* during trichome growth and development.

3.1. Association of CsGeBP Gene Family with High-Dense Trichome Phenotypes in Tea Plants

Tea trichomes are a unicellular, unbranched, and non-glandular type of surface hairs [1,37]. Many studies have demonstrated that fresh buds and tender leaves had higher trichome density than older leaves [1,6,37]. Apart from the impact of tenderness degree, tea plant variety and cultivar also affected tea trichome density [4]. Most modern tea plant cultivars, including both assamica and sinensis types of *C. sinensis*, have a higher density of trichomes than many wild tea relative plants in Section *Thea*, such as *C. tachangensis*, *C. taliensis*, and *C. angustifolia*, most of which have glabrous leaves or low trichome density [1,4,59]. Therefore, densely spaced trichomes on the apical buds and young leaves were considered to be one of the most important domestication traits for better tea flavor and plant development during the millennia-long history of tea plant cultivation. In the present study, a morphological investigation showed that the density index of tea trichomes was continuously distributed among 108 different cultivars of Yunwu Tribute Tea (Figure 1A). The trichome density was much higher in the hairy cultivars in comparison

with hairless cultivars (Figure 1B,C), indicating the cultivar-specific trichome phenotypes existed within Yunwu Tribute Tea, and they are considered to be suitable materials for functional studies on tea plant trichome formation.

Tea plant trichome initiation, growth, and development are regulated by numerous genes, especially transcription factors. The previous comparative transcriptomics studies conducted in hairy and hairless tea plants revealed that a wide array of transcription factors display differential expression patterns and are thought to participate in the regulation of tea trichome formation [2,5]. These transcription factors belong to various gene families, mainly AP2/ERF, bHLH, bZIP, C2H2, GRAS, GRF, HB, MYB, NAC, SPL, TCP, Trihelix, WD40, WRKY, etc. [2,5]. Among these, the regulatory functions of transcription factors from MYB, bHLH, and WD40 gene families have been widely studied. Some members of these families can form an MYB-bHLH-WD40 protein complex to activate the downstream regulator genes *GL2*, *CPC*, and other effectors, thereby regulating trichome growth and development [4,5,23]. In this study, a comparative transcriptomics analysis was conducted for two cultivars of Yunwu Tribute Tea. Great differences in trichome density revealed that *CsGeBP4* was among the top 30 up-regulated transcription factor genes with the highest fold changes in expression levels in the hairy tea cultivar compared with the hairless cultivar (Figure 2). The potential relationship between the *GeBP* family and plant trichome formation has been reported in other plants, such as *Arabidopsis* and soybean [42,43,50]. However, the function of *CsGeBPs* in tea trichome growth and development is still elusive. Therefore, further analysis of *CsGeBP* gene family characteristics is necessary for uncovering the regulatory functions in tea plant trichome formation.

3.2. Genome-Wide Analysis of GeBP Gene Family and Their Involvement in Tea Plant Growth and Environmental Responses

GeBP gene family has been identified and characterized in soybean [50], mango [38], moso bamboo [51], tomato [52], and nine Gramineae crops [45]. However, so far, the systematic analysis of *GeBPs* has not been reported in tea plants. Herein, the genome-wide identification and characterization of *GeBP* family members were performed to provide clues about their potential functions in tea trichome formation. A total of six *CsGeBPs* distributed in four out of fifteen chromosomes were classified into two main phylogenetic groups (Figures 3 and 4). The multi-species phylogenetic analysis showed that the number of *GeBPs* found in the genomes of different plant species varied in each group. *GeBPs* in Group I were expected to be conserved across all four species due to the findings that this group included *GeBPs* from all the tested species (Figure 3). *GeBPs* in Group II, III, and VI exhibited dicot-, *Arabidopsis*-, and rice-specific phylogenetic clustering patterns, indicating that the *GeBP* gene family evolved in multiple directions among these species (Figure 3). The investigation of gene structure features revealed that the number of introns in *CsGeBPs* ranged from zero to one (Figure 5). It was reported that *GeBPs* in soybean also contained no more than one intron [50]. Our results indicate that the gene structure of *GeBPs* was relatively stable and not prone to experience alternative splicing during gene replication. In addition, the analysis of conserved protein motifs revealed that closely related CsGeBP proteins on adjoining branches of the phylogenetic tree had similar motif constituent (Figure 6), which was consistent with the findings in soybean [50] and moso bamboo [51].

The *cis*-regulatory elements in the gene promoters may provide valuable information for further investigating the function of the *CsGeBP* gene family due to their critical roles in the regulation of developmental- and environmental-related gene expression patterns. In this study, the *cis*-regulatory elements of *CsGeBPs* were found to participate in the responses to various environmental factors, including light, drought, low temperature, and anoxic induction (Figure 7A,C). Moreover, the analysis of phytohormone-responsive elements indicated that MeJA, ABA, auxin, and salicylic acid could also impact the expression patterns of *CsGeBPs* (Figure 7B). In addition, the presence of tea-plant-development-related elements in the *CsGeBPs* promoters indicated their potential roles in the regulation of

tea plant growth and development (Figure 7D). The tissue-specific and environmental-responsive expression patterns of *CsGeBP* genes revealed using RNA-seq data provided further evidence for these speculations (Figures 8 and 9). *CsGeBP3* and *CsGeBP4* were up-regulated in response to both drought and cold treatments and showed relatively high levels in almost all tested tissues, which could be partially attributed to drought- and cold-responsive *cis*-elements, as well as development-related *cis*-elements in their gene promoters (Figures 7–9, Table S2). These results facilitate a further understanding of the regulatory roles of *CsGeBPs* in tea trichome formation and stress responses.

3.3. CsGeBP4 Is Required for Tea Plant Trichome Formation

Previous studies reported that an AtGeBP could bind to the *cis*-regulatory elements in the promoter of trichome-formation-related gene *AtGL1* in yeast and in vitro, and was regulated by the meristem establishment gene, *KNAT1* [42,43], suggesting its possible regulatory roles in plant trichome formation. In order to further verify the involvement of *CsGeBPs* in tea plant trichome formation, the association analyses of expression patterns of all *CsGeBP* genes in the tea plant genome with trichome phenotypes among different cultivars of Yunwu Tribute Tea revealed that *CsGeBP4*, with the highest positive correlation efficiency, was strongly related to the high-density trichome phenotype (Figure 10A,B). Moreover, the expression level of *CsGeBP4* was much higher in the apical buds than that in the first and second leaves, which was well-supported by previous findings that apical buds in tea plants had more trichomes than older leaves (Figure 10C) [1,6,37]. These results indicate that *CsGeBP4* might play a critical and potential positive regulatory role in tea trichome growth and development. To verify this speculation, the gene expression of *CsGeBP4* was silenced in the tea plants via the newly developed VIGS strategy [40]. Silencing of *CsGeBP4* led to the repressed gene expression and defect trichome development, as well as the significantly decreased trichome density in the *CsGeBP4*-silenced tea plants, which accurately supported our previous hypothesis and further demonstrated that *CsGeBP4* was required for tea trichome formation. However, the expression of *GmGeBP4* identified in soybean was found to be induced in abnormal trichome soybean with fewer trichomes in comparison to control plants [50], suggesting that the regulatory patterns of *GeBP* genes were distinct across plant species. Therefore, more functional investigations are needed to identify the upstream and downstream genes and/or proteins associated with *GeBPs* and uncover the specific regulatory mechanisms in both tea plants and other plant species.

4. Materials and Methods

4.1. Plant Materials and Morphological Observations

Yunwu Tribute Tea (*Camellia sinensis* (L.) Kuntze var. niaowangensis Q. H. Chen) plants were grown under standard field conditions at a tea plantation in Yunwu town in Guizhou Province in China (latitude 26°17′ N, longitude 107°03′ E, and altitude 1200 m above mean sea level). Tea plant variety 'Fudingdabai' were grown under standard field conditions at the experimental farm of Guizhou Academy of Agricultural Sciences (latitude 26°11′ N, longitude 106°27′ E, and altitude 1185 m above mean sea level, Guiyang, China). Apical buds, first leaves, and second leaves of five plants were sampled as one biological replicate in early spring of 2022. Three biological replicates of different tissues were either used fresh or immediately frozen in liquid nitrogen and stored at −80 °C. The trichome phenotype of apical buds of the selected cultivars was observed using a digital microscope system (VHX-6000, KEYENCE, Osaka, Japan). The density index of trichomes was further analyzed using ImageJ software (https://imagej.en.softonic.com/, accessed on 10 April 2022).

4.2. Identification of GeBP Family Genes in the Tea Plant Genome

To identify the GeBP gene family in tea plants, we used *Arabidopsis* AtGeBP amino acid sequences to search against tea proteomes downloaded from Tea Plant Information Archive (TPIA) (http://tpia.teaplant.org/, accessed on 13 June 2022) using a Basic Local Alignment Search Tool (BLAST-P) (E-value < 10^{-5}) and deleting redundant sequences. Candidate

CsGeBP protein sequences were submitted to the online bioinformatics tool, the National Center of Biotechnology Information Conserved Domain-Search Tool (NCBI CD-Search Tool) (https://www.ncbi.nlm.nih.gov/Structure/index.shtml, accessed on 13 June 2022) (E-value < 10^{-2}) to verify their conserved domains. The number of amino acid residues, gene structure (exon/intron arrangement), and start-to-end position of each gene in the genome were analyzed based on the annotated information downloaded from the tea plant genome database. Physical parameters of each gene product, such as molecular weight and isoelectric point, were predicted using ExPASy (http://web.expasy.org/compute_pi/, accessed on 15 June 2022). The subcellular localization of each CsGeBP was predicted by WoLF PSORT (https://wolfpsort.hgc.jp/, accessed on 15 June 2022).

4.3. Phylogenetic Analysis of GeBPs among Different Plant Species

An evolutionary relationships analysis of GeBP proteins from tea plants, *Arabidopsis*, rice, and soybean was carried out using MEGA X (Mega Limited, Auckland, New Zealand) [60] based on the neighbor-joining method [61]. Bootstrap test replicates were set to 1000 times [62]. The phylogenetic tree was displayed using FigTree software (version 1.4.2). Protein sequences of *Arabidopsis*, rice, and soybean were obtained from The *Arabidopsis* Information Resource (TAIR) (https://www.arabidopsis.org/, accessed on 13 June 2022), Rice Genome Annotation Project (RGAP) (http://rice.plantbiology.msu.edu/index.shtml, accessed on 13 June 2022), and SoyBase (http://www.soybase.org/, accessed on 13 June 2022), respectively.

4.4. Chromosomal Distribution, Gene Structure, and Conserved Motif Analysis

The chromosomal localization information (including chromosomal distribution, length, as well as the start and end positions) of *CsGeBP* genes was downloaded from TPIA (http://tpia.teaplant.org/, accessed on 13 June 2022) and visualized by TBtools. The exon–intron structures of the CsGeBP genes based on the genome annotations were visualized using TBtools software. The conserved protein motifs of CsGeBP proteins were analyzed through the Multiple Em for Motif Elicitation (MEME) program (http://meme-suite.org/, accessed on 14 June 2022) and visualized using TBtools.

4.5. Identification of Putative Cis-Regulatory Elements in the Promoters of CsGeBPs

The promoter sequences (2000 bp upstream of the start codon) of all *CsGeBP* genes were obtained from TPIA (http://tpia.teaplant.org/, accessed on 13 June 2022), analyzed by using the PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 14 June 2022), and visualized by TBtools.

4.6. Expression Analysis of CsGeBPs in Different Tissues and in Response to Environmental Stresses

The transcriptome data (TPM value) of *CsGeBPs* in eight tissues (apical bud, young leaf, mature leaf, old leaf, stem, root, flower, and fruit) of tea plant cultivar 'Shuchazao' and transcriptome data in 'Longjing43' and 'Tieguanyin' in response to drought and cold stresses, respectively, were downloaded from TPIA (http://tpia.teaplant.org, accessed on 13 June 2022). The log2-based fold changes were used to create heatmaps by TBtools.

4.7. RNA Extraction and qRT-PCR Analysis

Primers for *CsGeBP* gene cloning were designed using the online program Integrated DNA Technologies (IDT) (https://sg.idtdna.com/pages, accessed on 5 July 2022) (Supplementary Table S3), and the primer sequences were synthesized by Beijing Qingke Biotechnology limited company. Total RNA was extracted from apical buds, and the second leaves were extracted using a cetyltrimethylammonium bromide (CTAB) method [63]. The RNA was reverse transcribed into first-strand cDNA using the PrimeScript^{TM} II first-strand cDNA Synthesis Kit (Solarbio Technology, Beijing, China). The cDNA was subsequently employed as a template for qRT-PCR analysis using SYBR Green qPCR Mix (Genenode,

Wuhan, China) reagent. Each reaction system contained 10 μL SYBR Green qPCR Mix, 0.8 μL primers, 1.5 μL cDNA template, 7.7 μL H_2O. The reaction process is as follows: 95 °C for 3 min; 40 cycles of 95 °C for 10 s and 60 °C for 20 s; 72 °C for 30 s. *CsGAPDH* was used as the reference gene. The relative expression level was calculated using the $2^{-\Delta\Delta CT}$ method [64].

4.8. Differentially Expressed Gene Analysis by Transcriptome Sequencing

Total RNA was extracted from apical buds of two cultivars (No. 43 and No. 36) of Yunwu Tribute Tea using the RNeasy Plus Mini kit (Qiagen, Valencia, CA, USA). Three independent biological replicates were used for RNA sequencing. RNA integrity was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The samples with RNA integrity number (RIN) $\geq$ 7 were submitted to enrich mRNA and construct cDNA libraries using TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, USA) according to the manufacturer's instructions. The libraries were sequenced using the Illumina HiSeq™ 2000 platform (Illumina, San Diego, USA). In order to obtain high-quality clean reads, adaptor sequences, empty reads, and low-quality bases (Q < 30) were removed. The resulting clean reads were subsequently used for transcriptome de novo assembly by mapping to the tea plant reference genome (http://tpia.teaplant.org/, accessed on 20 May 2022) [54,58]. Fragments per kilobase of transcript per million (FPKM) of each gene and read counts value of each transcript (protein_coding) were calculated using bowtie2 and eXpress. The differential expressions of genes between the two tea plant cultivars were analyzed using the DESeq (2012) R package. The FPKM values between two cultivars were compared using a threshold of FDR < 0.001 and $|\log2ratio| > 1$ to investigate differentially expressed genes.

4.9. VIGS-Based Gene Silencing in Tea Plants

Gene silencing of *CsGeBP4* using VIGS technology was performed in the tea plant variety 'Fudingdabai' as described previously [40]. Briefly, a 292 bp fragment of *CsGeBP4* used for VIGS was assembled into the pTRV2 virus vector to construct the pTRV2-*CsGeBP4* vector. Then, pTRV1, pTRV2, and pTRV2-*CsGeBP4* were transformed into *Agrobacterium tumefaciens* strain GV3101, respectively. After cultivation and resuspension, *Agrobacterium* harboring pTRV1 were mixed with those harboring pTRV2 or pTRV2-*CsGeBP4*, respectively, and they were both infiltrated into tea plant cuttings via vacuum infiltration, respectively. The inoculated tea cuttings were kept in the dark for three days and then grown in a greenhouse at 25 °C under a 16 h/8 h light/dark cycle.

4.10. Statistical Analysis

All results are presented as means ± standard deviation of at least three biological replicates. The data were subjected to one-way analysis of variance using SPSS software (version 26.0). *p*-values less than 0.05 were considered statistically significant.

5. Conclusions

In this study, we reported the association of *CsGeBP* family members with high-density trichome phenotypes in tea plants. Our genome-wide analysis of the *CsGeBP* gene family in tea plants and the expression patterns of *CsGeBPs* in multiple tissues in response to environmental factors revealed valuable information for understanding the potential biological roles of *CsGeBPs* in tea trichome formation. More importantly, we demonstrated that *CsGeBP4* was required for tea trichome formation, potentially for positive regulation. This study provides new insights into the understanding of tea trichome formation and lays a foundation for future research to improve the flavor and quality of tea products and breed stress-tolerant tea plant cultivars.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24065207/s1.

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

Funding: This research was funded by the Guizhou Province Science and Technology Planning Project, grant number Qiankehe Support (2021) General 111; Natural Science Foundation of China, grant number 3226180451; Education Department of Guizhou Province Scientific Research Program of Higher Education Institutions (Young Scientific Program), grant number Qianjiaoji (2022) No. 118; Research Funds for Introduced Talents of Guizhou University, grant number Guizhou University renjihezi (2021) No. 34.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Li, P.; Xu, Y.; Zhang, Y.; Fu, J.; Yu, S.; Guo, H.; Chen, Z.; Chen, C.; Yang, X.; Wang, S.; et al. Metabolite profiling and transcriptome analysis revealed the chemical contributions of tea trichomes to tea flavors and tea plant defenses. *J. Agric. Food Chem.* **2020**, *68*, 11389–11401. [CrossRef] [PubMed]
2. Yue, C.; Cao, H.-L.; Chen, D.; Lin, H.-Z.; Wang, Z.; Hu, J.; Yang, G.-Y.; Guo, Y.-Q.; Ye, N.-X.; Hao, X.-Y. Comparative transcriptome study of hairy and hairless tea plant (*Camellia sinensis*) shoots. *J. Plant Physiol.* **2018**, *229*, 41–52. [CrossRef]
3. Sun, B.; Zhu, Z.; Liu, R.; Wang, L.; Dai, F.; Cao, F.; Liu, S. TRANSPARENT TESTA GLABRA1 (TTG1) regulates leaf trichome density in tea *Camellia sinensis*. *Nord. J. Bot.* **2020**, *38*, 1–10. [CrossRef]
4. Li, P.; Fu, J.; Xu, Y.; Shen, Y.; Zhang, Y.; Ye, Z.; Tong, W.; Zeng, X.; Yang, J.; Tang, D.; et al. CsMYB1 integrates the regulation of trichome development and catechins biosynthesis in tea plant domestication. *New Phytol.* **2022**, *234*, 902–917. [CrossRef] [PubMed]
5. Chen, L.; Tian, N.; Hu, M.; Sandhu, D.; Jin, Q.; Gu, M.; Zhang, X.; Peng, Y.; Zhang, J.; Chen, Z.; et al. Comparative transcriptome analysis reveals key pathways and genes involved in trichome development in tea plant (*Camellia sinensis*). *Front. Plant Sci.* **2022**, *13*, 997778. [CrossRef] [PubMed]
6. Li, J.; Zeng, L.; Liao, Y.; Tang, J.; Yang, Z. Evaluation of the contribution of trichomes to metabolite compositions of tea (*Camellia sinensis*) leaves and their products. *LWT* **2020**, *122*, 109023. [CrossRef]
7. Balkunde, R.; Pesch, M.; Hülskamp, M. Trichome Patterning in *Arabidopsis thaliana* from genetic to molecular models. *Curr. Top. Dev. Biol.* **2010**, *91*, 299–321. [CrossRef]
8. Bleeker, P.M.; Mirabella, R.; Diergaarde, P.J.; VanDoorn, A.; Tissier, A.; Kant, M.R.; Prins, M.; de Vos, M.; Haring, M.A.; Schuurink, R.C. Improved herbivore resistance in cultivated tomato with the sesquiterpene biosynthetic pathway from a wild relative. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20124–20129. [CrossRef]
9. Gonzales-Vigil, E.; Hufnagel, D.E.; Kim, J.; Last, R.L.; Barry, C.S. Evolution of TPS20-related terpene synthases influences chemical diversity in the glandular trichomes of the wild tomato relative *Solanum habrochaites*. *Plant J.* **2012**, *71*, 921–935. [CrossRef]
10. Luu, V.T.; Weinhold, A.; Ullah, C.; Dressel, S.; Schoettner, M.; Gase, K.; Gaquerel, E.; Xu, S.; Baldwin, I.T. *O*-Acyl sugars protect a wild tobacco from both native fungal pathogens and a specialist herbivore. *Plant Physiol.* **2017**, *174*, 370–386. [CrossRef]
11. Schilmiller, A.L.; Last, R.L.; Pichersky, E. Harnessing plant trichome biochemistry for the production of useful compounds. *Plant J.* **2008**, *54*, 702–711. [CrossRef] [PubMed]
12. Ning, P.; Wang, J.; Zhou, Y.; Gao, L.; Wang, J.; Gong, C. Adaptional evolution of trichome in *Caragana korshinskii* to natural drought stress on the Loess Plateau, China. *Ecol. Evol.* **2016**, *6*, 3786–3795. [CrossRef] [PubMed]
13. Wang, X.; Shen, C.; Meng, P.; Tan, G.; Lv, L. Analysis and review of trichomes in plants. *BMC Plant Biol.* **2021**, *21*, 70. [CrossRef] [PubMed]
14. Lloyd, A.M.; Schena, M.; Walbot, V.; Davis, R.W. Epidermal cell fate determination in *Arabidopsis*: Patterns defined by a steroid-inducible regulator. *Science* **1994**, *266*, 436–439. [CrossRef]
15. Chien, J.C.; Sussex, I.M. Differential regulation of trichome formation on the adaxial and abaxial leaf surfaces by gibberellins and photoperiod in *Arabidopsis thaliana* (L.) Heynh. *Plant Physiol.* **1996**, *111*, 1321–1328. [CrossRef]
16. Larkin, J.; Young, N.; Prigge, M.; Marks, M.D. The control of trichome spacing and number in Arabidopsis. *Development* **1996**, *122*, 997–1005. [CrossRef]
17. Telfer, A.; Bollman, K.; Poethig, R. Phase change and the regulation of trichome distribution in *Arabidopsis thaliana*. *Development* **1997**, *124*, 645–654. [CrossRef]
18. Perazza, D.; Vachon, G.; Herzog, M. Gibberellins promote trichome formation by up-regulating *GLABROUS1* in Arabidopsis. *Plant Physiol.* **1998**, *117*, 375–383. [CrossRef]

19. Cui, J.-Y.; Miao, H.; Ding, L.-H.; Wehner, T.C.; Liu, P.-N.; Wang, Y.; Zhang, S.-P.; Gu, X.-F. A new glabrous gene (*CsGl3*) identified in trichome development in cucumber (*Cucumis sativus* L.). *PLoS ONE* **2016**, *11*, e0148422. [CrossRef]
20. Yang, C.; Li, H.; Zhang, J.; Luo, Z.; Gong, P.; Zhang, C.; Li, J.; Wang, T.; Zhang, Y.; Lu, Y.; et al. A regulatory gene induces trichome formation and embryo lethality in tomato. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11836–11841. [CrossRef]
21. Zhang, J.; Lu, Y.; Yuan, Y.; Zhang, X.; Geng, J.; Chen, Y.; Cloutier, S.; McVetty, P.B.E.; Li, G. Map-based cloning and characterization of a gene controlling hairiness and seed coat color traits in *Brassica rapa*. *Plant Mol. Biol.* **2009**, *69*, 553–563. [CrossRef] [PubMed]
22. Gruber, M.; Alahakoon, U.; Taheri, A.; Nagubushana, N.; Zhou, R.; Aung, B.; Sharpe, A.; Hannoufa, A.; Bonham-Smith, P.; Hegedus, D.D.D. The biochemical composition and transcriptome of cotyledons from *Brassica napus* lines expressing the *AtGL3* transcription factor and exhibiting reduced flea beetle feeding. *BMC Plant Biol.* **2018**, *18*, 64. [CrossRef] [PubMed]
23. Wang, S.; Chen, J.-G. Arabidopsis transient expression analysis reveals that activation of *GLABRA2* may require concurrent binding of *GLABRA1* and *GLABRA3* to the promoter of *GLABRA2*. *Plant Cell Physiol.* **2008**, *49*, 1792–1804. [CrossRef] [PubMed]
24. Johnson, C.S.; Kolevski, B.; Smyth, D.R. *TRANSPARENT TESTA GLABRA2*, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *Plant Cell* **2002**, *14*, 1359–1375. [CrossRef]
25. Zhao, M.; Morohashi, K.; Hatlestad, G.; Grotewold, E.; Lloyd, A. The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. *Development* **2008**, *135*, 1991–1999. [CrossRef]
26. Tominaga-Wada, R.; Nukumizu, Y.; Sato, S.; Kato, T.; Tabata, S.; Wada, T. Functional divergence of MYB-Related genes, *WEREWOLF* and *AtMYB23* in Arabidopsis. *Biosci. Biotechnol. Biochem.* **2012**, *76*, 883–887. [CrossRef] [PubMed]
27. Zhao, H.; Wang, X.; Zhu, D.; Cui, S.; Li, X.; Cao, Y.; Ma, L. A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor *AtMYC1* leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in arabidopsis. *J. Biol. Chem.* **2012**, *287*, 14109–14121. [CrossRef]
28. Li, S.F.; Milliken, O.N.; Pham, H.; Seyit, R.; Napoli, R.; Preston, J.; Koltunow, A.M.; Parish, R.W. The Arabidopsis MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *Plant Cell* **2009**, *21*, 72–89. [CrossRef]
29. Kirik, V.; Lee, M.M.; Wester, K.; Herrmann, U.; Zheng, Z.; Oppenheimer, D.; Schiefelbein, J.; Hulskamp, M. Functional diversification of *MYB23* and *GL1* genes in trichome morphogenesis and initiation. *Development* **2005**, *132*, 1477–1485. [CrossRef]
30. Payne, C.T.; Zhang, F.; Lloyd, A.M. *GL3* encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with *GL1* and *TTG1*. *Genetics* **2000**, *156*, 1349–1362. [CrossRef]
31. Zhang, F.; Gonzalez, A.; Zhao, M.; Payne, C.T.; Lloyd, A. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development* **2003**, *130*, 4859–4869. [CrossRef] [PubMed]
32. Bernhardt, C.; Lee, M.M.; Gonzalez, A.; Zhang, F.; Lloyd, A.; Schiefelbein, J. The bHLH genes *GLABRA3 (GL3)* and *ENHANCER OF GLABRA3 (EGL3)* specify epidermal cell fate in the *Arabidopsis* root. *Development* **2003**, *130*, 6431–6439. [CrossRef]
33. Chopra, D.; Mapar, M.; Stephan, L.; Albani, M.C.; Deneer, A.; Coupland, G.; Willing, E.-M.; Schellmann, S.; Schneeberger, K.; Fleck, C.; et al. Genetic and molecular analysis of trichome development in *Arabis alpina*. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 12078–12083. [CrossRef] [PubMed]
34. Zhang, B.; Chopra, D.; Schrader, A.; Hülskamp, M. Evolutionary comparison of competitive protein-complex formation of MYB, bHLH, and WDR proteins in plants. *J. Exp. Bot.* **2019**, *70*, 3197–3209. [CrossRef] [PubMed]
35. Gao, Y.; Gong, X.; Cao, W.; Zhao, J.; Fu, L.; Wang, X.; Schumaker, K.S.; Guo, Y. SAD2 in *Arabidopsis* functions in trichome initiation through mediating *GL3* function and regulating *GL1*, *TTG1* and *GL2* Expression. *J. Integr. Plant Biol.* **2008**, *50*, 906–917. [CrossRef] [PubMed]
36. Szymanski, D.B.; Lloyd, A.M.; Marks, M. Progress in the molecular genetic analysis of trichome initiation and morphogenesis in Arabidopsis. *Trends Plant Sci.* **2000**, *5*, 214–219. [CrossRef]
37. Das, P.; Chettri, V.; Ghosh, S.; Ghosh, C. Micromorphological studies of the leaf and stem of *Camellia sinensis* (L.) Kuntze with reference to their taxonomic significance. *Microsc. Res. Tech.* **2022**. [CrossRef]
38. Ruiqing, S.; Nan, Y.; Zhixin, L.; Yuqi, X.; Yu, S.; Qingyuan, G.; Jinji, P.; Zhiguo, D.; He, Z. Identification and expression analysis of mango GeBP transcription factor family genes. *Non-wood For. Res.* **2021**, *39*, 137–147. [CrossRef]
39. Wakamatsu, J.; Wada, T.; Tanaka, W.; Fujii, S.; Fujikawa, Y.; Sambongi, Y.; Tominaga, R. Identification of six CPC-like genes and their differential expression in leaves of tea plant, *Camellia sinensis*. *J. Plant Physiol.* **2021**, *263*, 153465. [CrossRef]
40. Li, G.; Li, Y.; Yao, X.; Lu, L. Establishment of a virus-induced gene-silencing (VIGS) system in tea plant and its use in the functional analysis of *CsTCS1*. *Int. J. Mol. Sci.* **2022**, *24*, 392. [CrossRef]
41. Perazza, D.; Laporte, F.; Balagué, C.; Chevalier, F.; Remo, S.; Bourge, M.; Larkin, J.; Herzog, M.; Vachon, G. *GeBP/GPL* transcription factors regulate a subset of *CPR5*-dependent processes. *Plant Physiol.* **2011**, *157*, 1232–1242. [CrossRef] [PubMed]
42. Curaba, J.; Herzog, M.; Vachon, G. *GeBP*, the first member of a new gene family in Arabidopsis, encodes a nuclear protein with DNA-binding activity and is regulated by *KNAT1*. *Plant J.* **2003**, *33*, 305–317. [CrossRef] [PubMed]
43. Chuck, G.; Lincoln, C.; Hake, S. *KNAT1* induces lobed leaves with ectopic meristems when overexpressed in Arabidopsis. *Plant Cell* **1996**, *8*, 1277–1289. [CrossRef] [PubMed]
44. Dill, A.; Jung, H.-S.; Sun, T.-P. The DELLA motif is essential for gibberellin-induced degradation of RGA. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14162–14167. [CrossRef]

45. Huang, J.; Zhang, Q.; He, Y.; Liu, W.; Xu, Y.; Liu, K.; Xian, F.; Li, J.; Hu, J. Genome-wide identification, expansion mechanism and expression profiling analysis of *GLABROUS1* enhancer-binding protein *(GeBP)* gene family in Gramineae crops. *Int. J. Mol. Sci.* **2021**, *22*, 8758. [CrossRef]
46. Jasinski, S.; Piazza, P.; Craft, J.; Hay, A.; Woolley, L.; Rieu, I.; Phillips, A.; Hedden, P.; Tsiantis, M. KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr. Biol.* **2005**, *15*, 1560–1565. [CrossRef]
47. Yanai, O.; Shani, E.; Dolezal, K.; Tarkowski, P.; Sablowski, R.; Sandberg, G.; Samach, A.; Ori, N. Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Curr. Biol.* **2005**, *15*, 1566–1571. [CrossRef]
48. Chevalier, F.; Perazza, D.; Laporte, F.; Le Hénanff, G.; Hornitschek, P.; Bonneville, J.-M.; Herzog, M.; Vachon, G. GeBP and GeBP-Like proteins are noncanonical leucine-zipper transcription factors that regulate cytokinin response in Arabidopsis. *Plant Physiol.* **2008**, *146*, 1142–1154. [CrossRef]
49. Khare, D.; Mitsuda, N.; Lee, S.; Song, W.; Hwang, D.; Ohme-Takagi, M.; Martinoia, E.; Lee, Y.; Hwang, J. Root avoidance of toxic metals requires the GeBP-LIKE 4 transcription factor in *Arabidopsis thaliana*. *New Phytol.* **2017**, *213*, 1257–1273. [CrossRef]
50. Liu, S.; Liu, Y.; Liu, C.; Zhang, F.; Wei, J.; Li, B. Genome-wide characterization and expression analysis of GeBP family genes in Soybean. *Plants* **2022**, *11*, 1848. [CrossRef]
51. Xuemeng, S.; Kebin, Y.; Jingjing, S. Genome wide identification and expression analysis of GeBP transcription factor gene family in moso bamboo. *J. Nanjing For. Univ. (Nat. Sci. Ed.)* **2020**, *44*, 41–48. [CrossRef]
52. Kai, C.; Jinqiu, L.; Haihui, S.; Zhenzhu, Z.; Xiuling, C.; Wang, A. Identification, evolution and expression analysis of GeBP transcription factors family in Tomato. *Mol. Plant Breed.* **2017**, *15*, 3438–3445. [CrossRef]
53. Wang, Y.; Wan, C.; Li, L.; Xiang, Z.; Wang, J.; Li, Y.; Zhao, D. Transcriptomic analysis of Yunwu Tribute Tea leaves under cold stress. *Curr. Issues Mol. Biol.* **2023**, *45*, 699–720. [CrossRef] [PubMed]
54. Xia, E.; Li, F.; Tong, W.; Li, P.; Wu, Q.; Zhao, H.; Ge, R.; Li, R.; Li, Y.; Zhang, Z.; et al. Tea plant information archive: A comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol. J.* **2019**, *17*, 1938–1953. [CrossRef] [PubMed]
55. Wei, C.; Yang, H.; Wang, S.; Zhao, J.; Liu, C.; Gao, L.; Xia, E.; Lu, Y.; Tai, Y.; She, G.; et al. Draft genome sequence of *Camellia sinensis* var. sinensis provides insights into the evolution of the tea genome and tea quality. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4151–E4158. [CrossRef]
56. Zhang, Q.; Cai, M.; Yu, X.; Wang, L.; Guo, C.; Ming, R.; Zhang, J. Transcriptome dynamics of *Camellia sinensis* in response to continuous salinity and drought stress. *Tree Genet. Genomes* **2017**, *13*, 78. [CrossRef]
57. Wang, X.-C.; Zhao, Q.-Y.; Ma, C.-L.; Zhang, Z.-H.; Cao, H.-L.; Kong, Y.-M.; Yue, C.; Hao, X.-Y.; Chen, L.; Ma, J.-Q.; et al. Global transcriptome profiles of *Camellia sinensis* during cold acclimation. *BMC Genom.* **2013**, *14*, 415. [CrossRef]
58. Xia, E.; Tong, W.; Hou, Y.; An, Y.; Chen, L.; Wu, Q.; Liu, Y.; Yu, J.; Li, F.; Li, R.; et al. The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation. *Mol. Plant* **2020**, *13*, 1013–1026. [CrossRef]
59. Liu, Y.; Zhao, G.; Li, X.; Shen, Q.; Wu, Q.; Zhuang, J.; Zhang, X.; Xia, E.; Zhang, Z.; Qian, Y.; et al. Comparative analysis of phenolic compound metabolism among tea plants in the section Thea of the genus *Camellia*. *Food Res. Int.* **2020**, *135*, 109276. [CrossRef]
60. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
61. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef] [PubMed]
62. Felsenstein, J. Phylogenies and the comparative method. *Am. Nat.* **1985**, *125*, 1–15. [CrossRef]
63. Wang, L.; Stegemann, J.P. Extraction of high quality RNA from polysaccharide matrices using cetlytrimethylammonium bromide. *Biomaterials* **2010**, *31*, 1612–1618. [CrossRef] [PubMed]
64. 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] [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Transcriptome and Metabolome Reveal Distinct Sugar Accumulation Pattern between PCNA and PCA Mature Persimmon Fruit

Weijuan Han †, Yiru Wang †, Huawei Li, Songfeng Diao, Yujing Suo, Taishan Li, Peng Sun, Fangdong Li * and Jianmin Fu *

Research Institute of Non-Timber Forestry, Chinese Academy of Forestry, Zhengzhou 450003, China; hanweijuan2013@163.com (W.H.); wangyiru199702@163.com (Y.W.); lihuaweicaf@163.com (H.L.); dsf@caf.ac.cn (S.D.); suoyujing1988@126.com (Y.S.); litaishan910706@gmail.com (T.L.); sunpeng1017@126.com (P.S.)

* Correspondence: lifangdong@caf.ac.cn (F.L.); fujianmin@caf.ac.cn (J.F.)

† These authors contribute equally to this work.

Abstract: Persimmon (*Diospyros kaki*) fruit have significant variation between pollination-constant non-astringent (PCNA) and pollination-constant astringent (PCA) persimmons. The astringency type affects not only the soluble tannin concentration but also the accumulation of individual sugars. Thus, we comprehensively investigate the gene expression and metabolite profiles of individual sugars to resolve the formation of flavor differences in PCNA and PCA persimmon fruit. The results showed that soluble sugar, starch content, sucrose synthase, and sucrose invertase were significantly different between PCNA and PCA persimmon fruit. The sucrose and starch metabolism pathway was considerably enriched, and six sugar metabolites involving this pathway were significantly differentially accumulated. In addition, the expression patterns of diferentially expressed genes (such as *bglX*, *eglC*, *Cel*, *TPS*, *SUS*, and *TREH* genes) were significantly correlated with the content of deferentially accumulated metabolites (such as starch, sucrose, and trehalose) in the sucrose and starch metabolism pathway. These results indicated that the sucrose and starch metabolism pathway maintained a central position of sugar metabolism between PCNA and PCA persimmon fruit. Our results provide a theoretical basis for exploring functional genes related to sugar metabolism and provide useful resources for future studies on the flavor differences between PCNA and PCA persimmon fruit.

Keywords: persimmon; soluble sugar; starch and sucrose metabolism; transcriptome; metabolome

1. Introduction

The persimmon (*Diospyros kaki*), a plant of the family Ebenaceae, has a long history of cultivation [1]. The persimmon, as a major fruit variety, has a high commercial value in Asian countries. In China, the production of persimmon fruit was approximately 3,429,438 tonnes in 2021, which accounted for 79.16% of the total production around the world. The genetic characteristics of fruit de-astringency allow for the categorization of persimmon into PCNA (pollination-constant non-astringent), PCA (pollination-constant astringent), pollination-variant non-astringent (PVNA), and pollination-variant astringent (PVA) types [2]. More than 950 cultivars are now recognized in China, practically all of which belong to the PCA type; no PVA and PVNA varieties have been identified [3].

Soluble sugars mainly comprise fructose, sucrose, and glucose in fruit [4]. Fruit can be categorized into the hexose-accumulating type, sucrose-accumulating type, and intermediate type according to the sugar composition in mature fruit [5,6]. Fruit sugar metabolic processes are complex [7]; according to the different substrates, sugar metabolic processes can be divided into sorbitol, sucrose, hexose-type sugar, and starch metabolism [8]. Sucrose

metabolism is the basis of the fruit sugar metabolism; it directly affects fruit sugar accumulation [4] and is vital in forming fruit quality [9]. In the fruit of Rosaceae trees, sorbitol is important in translocating photosynthate [10]. Sucrose and sorbitol can be converted to glucose, fructose, and starch after entering the fruit, catalyzed by a range of enzymes implicated within sugar metabolic processes [4,11]. An essential enzyme in sucrose production, sucrose phosphate synthase (SPS), transforms fructose-6-phosphate into sucrose phosphate [12]. Sucrose synthase (SS) is a bifunctional enzyme that synthesizes/hydrolyzes sucrose [13]. The enzymes are classified according to their optimum pH into neutral invertase (NI) and acid invertase (AI) [14], which catalyze the sucrose decomposition into fructose and glucose [15].

Persimmon fruit have significant variation between PCNA and PCA types [16,17]. The astringency type influences flavonoid metabolite content, soluble tannin concentration, individual sugar accumulation, and antioxidant capacity [17,18]. Thus, the differences in sugar metabolism are crucial for resolving flavor formation between PCNA and PCA persimmon fruit. This investigation probed sugar accumulation variation across PCNA/PCA fully matured fruit by analyzing sugar content and enzyme activities between mature PCNA varieties ('Yohou' and 'Jiro') and PCA fruit varieties ('Zhongshi5'). We also examined the gene expression patterns and the metabolites' accumulation patterns in the sucrose and starch metabolism pathway through transcriptomic and metabolomic analyses. The findings of this study may provide basic information on the sugar accumulation of persimmons and facilitate a flavor formation analysis between PCNA and PCA persimmons.

2. Results

2.1. Soluble Sugar Content, Starch Content, and Sucrose Synthase and Invertase Activity

For evaluating the differences in sugar accumulation between PCNA and PCA persimmons, soluble sugar fructose and glucose concentration, starch concentration, and sucrose synthase (SPS and SS) and invertase (SS-I, AI, and NI) activities were compared in mature PCNA (varieties 'Jiro' and 'Youhou') and PCA (variety 'Zhongshi No.5') persimmon fruit (Figure 1). The fructose, glucose, SS-I, and AI activity levels in the PCNA persimmon fruit were markedly lower than those in the PCA persimmon fruit. The SS activity was 3.2-fold and ~3.4-fold higher within PCNA persimmon fruit compared to PCA persimmon fruit. The starch content was slightly higher within PCNA persimmon fruit compared to PCA fruit. Overall, the soluble sugar, starch content, and sucrose synthase and invertase of PCNA and PCA persimmon fruit were significantly different; thus, soluble sugar accumulation may be one of the important reasons for the flavor differences between PCNA and PCA persimmon fruit.

2.2. RNA-Seq of PCNA and PCA Persimmon Mature Fruit

'Jiro', 'Youhou', and 'Zhongshi No.5' matured persimmon fruit transcriptome sequencing produced 62.88 GB of raw data. Each sample had 6.99 GB of high-quality data with a Q30 score at 92.52% after the low-quality reads were removed. A total of 4417 new unique transcripts were also discovered, and 86.05% of reads could be mapped onto the reference *D. kaki* genome. This demonstrates that the sequencing data's precision and quality were good enough for further investigation. The samples were segregated into three distinct groups using PCA depending upon FPKM values, with each sample creating a separate group with its replicates. This revealed strong correlations within sample replicates and variations among different samples (Figure 2a).

To examine expression-profile differences linked to sugar accumulation between PCNA and PCA persimmons, the genes in nine libraries (Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5) were compared. A total of 11,088 genes were substantially different in pairwise comparisons, with 9507 DEGs in Jiro vs. Zhongshi No.5 and 9439 DEGs in Youhou vs. Zhongshi No.5. A total of 7858 DEGs were differently expressed in both comparisons, according to the Venn diagram (Figure 2b–d). By comparing the RT-qPCR assessment of nine sugar-accumulation-associated DEGs together with transcriptomic

FPKM datasets, a gene expression pattern consistent with the transcriptomic data was shown (Figure S1).

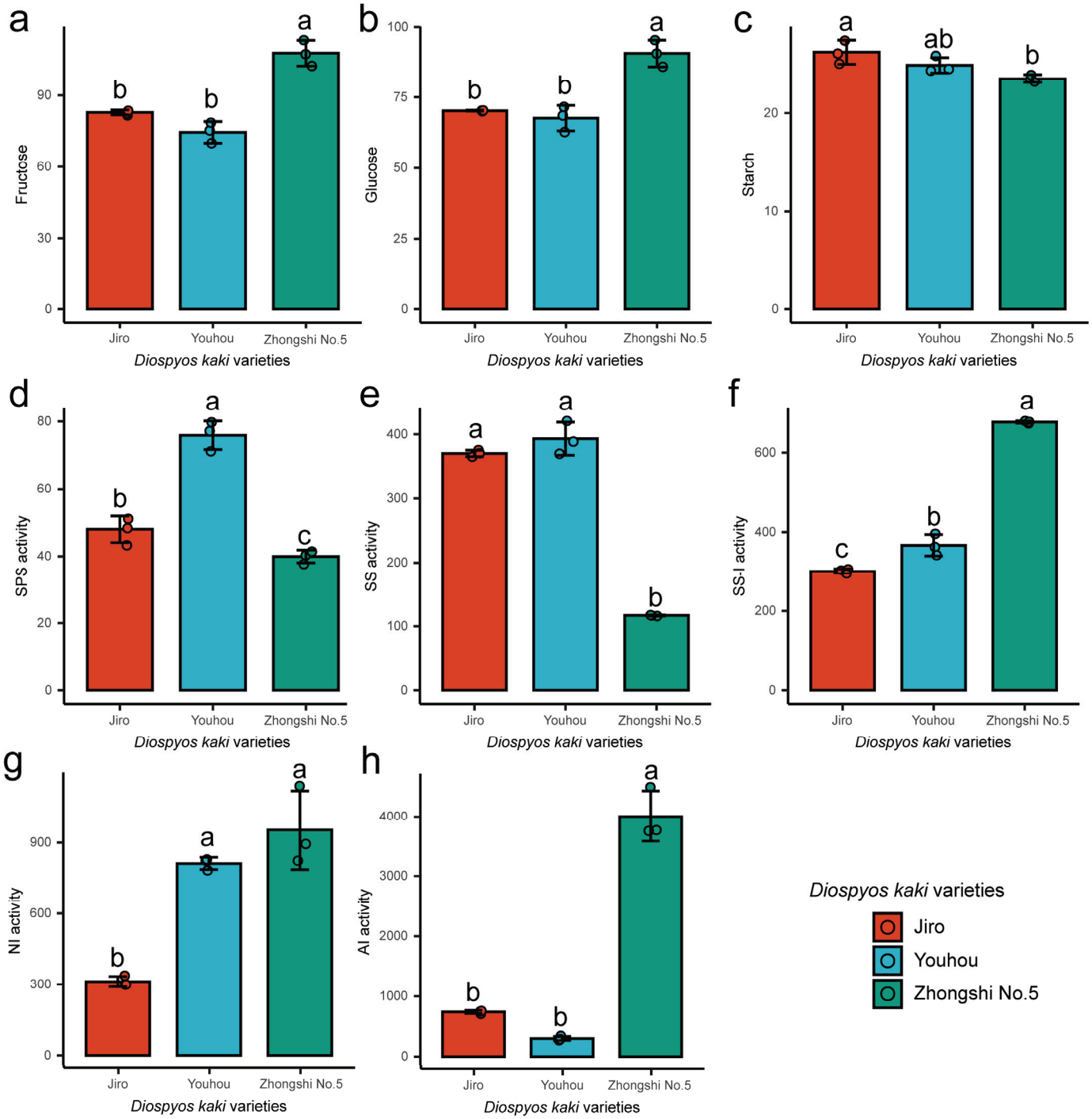


Figure 1. Soluble sugar content, starch content, and key enzymes within sugar metabolism in mature PCNA ('Jiro' and 'Youhou') and PCA ('Zhongshi No.5') persimmon fruit. (**a**–**h**) Fructose, glucose, starch, SPS, SS, SS-I, AI, and NI, accordingly. Significant variations ($p < 0.05$) are represented as lowercase letters. All error bars illustrate SD for the mean ($n = 3$).

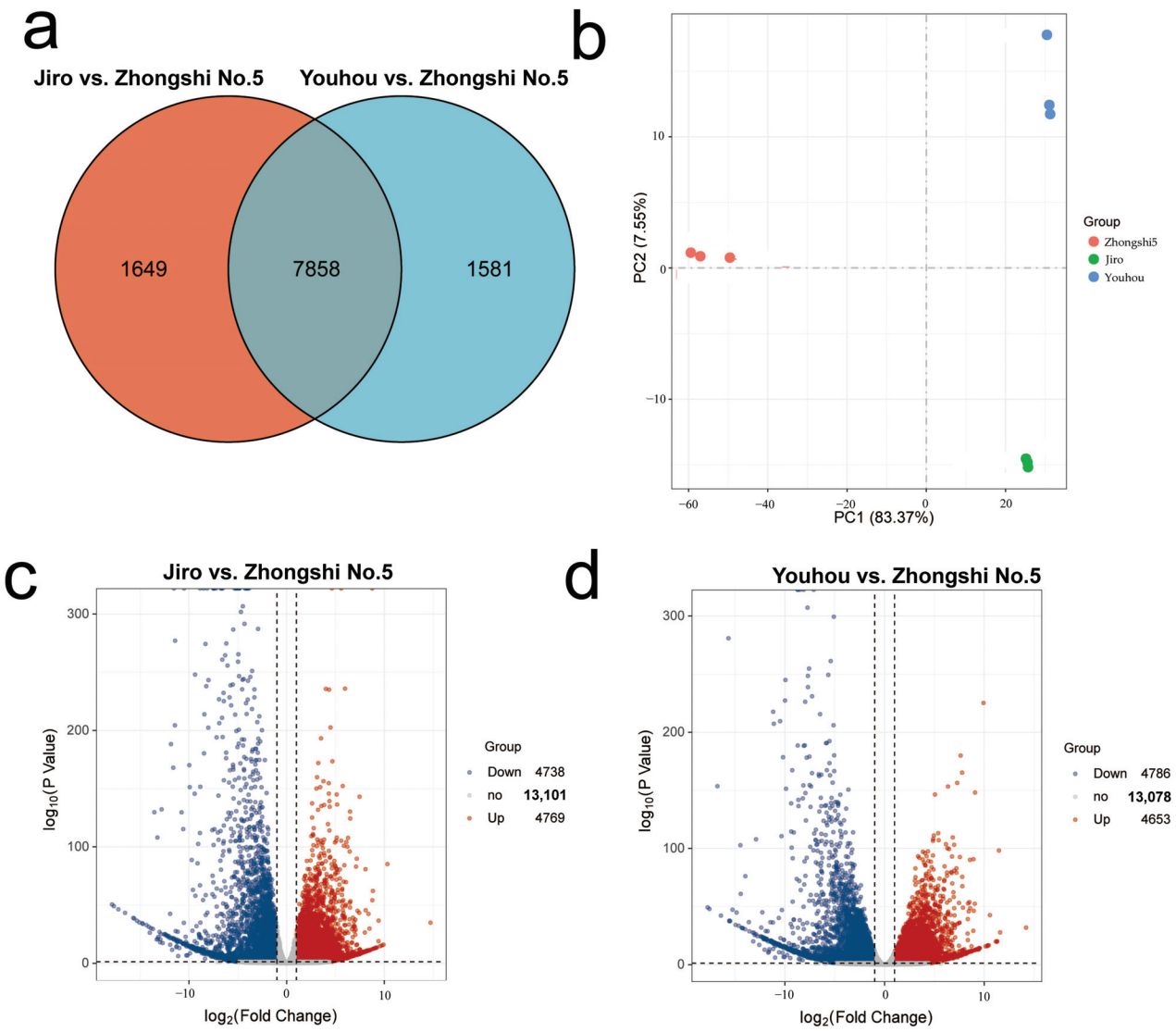


Figure 2. PCNA and PCA persimmon mature fruit transcriptome. (**a**) Venn diagram of unique/common DEGs between Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5. (**b**) Principal component analysis (PCA) of Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5. The samples from the same varieties were grouped. (**c**) The volcano plot shows the number of DEGs in Jiro vs. Zhongshi No.5 (|log2-fold change| > 1 and padj < 0.05). (**d**) DEG number in Youhou vs. Zhongshi No.5 was shown by the volcano plot (|log2-fold change| > 1 and padj < 0.05).

2.3. GO and KEGG Enrichment Analysis for DEGs

Further analysis of DEGs in the comparison groups Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5 was performed using the GO and KEGG databases; padj < 0.05 represented a significant difference. The DEGs in the comparison of Jiro vs. Zhongshi No.5 were mainly enriched in two GO terms: DNA-binding transcription regulator activity (GO:0140110) and transcription factor activity (GO:0003700). The DEGs in the comparison of Youhou vs. Zhongshi No.5 were mainly enriched in several different terms, such as the glucan metabolic process (GO:0044042), cellular glucan metabolic process (GO:0006073), cellular polysaccharide metabolic process (GO:0044264), cellular carbohydrate metabolic process (GO:0044262), and carbohydrate metabolic process (GO:0005975) (Figure 3a,b).

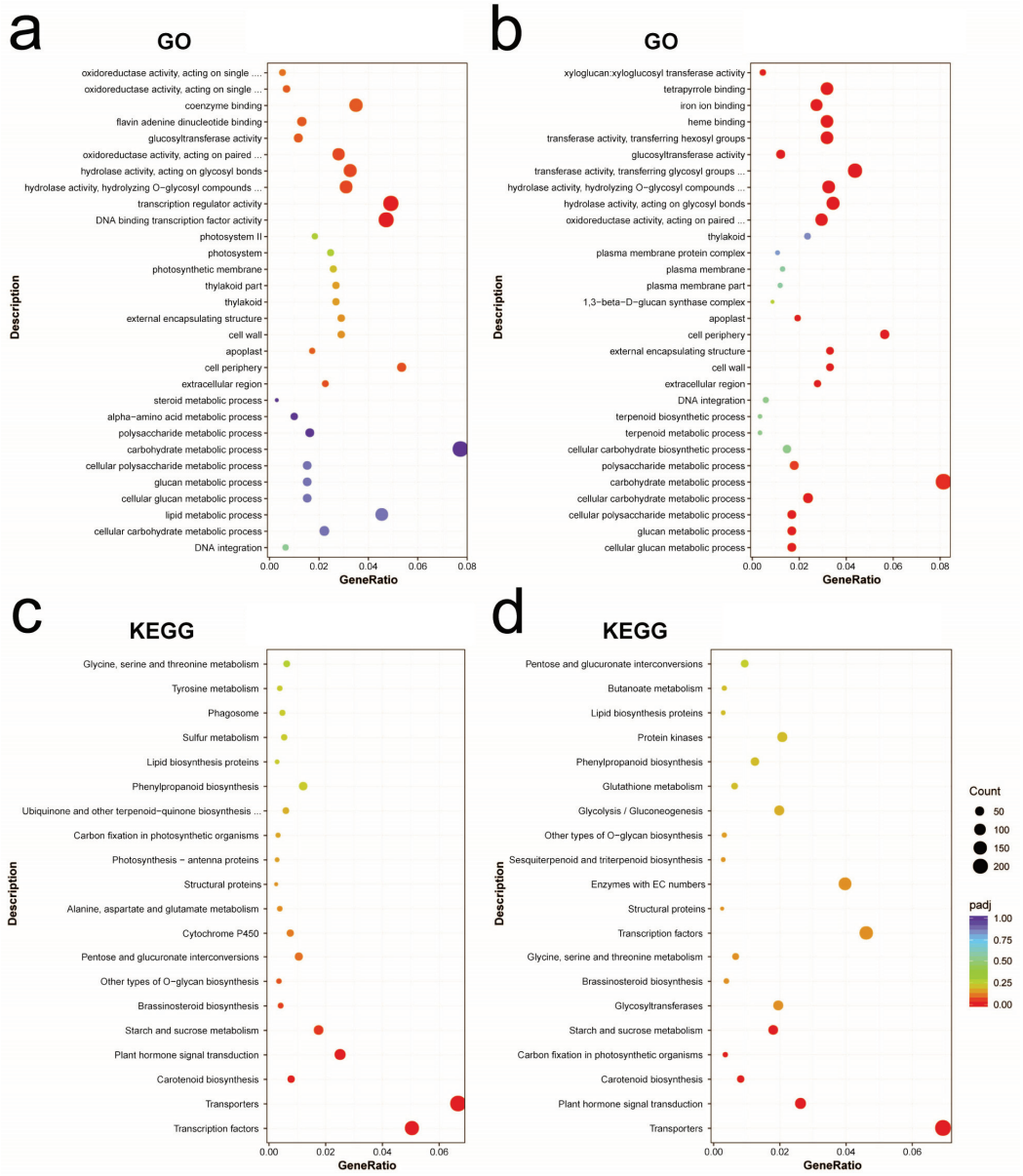


Figure 3. DEGs in PCNA and PCA. DEG enrichment GO enrichment scatter plot for (**a**) Jiro vs. Zhongshi No.5 and (**b**) Youhou vs. Zhongshi No.5. DEG enrichment KEGG scatter plot for (**c**) Jiro vs. Zhongshi No.5 and (**d**) Youhou vs. Zhongshi No.5.

Furthermore, KEGG enrichment analysis showed that the DEGs in Jiro vs. Zhongshi No.5 were significantly enriched in transcription factors, transporters, plant hormone signal transduction, carotenoid biosynthesis, and sucrose and starch metabolism. The DEGs in the comparison between Youhou and Zhongshi No.5 were significantly enriched in transporters, plant hormone signal transduction, carotenoid biosynthesis, carbon fixation in photosynthetic organisms, starch and sucrose metabolism, and biological processes (Figure 3c,d). Though the enrichment analysis, sucrose and starch metabolic processes were highly enriched between the two combinations. To understand the differences in sugar accumulation between PCNA and PCA persimmon fruit, the starch and sucrose metabolism were considered for downstream analysis (Figure 3).

2.4. Metabolic Features of Starch and Sucrose Metabolism

A total of 728 compounds were identified throughout quasi-targeted metabolome analysis, allowing us to characterize metabolic changes in mature PCNA ('Jiro' and 'Youhou')

and PCA ('Zhongshi No.5') persimmon fruit (Table S2). Organic acids and derivatives (197); phenylpropanoids and polyketides (126); lipids and lipid-like molecules (119); organic oxygen compounds (93); organic heterocyclic compounds (79); nucleosides, nucleotides, and analogs (60); benzenoids (34); and organic nitrogen compounds (13) were all included in the eight major categories of metabolites (Figure 4a). There were apparent similarities within sample replicates and differences between the samples, as shown by the PCA analysis results based on intensity values for metabolites, which clustered samples into three groups (Figure 4b).

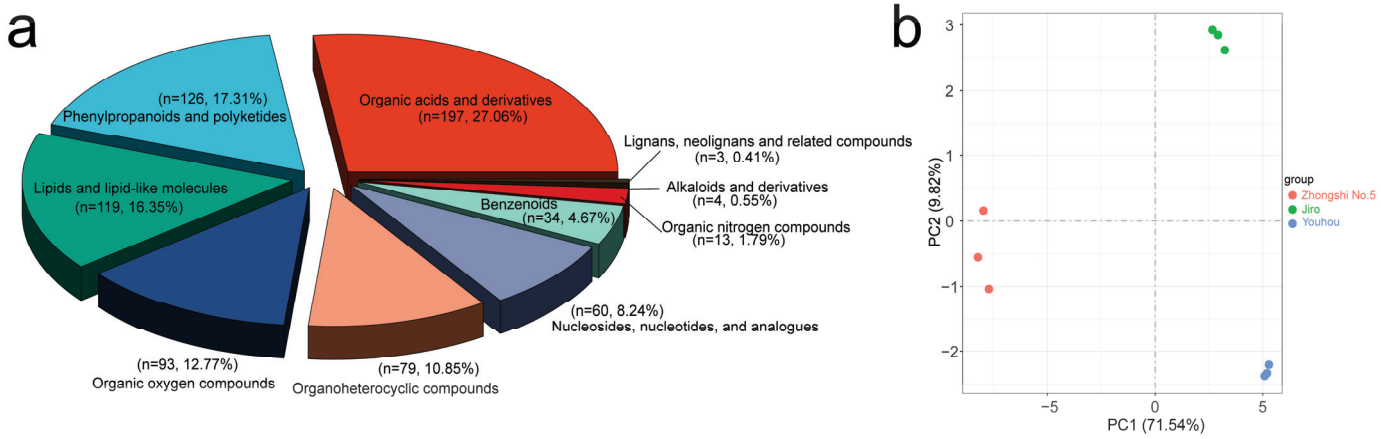


Figure 4. Analysis of metabolites: (**a**) statistics on the types of metabolites; (**b**) PCA of Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5. The samples from the same varieties were grouped.

For starch and sucrose metabolism, 11 metabolites were identified, including UDP-D-glucose, GDP-α-D-glucose, sucrose, isomaltose, maltose, D-glucose 1-phosphate, α,α-trehalose, trehalose 6-phosphate, D-glucose 6-phosphate, D-(+)-Cellobiose, D-glucopyranose, and D-glucose (Table S3). PCNA ('Jiro' and 'Youhou') persimmon fruit differed from PCA ('Zhongshi No.5') persimmon fruit in sucrose, maltose, D-(+)-Cellobiose, D-glucose 1-phosphate, α,α-trehalose, and trehalose 6-phosphate content. The contents of sucrose, maltose, and trehalose 6-phosphate in PCA persimmon fruit were markedly higher than compared to the two PCNA varieties of fruit, while D-glucose 1-phosphate contents were markedly reduced compared to PCNA fruit. Most of these differentially expressed sucrose metabolites were related to trehalose synthesis, a module of the starch and sucrose metabolism pathway, indicating that the trehalose synthesis pathway might influence the sugar accumulation between PCNA and PCA persimmon fruit.

2.5. Expression of Genes Implicated within Sucrose and Starch Metabolic Pathway

Fifty-eight DEGs related to the starch and sucrose metabolic pathway that encode 11 key enzymes were identified (Figure 5). These enzymes include granule-bound starch synthase (EC:2.4.1.242), 1,4-alpha-glucan branching enzyme (EC:2.4.1.18), trehalose 6-phosphate synthase (2.4.1.15), Alpha-amylase and beta amylase (EC:3.2.1.1; 3.2.1.2), beta-glucosidase (EC:3.2.1.21), glucan endo-1,3-beta-Dglucosidase (EC:3.2.1.39), sucrose synthetase (EC 2.4.1.13), alpha,alpha-trehalase (EC:3.2.1.28), trehalose 6-phosphate phosphatase (EC:3.1.3.12), and 4-alpha-glucanotransferase (EC:2.4.1.25). Out of 58 genes, 42 were differentially expressed in both comparison groups, Jiro vs. Zhongshi No.5 and Youhou vs. Zhongshi No.5. Ten bglX, five eglC, four Cel, four TPS, one GBE1, one GYG1, one malQ, one SUS, and one TREH genes were upregulated in Jiro and Youhou in comparison with Zhongshi No.5, though the other genes had reduced expression levels. Differing transcripts from identical genes were dysregulated, suggesting that intermediate products may be being converted between each other (Figure 5a). Sugar accumulation may be facilitated by gene expression implicated within the sucrose and

starch metabolic pathway since there was a robust association between their expression and metabolite levels in PCNA and PCA persimmon fruit (Figure 5b).

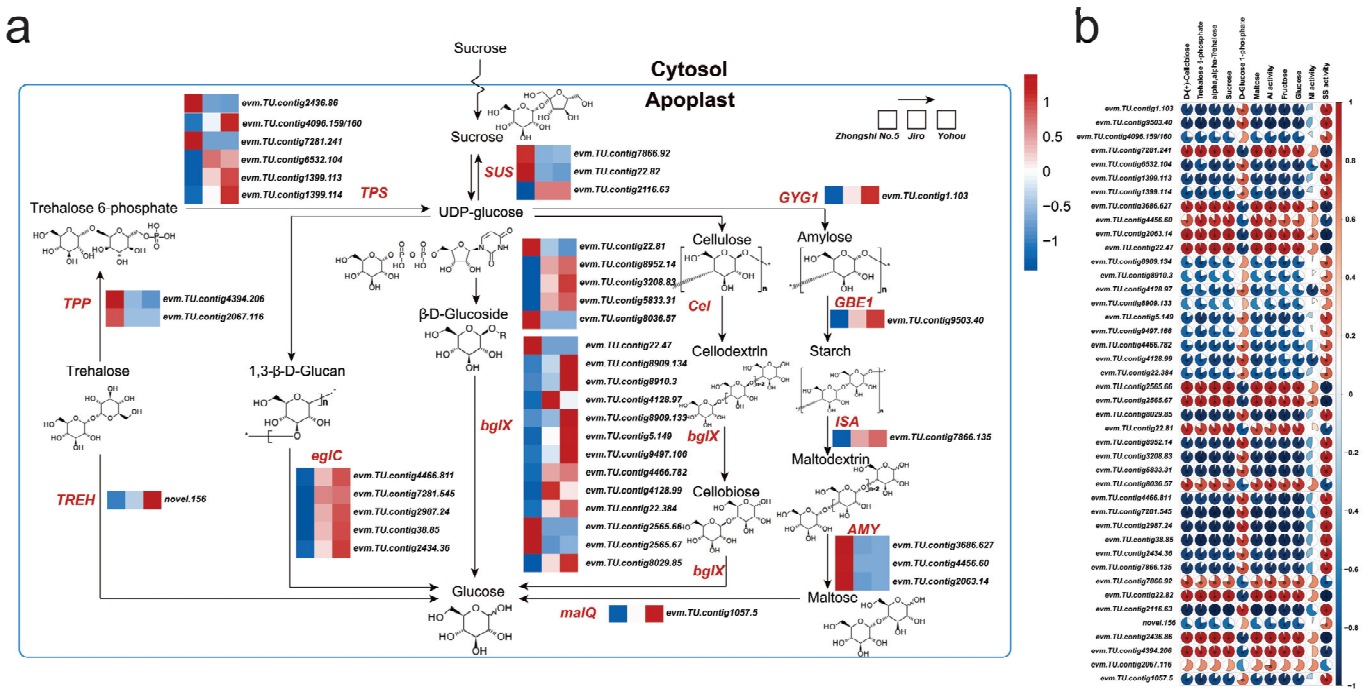


Figure 5. DEG analysis of sucrose and starch metabolic processes between PCNA and PCA: (**a**) diagram of DEGs implicated within sucrose and starch metabolic processes between PCNA and PCA; (**b**) correlation analysis between sucrose and starch metabolic pathway DEGs and metabolite content in PCNA and PCA persimmon fruit.

Combined KEGG analysis and gene expression patterns showed that widespread shifts within sucrose and starch metabolic processes mainly concentrated on converting sucrose into glucose. These changes indicate that converting sucrose into glucose might influence sugar accumulation between PCNA and PCA persimmon fruit.

3. Discussion

Quality improvement and efficiency are currently important issues for the persimmon industry, and flavor is the most important component of persimmon fruit quality, which has an important impact on both the fresh and processed quality of persimmon fruit. The fruit flavor quality varies greatly among different types and varieties, and there are few germplasm resources with excellent flavor. In this study, we analyzed the differences in the accumulation of various sugars and the differences in gene expression and enzyme activities related to sugar metabolism in different PCNA and PCA types of persimmon varieties, to further provide basic information on the sugar accumulation between PCA and PCNA persimmon fruit and facilitate a flavor formation analysis between PCNA and PCA persimmons.

Glucose and fructose are generally considered the main sugars, while sucrose is present as a minor component in mature persimmon fruits [19,20]. The HPLC method is less sensitive and has a higher detection limit, while the ultra-performance liquid chromatography coupled with triple quadrupole mass spectrometer has a lower detection limit and can detect compounds generally greater than 10 ng/mL, allowing a broad-spectrum determination of compounds in plant samples. Thus, the fructose and glucose contents were determined by HPLC, and the sucrose was determined by LC/MS-MS due to its low content in this study. The fructose and glucose levels of PCNA cultivars were lower than those of non-PCNA cultivars in tree maturity, which agrees with the work by Yildiz et al. [17]. Changes in the content of soluble sugars in fruits were closely related to their metabolic

enzyme activities. By measuring the content of sugar fractions and enzyme activities related to sucrose metabolism during fruit ripening in different astringent-type persimmon varieties, it was shown that SS, SS-I, and AI activities differed significantly in PCNA and PCA persimmons. Sucrose metabolism in red Fuji apple fruit was mainly regulated by AI and SS enzyme activities [21]. Studies on 'Malaysia 1' pineapple honey found that sucrose accumulation was significantly negatively correlated with AI activity and significantly positively correlated with SS and SPS activities [22], in agreement with studies on grapes [23]. Bubba et al. [24] showed that sucrose content peaked at the beginning of October and then gradually decreased, while glucose and fructose content showed an increasing trend throughout development. The above studies suggest that sucrose synthesis decreases and catabolism increases in PCNA and PCA persimmons before ripening, which results in an increase in soluble sugar content in ripe persimmon fruit. In addition, the difference in resource enzyme activity changes between PCNA and PCA persimmons eventually led to a lower soluble sugar content in PCNA persimmon fruit than in PCA persimmons.

The post-harvest conversion of starch to sugar improves fruit quality and flavor [25–28]. KEGG analysis showed that the two comparative combinations of sweet and astringent persimmons were significantly enriched in starch and sucrose metabolism. Therefore, analysis for differences in metabolites on the starch and sucrose pathways revealed that PCNA ('Jiro' and 'Youhou') persimmon fruits differed from PCA ('Zhongshi No.5') persimmon fruit in trehalose and tre6p content. The trehalose metabolic pathway is involved in processes such as cell wall cellulose synthesis, energy release from respiration, and the carbon skeleton composition of amino acids and fatty acids, and it is central to cellular metabolism [29]. Tre6P is a precursor of trehalose metabolism, employing UDP-glucose and glucose-6-phosphate as substrates, which *TPS* catalyzes to produce Tre6P; subsequently, Tre6P is dephosphorylated by *TPP* to produce trehalose [30]. Tre6P regulates various vital activities of tissues by effectively regulating the content of UDPG and G6P, reflects a vital signal for controlling plant growth and development [31], and is implicated within metabolic/genomic expression regulation within plants [32]. Tre6P acts as a signaling metabolite for sucrose; its content exists in proportion to sucrose content and affects starch synthesis [33,34]. Tao et al. [35] found that the possible reasons for the increase in sugar content during apple fruit development were related to the synthesis of Tre6P and trehalose. In this study, we found that Trehalose and Tre6P were both significantly higher in mature PCA persimmons than in PCNA persimmons, that Tre6P may act in key signaling to regulate the metabolic pathways of sucrose and starch in persimmon fruit, and that trehalose can be degraded to produce large amounts of glucose under the catalysis of Treh.

RNA-seq data showed that the gene expressions encoding the starch synthesis enzymes (i.e., AGPase and GBE1) in PCNA and PCA persimmon mature fruits were significantly differentially expressed. From the summary of differential genes, it can be concluded that the conversion from sucrose to glucose may be the main difference in sugar metabolism between PCNA and PCA mature fruits by means of the trehalose pathway, cellulose degradation, β-D-glucosidase conversion, and glucose hydrolysis. β-glucosidase is a hydrolytic enzyme and is the rate-limiting enzyme in the cellulose degradation process [36]. During the metabolism of starch and sucrose in dragon fruit pulp, β-glucosidase mainly catalyzes the generation of cellobiose from fibrous dextrin, and cellobiose generates glucose under the catalysis of β-glucosidase [37]; in addition, β-glucosidase also catalyzes the generation of glucose from β-D-glucoside [38]. The number of downregulated genes encoding bglX (10 DEGs) was obviously higher than the number of upregulated genes (3 DEGs) between PCNA and PCA persimmons; In addition, glucan hydrolysis may play a role in the difference between PCNA and PCA persimmons. The eglC gene encodes an endoglucanase, and the expressions of the eglC gene were also upregulated in PCNA persimmons; starch degraded to dextrin, and then dextrin was converted to maltose and glucose [39]. The expression of GYG1, GBE1, ISA, and malQ se in PCNA persimmons was obviously downregulated (Figure 5b). Taken together, gene expressions that related to sucrose and starch metabolism (sucrose hydrolyses) in PCNA persimmon were significantly

upregulated compared to PCA persimmons, and the results confirmed the downregulation of sucrose content. Interestingly, we found that PCNA persimmon glucose content was not higher than that of PCA under sucrose hydrolysis, probably because sucrose content is very low even if hydrolysis is not the main pathway of glucose synthesis, but this still needs further investigation. In conclusion, we suggest that sucrose catabolism in the starch and sucrose metabolic pathways may be the main difference in sugar metabolism between PCNA and PCA persimmons.

4. Materials and Methods

4.1. Plant Material

PCNA (varieties 'Jiro' and 'Youhou') and PCA (variety 'Zhongshi No.5') persimmons were planted in the forest planting base managed by the Research Institute of Non-timber Forestry (34°55′18″–34°56′27″ N, 113°46′14″–113°47′35″ E), Yuanyang County, Henan Province, China. These 10-year-old cultivars were managed using conventional cultivation measures, with a row spacing of 3 × 4 m. Persimmon fruit without astringency were collected during the matured-fruit phase. To ensure sample consistency, completely matured PCA persimmons on the trees were collected when they lost enough astringency to be eaten. The fruits of three varieties were randomly collected from the three clones with each replicate consisting of ten fruits. Until they could be processed for metabolic detection and RNA extraction, the flesh in the equatorial plane was flash-frozen within liquid nitrogen and kept at −80 °C in a refrigerator.

4.2. Soluble Sugar, Starch, and Sucrose Synthase and Invertase Activity Measurement

The soluble sugar content in persimmon fruit was measured according to previous studies with some modifications [40]. Approximately 1 g of fruit was grounded into powder and then extracted in 80 °C water using 4.5 mL of Milli-Q® ultrapure water for 30 min. After being chilled, the sample was centrifuged at 10,000 rpm for 20 min. The pellet extraction was performed once again, and the supernatant was saved. After combining the supernatants, the final volume was 10 mL. High-performance liquid chromatography (HPLC) with a CNW Athena NH2-RP column (4.6 × 250 mm, 5 μm) was used to determine the concentration of glucose and fructose together with the following parameters: injection volume = 10 μL, mobile phase = 75% acetonitrile, flow rate = 1.0 μL/min, and column heater temperature = 40 °C. An external reference fructose and glucose solution acquired from Beijing Solarbio Science & Technology Co., Ltd. (Beijing, China) was employed for determining the total amount of sugar in all samples. By comparing each sample's peak area and retention duration to that of a calibrated sugar solution, we could determine each sample's concentration. Each experiment had four independent replications.

Starch content was estimated using anthrone colorimetry by Cao et al. [41]. A fruit sample of approximately 1.0 g was placed in a prechilled mortar, ground into a homogenate with 80% (*v*/*v*) ethanol on an ice bath and transferred into a stoppered test tube. Then, 9 mL of 80% (*v*/*v*) ethanol was added, and the mixture was boiled in a boiling water bath for 30 min. Afterwards, it was removed and cooled in ice water, then it was centrifuged at a low temperature of 8000 r/min for 20 min, and the supernatant was discarded. An amount of 9 mL of 80% (*v*/*v*) ethanol was added again and the above steps were repeated once. The residue was dried, then 2 mL of distilled water was added. It was boiled in a boiling water bath for 15 min and cooled in ice-cold water. Then, 2 mL of cold 9.2 mol/L perchloric acid ($HClO_4$) was added for 15 min, followed by an addition of 6 mL of distilled water, mixing and centrifuging at a low temperature of 8000 r/min for 10 min, and the transference of the supernatant to a 25 mL volumetric flask. An amount of 2 mL of cold 4.6 mol/L $HClO_4$ was added to the filter residue for 15 min, then 6 mL of distilled water was added, followed by centrifugation for 20 min and the transference of the supernatant to a 25 mL volumetric flask. The precipitate was washed twice with 2 mL of distilled water, followed by centrifugation and the transference of the supernatant to the volumetric flask. Finally, the volumetric flask was filled with distilled water to 25 mL, resulting in the

starch extraction solution. The detection of starch was performed at 620 nm with a visible ultraviolet spectrophotometer UV. Starch content was determined according to an external sucrose standard solution (Beijing Solarbio Science & Technology Co., Ltd.).

The enzyme activities of NI, AI, sucrose synthase-invertase (SS-I), SS, and SPS were determined by using the kit purchased from Beijing Solarbio Science & Technology Co., Ltd. (BC0565, BC0575, BC0585, BC0605, and BC4315). These assays were conducted according to the protocol of the correlated kit.

4.3. Transcriptome Sequencing

Total RNA was isolated with TRIzol Reagent (B511321; Sangon, Shanghai, China). Paired-end sequencing libraries were prepared with three biological replicates for each sample and subsequently placed for sequencing through the Illumina™ NovaSeq® platform (Illumina™, San Diego, CA, USA). We used the hexaploid persimmon genome (*D. kaki* (variety 'Xiaoguo-tianshi', unpublished) as a reference sequence for alignment and subsequent analysis. Sequencing reads and read alignments were compared using HISAT2 [42] and assembled using StringTie. To determine how many reads were mapped to each gene, we utilized FeatureCounts® v1.5.0-p3 to further predict new transcripts [43]. The Fragments Per Kilobase of Transcript per Million Mapping Reads (FPKM) was then calculated by multiplying its length by the number of mapped reads. We used DESeq2 (1.18.0) to compare the expression levels of two distinct groups [44]. Genes that had a padj $\leq$0.05 and |log2-fold change| $\geq$ 1 were considered DEGs.

4.4. Metabolite Profiling Analysis

Extracts of 100 mg of persimmon fruit powder were made by vortexing 500 µL of 80% (v/v) prechilled methanol. There were 3 biological replicates for each sample. Then, 500 µL of the supernatant from centrifugation (15,000× g, 4 °C/20 min) was diluted into 53% methanol using Milli-Q® ultrapure water. After that, the samples were filtered (0.22 µm membrane filter) and subjected to centrifuging (15,000× g for 20 min at 4 °C). The persimmon fruit extract samples were separated through the ExionLCTM AD system (SCIEX™) connected to a QTRAP®6500+ mass spectrometer (SCIEX™) and fitted with an Xselect HSS T3 column (2.1 × 150 mm, 2.5 µm). The column temperature was 50 °C, the injection volume was 1.5 µL, and the flow rate was 0.4 mL/min; these were the settings used for the analysis. Water was the mobile phase. The gradient program for phase A/phase B was 98:2 (v/v) at 0 min, 98:2 (v/v) at 2 min, 0:100 (v/v) at 15 min, 0:100 (v/v) at 17 min, 98:2 (v/v) at 17.1 min, and 98:2 (v/v) at 20 min.

To check the system's consistency and the experimental data's accuracy, samples were placed into quality control (QC) within a queue mode. Electrospray ionization (ESI) source settings allowed each sample to be run in both negative and positive ion modes. Mass spectrum databases from the MRM (Multiple Reaction Monitoring) of the Novogene in-house database were consulted to compare the spectra and the retention index (RI) with the reference compounds previously examined with the same system. KEGG (http://www.genome.jp/kegg/) [45], HMDB (http://www.hmdb.ca/) [46], and Lipidmaps databases (http://www.lipidmaps.org/) [47] were used for metabolite annotation. Principal components analysis (PCA) was performed at metaX [48]. Differential metabolites were defined as those with a p-value < 0.05 and fold change $\geq$ 2.

4.5. Quantitative RT-PCR Assessments

The cDNA used in the RNA-seq study was converted from total RNA through a TRUEscript First-Strand® cDNA Synthesis Kit (Kemix™, Beijing, China). RT-qPCR runs were conducted with the LightCycler 480 II (Roche), using 96-well plates. Each gene underwent a three-minute reaction at 95 °C, with 45 subsequent cycles of 5 s at that temperature and 30 cycles at 55–60 °C. Using glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as a reference gene [49], the $2^{-\Delta\Delta Ct}$ technique determined the relative expression for each gene, and three replicates of each reaction were performed. A Pearson's correlation assessment

was conducted through SPSS® (v 24.0; SPSS Inc.™, Chicago, IL, USA). The RT-qPCR primers are listed in Table S1.

4.6. Statistical Analysis

ANOVA and the post hoc test were performed using Excel 2019 and SPSS 24.0 software. The Pearson correlation analysis was carried out through Hiplot [50].

5. Conclusions

In conclusion, a comprehensive transcriptomic and metabolomic analysis of PCNA ('Jiro' and 'Yohou') and PCA ('Zhongshi No.5') persimmon fruit was conducted. The concentration of soluble sugar, starch content, and sucrose synthase and invertase of PCNA and PCA persimmons were significantly different. For example, the fructose, glucose, SS-I, and AI activity levels in the PCNA persimmon fruit were markedly lower than those in the PCA persimmon fruit. Through the KEGG enrichment analysis, it was found that the sucrose and starch metabolic pathways were highly enriched in the PCNA and PCA fruit. For the starch and sucrose metabolism pathway, 11 metabolites were identified. Moreover, most of these differentially accumulated sucrose metabolites were related to trehalose synthesis, indicating that the trehalose synthesis pathway might influence the sugar variation between PCNA and PCA persimmons. In addition, the expression patterns of deferentially expressed genes (such as *bglX*, *eglC*, *Cel*, *TPS*, *SUS*, and *TREH* genes) were significantly correlated with the content of deferentially accumulated metabolites (such as starch, sucrose, and trehalose). These results showed that the sucrose and starch metabolism pathway maintained a central position of sugar variation between PCNA and PCA fruit. This study provides basic information and useful resources for future studies on the influence of astringency types on sugar differences between PCNA and PCA persimmon fruit.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24108599/s1.

Author Contributions: Conceptualization, F.L. and J.F.; methodology, W.H.; software, Y.W.; validation, H.L., S.D., Y.S., T.L. and P.S.; writing—original draft preparation, W.H. and Y.W.; writing—review and editing, F.L. and J.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China (2022YFD2200400) and National Natural Science Foundation of China (32071801).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The transcriptome sequencing raw data were deposited in the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) under the Bioproject ID PRJNA897083.

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

References

1. Guan, C.; Zhang, P.; Hu, C.; Chachar, S.; Riaz, A.; Wang, R.; Yang, Y. Genetic diversity, germplasm identification and population structure of *Diospyros kaki* Thunb. from different geographic regions in China using SSR markers. *Sci. Hortic.* **2019**, *251*, 233–240. [CrossRef]
2. Yonemori, K.; Sugiura, A.; Yamada, M. Persimmon genetics and breeding. *Plant Breed. Rev.* **2000**, *19*, 191–225. [CrossRef]
3. Du, X.; Zhang, Q.; Luo, Z. Comparison of four molecular markers for genetic analysis in *Diospyros* L. (Ebenaceae). *Plant Syst. Evol.* **2009**, *281*, 171. [CrossRef]
4. Li, M.; Li, P.; Ma, F.; Dandekar, A.M.; Cheng, L. Sugar metabolism and accumulation in the fruit of transgenic apple trees with decreased sorbitol synthesis. *Hortic. Res.* **2018**, *5*, 60. [CrossRef]
5. Ayre, B.G. Membrane-transport systems for sucrose in relation to whole-plant carbon partitioning. *Mol. Plant* **2011**, *4*, 377–394. [CrossRef]
6. Kühn, C.; Grof, C.P.L. Sucrose transporters of higher plants. *Curr. Opin. Plant Biol.* **2010**, *13*, 287–297. [CrossRef]
7. Cirilli, M.; Bassi, D.; Ciacciulli, A. Sugars in peach fruit: A breeding perspective. *Hortic. Res.* **2016**, *3*, 15067. [CrossRef]

8. Huang, Y.H.; Zeng, M. Research advances in sugar metabolism and regulatory factors in pear fruits. *Plant Physiol. J.* **2013**, *49*, 709–714.
9. Koch, K. Sucrose metabolism: Regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Curr. Opin. Plant Biol.* **2004**, *7*, 235–246. [CrossRef]
10. Moing, A.; Langlois, N.; Svanella, L.; Zanetto, A.; Gaudillère, J.P. Variability in sorbitol: Sucrose ratio in mature leaves of different prunus species. *J. Am. Soc. Hortic. Sci.* **1997**, *122*, 83–90. [CrossRef]
11. Li, M.; Feng, F.; Cheng, L. Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. *PLoS ONE* **2012**, *7*, e33055. [CrossRef] [PubMed]
12. Ma, P.; Zhang, X.; Chen, L.; Zhao, Q.; Zhang, Q.; Hua, X.; Wang, Z.; Tang, H.; Yu, Q.; Zhang, M.; et al. Comparative analysis of sucrose phosphate synthase (*SPS*) gene family between Saccharum officinarum and Saccharum spontaneum. *BMC Plant Biol.* **2020**, *20*, 422. [CrossRef] [PubMed]
13. Verma, A.K.; Upadhyay, S.K.; Verma, P.C.; Solomon, S.; Singh, S.B. Functional analysis of sucrose phosphate synthase (SPS) and sucrose synthase (SS) in sugarcane (Saccharum) cultivars. *Plant Biol.* **2011**, *13*, 325–332. [CrossRef] [PubMed]
14. Gayler, K.R.; Glasziou, K.T. Physiological functions of acid and neutral invertases in growth and sugar storage in sugar cane. *Physiol. Plant* **1972**, *27*, 25–31. [CrossRef]
15. Tymowska-Lalanne, Z.; Kreis, M. The plant invertases: Physiology, biochemistry and molecular biology. In *Advances in Botanical Research*, 1st ed.; Callow, J.A., Ed.; Academic Press: San Diego, CA, USA, 1998; Volume 28, pp. 71–117.
16. Novillo, P.; Salvador, A.; Crisosto, C.; Besada, C. Influence of persimmon astringency type on physico-chemical changes from the green stage to commercial harvest. *Sci. Hortic.* **2016**, *206*, 7–14. [CrossRef]
17. Yildiz, E.; Kaplankiran, M. Changes in sugars content and some biochemical substances during fruit development in different persimmon cultivars. *J. Agric. Fac. Mustafa Kemal Univ.* **2018**, *23*, 12–23.
18. Wang, Y.; Suo, Y.; Han, W.; Li, H.; Wang, Z.; Diao, S.; Sun, P.; Fu, J. Comparative transcriptomic and metabolomic analyses reveal differences in flavonoid biosynthesis between PCNA and PCA persimmon fruit. *Front. Plant Sci.* **2023**, *14*, 1130047. [CrossRef]
19. Baltacıoğlu, H.; Artık, N. Study of postharvest changes in the chemical composition of persimmon by HPLC. *Turk. J. Agric. For.* **2013**, *37*, 568–574. [CrossRef]
20. Veberic, R.; Jurhar, J.; Mikulic-Petkovsek, M.; Stampar, F.; Schmitzer, V. Comparative study of primary and secondary metabolites in 11 cultivars of persimmon fruit (*Diospyros kaki* L.). *Food Chem.* **2010**, *119*, 477–483. [CrossRef]
21. Wang, Y.; Zhang, D. A Study on the relationships between acid invertase, sucrose synthase and sucrose metabolism in 'Red Fuji' apple fruit. *Acta Hortic. Sin.* **2001**, *28*, 256–261.
22. Gutiérrez-Miceli, F.A.; Rodríguez-Mendiola, M.A.; Ochoa-Alejo, N.; Méndez-Salas, R.; Dendooven, L.; Arias-Castro, C. Relationship between sucrose accumulation and activities of sucrose-phosphatase, sucrose synthase, neutral invertase and soluble acid invertase in micropropagated sugarcane plants. *Acta Physiol. Plant* **2002**, *24*, 441–446. [CrossRef]
23. Wu, B.; Liu, H.; Guan, L.; Fan, P.; Li, S. Carbohydrate metabolism in grape cultivars that differ in sucrose accumulation. *Vitis* **2011**, *50*, 51–57. [CrossRef]
24. Del Bubba, M.; Giordani, E.; Pippucci, L.; Cincinelli, A.; Checchini, L.; Galvan, P. Changes in tannins, ascorbic acid and sugar content in astringent persimmons during on-tree growth and ripening and in response to different postharvest treatments. *J. Food Compos.* **2009**, *22*, 668–677. [CrossRef]
25. Harvey, W.J.; Grant, D.G.; Lammerink, J.P. Physical and sensory changes during the development and storage of buttercup squash. *N. Z. J. Crop Hortic.* **1997**, *25*, 341–351. [CrossRef]
26. Irving, D.E.; Hurst, P.L.; Ragg, J.S. Changes in carbohydrates and carbohydrate metabolizing enzymes during the development, maturation, and ripening of buttercup squash (*Cucurbita maxima* D. 'Delica'). *J. Am. Soc. Hortic. Sci.* **1997**, *122*, 310–314. [CrossRef]
27. Phillips, T.G. Changes in the composition of squash during storage1. *Plant Physiol.* **1946**, *21*, 533–541. [CrossRef]
28. Yu, J.; Xu, S.; Liu, X.; Li, T.; Zhang, D.; Teng, N.; Wu, Z. Starch degradation and sucrose accumulation of lily bulbs after cold storage. *Int. J. Mol. Sci.* **2022**, *23*, 4366. [CrossRef]
29. Wingler, A. The function of trehalose biosynthesis in plants. *Phytochemistry* **2002**, *60*, 437–440. [CrossRef]
30. Ponnu, J.; Wahl, V.; Schmid, M. Trehalose-6-phosphate: Connecting plant metabolism and development. *Front. Plant Sci.* **2011**, *2*, 70. [CrossRef]
31. Paul, M.J.; Primavesi, L.F.; Jhurreea, D.; Zhang, Y. Trehalose Metabolism and signaling. *Annu. Rev. Plant Biol.* **2008**, *59*, 417–441. [CrossRef]
32. O'Hara, L.E.; Paul, M.J.; Wingler, A. How do sugars regulate plant growth and development? New insight into the role of Trehalose-6-Phosphate. *Mol. Plant* **2013**, *6*, 261–274. [CrossRef]
33. Kolbe, A.; Tiessen, A.; Schluepmann, H.; Paul, M.; Ulrich, S.; Geigenberger, P. Trehalose 6-phosphate regulates starch synthesis via posttranslational redox activation of ADP-glucose pyrophosphorylase. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11118–11123. [CrossRef] [PubMed]
34. Yadav, U.P.; Ivakov, A.; Feil, R.; Duan, G.Y.; Walther, D.; Giavalisco, P.; Piques, M.; Carillo, P.; Hubberten, H.M.; Stitt, M.; et al. The sucrose–trehalose 6-phosphate (Tre6P) nexus: Specificity and mechanisms of sucrose signalling by Tre6P. *J. Exp. Bot.* **2014**, *65*, 1051–1068. [CrossRef]
35. Tao, H.; Sun, H.; Wang, Y.; Song, X.; Guo, Y. New insights on 'GALA' apple fruit development: Sugar and acid accumulation: A transcriptomic approach. *J. Plant Growth Regul.* **2020**, *39*, 680–702. [CrossRef]

36. Naraian, R.; Gautam, R.L. Penicillium enzymes for the saccharification of lignocellulosic feedstocks. In *New and Future Developments in Microbial Biotechnology and Bioengineering*, 2nd ed.; Gupta, V.K., Rodriguez-Couto, S., Eds.; Elsevier: Amsterdam, The Netherlands, 2018; Volume 6, pp. 121–136. [CrossRef]
37. Yang, Y.; Li, J.; Ma, G. Differential expression analysis of genes related to starch and sucrose metabolism pathway in different developmental stages of dragon fruit pulp based on transcriptome. *Chin. J. Trop. Crop.* **2021**, *42*, 1520–1530. [CrossRef]
38. Bhatia, Y.; Mishra, S.; Bisaria, V.S. Microbial β-glucosidases: Cloning, properties, and applications. *Crit. Rev. Biotechnol.* **2002**, *22*, 375–407. [CrossRef]
39. Tomasik, P.; Horton, D. Enzymatic conversions of starch. In *Advances in Carbohydrate Chemistry and Biochemistry*, 2nd ed.; Horton, D., Ed.; Academic Press: San Diego, CA, USA, 2012; Volume 2, pp. 59–436. [CrossRef]
40. Zhang, Y.; Li, P.; Cheng, L. Developmental changes of carbohydrates, organic acids, amino acids, and phenolic compounds in 'Honeycrisp' apple flesh. *Food Chem.* **2010**, *123*, 1013–1018. [CrossRef]
41. Cao, J.; Jiang, W.; Zhao, Y. *Experiment Guidance of Postharvest Physiology and Biochemistry of Fruits and Vegetables*, 1st ed.; China Light Industry Press Ltd.: Beijing, China, 2007; pp. 76–84. [CrossRef]
42. Zhang, Y.; Park, C.; Bennett, C.; Thornton, M.; Kim, D. Rapid and accurate alignment of nucleotide conversion sequencing reads with HISAT-3N. *Genome Res.* **2021**, *31*, 1290–1295. [CrossRef]
43. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923–930. [CrossRef]
44. 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]
45. Kanehisa, M.; Furumichi, M.; Sato, Y.; Kawashima, M.; Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **2022**, *51*, D587–D592. [CrossRef] [PubMed]
46. Wishart, D.S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B.L.; et al. HMDB 5.0: The Human Metabolome Database for 2022. *Nucleic Acids Res.* **2022**, *50*, D622–D631. [CrossRef]
47. Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E.A.; Glass, C.K.; Merrill, A.H.; Murphy, R.C., Jr.; Raetz, C.R.H.; Russell, D.W.; et al. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **2007**, *35*, D527–D532. [CrossRef] [PubMed]
48. Wen, B.; Mei, Z.; Zeng, C.; Liu, S. metaX: A flexible and comprehensive software for processing metabolomics data. *BMC Bioinform.* **2017**, *18*, 183. [CrossRef]
49. Du, G.; Wang, L.; Li, H.; Sun, P.; Fu, J.; Suo, Y.; Han, W.; Diao, S.; Mai, Y.; Li, F. Selection and validation of reference genes for quantitative gene expression analyses in persimmon (*Diospyros kaki* Thunb.) using real-time quantitative PCR. *Biol. Futur.* **2019**, *70*, 261–267. [CrossRef] [PubMed]
50. Li, J.; Miao, B.; Wang, S.; Dong, W.; Xu, H.; Si, C.; Wang, W.; Duan, S.; Lou, J.; Bao, Z.; et al. Hiplot: A comprehensive and easy-to-use web service for boosting publication-ready biomedical data visualization. *Brief. Bioinform.* **2022**, *23*, bbac261. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

FaMYB5 Interacts with FaBBX24 to Regulate Anthocyanin and Proanthocyanidin Biosynthesis in Strawberry (*Fragaria* × *ananassa*)

Lianxi Zhang †, Yiping Wang †, Maolan Yue, Leiyu Jiang, Nating Zhang, Ya Luo, Qing Chen, Yong Zhang, Yan Wang, Mengyao Li, Yunting Zhang, Yuanxiu Lin and Haoru Tang *

College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China; zlx981027@163.com (L.Z.); wangyp_sicau@163.com (Y.W.); maolanyue@outlook.com (M.Y.); jiangleiyusicau@outlook.com (L.J.); zhangnating714@163.com (N.Z.); luoya945@sicau.edu.cn (Y.L.); supnovel@sicau.edu.cn (Q.C.); zhyong@sicau.edu.cn (Y.Z.); wangyanwxy@163.com (Y.W.); limy@sicau.edu.cn (M.L.); asyunting@sicau.edu.cn (Y.Z.); linyx@sicau.edu.cn (Y.L.)
* Correspondence: htang@sicau.edu.cn; Tel.: +86-136-0826-4028
† These authors contributed equally to this work.

Abstract: MYB and BBX transcription factors play important roles in flavonoid biosynthesis. Here, we obtained transgenic woodland strawberry with stable overexpression of *FaMYB5*, demonstrating that FaMYB5 can increase anthocyanin and proanthocyanidin content in roots, stems and leaves of woodland strawberry. In addition, bimolecular fluorescence complementation assays and yeast two-hybridization demonstrated that the N-terminal (1-99aa) of FaBBX24 interacts with FaMYB5. Transient co-expression of *FaBBX24* and *FaMYB5* in cultivated strawberry 'Xiaobai' showed that co-expression strongly promoted the expression of *F3′H*, *4CL-2*, *TT12*, *AHA10* and *ANR* and then increased the content of anthocyanin and proanthocyanidin in strawberry fruits. We also determined that FaBBX24 is also a positive regulator of anthocyanin and proanthocyanidin biosynthesis in strawberry. The results reveal a novel mechanism by which the FaMYB5–FaBBX24 module collaboratively regulates anthocyanin and proanthocyanidin in strawberry fruit.

Keywords: strawberry; transcription factor; anthocyanin; proanthocyanidin

1. Introduction

Strawberry (*Fragaria* × *ananassa* Duch.) is full of macronutrients and bioactive ingredients, such as vitamins, minerals and flavonoids. It has high nutritional and healthcare value, especially its health effect of preventing cardiovascular disease, cancer and other diseases [1], which makes it popular with consumers. The flesh of the strawberry is actually enlarged from the receptacle, and the real fruit (achene) is the seed-like object on the surface, so we often say that the strawberry fruit actually refers to the edible fleshy part of the strawberry. Flavonoid compounds are one of the important secondary metabolites of plants and are widely distributed in the plantae. Because of flavonoids' strong biological activity, they often act as an antioxidant or signaling molecule to regulate cell function [2]. At present, there have been extensive studies on flavonoid synthesis pathways in horticultural crops [3,4]. There are two main types of genes involved in flavonoid metabolism. One is structural genes that are involved in flavonoid synthesis by directly encoding various biosynthetic enzymes. These include *PAL* (phenylalanine lyase), *4CL* (4-cinnamate-CoA ligase), *CHS* (Chalcone synthase), *CHI* (Chalcone isomerase), *F3H* (flavanone 3-hydroxylase), *F3′H* (flavonoid 3′-hydroxylase), *DFR* (dihydroflavonol 4-reductase), *ANS* (anthocyanin synthase), *ANR* (Anthocyanidin reductase), *LAR* (leucoanthocyanidin reductase) and *UFGT* (UDP-glucose: flavonoid 3-o-glucosyltransferase) [5], some genes that control transport and

modify enzymes such as *TT10* (laccase 15), *TT12* (MATE transporter), *TT19* (glutathione-S-transferase) and *AHA10* (H^+-ATPase) [6]. The other type of gene is transcription factors (TFs) which regulate flavonoid metabolism by regulating structural genes [7].

V-myb myeloblastosis viral oncogene homolog (MYB) proteins are one of the largest families of plant transcription factors [8]. MYB transcription factors have regulatory effects on metabolic pathways such as anthocyanin and proanthocyanidin in flavonoids and can affect flower organ development and fruit coloring in plants [9]. Anthocyanin synthesis in plants is mainly regulated by the transcription complex (MBW complex) composed of MYB, basic helical loop helix (bHLH) and WD-repeat (WDR) proteins [10]. Since the discovery of ZmC1 [11], the first MYB transcription factor regulating anthocyanin in maize, regulation of anthocyanin by MYB transcription factors has been reported in many plants. The expression level of anthocyanin biosynthesis structural gene in transgenic *Arabidopsis thaliana* overexpressed *AgMYB2*, which was significantly up-regulated [12]; carrot DcMYB6 can directly activate *DcUCGXT1* and *DcSAT1* to regulate the glycosylation and acylation of anthocyanin [13]. Apple MdMYBA specifically bound to the *MdANS* promoter, increasing anthocyanin accumulation in apple [14]; MdMYB3, MdMYB9, MdMYB10 and MdMYB114 can promote apple fruit coloring by interacting with bHLH3 and WD40 [15–17]; MdMYB306 interacted with MdbHLH33 to regulate structural genes *F3H*, *DFR* and *UFGT* to regulate anthocyanin synthesis [18]. The complex formed by pear PyMYB10, PybHLH and PyWD40 transcription factors positively regulates pear anthocyanin synthesis [19]. Peach PpMYB10.1 and PpMYB10.3 interacted with PpbHLH proteins to activate the expression of *CHS*, *F3′H* and *UFGT*, thereby promoting the synthesis of anthocyanin [20]. In strawberry, FaMYB5/FaMYB10-FaEGL3 (bHLH)-FaLWD1/FaLWD1-like (WD40) formed an 'MBW' complex to positively regulate anthocyanin synthesis in strawberry fruit [21,22], FaMYB10 promotes anthocyanin synthesis in strawberry, and *FaMYB10* mutation is the main cause of color difference between yellow woodland strawberry 'Yellow wonder' and red woodland strawberry 'Ruegen' [23,24]. The interaction between FaMYB123 and FabHLH3 in strawberry regulates anthocyanin synthesis by regulating late biosynthetic genes (LBGs) [25]. FaMYB1 inhibited anthocyanin synthesis, while FaMYB9 and FaMYB11 promoted proanthocyanidin synthesis [26].

The B-BOX protein family plays an important role in plant photomorphogenesis [27,28], the flowering cycle [29,30] and stress of adversity [31–33]. In recent years, many studies on the regulation of plant anthocyanin metabolism by B-BOX proteins have been reported. Strawberry FaBBX22 can promote the expression of anthocyanin structural genes, and the interaction between FaHY5 and FaBBX22 can enhance this effect [34]. In apple, MdBBX1 regulated anthocyanin accumulation by regulating *MdMYB10* and *DFR* [35], and MdBBX21 interacted with MdHY5 to promote *MdMYB1* and thus regulated anthocyanin synthesis [36]. PpBBX18 formed a heterodimer with PpHY5 to regulate anthocyanin accumulation in pear fruit [37]. Overexpression of *PpBBX16* promoted anthocyanin accumulation in pear callus and was a positive regulator of photoinduced anthocyanin accumulation, and the presence of PpHY5 was required to activate its full function [38]; overexpression of *VvBBX44* inhibited *UFGT* expression and anthocyanin accumulation in grapes callus [39]; and PavBBX6 and PavBBX9 positively regulate light-induced anthocyanin and ABA biosynthesis by promoting *PavUFGT* and *PavNCED1* expression in sweet cherry [40]. Overexpression of *PtrBBX23* in poplar activated expression of MYB TFs and structural genes in the flavonoid pathway, thereby promoting the accumulation of proanthocyanidin and anthocyanin in poplar [41]. Although the functions of some transcription factors have been clearly defined, their interaction networks and mechanisms remain unclear.

In previous studies, we evaluated FaMYB5 as playing an important role in anthocyanin and proanthocyanidin metabolism through transcriptomics and metabolomics [21,22]. In this study, FaMYB5 was stably overexpressed in woodland strawberry to confirm that FaMYB5 can promote anthocyanin and proanthocyanidin accumulation in strawberry plants. We found that FaBBX24, a member of the strawberry BBX family, interacts with

FaMYB5 to synergistically regulate anthocyanin accumulation. The aim of this study is to improve the flavonoid metabolic network regulated by FaMYB5 and enrich the understanding of the flavonoid metabolic regulatory network in strawberry.

2. Results

2.1. Overexpression of FaMYB5 Promoted the Accumulation of Anthocyanin and Proanthocyanidin in Woodland Strawberry

In our previous studies, transient overexpression of cultivated strawberries showed that FaMYB5 positively regulates flavonoid metabolism in strawberry fruits [21]. To reveal the regulatory role of FaMYB5 in the whole life cycle of strawberry, the pCAMBIA1301-FaMYB5-3XFlag vector was constructed, and the cotyledon of diploid strawberry 'Ruegen' was infected with *Agrobacterium tumefaciens*. Three independent transgenic lines #2, #3 and #5, with high expression levels were identified using GUS staining (Figure 1A), PCR identification (Figure 1B) and RT–qPCR identification (Figure 1C) detection. Anthocyanin and proanthocyanidin in roots, stems and leaves of transgenic plants were measured, and wild-type woodland strawberry with similar growth was used as a control (Figure 1D). The results showed that overexpression of *FaMYB5* increased the anthocyanin content in stems (Figure 1E) and the proanthocyanidin content in roots, stems and leaves (Figure 1F) of woodland strawberry.

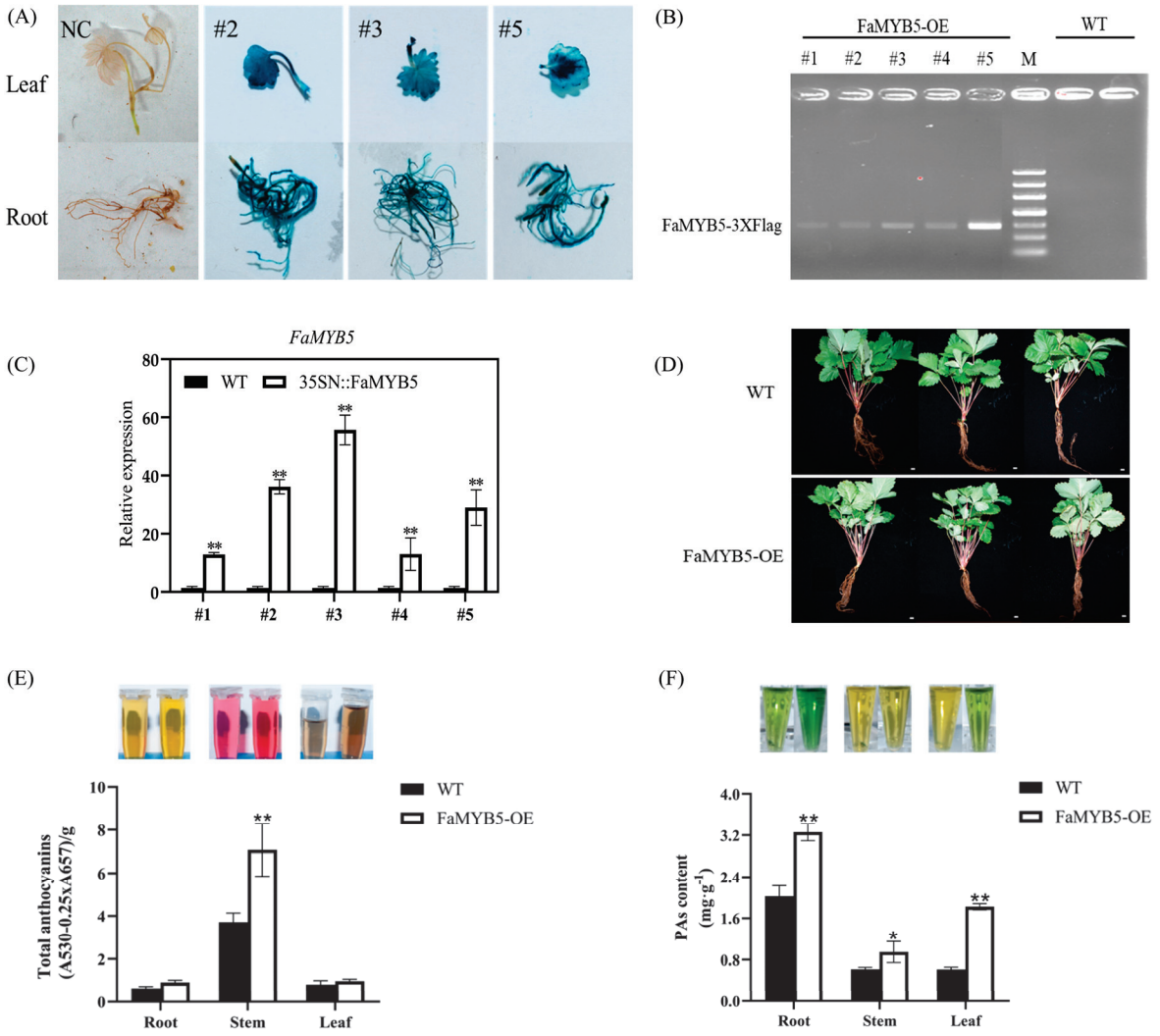


Figure 1. Transgenic woodland strawberry with stable overexpression of *FaMYB5*. (**A**) *FaMYB5* was expressed in roots, stems and leaves with GUS staining; roots and leaves of wild-type woodland

strawberries, which could not be stained by GUS, were used as negative controls. (**B**) Compared with the wild-type, the transgenic plants could amplify the special bands of FaMYB5-3XFLag. (**C**) The relative expression levels of *FaMYB5* in transgenic lines using RT–qPCR. (**D**) The wild-type and FaMYB5-overexpressed plants with the same growth status were selected to determine the content of anthocyanin and proanthocyanidin in leaves, stems and roots; with three biological replicates, bar = 10 cm. (**E**,**F**) Overexpression of FaMYB5 significantly increased the anthocyanin content in stems and proanthocyanidin in roots, stems and leaves compared with wild-type. The red color of the solution indicates the content of anthocyanins; blue solution indicates the content of proanthocyanidin. The proanthocyanidin solution was diluted 20× at the time of determination, and the color was light. Error bars are SEs for three replicates. Significant differences with wild-type were compared using Student's *t* test (** $p < 0.01$; * $p < 0.05$).

2.2. B-BOX Protein FaBBX24 Interacts with FaMYB5

The yeast two-hybrid screening library was used to search for proteins that can interact with FaMYB5. Full-length FaMYB5 showed strong trans-acting activity; then, the *FaMYB5* sequence was cut and a longer fragment FaMYB5^{-615} was selected for further library screening (Figure 2A). A total of 189 potential interacting proteins were screened (Table S1). The potential protein FaBBX24 was inserted into pGBKT7, and BD-FaBBX24 has trans-acting activity. According to the B-BOX domains, FaBBX24 is divided into FaBBX24^{N99} and FaBBX24^{C139} for yeast two-hybrid (Figure 2B). The results show that BD-FaBBX24^{N99} interacts with AD-FaMYB5 (Figure 2C). BiFC assay also demonstrated that FaBBX24 interacts with FaMYB5 in vivo (Figure 2D).

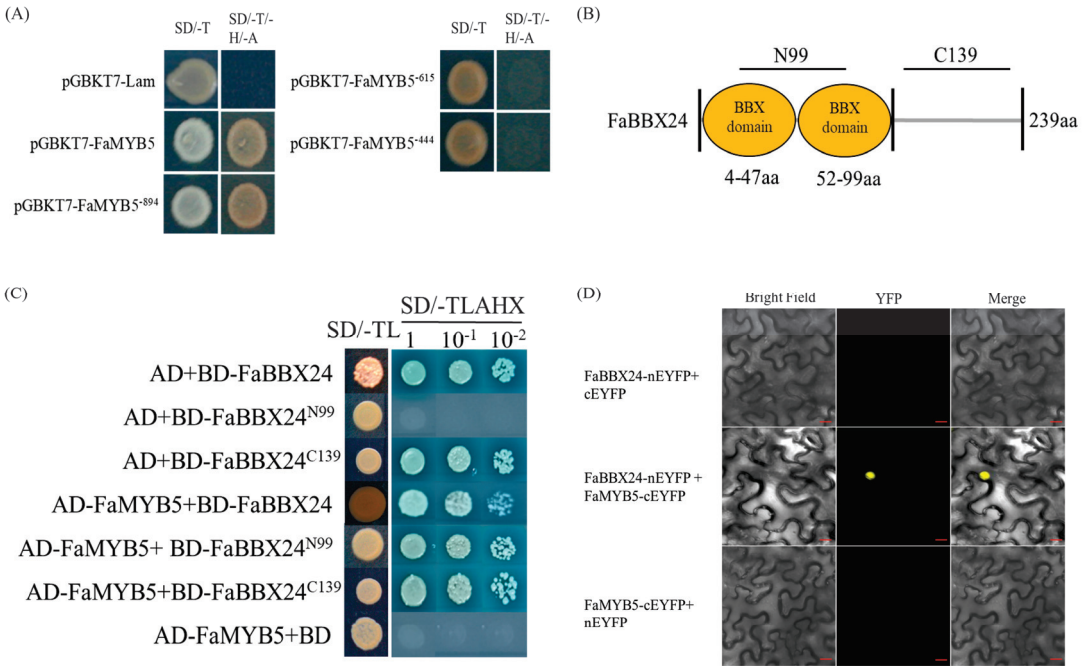


Figure 2. FaMYB5 interacts with FaBBX24. (**A**) FaMYB5 trans-acting activity verification. SD/-T, SD/-Trp medium, SD/-T/-H/-A, SD/-Trp/-His/-Ade medium. n yeast cell. The yeast cells transformed with pGBKT7-Lam vector were used as a negative control. (**B**) The 4-47aa of FaBBX24 belong to B-BOX domain 1, and the 52-99aa belong to B-BOX domain 2, both of which constitute the N terminus of FaBBX24. (**C**) BD-FaBBX24^{N99} interacted with AD-FaMYB5 using yeast two-hybrid, SD/-TL, SD/-Trp/-Leu medium, SD/−TLAHX, SD/-Trp/-Leu/-Ade/-His/X-α-gal. (**D**) Physical association between FaMYB5 and FaBBX24 confirmed with BiFC assay. FaBBX24-nEYFP/ FaMYB5-cEYFP co-transformed tobacco leaves, with FaBBX24-nEYFP/cEYFP and FaMYB5-cEYFP/nEYFP as negative controls; bars = 20 μm.

2.3. Bioinformatics Analysis and Tissue Expression Pattern of FaBBX24

FaBBX24 was cloned from cultivated strawberry 'Benihoppe'. FaBBX24 contains a 717 bp open reading frame (ORF) encoding 238 amino acids, with a predicted protein size of 26.2 kDa and an average isoelectric point (PI) of 4.624. Phylogenetic tree analysis showed that FaBBX24 was closely related to FvBBX24 and clustered with RcBBX24 and PaBBX24 (Figure 3A). Multiple protein sequence alignment showed that FaBBX24 was similar to *Arabidopsis thaliana* AtBBX24, containing B-BOX domain 1 and B-BOX domain 2 and belonging to Group IV B-BOX proteins [42] (Figure 3B). The temporal and spatial expression patterns of *FaBBX24* in different tissues of strawberry 'Benihoppe' and 'Xiaobai' were determined with RT–qPCR (Figure 3C,D). The results showed that *FaBBX24* was expressed in both the fruit development stages and vegetative organs of cultivated strawberry 'Benihoppe', with the highest expression level in the white stage and functional leaf and the lowest expression level in the root and stem. In cultivated strawberry 'Xiaobai', the expression level of *FaBBX24* gradually increased with fruit ripening and reached the maximum in the full red stage, while the expression level was low in the root, stem, leaf and flower.

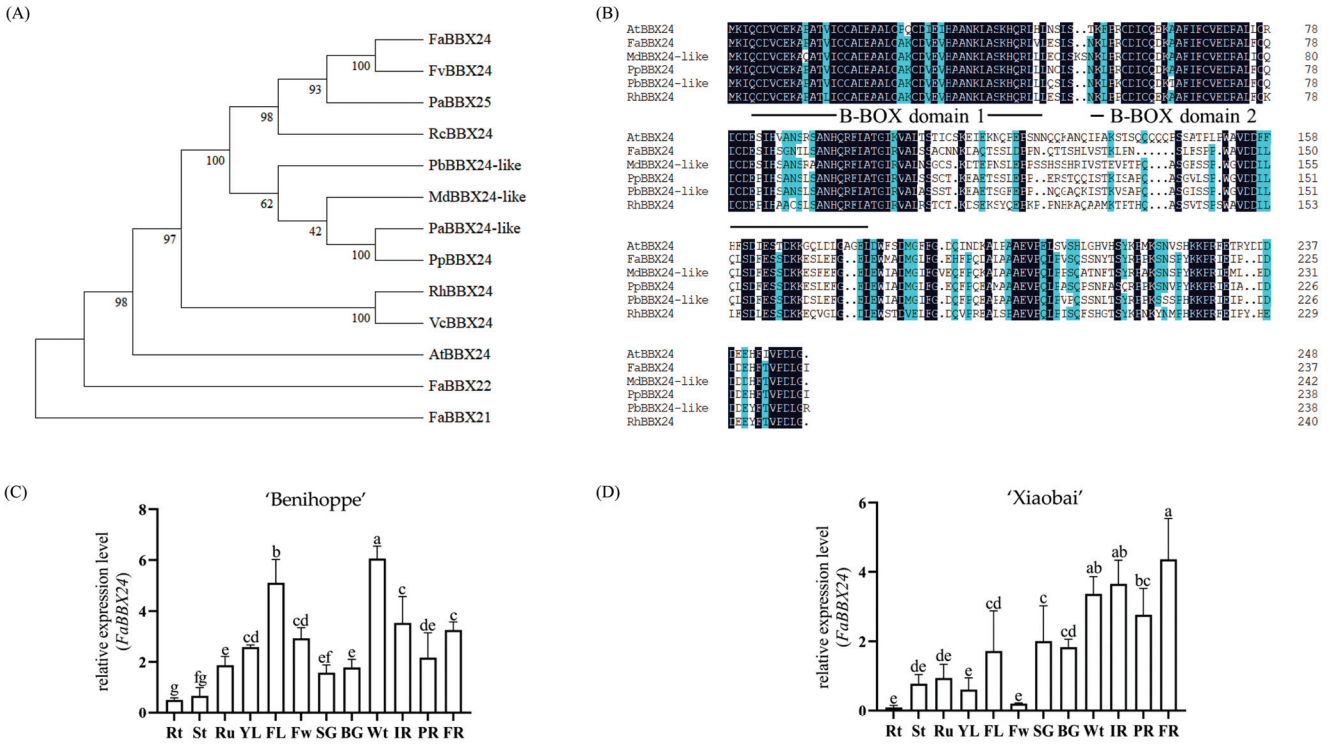


Figure 3. Bioinformatics analysis and spatiotemporal expression patterns in strawberry tissues of *FaBBX24*. (**A**) Phylogenetic analysis of FaBBX24 with its homologous proteins from other species. The phylogenetic tree was constructed using the Neighbor-Joining Tree Method in MEGA6, and the bootstrap was 1000 replicates. (**B**) Multiple protein sequence alignment of FaBBX24 with other BBX24 proteins. The black line marks the B-BOX domain. The gene accession numbers that have appeared above are as follows: FvBBX24(XM_004309761.2), PaBBX25(XM_050517417.1), RcBBX24(XM_024337817.2), PbBBX24-like(XM_048579829.1), MdBBX24-like(NM_001328919.1), PaBBX24-like(XM_021967653.1), PpBBX24(XM_007223749.2), RhBBX24(OL690502.1), VcBBX24(OP957064.1), LbBBX24(MH813941.1) and AtBBX24(NM_100484.4). (**C**,**D**) Expression pattern of *FaBBX24* in different tissues of cultivated strawberry 'Benihoppe' and 'Xiaobai'. Rt—root; St—stem; Ru—runner; YL—young leaf; FL—functional leaf; Fw—flower; SG—small green; BG—big green; Wt—white; IR—initially red; PR—partially red; FR—full red. Error bars are SEs for three replicates. Different letters above the bars indicate significantly different values ($p < 0.05$) according to a Least Significant Difference (LSD) test.

2.4. FaMYB5 Interacted with FaBBX24 to Promote Anthocyanin and Proanthocyanidin Accumulation in Strawberry Fruits

In previous studies, we found that FaMYB5 can up-regulate structural genes *PAL*, *C4H*, *F3′H* and *LAR* and can directly regulate *F3′H* and *LAR* promoters to participate in anthocyanin and proanthocyanidin metabolism [21]. In this study, *FaMYB5*, *FaBBX24* and co-expression of *FaMYB5* and *FaBBX24* were transiently overexpressed in cultivated strawberry 'Xiaobai' (Figure 4A). Transient overexpression of *FaBBX24* was able to induce anthocyanin accumulation in strawberry fruits (Figure 4A). RT–qPCR results show that FaBBX24 was able to significantly up-regulate the anthocyanin structural gene *ANS* (Figure 4D). Co-expression of *FaMYB5* and *FaBBX24* significantly increased anthocyanin and proanthocyanidin content in fruit (Figure 4B,C); moreover, co-expression significantly increased the expression levels of structural genes *F3′H*, *4CL-2*, *TT12* and *AHA10* (Figure 4D,E).

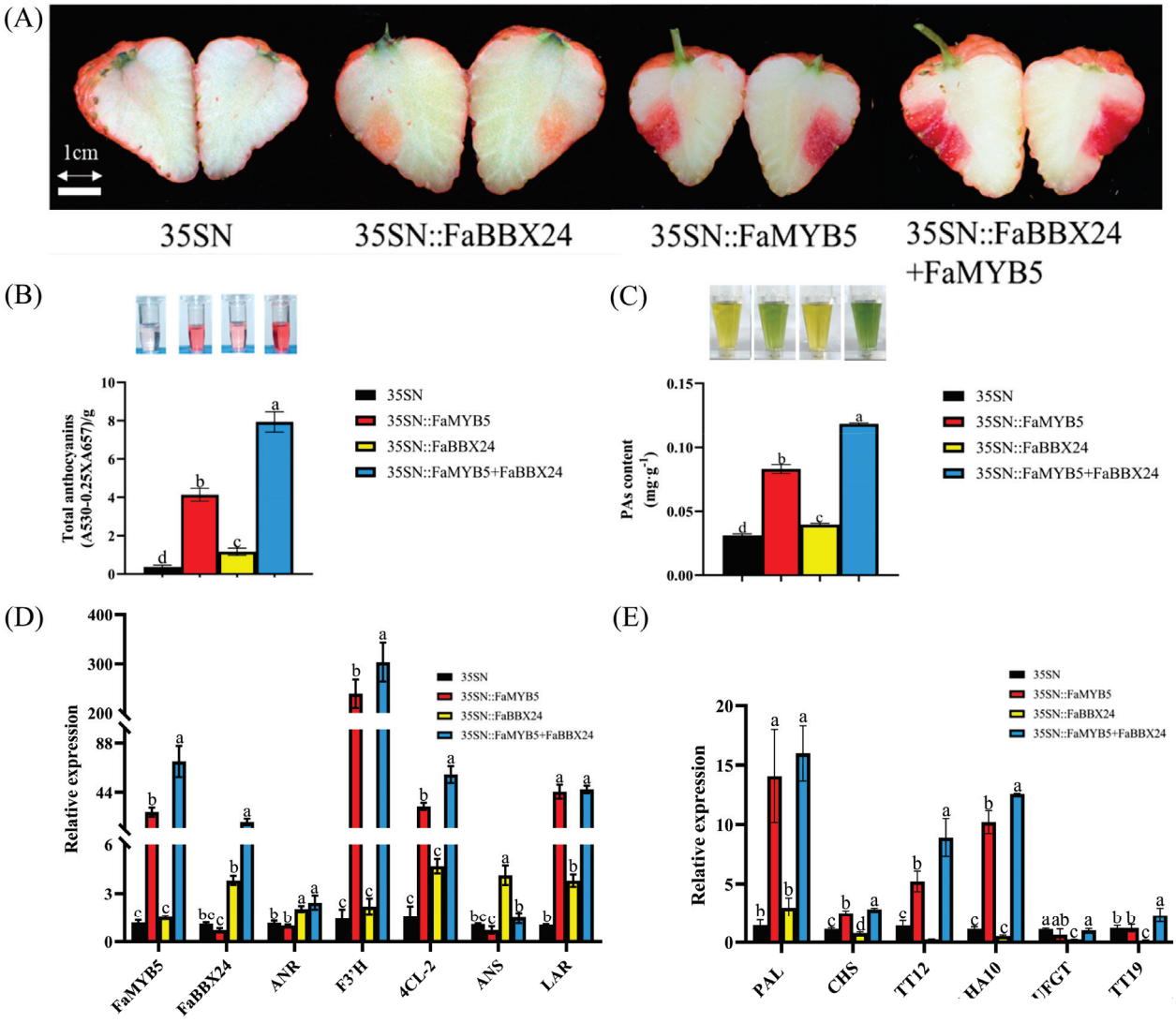


Figure 4. Co-expression of *FaMYB5* and *FaBBX24* in strawberry fruits. (**A**) Transient overexpression of *FaMYB5* and *FaBBX24* in cultivated strawberry 'Xiaobai'. Bars = 1 cm. (**B**,**C**) Anthocyanin and proanthocyanidin content in strawberry fruits with overexpression of *FaBBX24* and *FaMYB5*. The red color of the solution indicates the content of anthocyanins; blue solution indicates the content of proanthocyanidin. (**D**,**E**) The relative expression levels of anthocyanin and proanthocyanidin structural genes in strawberry fruits overexpressing *FaMYB5* and *FaBBX24* were determined with RT–qPCR. Error bars are SEs for three replicates. Different letters above the bars indicate significantly different values ($p < 0.05$) compared to 35SN according to a Least Significant Difference (LSD) test.

3. Discussion

In previous studies, transient overexpression of *FaMYB5* in strawberry fruits has been shown to promote anthocyanin and proanthocyanidin accumulation in strawberry fruit. In this study, we obtained stable *FaMYB5* overexpression plants and measured the anthocyanin content in each organ. The results show that compared with wild-type, overexpression of *FaMYB5* promoted the accumulation of anthocyanin and proanthocyanidin in roots, stems and leaves, which may help plants resist adversity [43].

The MYB protein family plays an important role in plant phenylpropane metabolism [44]. According to the number and position of its domains, it has been divided into four subclasses: 1R (R1/R2, R3-MYB), 2R (R2R3-MYB), 3R (R1R2R3-MYB) and 4R (containing four similar R1/R2 repeats) [45], with the 2R protein being the largest class of MYB factors in plants. They are grouped into 22 subgroups on the basis of the conserved amino acid sequence motifs present in their most C-terminal MYB domain [46], which are involved in primary and secondary metabolism and determine cell life and properties, hormone signaling, developmental regulation and responses to biological and abiotic stresses [47–49]. In recent years, more and more studies on the regulation of flavonoid metabolism by R2R3-MYB have been reported, including studies on PtMYB6 [50], NtMYB330 [51], AN2like [52] and CaMYB39 [53]. The lack of R2 motif in FaMYB5 suggests that FaMYB5 is a potential negative regulator of anthocyanin synthesis [26]. In previous studies [21,22] and this study, we have demonstrated that FaMYB5 is a typical R2R3-MYB protein and positively regulates anthocyanin and proanthocyanidin biosynthesis in strawberry.

Flavonoids such as anthocyanin and proanthocyanidin are important products of phenylpropanoid metabolism. In addition to directly regulating structural genes of flavonoid biological metabolism, MYB proteins usually form complexes with other proteins to perform functions, such as the MYB-bHLH-WD40 complex [54,55]. It has also been reported that proteins other than bHLH protein interact with MYB transcription factors to regulate flavonoid metabolism. For example, Pp4ERF24 and Pp12ERF96 in pear were able to enhance the expression of *PpUFGT* and promoted anthocyanin synthesis through interaction with PpMYB114 [56], and PyERF3-PyMYB114-PybHLH3 complex promoted anthocyanin accumulation in pear peel by regulating target gene *ANS* [57]. Jasmonic acid domain protein interacted with AtMYB75 and bHLH protein TT8/GL3/EGL3, suggesting that it may promote anthocyanin synthesis and trichome initiation by regulating MBW complex and activating downstream signaling cascade [58]. The GARP-type transcription factor GLK1 regulated anthocyanin biosynthesis in *Arabidopsis thaliana* by interacting with MYB75, MYB90 and MYB113 [59]. COP1/SPA interacted with PAP1 and PAP2, which can affect their functions at the transcription and translation levels, and then regulated anthocyanin synthesis in *Arabidopsis thaliana* [60]. MYB75 interacted with MPK4, which regulates phosphorylation of MYB75 through photoinduction to maintain its stability, and played an important role in photoinduced anthocyanin synthesis in *Arabidopsis thaliana* [61]. In apples, MdBT2 interacted with MdMYB1 and target ubiquitination to degrade MdMYB1 and regulated the biological metabolism of anthocyanin [62]. MdNAC42-MdMYB10 was an important component of the regulatory network controlling anthocyanin accumulation in red-flesh apples [63]. MdbZIP44 and MdERF78 partnered with MdMYB1 to activate downstream target genes and promoted anthocyanin accumulation [64,65]. Here, we reveal the BBX protein and MYB protein interaction module and speculate that FaBBX24 interacted with FaMYB5 to enhance the transcriptional activation of FaMYB5 on *F3′H*, *TT12*, *AHA10* and *4CL-1* and promoted anthocyanin accumulation in strawberry fruits.

In this study, we screened the B-BOX protein FaBBX24, whose N-terminal contains two B-BOX domains that are involved in protein–protein interactions [66]. B-BOX proteins, as an important participant in the plant light-signaling pathway, have been shown to play an important role in light-mediated anthocyanin accumulation [67,68]. Among them, AtBBX20 [69], AtBBX21 [70], MdBBX20 [71], MdBBX21 [36], MdBBX22 [72], PpBBX16 [38], PpBBX18 [37], PavBBX6/9 [40] and FaBBX22 [34] have been proven to be positive regulators of anthocyanin, and AtBBX19 [73], AtBBX24 [74], AtBBX25 [75], AtBBX32 [76],

PpBBX21 [77], PpBBX24 [78], MdBBX37 [79], MdCOL4 [80] and VvBBX44 inhibit anthocyanin biosynthesis. The same transcription factor in different species may perform opposite functions. MdMYB1 plays a central role in promoting anthocyanin accumulation and directly regulates anthocyanin structural genes [81,82], while MYB1 in strawberry is a repressor of flavonoid metabolism [26,83,84]. It has been reported that PpBBX24 causes the red skin of 'Zaosured' due to the deletion of 14 nucleotides [78], and these results indicated that PpBBX24 negatively regulated anthocyanin in pear. The amino acid sequence of FaBBX24 and PpBBX24 is highly similar (80.91%), but this study showed that FaBBX24 plays a role in promoting anthocyanin accumulation in strawberry fruits. Previous studies have found that FaMYB5 can promote the biosynthesis of proanthocyanidin in strawberry fruits by combining with *LAR* promoter. This research proves that the interaction between FaMYB5 and FaBBX24 did not further increase the *LAR* expression level. It is speculated that the increase in proanthocyanidin content after co-expression was due to the regulation of proanthocyanidin structural gene *ANR* by FaBBX24. Therefore, FaBBX24 is a positive regulator of anthocyanin and proanthocyanidin biosynthesis in strawberry fruits.

B-BOX proteins usually interact with HY5 and act as upstream MYB transcription factors such as MYB1 or MYB10 to regulate anthocyanin biosynthesis [28,79]. FaMYB5 has been proven to be a key transcription factor for anthocyanin metabolism in strawberry. However, RT–qPCR data showed that the expression level of *FaMYB5* did not change in strawberry fruits with *FaBBX24* overexpression. Therefore, whether FaBBX24 can directly regulate FaMYB5 remains to be further explored.

4. Materials and Methods

4.1. Plant Materials and Growth Conditions

The plant materials were woodland strawberry (*Fragaria vesca*) 'Ruegen' and its aseptic tissue culture seedlings; tobacco (*Nicotiana benthamiana*), which was grown in a greenhouse of the Horticulture College of Sichuan Agricultural University; and the cultivated strawberry (*Fragaria* × *ananassa*) 'Benihoppe' and 'Xiaobai' fruits which were picked from Xin Yue Strawberry Garden (Chengdu, China). Environmental conditions were controlled at 22 ± 2 °C, 80–90% relative humidity and 16/8 h light–dark cycle with 220 μmol m^{-2} s^{-1}. The developmental stages of fruit are defined as small green, big green, white, partially red and full red [21]. All collected samples were immediately snap-frozen with liquid nitrogen and stored at −80 °C for further use.

4.2. Gene Clone and RT–qPCR Analysis

The total RNA of fruit samples was extracted using the improved CTAB method [85] and then reversed using RT Easy™ II (with gDNase) (Foregene, Chengdu, China). cDNA was diluted tenfold and used as a template for gene cloning. The FaBBX24 sequences are shown in Figure S1. Cloning primers were designed using SnapGene software, as shown in Table S2 (all primers used in this article are in this table).

RT–qPCR was performed by a CFX Connect real-time system ((Bio-rad, Hercules, CA, USA), and the interspace 26S-18S RNA was used as a reference gene [86]. Primers were designed in NCBI (https://www.ncbi.nlm.nih.gov/) [Accessed on 20 April 2020], and the relative expression level was analyzed using the $2^{-\Delta\Delta Ct}$ method [87]. Hieff® qPCR SYBR Green Master Mix (No Rox) (Yeasen Biotechnology Co., Ltd., Shanghai, China) was used for detecting the PCR products on a CFX96 Real-time reaction system (Bio-Rad, Hercules, CA, USA). Three biological replicates were set, and each biological replicate was set with three technical replicates.

4.3. Strawberry Genetic Transformation

FaMYB5 coding sequence (previously cloned in the laboratory) was inserted into pCAMBIA1301-Chip (with 3X Flag tag), and strawberry leaf rounds were infected with *Agrobacterium*-mediated genetic transformation. The specific method is as follows: strawberry leaf rounds were infected with *Agrobacterium tumefaciens* GV3101 containing pCAMB-

IA1301-FaMYB5-Chip (with 3X Flag tag); then, the leaf rounds were placed in 25 mL medium (MS salts + B5 vitamins +6-BA 3.0 mg·L^{-1} + IBA 0.2 mg·L^{-1} + 2% Sucrose + 1% Glucose + 0.25% Phytagel + 100 μM Acetosyringone, pH 5.5) in a petri dish for dark culture at 25 °C for 2 days. The leaf rounds were transferred to fresh culture bottles (MS salts + B5 vitamins + 6-BA 3.0 mg/L + IBA 0.2 mg/L + 3% Sucrose + 0.25% Phytagel + 250 mg/L Timentin + 250 mg/L Carbenicillin + 5 mg/L Hygromycin B, pH 5.8), 22 °C, 16 h (light) /8 h (dark), light intensity of 40 μmol·m^{-2}·s^{-1} (about 2200 lx) for 15 days. The culture medium was then transferred every 20 days, during which the necrotic tissue was carefully excised and part of the green callus was transferred to a new screening medium until adventitious buds developed. When the indeterminate bud grew to 0.5–1 cm, it was removed from the base and transferred to a secondary medium (MS + 6-BA 0.2 mg/L + GA3 0.2 mg/L + 5 mg/L Hygromycin B + 3% Sucrose + 0.7% Agar, pH 5.8) for further culture. After identification of positive plants, 2–3 cm adventitive buds were cut off and transferred to rooting medium (1/2 MS +5 mg/L Hygromycin B + 3% Sucrose + 0.7% Agar, pH 5.8) under the same culture conditions as above to obtain intact plants.

4.4. Yeast Two-Hybrid Assay

The yeast two-hybrid assays were performed, according to the manufacturer's instructions using the Matchmaker Gold Yeast Two-Hybrid System kit (Takara, Beijing, China). The coding sequence encoding the N-terminal of FaBBX24 (amino acids 1–99) and the C-terminal of FaBBX24 (amino acids 100–238) was cloned into pGBKT7 (Clontech) to generate the bait vector (BD-FaBBX24^{N99}, BD-FaBBX24^{C139}) containing the GAL4 DNA-BD sequence. The full-length coding sequence of FaMYB5 was individually cloned into pGADT7 vector (Clontech) to produce the prey vectors (AD-FaMYB5) containing the sequence encoding the GAL4 activation domain (AD). Yeast strain AH109 was co-transformed with the bait and prey vectors, and then protein interactions were evaluated on the basis of the ability of the cells to grow on synthetic defined (SD) medium lacking Leu, Trp, His, Ade and +X-α-gal after 4 days of growth at 28 °C.

4.5. BiFC Assay

The candidate genes were amplified with PCR using primers with restriction sites, and the fluorescent complementary vectors pXY104-cEYFP-FaMYB5 and pXY103-nEYFP-FaBBX24 were constructed. The recombinant plasmids were transformed into *Agrobacterium tumefaciens* GV3101, and the transformed positive and negative plasmids were used as control plasmids and then injected into tobacco leaves. Next, 48 h later, Laser Scanning Confocal Microscope (FV3000, OLYMPUS, Tokyo, Japan) was used to observe fluorescence signals in tobacco cells.

4.6. Transient Overexpression in Strawberry Fruit

The CDS of *FaBBX24* and *FaMYB5* were inserted into pCAMBIA1301 vector and transferred into *Agrobacterium tumefaciens* GV3101. A 1 mL sterile microsyringe was used to inject 300–500 μL subcutaneously into the fruit pedicles of strawberries at white stage. The strawberries after injection were dark cultured overnight at 18 °C and then transferred to normal material culture conditions; the steps of transfection are described in detail elsewhere [88].

4.7. Anthocyanin and Proanthocyanidin Content Measurement

The measurement of anthocyanin has been reported previously [89]. Specifically, 0.2 g plant material (extract: fresh weight of plant = 10:1) was ground with liquid nitrogen, and 2 mL anthocyanin extract (methanol:H_2O:formic acid:trifluoroacetic acid, 70:27:2:1) was placed in a 40 °C water bath to avoid light for 4 h, followed by centrifugal collection of supernatants with UV spectrophotometer (UV-530pc, MAPADA, Shanghai, China) to measure the absorbance of A530, A657. The calculation formula is total anthocyanin = [A530

− (0.25 × A657)]/M, where A530 and A657 are absorbance at corresponding wavelengths, and M is fresh weight of plants.

Proanthocyanidin content measurement was described previously [90]. Specifically, 0.15 g of the sample was weighed (fully ground with liquid nitrogen), and 3 mL of proanthocyanidin extract (acetone:water:glacial acetic acid = 150:49:1) was added and shaken at 150 rpm for 1 h at 20 °C. The samples were centrifuged at 12 °C for 20 min at 10,000 rpm, and the supernatant was retained as the proanthocyanidin extract. After the extraction solution, 80% ethanol and 1% DMACA solution were mixed at 1:9:30, substance. Varioskan™ LUX (Thermo Fisher Scientific, Waltham, MA, USA) was used to determine OD value at 640 nm. Proanthocyanidin B2 is the standard.

4.8. Statistical Analysis

All the experimental data are expressed as the mean ± standard deviation from the mean (SD). The statistical analysis was performed using Student's *t* test (** $p < 0.01$) and Least Significant Difference (LSD) test in IBM SPSS Statistics software, version 28.0 (IBM, Chicago, IL, USA).

5. Conclusions

In summary, by obtaining *FaMYB5* overexpression plants and transient overexpression of *FaMYB5* and *FaBBX24*, we determined that FaMYB5 and FaBBX24 are promoting factors for anthocyanin and proanthocyanidin biosynthesis of strawberry. It was also proven that FaMYB5 can further promote the expression of anthocyanin and proanthocyanidin structural genes *F3′H*, *TT12*, *AHA10*, *4CL-2* and *ANR* by forming a complex with FaBBX24 and can regulate the biosynthesis of anthocyanidin and proanthocyanidin in strawberry fruit. This study provides new insight into the transcriptional regulation of anthocyanin and proanthocyanidin by FaMYB5 and FaBBX24 in strawberry.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms241512185/s1.

Author Contributions: Conceptualization, L.Z. and H.T.; formal analysis, Y.W. (Yiping Wang), Y.Z. (Yunting Zhang) and Q.C.; performed the experiments, L.Z., Y.W. (Yiping Wang) and N.Z.; methodology, M.Y., L.J., N.Z. and Y.W. (Yan Wang); software, L.J., M.Y. and M.L.; validation, N.Z., L.Z. and Y.Z. (Yong Zhang); writing—original draft, L.Z. and Y.W. (Yiping Wang); writing—review and editing, H.T., Y.L. (Yuanxiu Lin) and Y.L. (Ya Luo). 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, grant number 31872083.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Capocasa, F.; Diamanti, J.; Tulipani, S.; Battino, M.; Mezzetti, B. Breeding strawberry (*Fragaria* × *ananassa* Duch) to increase fruit nutritional quality. *Biofactors* **2008**, *34*, 67–72. [CrossRef] [PubMed]
2. Tanaka, Y.; Sasaki, N.; Ohmiya, A. Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *Plant J.* **2008**, *54*, 733–749. [CrossRef]
3. Hanhineva, K.; Rogachev, I.; Kokko, H.; Mintz-Oron, S.; Venger, I.; Karenlampi, S.; Aharoni, A. Non-targeted analysis of spatial metabolite composition in strawberry (*Fragaria* × *ananassa*) flowers. *Phytochemistry* **2008**, *69*, 2463–2481. [CrossRef] [PubMed]
4. Pillet, J.; Yu, H.-W.; Chambers, A.H.; Whitaker, V.M.; Folta, K.M. Identification of candidate flavonoid pathway genes using transcriptome correlation network analysis in ripe strawberry (*Fragaria* × *ananassa*) fruits. *J. Exp. Bot.* **2015**, *66*, 4455–4467. [CrossRef] [PubMed]

5. 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]
6. Xu, W.; Grain, D.; Bobet, S.; Le Gourrierec, J.; Thevenin, J.; Kelemen, Z.; Lepiniec, L.; Dubos, C. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in *Arabidopsis* seed. *New Phytol.* **2014**, *202*, 132–144. [CrossRef]
7. Nix, A.; Paull, C.A.; Colgrave, M. The flavonoid profile of pigeonpea, *Cajanus cajan*: A review. *SpringerPlus* **2015**, *4*, 125. [CrossRef]
8. Liu, J.; Osbourn, A.; Ma, P. MYB Transcription Factors as Regulators of Phenylpropanoid Metabolism in Plants. *Mol. Plant* **2015**, *8*, 689–708. [CrossRef]
9. Xu, P.; Wu, L.; Cao, M.; Ma, C.; Xiao, K.; Li, Y.; Lian, H. Identification of MBW Complex Components Implicated in the Biosynthesis of Flavonoids in Woodland Strawberry. *Front. Plant Sci.* **2021**, *12*, 774943. [CrossRef]
10. Goswami, G.; Nath, U.K.; Park, J.-I.; Hossain, M.R.; Biswas, M.K.; Kim, H.-T.; Kim, H.R.; Nou, I.-S. Transcriptional regulation of anthocyanin biosynthesis in a high-anthocyanin resynthesized *Brassica napus* cultivar. *J. Biol. Res. Thessalon.* **2018**, *25*, 19. [CrossRef]
11. Pazares, J.; Ghosal, D.; Wienand, U.; Peterson, P.A.; Saedler, H. The regulatory c1 locus of *Zea mays* encodes a protein with homology to MYB protooncogene products and with structural similarities to transcriptional activators. *EMBO J.* **1987**, *6*, 3553–3558. [CrossRef] [PubMed]
12. Feng, K.; Liu, J.-X.; Duan, A.-Q.; Li, T.; Yang, Q.-Q.; Xu, Z.-S.; Xiong, A.-S. AgMYB2 transcription factor is involved in the regulation of anthocyanin biosynthesis in purple celery (*Apium graveolens* L.). *Planta* **2018**, *248*, 1249–1261. [CrossRef]
13. Xu, Z.-S.; Feng, K.; Que, F.; Wang, F.; Xiong, A.-S. A MYB transcription factor, DcMYB6, is involved in regulating anthocyanin biosynthesis in purple carrot taproots. *Sci. Rep.* **2017**, *7*, 45324. [CrossRef] [PubMed]
14. Ban, Y.; Honda, C.; Hatsuyama, Y.; Igarashi, M.; Bessho, H.; Moriguchi, T. Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. *Plant Cell Physiol.* **2007**, *48*, 958–970. [CrossRef] [PubMed]
15. Vimolmangkang, S.; Han, Y.; Wei, G.; Korban, S.S. An apple MYB transcription factor, *MdMYB3*, is involved in regulation of anthocyanin biosynthesis and flower development. *BMC Plant Biol.* **2013**, *13*, 176. [CrossRef]
16. Jiang, S.; Sun, Q.; Zhang, T.; Liu, W.; Wang, N.; Chen, X. MdMYB114 regulates anthocyanin biosynthesis and functions downstream of MdbZIP4-like in apple fruit. *J. Plant Physiol.* **2021**, *257*, 153353. [CrossRef]
17. Espley, R.V.; Hellens, R.P.; Putterill, J.; Stevenson, D.E.; Kutty-Amma, S.; Allan, A.C. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* **2007**, *49*, 414–427. [CrossRef]
18. Zhang, S.; Wang, H.; Wang, T.; Liu, W.; Zhang, J.; Fang, H.; Zhang, Z.; Peng, F.; Chen, X.; Wang, N. MdMYB305-MdbHLH33-MdMYB10 regulates sugar and anthocyanin balance in red-fleshed apple fruits. *Plant J.* **2023**, *113*, 1062–1079. [CrossRef]
19. Cui, D.; Zhao, S.; Xu, H.; Allan, A.C.; Zhang, X.; Fan, L.; Chen, L.; Su, J.; Shu, Q.; Li, K. The interaction of MYB, bHLH and WD40 transcription factors in red pear (*Pyrus pyrifolia*) peel. *Plant Mol. Biol.* **2021**, *106*, 407–417. [CrossRef]
20. Rahim, M.A.; Busatto, N.; Trainotti, L. Regulation of anthocyanin biosynthesis in peach fruits. *Planta* **2014**, *240*, 913–929. [CrossRef]
21. Jiang, L.; Yue, M.; Liu, Y.; Zhang, N.; Lin, Y.; Zhang, Y.; Wang, Y.; Li, M.; Luo, Y.; Zhang, Y.; et al. A novel R2R3-MYB transcription factor FaMYB5 positively regulates anthocyanin and proanthocyanidin biosynthesis in cultivated strawberries (*Fragaria* × *ananassa*). *Plant Biotechnol. J.* **2023**, *21*, 1140–1158. [CrossRef] [PubMed]
22. Yue, M.; Jiang, L.; Zhang, N.; Zhang, L.; Liu, Y.; Lin, Y.; Zhang, Y.; Luo, Y.; Zhang, Y.; Wang, Y.; et al. Regulation of flavonoids in strawberry fruits by FaMYB5/FaMYB10 dominated MYB-bHLH-WD40 ternary complexes. *Front. Plant Sci.* **2023**, *14*, 1145670. [CrossRef] [PubMed]
23. Chen, G.; Xu, P.; Pan, J.; Li, Y.; Zhou, J.; Kuang, H.; Lian, H. Inhibition of FvMYB10 transcriptional activity promotes color loss in strawberry fruit. *Plant Sci.* **2020**, *298*, 110578. [CrossRef]
24. Medina-Puche, L.; Cumplido-Laso, G.; Amil-Ruiz, F.; Hoffmann, T.; Ring, L.; Rodriguez-Franco, A.; Luis Caballero, J.; Schwab, W.; Munoz-Blanco, J.; Blanco-Portales, R. MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria ananassa* fruits. *J. Exp. Bot.* **2014**, *65*, 401–417. [CrossRef]
25. Martinez-Rivas, F.J.; Blanco-Portales, R.; Serratosa, M.P.; Ric-Varas, P.; Guerrero-Sanchez, V.; Medina-Puche, L.; Moyano, L.; Mercado, J.A.; Alseekh, S.; Caballero, J.L.; et al. FaMYB123 interacts with FabHLH3 to regulate the late steps of anthocyanin and flavonol biosynthesis during ripening. *Plant J.* **2023**, *114*, 683–698. [CrossRef]
26. Schaart, J.G.; Dubos, C.; De La Fuente, I.R.; van Houwelingen, A.M.M.L.; de Vos, R.C.H.; Jonker, H.H.; Xu, W.; Routaboul, J.-M.; Lepiniec, L.; Bovy, A.G. Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (*Fragaria* × *ananassa*) fruits. *New Phytol.* **2013**, *197*, 454–467. [CrossRef] [PubMed]
27. Yadav, A.; Bakshi, S.; Yadukrishnan, P.; Lingwan, M.; Dolde, U.; Wenkel, S.; Masakapalli, S.K.; Datta, S. The B-Box-Containing MicroProtein miP1a/BBX31 Regulates Photomorphogenesis and UV-B Protection. *Plant Physiol.* **2019**, *179*, 1876–1892. [CrossRef]
28. Xu, D. COP1 and BBXs-HY5-mediated light signal transduction in plants. *New Phytol.* **2020**, *228*, 1748–1753. [CrossRef]
29. Ouyang, Y.; Zhang, X.; Wei, Y.; He, Y.; Zhang, X.; Li, Z.; Wang, C.; Zhang, H. AcBBX5, a B-box transcription factor from pineapple, regulates flowering time and floral organ development in plants. *Front. Plant Sci.* **2022**, *13*, 1060276. [CrossRef]
30. Ye, Y.; Liu, Y.; Li, X.; Wang, G.; Zhou, Q.; Chen, Q.; Li, J.; Wang, X.; Tang, H. An Evolutionary Analysis of B-Box Transcription Factors in Strawberry Reveals the Role of FaBBx28c1 in the Regulation of Flowering Time. *Int. J. Mol. Sci.* **2021**, *22*, 11766. [CrossRef]

31. Li, Y.; Shi, Y.; Li, M.; Fu, D.; Wu, S.; Li, J.; Gong, Z.; Liu, H.; Yang, S. The CRY2-COP1-HY5-BBX7/8 module regulates blue light-dependent cold acclimation in *Arabidopsis*. *Plant Cell* **2021**, *33*, 3555–3573. [CrossRef]
32. Dai, Y.; Lu, Y.; Zhou, Z.; Wang, X.; Ge, H.; Sun, Q. B-box containing protein 1 from *Malus domestica* (MdBBX1) is involved in the abiotic stress response. *PeerJ* **2022**, *10*, e12852. [CrossRef]
33. Bu, X.; Wang, X.; Yan, J.; Zhang, Y.; Zhou, S.; Sun, X.; Yang, Y.; Ahammed, G.J.; Liu, Y.; Qi, M.; et al. Genome-Wide Characterization of B-Box Gene Family and Its Roles in Responses to Light Quality and Cold Stress in Tomato. *Front. Plant Sci.* **2021**, *12*, 698525. [CrossRef] [PubMed]
34. Liu, Y.; Ye, Y.; Wang, Y.; Jiang, L.; Yue, M.; Tang, L.; Jin, M.; Zhang, Y.; Lin, Y.; Tang, H. B-Box Transcription Factor FaBBX22 Promotes Light-Induced Anthocyanin Accumulation in Strawberry (*Fragaria* × *ananassa*). *Int. J. Mol. Sci.* **2022**, *23*, 7757. [CrossRef]
35. Plunkett, B.J.; Henry-Kirk, R.; Friend, A.; Diack, R.; Helbig, S.; Mouhu, K.; Tomes, S.; Dare, A.P.; Espley, R.V.; Putterill, J.; et al. Apple B-box factors regulate light-responsive anthocyanin biosynthesis genes. *Sci. Rep.* **2019**, *9*, 17762. [CrossRef]
36. Zhang, B.; Zhu, Z.-Z.; Qu, D.; Wang, B.-C.; Hao, N.-N.; Yang, Y.-Z.; Yang, H.-J.; Zhao, Z.-Y. MdBBX21, a B-Box Protein, Positively Regulates Light-Induced Anthocyanin Accumulation in Apple Peel. *Front. Plant Sci.* **2021**, *12*, 774446. [CrossRef] [PubMed]
37. Bai, S.; Tao, R.; Yin, L.; Ni, J.; Yang, Q.; Yan, X.; Yang, F.; Guo, X.; Li, H.; Teng, Y. Two B-box proteins, PpBBX18 and PpBBX21, antagonistically regulate anthocyanin biosynthesis via competitive association with *Pyrus pyrifolia* ELONGATED HYPOCOTYL 5 in the peel of pear fruit. *Plant J.* **2019**, *100*, 1208–1223. [CrossRef]
38. Bai, S.; Tao, R.; Tang, Y.; Yin, L.; Ma, Y.; Ni, J.; Yan, X.; Yang, Q.; Wu, Z.; Zeng, Y.; et al. BBX16, a B-box protein, positively regulates light-induced anthocyanin accumulation by activating MYB10 in red pear. *Plant Biotechnol. J.* **2019**, *17*, 1985–1997. [CrossRef]
39. Liu, W.; Tang, R.; Zhang, Y.; Liu, X.; Gao, Y.; Dai, Z.; Li, S.; Wu, B.; Wang, L. Genome-wide identification of B-box proteins and VvBBX44 involved in light-induced anthocyanin biosynthesis in grape (*Vitis vinifera* L.). *Planta* **2021**, *253*, 114. [CrossRef] [PubMed]
40. Wang, Y.; Xiao, Y.; Sun, Y.; Zhang, X.; Du, B.; Turupu, M.; Yao, Q.; Gai, S.; Tong, S.; Huang, J.; et al. Two B-box proteins, PavBBX6/9, positively regulate light-induced anthocyanin accumulation in sweet cherry. *Plant Physiol.* **2023**, *192*, 2030–2048. [CrossRef]
41. Li, C.F.; Pei, J.L.; Yan, X.; Cui, X.; Tsuruta, M.; Liu, Y.; Lian, C.L. A poplar B-box protein PtrBBX23 modulates the accumulation of anthocyanins and proanthocyanidins in response to high light. *Plant Cell Environ.* **2021**, *44*, 3015–3033. [CrossRef]
42. Khanna, R.; Kronmiller, B.; Maszle, D.R.; Coupland, G.; Holm, M.; Mizuno, T.; Wu, S.-H. The *Arabidopsis* B-Box Zinc Finger Family. *Plant Cell* **2009**, *21*, 3416–3420. [CrossRef]
43. Naing, A.H.; Kim, C.K. Abiotic stress-induced anthocyanins in plants: Their role in tolerance to abiotic stresses. *Physiol. Plant.* **2021**, *172*, 1711–1723. [CrossRef] [PubMed]
44. Feller, A.; Machemer, K.; Braun, E.L.; Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **2011**, *66*, 94–116. [CrossRef] [PubMed]
45. Du, H.; Feng, B.-R.; Yang, S.-S.; Huang, Y.-B.; Tang, Y.-X. The R2R3-MYB Transcription Factor Gene Family in Maize. *PLoS ONE* **2012**, *7*, e37463. [CrossRef] [PubMed]
46. Stracke, R.; Werber, M.; Weisshaar, B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [CrossRef]
47. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **2010**, *15*, 573–581. [CrossRef]
48. Xue, L.; Huang, X.; Zhang, Z.; Lin, Q.; Zhong, Q.; Zhao, Y.; Gao, Z.; Xu, C. An Anthocyanin-Related Glutathione S-Transferase, MrGST1, Plays an Essential Role in Fruit Coloration in Chinese Bayberry (*Morella rubra*). *Front. Plant Sci.* **2022**, *13*, 903333. [CrossRef]
49. Shen, X.-J.; Wang, Y.-Y.; Zhang, Y.-X.; Guo, W.; Jiao, Y.-Q.; Zhou, X.-A. Overexpression of the Wild Soybean R2R3-MYB Transcription Factor *GsMYB15* Enhances Resistance to Salt Stress and *Helicoverpa armigera* in Transgenic *Arabidopsis*. *Int. J. Mol. Sci.* **2018**, *19*, 3958. [CrossRef]
50. Wang, L.; Lu, W.; Ran, L.; Dou, L.; Yao, S.; Hu, J.; Fan, D.; Li, C.; Luo, K. R2R3-MYB transcription factor MYB6 promotes anthocyanin and proanthocyanidin biosynthesis but inhibits secondary cell wall formation in *Populus tomentosa*. *Plant J.* **2019**, *99*, 733–751. [CrossRef]
51. Zhao, L.; Song, Z.; Wang, B.; Gao, Y.; Shi, J.; Sui, X.; Chen, X.; Zhang, Y.; Li, Y. R2R3-MYB Transcription Factor NtMYB330 Regulates Proanthocyanidin Biosynthesis and Seed Germination in Tobacco (*Nicotiana tabacum* L.). *Front. Plant Sci.* **2022**, *13*, 819247. [CrossRef]
52. Menconi, J.; Perata, P.; Gonzali, S. Novel R2R3 MYB transcription factors regulate anthocyanin synthesis in *Aubergine* tomato plants. *BMC Plant Biol.* **2023**, *23*, 148. [CrossRef] [PubMed]
53. Saxena, S.; Pal, L.; Naik, J.; Singh, Y.; Verma, P.K.; Chattopadhyay, D.; Pandey, A. The R2R3-MYB-SG7 transcription factor CaMYB39 orchestrates surface phenylpropanoid metabolism and pathogen resistance in chickpea. *New Phytol.* **2023**, *238*, 798–816. [CrossRef]
54. An, J.-P.; Wang, X.-F.; Zhang, X.-W.; Xu, H.-F.; Bi, S.-Q.; You, C.-X.; Hao, Y.-J. An apple MYB transcription factor regulates cold tolerance and anthocyanin accumulation and undergoes MIEL1-mediated degradation. *Plant Biotechnol. J.* **2020**, *18*, 337–353. [CrossRef] [PubMed]
55. Dubos, C.; Le Gourrierec, J.; Baudry, A.; Huep, G.; Lanet, E.; Debeaujon, I.; Routaboul, J.-M.; Alboresi, A.; Weisshaar, B.; Lepiniec, L. MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* **2008**, *55*, 940–953. [CrossRef]

56. Ni, J.; Bai, S.; Zhao, Y.; Qian, M.; Tao, R.; Yin, L.; Gao, L.; Teng, Y. Ethylene response factors Pp4ERF24 and Pp12ERF96 regulate blue light-induced anthocyanin biosynthesis in 'Red Zaosu' pear fruits by interacting with MYB114. *Plant Mol. Biol.* **2019**, *99*, 67–78. [CrossRef]
57. Sun, H.; Hu, K.; Wei, S.; Yao, G.; Zhang, H. ETHYLENE RESPONSE FACTORS 4.1/4.2 with an EAR motif repress anthocyanin biosynthesis in red-skinned pears. *Plant Physiol.* **2023**, *92*, 1892–1912. [CrossRef]
58. Qi, T.; Song, S.; Ren, Q.; Wu, D.; Huang, H.; Chen, Y.; Fan, M.; Peng, W.; Ren, C.; Xie, D. The Jasmonate-ZIM-Domain Proteins Interact with the WD-Repeat/bHLH/MYB Complexes to Regulate Jasmonate-Mediated Anthocyanin Accumulation and Trichome Initiation in *Arabidopsis thaliana*. *Plant Cell* **2011**, *23*, 1795–1814. [CrossRef]
59. Li, Y.; Lei, W.; Zhou, Z.; Li, Y.; Zhang, D.; Lin, H. Transcription factor GLK1 promotes anthocyanin biosynthesis via an MBW complex-dependent pathway in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* **2023**, *65*, 1521–1535. [CrossRef] [PubMed]
60. Maier, A.; Schrader, A.; Kokkelink, L.; Falke, C.; Welter, B.; Iniesto, E.; Rubio, V.; Uhrig, J.F.; Huelskamp, 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]
61. Li, S.; Wang, W.; Gao, J.; Yin, K.; Wang, R.; Wang, C.; Petersen, M.; Mundy, J.; Qiu, J.-L. MYB75 Phosphorylation by MPK4 Is Required for Light-Induced Anthocyanin Accumulation in *Arabidopsis*. *Plant Cell* **2016**, *28*, 2866–2883. [CrossRef]
62. Wang, X.-F.; An, J.-P.; Liu, X.; Su, L.; You, C.-X.; Hao, Y.-J. The Nitrate-Responsive Protein MdBT2 Regulates Anthocyanin Biosynthesis by Interacting with the MdMYB1 Transcription Factor. *Plant Physiol.* **2018**, *178*, 890–906. [CrossRef] [PubMed]
63. Zhang, S.; Chen, Y.; Zhao, L.; Li, C.; Yu, J.; Li, T.; Yang, W.; Zhang, S.; Su, H.; Wang, L. A novel NAC transcription factor, MdNAC42, regulates anthocyanin accumulation in red-fleshed apple by interacting with MdMYB10. *Tree Physiol.* **2020**, *40*, 413–423. [CrossRef] [PubMed]
64. Fang, X.; Zhang, L.; Wang, L. The Transcription Factor MdERF78 Is Involved in ALA-Induced Anthocyanin Accumulation in Apples. *Front. Plant Sci.* **2022**, *13*, 915197. [CrossRef] [PubMed]
65. An, J.-P.; Yao, J.-F.; Xu, R.-R.; You, C.-X.; Wang, X.-F.; Hao, Y.-J. Apple bZIP transcription factor MdbZIP44 regulates abscisic acid-promoted anthocyanin accumulation. *Plant Cell Environ.* **2018**, *41*, 2678–2692. [CrossRef]
66. Zhang, H.; Wang, Z.; Li, X.; Gao, X.; Dai, Z.; Cui, Y.; Zhi, Y.; Liu, Q.; Zhai, H.; Gao, S.; et al. The IbBBX24-IbTOE3-IbPRX17 module enhances abiotic stress tolerance by scavenging reactive oxygen species in sweet potato. *New Phytol.* **2022**, *233*, 1133–1152. [CrossRef]
67. Song, Z.; Bian, Y.; Liu, J.; Sun, Y.; Xu, D. B-box proteins: Pivotal players in light-mediated development in plants. *J. Integr. Plant Biol.* **2020**, *62*, 1293–1309. [CrossRef]
68. Liu, Y.; Tang, L.; Wang, Y.; Zhang, L.; Xu, S.; Wang, X.; He, W.; Zhang, Y.; Lin, Y.; Wang, Y.; et al. The blue light signal transduction module FaCRY1-FaCOP1-FaHY5 regulates anthocyanin accumulation in cultivated strawberry. *Front. Plant Sci.* **2023**, *14*, 1144273. [CrossRef]
69. Wei, C.-Q.; Chien, C.-W.; Ai, L.-F.; Zhao, J.; Zhang, Z.; Li, K.H.; Burlingame, A.L.; Sun, Y.; Wang, Z.-Y. The *Arabidopsis* B-box protein BZS1/BBX20 interacts with HY5 and mediates strigolactone regulation of photomorphogenesis. *J. Genet. Genom.* **2016**, *43*, 555–563. [CrossRef]
70. Xu, D.; Jiang, Y.; Li, J.; Lin, F.; Holm, M.; Deng, X.W. BBX21, an *Arabidopsis* B-box protein, directly activates HY5 and is targeted by COP1 for 26S proteasome-mediated degradation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 7655–7660. [CrossRef]
71. Fang, H.; Dong, Y.; Yue, X.; Hu, J.; Jiang, S.; Xu, H.; Wang, Y.; Su, M.; Zhang, J.; Zhang, Z.; et al. The B-box zinc finger protein MdBBX20 integrates anthocyanin accumulation in response to ultraviolet radiation and low temperature. *Plant Cell Environ.* **2019**, *42*, 2090–2104. [CrossRef] [PubMed]
72. An, J.-P.; Wang, X.-F.; Zhang, X.-W.; Bi, S.-Q.; You, C.-X.; Hao, Y.-J. MdBBX22 regulates UV-B-induced anthocyanin biosynthesis through regulating the function of MdHY5 and is targeted by MdBT2 for 26S proteasome-mediated degradation. *Plant Biotechnol. J.* **2019**, *17*, 2231–2233. [CrossRef] [PubMed]
73. Wang, C.-Q.; Sarmast, M.K.; Jiang, J.; Dehesh, K. The Transcriptional Regulator BBX19 Promotes Hypocotyl Growth by Facilitating COP1-Mediated EARLY FLOWERING3 Degradation in *Arabidopsis*. *Plant Cell* **2015**, *27*, 1128–1139. [CrossRef]
74. Jiang, L.; Wang, Y.; Li, Q.-F.; Bjoern, L.O.; He, J.-X.; Li, S.-S. *Arabidopsis* STO/BBX24 negatively regulates UV-B signaling by interacting with COP1 and repressing HY5 transcriptional activity. *Cell Res.* **2012**, *22*, 1046–1057. [CrossRef] [PubMed]
75. Gangappa, S.N.; Holm, M.; Botto, J.F. Molecular interactions of BBX24 and BBX25 with HYH, HY5 HOMOLOG, to modulate *Arabidopsis* seedling development. *Plant Signal. Behav.* **2013**, *8*, e25208. [CrossRef]
76. Holtan, H.E.; Bandong, S.; Marion, C.M.; Adam, L.; Tiwari, S.; Shen, Y.; Maloof, J.N.; Maszle, D.R.; Ohto, M.-a.; Preuss, S.; et al. BBX32, an *Arabidopsis* B-Box Protein, Functions in Light Signaling by Suppressing HY5-Regulated Gene Expression and Interacting with STH2/BBX21. *Plant Physiol.* **2011**, *156*, 2109–2123. [CrossRef] [PubMed]
77. Yao, G.; Ming, M.; Allan, A.C.; Gu, C.; Li, L.; Wu, X.; Wang, R.; Chang, Y.; Qi, K.; Zhang, S.; et al. Map-based cloning of the pear gene MYB114 identifies an interaction with other transcription factors to coordinately regulate fruit anthocyanin biosynthesis. *Plant J.* **2017**, *92*, 437–451. [CrossRef]
78. Ou, C.; Zhang, X.; Wang, F.; Zhang, L.; Zhang, Y.; Fang, M.; Wang, J.; Wang, J.; Jiang, S.; Zhang, Z. A 14 nucleotide deletion mutation in the coding region of the *PpBBX24* gene is associated with the red skin of "Zaosu Red" pear (*Pyrus pyrifolia* White Pear Group): A deletion in the *PpBBX24* gene is associated with the red skin of pear. *Hortic. Res.* **2020**, *7*, 39. [CrossRef]

79. An, J.-P.; Wang, X.-F.; Espley, R.V.; Lin-Wang, K.; Bi, S.-Q.; You, C.-X.; Hao, Y.-J. An Apple B-Box Protein MdBBX37 Modulates Anthocyanin Biosynthesis and Hypocotyl Elongation Synergistically with MdMYBs and MdHYS. *Plant Cell Physiol.* **2020**, *61*, 130–143. [CrossRef]
80. Fang, H.; Dong, Y.; Yue, X.; Chen, X.; He, N.; Hu, J.; Jiang, S.; Xu, H.; Wang, Y.; Su, M.; et al. MdCOL4 Interaction Mediates Crosstalk between UV-B and High Temperature to Control Fruit Coloration in Apple. *Plant Cell Physiol.* **2019**, *60*, 1055–1066. [CrossRef]
81. Hu, Y.; Cheng, H.; Zhang, Y.; Zhang, J.; Niu, S.; Wang, X.; Li, W.; Zhang, J.; Yao, Y. The MdMYB16/MdMYB1-miR7125-MdCCR module regulates the homeostasis between anthocyanin and lignin biosynthesis during light induction in apple. *New Phytol.* **2021**, *231*, 1105–1122. [CrossRef]
82. Hu, D.-G.; Sun, C.-H.; Ma, Q.-J.; You, C.-X.; Cheng, L.; Hao, Y.-J. MdMYB1 Regulates Anthocyanin and Malate Accumulation by Directly Facilitating Their Transport into Vacuoles in Apples. *Plant Physiol.* **2016**, *170*, 1315–1330. [CrossRef] [PubMed]
83. Aharoni, A.; De Vos, C.H.; Wein, M.; Sun, Z.; Greco, R.; Kroon, A.; Mol, J.N.; O'Connell, A.P. The strawberry *FaMYB1* transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant J.* **2001**, *28*, 319–332. [CrossRef] [PubMed]
84. Salvatierra, A.; Pimentel, P.; Alejandra Moya-Leon, M.; Herrera, R. Increased accumulation of anthocyanins in *Fragaria chiloensis* fruits by transient suppression of *FcMYB1* gene. *Phytochemistry* **2013**, *90*, 25–36. [CrossRef]
85. Chen, Q.; Yu, H.W.; Wang, X.R.; Xie, X.L.; Yue, X.Y.; Tang, H.R. An alternative cetyltrimethylammonium bromide-based protocol for RNA isolation from blackberry (*Rubus* L.). *Genet. Mol. Res.* **2012**, *11*, 1773–1782. [CrossRef]
86. Wei, W.; Chai, Z.; Xie, Y.; Gao, K.; Cui, M.; Jiang, Y.; Feng, J. Bioinformatics identification and transcript profile analysis of the mitogen-activated protein kinase gene family in the diploid woodland strawberry *Fragaria vesca*. *PLoS ONE* **2017**, *12*, e0178596. [CrossRef] [PubMed]
87. 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]
88. Lin, Y.; Jiang, L.; Chen, Q.; Li, Y.; Zhang, Y.; Luo, Y.; Zhang, Y.; Sun, B.; Wang, X.; Tang, H. Comparative Transcriptome Profiling Analysis of Red- and White-Fleshed Strawberry (*Fragaria* × *ananassa*) Provides New Insight into the Regulation of the Anthocyanin Pathway. *Plant Cell Physiol.* **2018**, *59*, 1844–1859. [CrossRef]
89. Luo, H.; Dai, C.; Li, Y.; Feng, J.; Liu, Z.; Kang, C. Reduced Anthocyanins in Petioles codes for a GST anthocyanin transporter that is essential for the foliage and fruit coloration in strawberry. *J. Exp. Bot.* **2018**, *69*, 2595–2608. [CrossRef]
90. Prior, R.L.; Fan, E.; Ji, H.; Howell, A.; Nio, C.; Payne, M.J.; Reed, J. Multi-laboratory validation of a standard method for quantifying proanthocyanidins in cranberry powders. *J. Sci. Food Agric.* **2010**, *90*, 1473–1478. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

Genome-Wide Exploration and Expression Analysis of the CNGC Gene Family in Eggplant (*Solanum melongena* L.) under Cold Stress, with Functional Characterization of *SmCNGC1a*

Zheng Jiang †, Lihui Du †, Lei Shen, Jie He, Xin Xia, Longhao Zhang and Xu Yang *

College of Horticulture and Landscape Architecture, Yangzhou University, Yangzhou 225009, China
* Correspondence: yangxu@yzu.edu.cn
† These authors contributed equally to this work.

Abstract: Eggplant (*Solanum melongena* L.) is an important economic crop, and to date, there has been no genome-wide identification and analysis of the cyclic nucleotide-gated channel (CNGC) gene family in eggplant. In this study, we identified the CNGC gene family in eggplant, and the results showed that 29 *SmCNGC* genes were classified into five groups, unevenly distributed across the 12 chromosomes of eggplant. The gene structure and motif analysis indicated that the *SmCNGC* family proteins may exhibit apparent preferences during evolution. Furthermore, our study revealed the presence of numerous light-responsive elements, hormone-responsive elements, and transcription factor binding sites in the promoter regions of *SmCNGC* genes, suggesting their significant role in environmental adaptability regulation. Finally, we analyzed the expression patterns of all *SmCNGC* genes under cold stress and found that *SmCNGC1a* was significantly upregulated under cold stress. Subcellular localization experiments indicated that this gene is located on the plasma membrane. Subsequently, its importance in the low-temperature response of eggplant was validated through virus-induced gene silencing (VIGS), and its protein interactome was predicted. In summary, our study provides a comprehensive understanding of the function and regulatory mechanisms of the CNGC gene family in eggplant, laying an important foundation for further research on cold adaptation in eggplant.

Keywords: eggplant; gene family; cold stress; expression pattern; functional analysis

1. Introduction

Eggplant (*Solanum melongena* L.), a thermophilic vegetable belonging to the Solanaceae family, is extensively cultivated on a global scale. It is an important vegetable variety during the off season and is also one of the most popular and widely consumed vegetables [1]. Eggplant contains rich nutritional components, such as vitamins, minerals, and especially high levels of vitamin E, P, and iron. It also contains various bioactive compounds, offering health benefits and medicinal value [2]. Calcium ions (Ca^{2+}) are indispensable for upholding the integrity of plant cell architecture [3]. They assume a pivotal role as both a critical cellular signal and a secondary messenger in numerous pathways, governing various responses to environmental stimuli, including plant hormones, temperature fluctuations, light exposure, and salt stress [4,5]. Cyclic nucleotide-gated channels (CNGCs) are non-selective cation channels situated on the cytoplasmic membrane. They play a crucial role in facilitating the transmembrane transport of monovalent cations, such as sodium and potassium, as well as divalent cations, like calcium and magnesium [6]. CNGCs in plants were first discovered in 1998 during the screening of a barley calcium-binding transporter (Hordeum vulgare CaM-binding transporter, HvCBT1), which revealed the existence of this non-selective cation channel [7]. CNGCs are widely distributed in animals and plants, serving as universal calcium ion channels in eukaryotes [8]. Prior research has substantiated the existence of CNGCs in both monocotyledonous and dicotyledonous plant

species [9]. Preliminary investigations have suggested that CNGCs function as ligand-gated ion channels, facilitating the passage of calcium ions. These channels can be activated by cyclic nucleotides (cNMP) and their activity can be inhibited through binding with Ca^{2+}/calmodulin (CaM) [10,11]. Structurally, CNGCs consist of six transmembrane domains (S1–S6), a pore region (P) located between the fifth and sixth domains, a C-terminal CaM-binding domain (CaMB), and a cyclic nucleotide-binding domain (CNBD) [12]. The CNBD, which represents a highly conserved region within CNGCs, is characterized by the presence of a cyclic nucleotide-binding cassette (PBC) and an adjacent unique "hinge" region. The PBC specifically interacts with the sugar and phosphate groups of the nucleotide ligand, while the adjacent "hinge" region contributes to the efficiency and selectivity of ligand binding [13]. Previous investigations have demonstrated that the carboxyl-terminal tails of specific plant CNGCs exhibit a Ca^{2+}-dependent interaction with CaM [14–16]. In plants, CNGCs contain a conserved cNMP-binding domain that is considered a specific domain for identifying CNGCs [17]. Early studies proposed that plant CNGCs are regulated by cNMP gating. However, as research on calcium signaling has progressed, many scholars have questioned whether cNMP activation is necessary for CNGCs [18]. Recent studies have found that in *Xenopus laevis* oocyte cells, cAMP and cGMP can hyperpolarize activate calcium channels through heterologous assembly, without requiring an increase in cNMP levels [19]. Therefore, cNMP may only act as an auxiliary factor for CNGC subunits or exert its effects by modulating the membrane voltage or other regulators [19]. Furthermore, with further research, an increasing number of experimental findings indicate that the activation and regulation of CNGC channels rely more on phosphorylation modifications and the binding of CaM. The mechanisms underlying cNMP regulation of CNGCs require further investigation [20].

CNGCs regulate the growth of root tips in plants. Recent studies have shown that *CNGC5*, *CNGC6*, *CNGC9*, and *CNGC14* in *Arabidopsis thaliana* are involved in regulating the growth of root hairs [21]. *CNGC14* plays a role in maintaining cell integrity during polar growth of root hairs, and its mutation leads to abnormal growth of root hairs, including swelling, branching, and bursting at the tip [22]. Additionally, *CNGC5*, *CNGC6*, *CNGC9*, and *CNGC14* are also involved in maintaining the stability of unidirectional cell proliferation and cytoplasmic Ca^{2+} oscillations. In *CNGC14* mutants, the stability of cytoplasmic Ca^{2+} oscillations are severely impaired, and *cngc14/cngc6* and *cngc14/cngc9* double mutants lose the typical 30 s Ca^{2+} oscillation cycle [22]. In the double mutants of *cngc6* and *cngc9*, as well as in the single mutants of *cngc9*, this oscillation cycle still exists but with reduced stability [23]. These findings indicate that *CNGC14* plays a crucial role in these processes. Another study found that *CNGC5*, *CNGC6*, and *CNGC9* are essential Ca^{2+} channels in the structural growth of root hairs and auxin signal transduction in Arabidopsis. Mutations in these three genes lead to root hair growth defects, including shorter and defective root hairs. However, expressing any one of the CNGC subtypes individually or providing a high concentration of exogenous Ca^{2+} can restore this growth defect, whereas the supply of exogenous K^+ cannot [24]. These three genes also exhibit Ca^{2+} permeation channel function in HEK293T cells. Cytosolic Ca^{2+} imaging and patch clamp data in root hairs indicate a significant reduction in the Ca^{2+} gradient and oscillation at the tip when *CNGC5*, *CNGC6*, and *CNGC9* are absent [24].

CNGCs play important roles in the growth and development of plant pollen tubes. The CNGC family in *Arabidopsis* comprises a total of 20 members [25]. Studies have shown that *AtCNGC16* contributes to the development of pollen grains in *Arabidopsis* under high-temperature conditions [26]. *AtCNGC7* and *AtCNGC8* share high amino acid sequence homology, and both have overlapping functions in the germination and development of pollen as the *cngc7/cngc8* double mutation leads to pollen sterility, confirming their roles in plant pollen development [27]. Further research has revealed that *CNGC7* or *CNGC8* interacts with *CNGC18* to form an inactive heterotetramer. However, when the Ca^{2+} concentration reaches its peak level, CaM dissociates from the *CNGC18-CNGC8* heterotetramer, relieving the inhibition of *CNGC18* by *CNGC8*. This pathway, regulated by

Ca^{2+}-CaM interaction with CNGC channels, facilitates autoregulatory feedback in calcium oscillations while facilitating pollen tube growth and plays an important role in fine-tuning the growth of pollen tubes [28]. Moreover, *CNGC18* exhibits asymmetric distribution on the plasma membrane at the growing tip of pollen tubes and possesses Ca^{2+} channel activity, which is crucial for pollen tube tip growth and guidance [29,30].

In recent years, with the increasing demand for year-round vegetables, particularly the rising requirements for eggplant production during the winter and spring seasons, eggplants frequently encounter cold stress during protected cultivation in winter and early spring, leading to cold damage, blossom and fruit drop, as well as poor fruit coloration. These phenomena severely affect the yield and quality of eggplants. Additionally, low-temperature stress hampers the growth and development of eggplants, resulting in stunted growth, yellowing and withering of leaves, and even plant death [31]. These occurrences not only directly impact the yield and quality of eggplants, but also cause economic losses and agricultural production instability for growers. Cold damage during the winter and spring seasons has become a bottleneck factor restraining the development of the eggplant industry, thereby imposing higher demands on the cold tolerance of eggplant cultivars during these seasons. Therefore, conducting research on the cold-tolerance mechanism of eggplants, exploring cultivation techniques to enhance their cold tolerance, and breeding new cold-tolerant eggplant varieties hold significant importance. In this study, we analyzed the genomic data of the eggplant CNGC gene family, identified 29 members of the *SmCNGCs* gene family, and conducted physicochemical property analysis, promoter analysis, phylogenetic tree construction, gene structure analysis, motif analysis, chromosome localization analysis, and collinearity analysis. We also investigated the expression patterns of *SmCNGCs* under different cold-stress conditions at different time points. Expression pattern analysis and subcellular localization were performed for the cold-stress-responsive gene *SmCNGC1a* in different tissues. Furthermore, the functionality of *SmCNGC1a* was validated using VIGS (virus-induced gene silencing) technology, and protein interaction networks were analyzed to explore the role of this gene under low-temperature stress in eggplants.

2. Results

2.1. Identification of the CNGC Genes in Eggplant

Based on the HMM file of the CNGC gene family, a search was conducted on the reference genome of eggplant. After manual screening and removal of redundant genes, a total of 29 *SmCNGC* genes were identified. Following the nomenclature guidelines for CNGC genes in *Arabidopsis* and referring to the information in the eggplant database, the identified 29 eggplant CNGC genes were named as *SmCNGC1a*-*SmCNGC24* (Table 1). The predicted lengths of *SmCNGC* proteins ranged from 482 amino acids (*SmCNGC9*) to 1074 amino acids (*SmCNGC7*). The majority of *SmCNGC* proteins had lengths concentrated in the range of 600–800 amino acids, with molecular weights varying from 55.40 kDa (*SmCNGC9*) to 122.28 kDa (*SmCNGC7*). The significant differences in amino acid length and molecular weight suggest potential structural variations among members of the *SmCNGC* gene family. Among the 29 *SmCNGC* proteins, seven had a theoretical pI value below 7, indicating a potential negative charge within the alkaline pH range. The remaining 22 *SmCNGC* proteins had theoretical pI values above 7, suggesting a positive charge within the acidic pH range. Additionally, 8 *SmCNGC* proteins had an instability index value below 40, indicating relative stability, while the remaining 21 proteins were relatively unstable. The predicted aliphatic index for *SmCNGC* proteins ranged from 88.78 (*SmCNGC18*) to 99.63 (*SmCNGC22*). The grand average of hydropathicity predicted that, with the exception of *SmCNGC13*, the remaining proteins were hydrophilic. The prediction of subcellular localization results indicated that the majority of *SmCNGC* proteins were located in the cell membrane (plasma membrane), consistent with their primary function as channel proteins. However, there were some special cases: *SmCNGC22* was found in the nucleus,

SmCNGC10 and *SmCNGC9* were located in the cytoplasm, and *SmCNGC23* was present in the chloroplast, suggesting potential diverse biological functions for these proteins.

Table 1. Identification and Physicochemical Analysis of *SmCNGC* Proteins.

Gene Name	Gene ID	Number of Amino Acid/aa	Molecular Weight	Theoretical pI	Instability Index	Aliphatic Index	Grand Average of Hydropathicity	Prediction of Subcellular Localization
SmCNGC1a	Smechr0500088.1	714	82,462.57	9.51	53.73	90.17	−0.208	cytomembrane
SmCNGC1b	Smechr0101357.1	708	81,926.22	9.32	48.72	93.4	−0.071	cytomembrane
SmCNGC1c	Smechr0302178.1	710	81,869.6	8.9	50.11	92.46	−0.132	cytomembrane
SmCNGC2	Smechr0202893.1	708	81,371.83	9.65	53.34	92.78	−0.034	cytomembrane
SmCNGC3	Smechr0400998.1	838	96,160.87	7.32	32.86	93.05	−0.198	cytomembrane
SmCNGC4a	Smechr1000491.1	685	80,107.48	8.28	46.96	90.31	−0.169	cytomembrane
SmCNGC4b	Smechr0400207.1	665	77,317.03	8.98	45.03	92.77	−0.075	cytomembrane
SmCNGC5	Smechr1200261.1	692	78,774.97	9.19	51.32	92.76	−0.114	cytomembrane
SmCNGC6	Smechr1100117.1	823	94,278.62	6.38	40.34	98.54	−0.131	cytomembrane
SmCNGC7	Smechr0700086.1	1074	122,281.25	9.3	53.27	90.8	−0.079	cytomembrane
SmCNGC8	Smechr0303133.1	689	79,669.91	8.78	43.35	91.13	−0.183	cytomembrane
SmCNGC9	Smechr0801568.1	482	55,399.66	8.21	37.08	96.24	−0.188	cytoplasm
SmCNGC10	Smechr0800622.1	714	82,222.56	5.82	38.97	95.03	−0.123	cytoplasm
SmCNGC11	Smechr0801458.1	859	96,677.4	6.6	40.65	98.43	−0.055	cytomembrane
SmCNGC12	Smechr0201396.1	655	75,003.7	7.84	39.83	93.45	−0.092	cytomembrane
SmCNGC13	Smechr0900702.1	612	69,109.84	7.32	46.02	99.1	0.033	cytomembrane
SmCNGC14	Smechr0801673.1	634	72,997.99	8.67	47.05	91.85	−0.097	cytomembrane
SmCNGC15a	Smechr0900248.1	659	76,020.09	8.73	49.81	93.38	−0.033	cytomembrane
SmCNGC15b	Smechr0600577.1	696	80,257.66	9.24	50.58	91.18	−0.19	cytomembrane
SmCNGC15c	Smechr1201893.1	704	80,655.83	9.26	53.81	88.82	−0.217	cytomembrane
SmCNGC16	Smechr0302884.1	674	77,929.8	8.31	47.48	89.39	−0.157	cytomembrane
SmCNGC17	Smechr0700014.1	720	83,011.82	9.32	41.97	93.74	−0.174	cytomembrane
SmCNGC18	Smechr0203033.1	689	79,438.12	7.17	49.98	88.78	−0.15	cytomembrane
SmCNGC19	Smechr0302451.1	840	95,490.13	6.12	39.05	97.96	−0.151	cytomembrane
SmCNGC20	Smechr0302478.1	771	88,894.76	9.32	49.87	89.29	−0.191	cytomembrane
SmCNGC21	Smechr0500154.1	827	94,588.85	6.62	39.78	94.79	−0.13	cytomembrane
SmCNGC22	Smechr0103753.1	629	72,010.19	8.16	44.82	99.63	−0.018	nucleus
SmCNGC23	Smechr1200076.1	884	99,814.77	6.47	39.65	96.24	−0.111	chloroplast
SmCNGC24	Smechr0100701.1	837	95,413.1	6.61	33.24	95.97	−0.13	cytomembrane

2.2. Phylogenetic Analysis of SmCNGC Proteins

To investigate the evolutionary relationships of the *SmCNGC* gene family, we constructed a phylogenetic tree of the CNGC protein family, including eggplant, Arabidopsis, and tomato (*Solanum lycopersicum* L.). The maximum likelihood (ML) method was employed to generate the phylogenetic tree with five groups (Figure 1). Group 1 consists of 6 eggplant CNGC proteins, 11 Arabidopsis CNGC proteins, and 9 tomato CNGC proteins. Group 2 comprises seven eggplant CNGC proteins, five Arabidopsis CNGC proteins, and five tomato CNGC proteins. Group 3 contains three eggplant CNGC proteins, two Arabidopsis CNGC proteins, and three tomato CNGC proteins. Group 4 consists of one eggplant CNGC protein, two Arabidopsis CNGC proteins, and one tomato CNGC protein. Group 5 exclusively includes 12 eggplant CNGC proteins, indicating a relatively distant evolutionary relationship with Arabidopsis and tomato. Based on the evolutionary analysis results, it can be inferred that CNGC maintains a high degree of stability throughout the process of species evolution, suggesting its conserved and indispensable role in organisms.

2.3. SmCNGC Gene Structures and the Conserved Motifs Analyses

Motif analysis was conducted on the CNGC gene family in eggplant. The results revealed that the encoded proteins of this family possess 10 conserved motifs (motif 1–10, Figure 2A). Members within the same subfamily exhibited similar distribution patterns of these conserved motifs. Motif 10 was identified as a shared motif among 28 *SmCNGC* members, while motif 1 was absent in *SmCNGC9*. Furthermore, it was observed that CNGC members in group 5 displayed unique regularity in the motifs they contained, with motif 6 being exclusive to this group (Figure 2B). Structural domain analysis indicated that most of the genes contained the PLN03192 superfamily domain. However, *SmCNGC1a* lacked this domain, which suggests that there might be functional differences associated with this gene (Figure 2C). Gene structure analysis, as shown in the diagram on the right, revealed that the number of coding sequences (CDSs) and untranslated regions (UTRs)

were similar within the same subfamily. Additionally, *SmCNGC13* displayed significant structural differences compared to other members, indicating possible variations in gene structure during the process of evolution (Figure 2D).

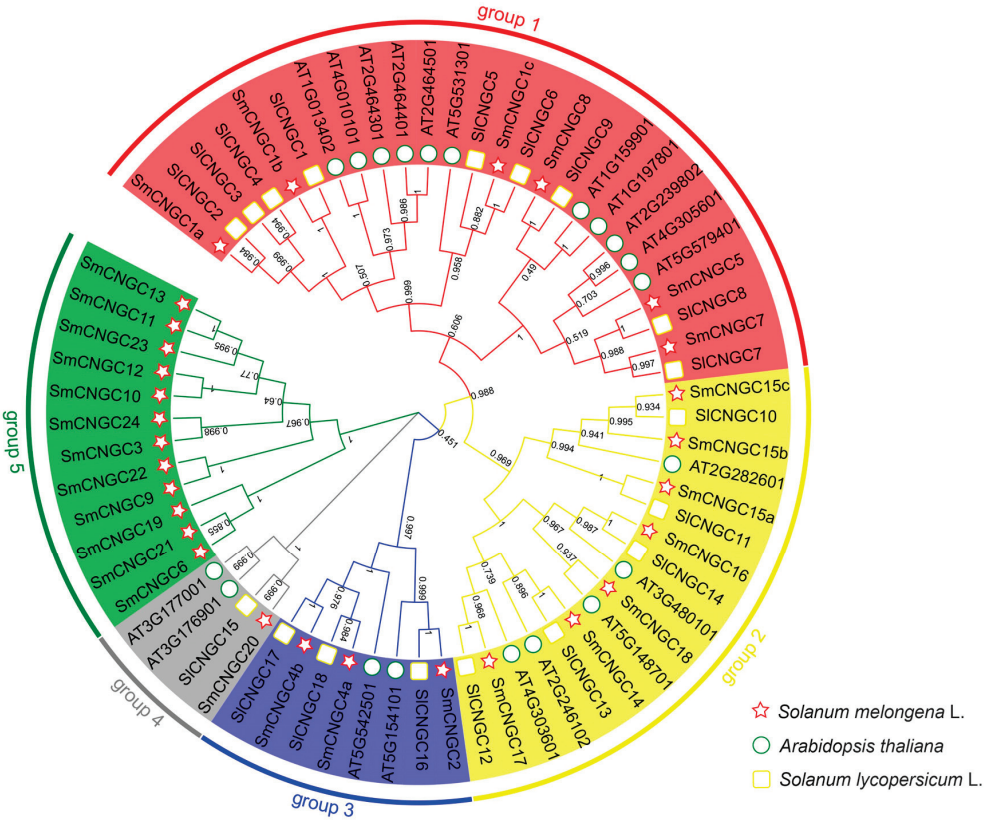


Figure 1. Phylogenetic relationships among the SmCNGCs, AtCNGCs, and SlCNGCs. The amino acid sequences of CNGC gene families in Arabidopsis, tomato, and eggplant were aligned using MEGA v7.0. After removing non-conserved gaps, the aligned sequences were used to construct a phylogenetic tree using the maximum likelihood (ML) method. Bootstrap values were set to 1000. Sm: eggplant; At: Arabidopsis; Sl: tomato. Different colors represent different groups.

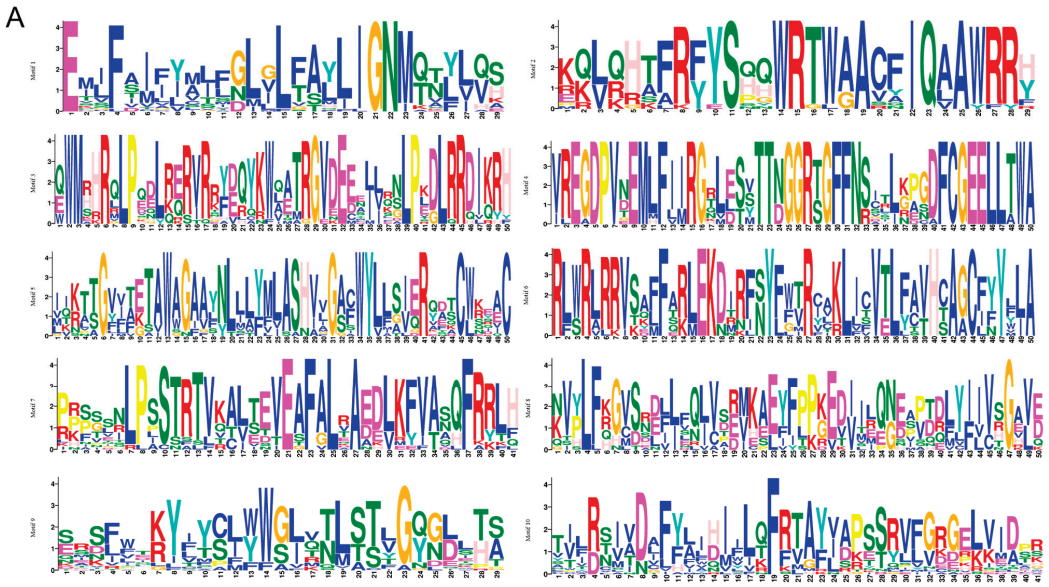


Figure 2. *Cont.*

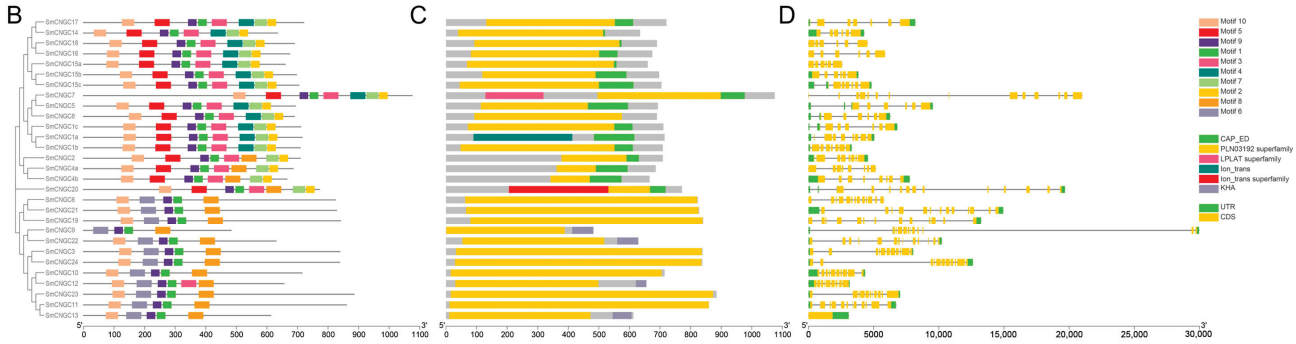


Figure 2. *SmCNGC* gene structures and the conserved motifs analyses. (**A**) Amino acid composition of each motif. (**B**) Motif compositions. (**C**) Conserved domains. (**D**) Gene structure. The motif logos were generated using the MEME Suite web server, while the remaining figures were generated using Tbtools.

2.4. Chromosome Localization and Cis-Acting Elements Prediction of SmCNGCs

Chromosomal localization analysis revealed that members of the *SmCNGC* gene family are distributed across all 12 chromosomes of eggplant. Most of these genes are located at the ends of the chromosomes. Chromosome 3 harbors the highest number of CNGC gene family members, with five genes. Chromosome 8 contains four CNGC gene family members. Chromosomes 1, 2, and 12 have three members each, while chromosomes 4, 5, 7, and 9 have two members each. Chromosomes 6, 10, and 11 have only one member of the CNGC gene family (Figure 3A). Promoter analysis predicted the presence of numerous cis-regulatory elements. The results indicate that the promoters of *SmCNGC* genes contain a substantial number of light-responsive elements, anaerobic-induction elements, hormone-responsive elements, and MYB transcription factor binding sites. All *SmCNGC* gene promoters contain light-responsive elements, suggesting their potential involvement in light signal regulation. Additionally, we found that a significant portion of *SmCNGC* genes contain methyl jasmonate, salicylic acid, or abscisic-acid-responsive elements, indicating their potential role in responding to these hormones and regulating plant growth and development. Nearly half of the *SmCNGC* genes contain MYB transcription factor binding sites, suggesting that SmCNGC may be directly regulated by MYB transcription factors, thus modulating ion channel activity and signal transduction (Figure 3B).

2.5. Collinearity Analysis of SmCNGCs

To elucidate the origin and evolutionary relationship of the *SmCNGC* gene family, we analyzed the gene duplication events of *SmCNGC*. The results showed that the *SmCNGC* gene family has a total of nine pairs of gene sequence duplication events, namely *SmCNGC13/SmCNGC15c*, *SmCNGC7/SmCNGC23/SmCNGC17*, and *SmCNGC15b/SmCNGC15c*, etc. (Figure 4A). This indicates that the *SmCNGC* genes have undergone multiple duplication events during the evolutionary process, which facilitated the rapid expansion of the *SmCNGC* gene family. In addition, we performed collinearity analysis of the genomes of eggplant, Arabidopsis, and tomato using MCScanX v1.5.1. The results showed that there are 22 pairs of collinearity relationships between eggplant and Arabidopsis CNGC genes, and 34 pairs of collinearity relationships between eggplant and tomato. It is noteworthy that eggplant's chromosome 3 (E03) has the highest number of collinear gene pairs (Figure 4B).

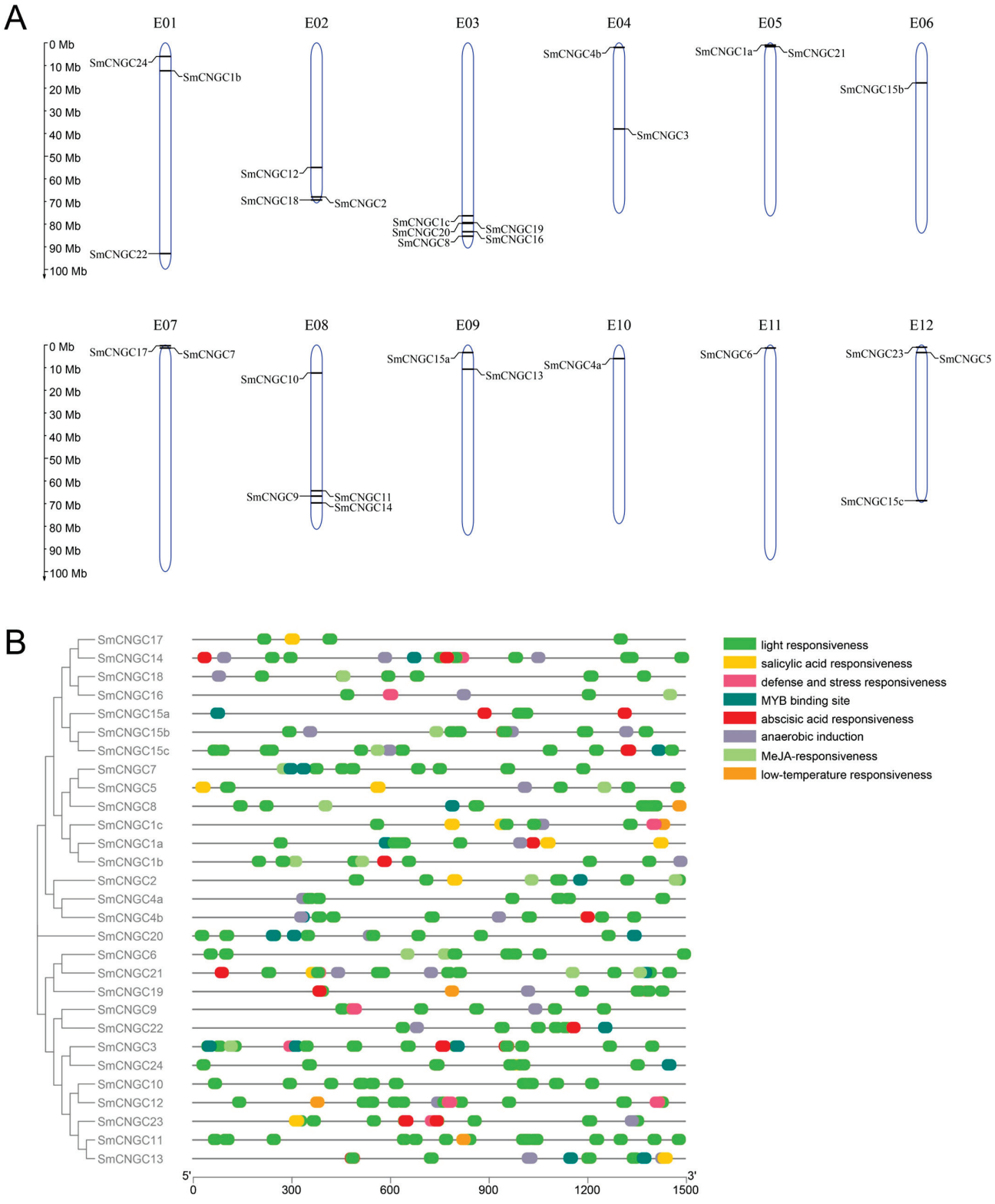


Figure 3. Analysis of Chromosomal Location and Cis-Regulatory Elements of the *SmCNGC* Gene. (**A**) Chromosomal localization of *SmCNGCs*. The blue bars represent the length of each chromosome, reflecting their relative sizes. The *SmCNGCs* are indicated on the chromosomes using their respective gene names. The placement of gene names along the chromosomes signifies their positions within the genome of eggplant. (**B**) Prediction of cis-regulatory elements in the upstream regions of *SmCNGC* genes. Different colors represent different cis-regulatory elements, with each color corresponding to a specific functional motif. The genes are arranged on the left side of the figure according to their evolutionary relationships. The color legend on the right side of the figure indicates the different cis-regulatory elements, with the colors arranged in descending order of their occurrence frequency in *SmCNGCs*.

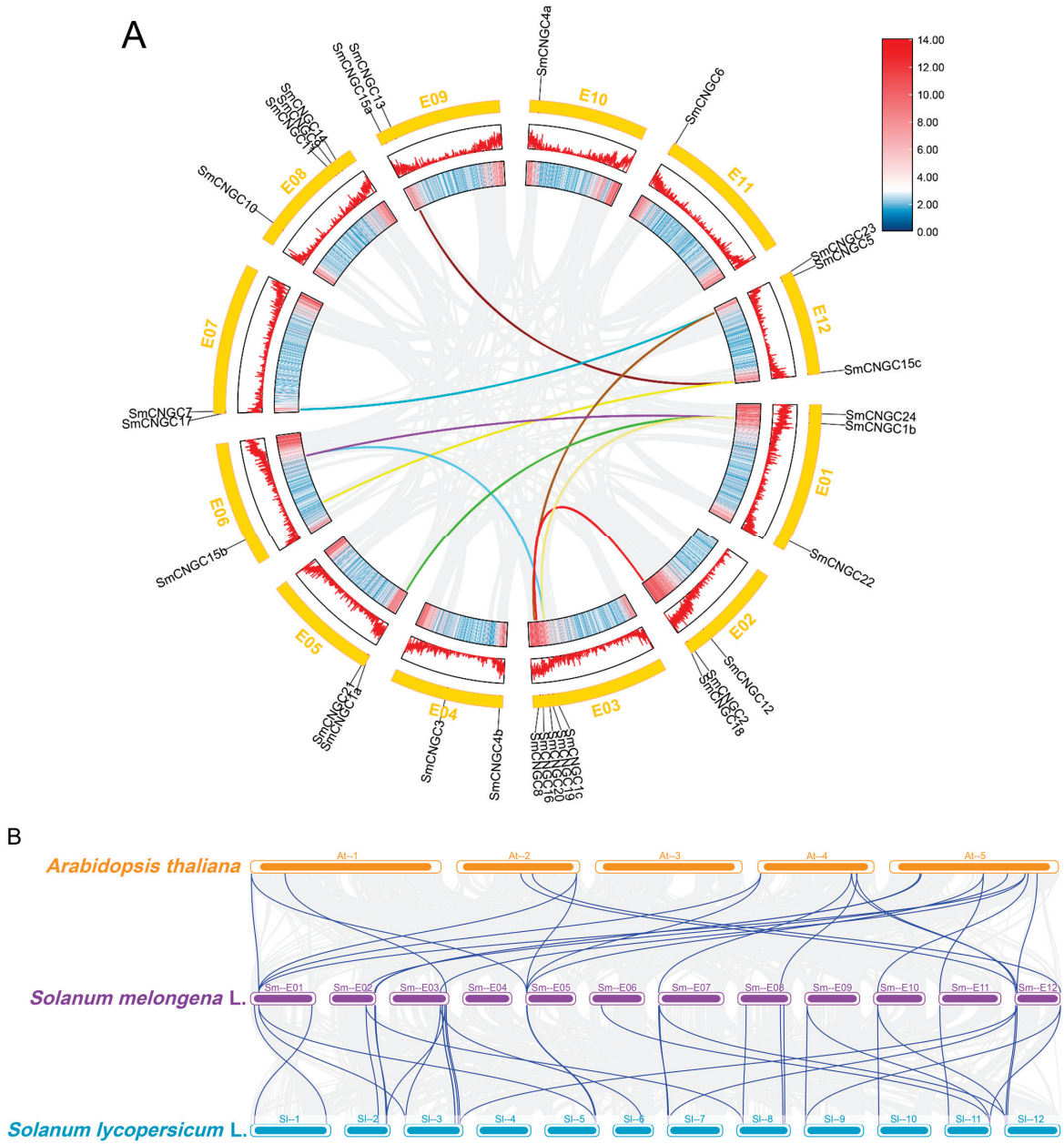


Figure 4. Collinearity analysis of *SmCNGCs*. (**A**) Chromosome locations and inter-chromosomal associations of *SmCNGC* genes. The colored lines represent different sets of collinear genes. The gene density in the chromosome is also depicted in the graph, with red indicating high density and blue indicating low density. (**B**) Collinearity analysis of CNGC genes between eggplant and the other two representative plants. The syntenic gene pairs were linked by blue lines.

2.6. Expression Patterns of SmCNGC Genes in Cold Stress

To investigate the expression patterns of the *SmCNGC* genes under cold stress in eggplant and evaluate their roles during cold stress, we employed qRT-PCR to detect the expression levels of *SmCNGC* genes at different time points (0, 0.25, 0.5, 1, 2, and 4 h) of cold stress. The results revealed that the *SmCNGC* genes exhibited distinct expression patterns in response to cold stress at different time intervals. Specifically, the expression levels of *SmCNGC1a*, *SmCNGC1c*, *SmCNGC7*, *SmCNGC12*, and *SmCNGC20* showed an initial increase followed by a decrease, suggesting their potential importance in the response to cold stress. This initial upregulation of expression might be associated with the rapid adaptation to environmental changes induced by cold stress and could be involved in

the regulation of ion channels and signal transduction within the cells. As the duration of cold stress continued, the expression levels of these genes gradually decreased, which could indicate cellular acclimation to the cold-stress conditions or involvement of negative feedback regulatory mechanisms. Furthermore, the expression levels of *SmCNGC4a*, *SmCNGC6*, *SmCNGC18*, and *SmCNGC24* decreased with prolonged cold stress, indicating their negative regulatory roles in eggplant's adaptation to cold stress. Interestingly, we observed a trend: initially decreased expression levels followed by increased expression levels for *SmCNGC2* and *SmCNGC21* (Figure 5). These experimental findings suggest that the *SmCNGC* genes may exhibit temporal specificity during the plant's response to cold stress. Of note, we observed a significant upregulation of *SmCNGC1a* at 0.5 h of cold stress, indicating its potential crucial role in the response of young eggplant seedlings to cold stress. Subsequent experiments will be conducted to further investigate this gene.

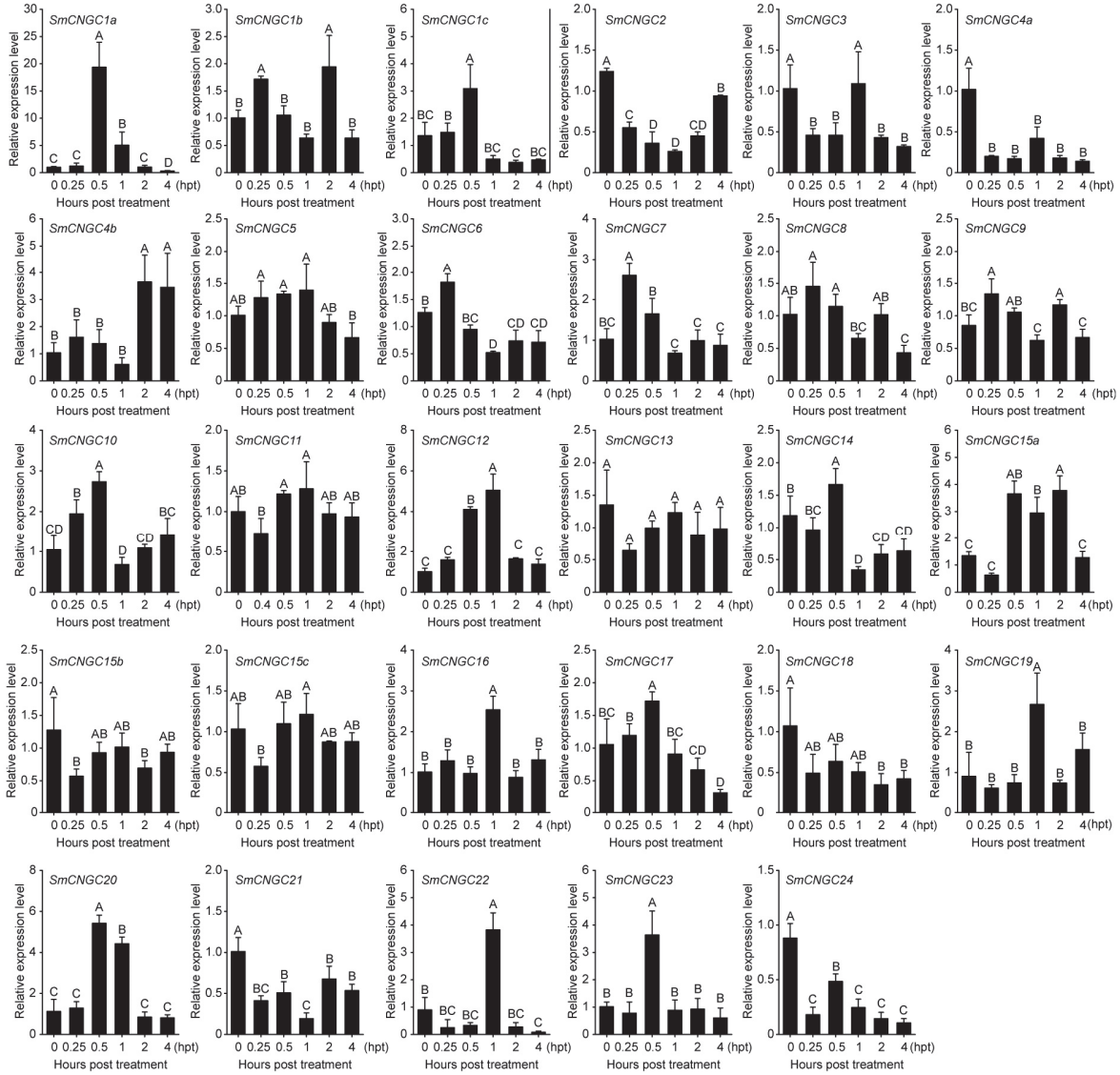


Figure 5. Expression patterns of *SmCNGC* genes under different time points of 4 °C cold stress. The bar graph represents the expression patterns of *SmCNGC* genes under different time points of cold stress. The y-axis represents the relative expression levels, and the x-axis represents the time points of cold stress. Data are means ± standard deviation from four biological replicates. Different capital letters between samples denote significant differences according to one-way ANOVA and Tukey's test ($p < 0.01$). The error bars represent the standard deviation.

2.7. Expression Patterns of SmCNGC1a under Various Stress Conditions

To comprehensively investigate the expression patterns of *SmCNGC1a* under various stress conditions, we conducted analyses of its response to cold stress, heat stress, and salt stress. The qPCR results revealed distinct expression profiles of *SmCNGC1a* under these diverse stress scenarios. Under cold stress, the transcription level of *SmCNGC1a* rapidly increased during the initial phase of exposure to cold stress, reaching a significant level after 0.5 h, followed by a gradual decline during sustained cold stress. This transient nature implies that *SmCNGC1a* may be involved in the early perception and response phases of cold stress. Under heat stress, the transcription level of *SmCNGC1a* showed a sustained upregulation, maintaining a significant elevation compared to the untreated control. This sustained upregulation suggests that *SmCNGC1a* could potentially play a role in mediating eggplant's response to heat stress. In contrast to cold and heat stress, the response of *SmCNGC1a* to salt stress exhibited a different expression pattern. Although there was a slight increase in gene expression, it was not statistically significant. This indicates a potential subtle role of *SmCNGC1a* in the regulation of salt-stress response (Figure 6).

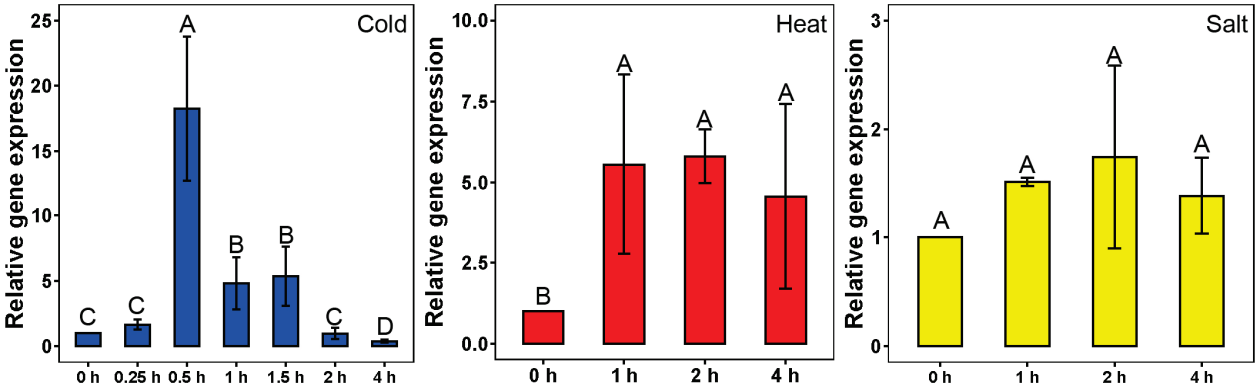


Figure 6. Expression patterns of SmCNGC1a under different stresses. Gene expression was analyzed using qPCR, with 0 h as the control. The blue columns represent cold stress, the red columns represent heat stress, and the yellow columns represent salt stress. Data are means ± standard deviation from four biological replicates. Different capital letters between samples denote significant differences according to one-way ANOVA and Tukey's test ($p < 0.01$). The error bars represent the standard deviation.

2.8. Subcellular Localization and Tissue-Specific Expression Patterns of SmCNGC1a

To investigate the subcellular localization and expression patterns of *SmCNGC1a* in different plant tissues, we cloned *SmCNGC1a* into the pCambia1300-35S-EGFP plant expression vector containing the 35S promoter and GFP reporter gene (Figure 7A). The recombinant constructs, 35S:*SmCNGC1a*-GFP and 35S:GFP (control), were transformed into *N. benthamiana* epidermal cells using *Agrobacterium*-mediated transformation and injected. After 48 h, the fluorescence signals of the genes in the leaves were observed using a laser scanning confocal microscope. The results indicate that the fluorescence signal of *SmCNGC1a*-GFP was localized to the plasma membrane of the cells, while the control group (GFP) exhibited a fluorescence signal distributed throughout the entire cell, suggesting that *SmCNGC1a* is localized to the plasma membrane (Figure 7B). Furthermore, we performed qRT-PCR to detect the expression patterns of *SmCNGC1a* in different tissues of eggplant (root, stem, leaf, and flower). The results revealed that *SmCNGC1a* exhibited relatively higher expression levels in the root and leaf, and relatively lower expression levels in the stem and flower, suggesting its possible association with environmental adaptation in roots and leaves (Figure 8).

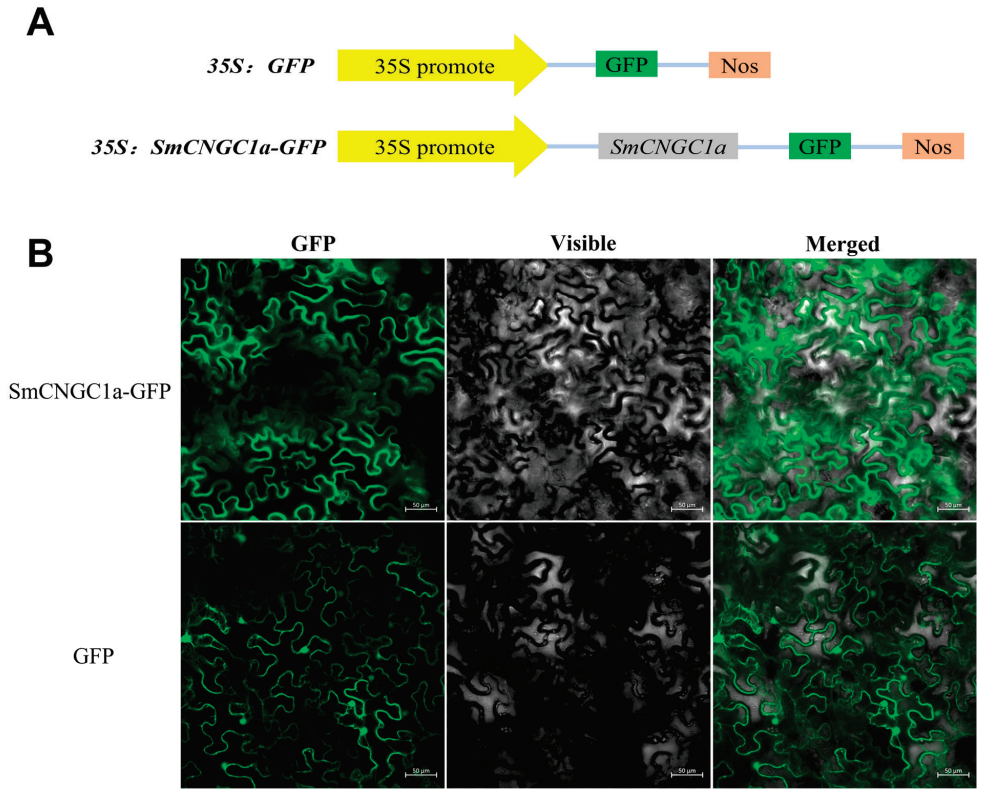


Figure 7. Subcellular localization of *SmCNGC1a*. (**A**) Construction of the *SmCNGC1a* vector. (**B**) 35S:*SmCNGC1a*-GFP and 35S:GFP (control) were transformed into *N. benthamiana* epidermal cells using *Agrobacterium*-mediated transformation and injected. Results were observed using confocal microscopy 48 h after transformation. Scale bars = 50 µm.

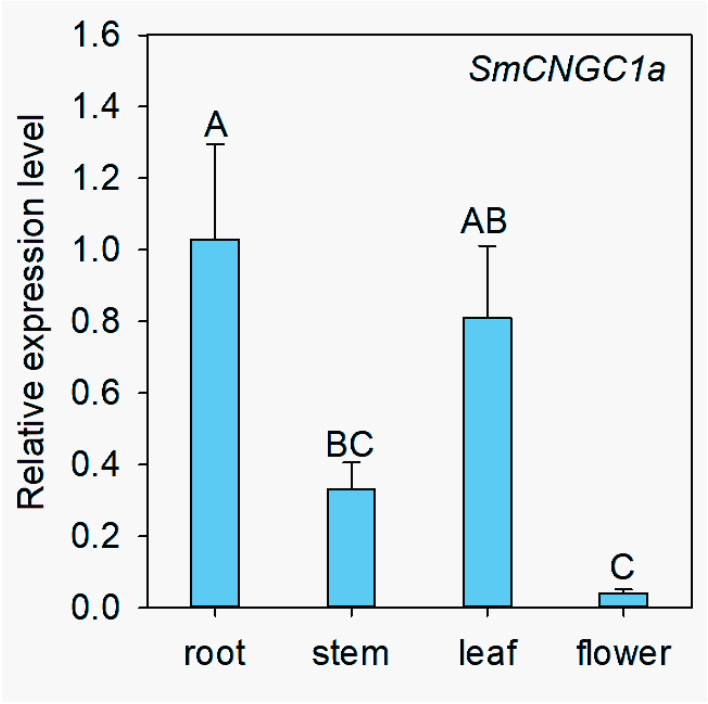


Figure 8. Expression levels of *SmCNGC1a* in different tissues of eggplant. Data are means ± standard deviation from four biological replicates. Different capital letters between samples denote significant differences according to one-way ANOVA and Tukey's test ($p < 0.01$), The error bars represent the standard deviation.

2.9. Silencing of SmCNGC1a Reduced Eggplant Tolerance to Cold Stress

To further investigate the function of *SmCNGC1a* in the cold-stress response of eggplant, we employed the VIGS method to silence *SmCNGC1a*. *Agrobacterium* infiltration solution was injected into the leaves of "JS221" seedlings, and 21 days later, the JS221 leaves injected with TRV:*SmPDS* exhibited pronounced chlorosis symptoms (Figure 9A), while no significant phenotype was observed in the control plants. Subsequently, we used qRT-PCR to examine the expression levels of the *SmCNGC1a* gene in the roots of silenced and control plants under cold stress to calculate the silencing efficiency. The results showed that the relative expression level of *SmCNGC1a* was significantly lower in the silenced plants compared to the control plants. In the control group, the expression of *SmCNGC1a* in the leaves was significantly upregulated under cold stress, while the expression level of *SmCNGC1a* in the silenced plants was suppressed (Figure 9B). These findings indicate successful silencing of *SmCNGC1a*. Furthermore, after 2 h of cold stress, the silenced plants exhibited more severe leaf wilting, stem bending, and loss of turgidity (Figure 9C), and the survival rate of the plants also significantly decreased (Figure 9D), suggesting that the silencing of *SmCNGC1a* reduced the cold tolerance of JS221, thus preliminarily confirming its positive regulatory role in the cold-stress response of eggplant. We also monitored the contents of chlorophyll a, chlorophyll b, carotenoids, proline, and malondialdehyde in the leaves of plants before and after silencing. The results revealed a significant decrease in the levels of photosynthetic pigments (chlorophyll a and chlorophyll b) compared to the control, accompanied by a substantial increase in proline and malondialdehyde content (Figure 9E). This suggests that the silencing of SmCNGC1a significantly affects the accumulation of proline and malondialdehyde in eggplant leaves, leading to a more pronounced degradation of chlorophyll a and chlorophyll b when compared to non-silenced plants.

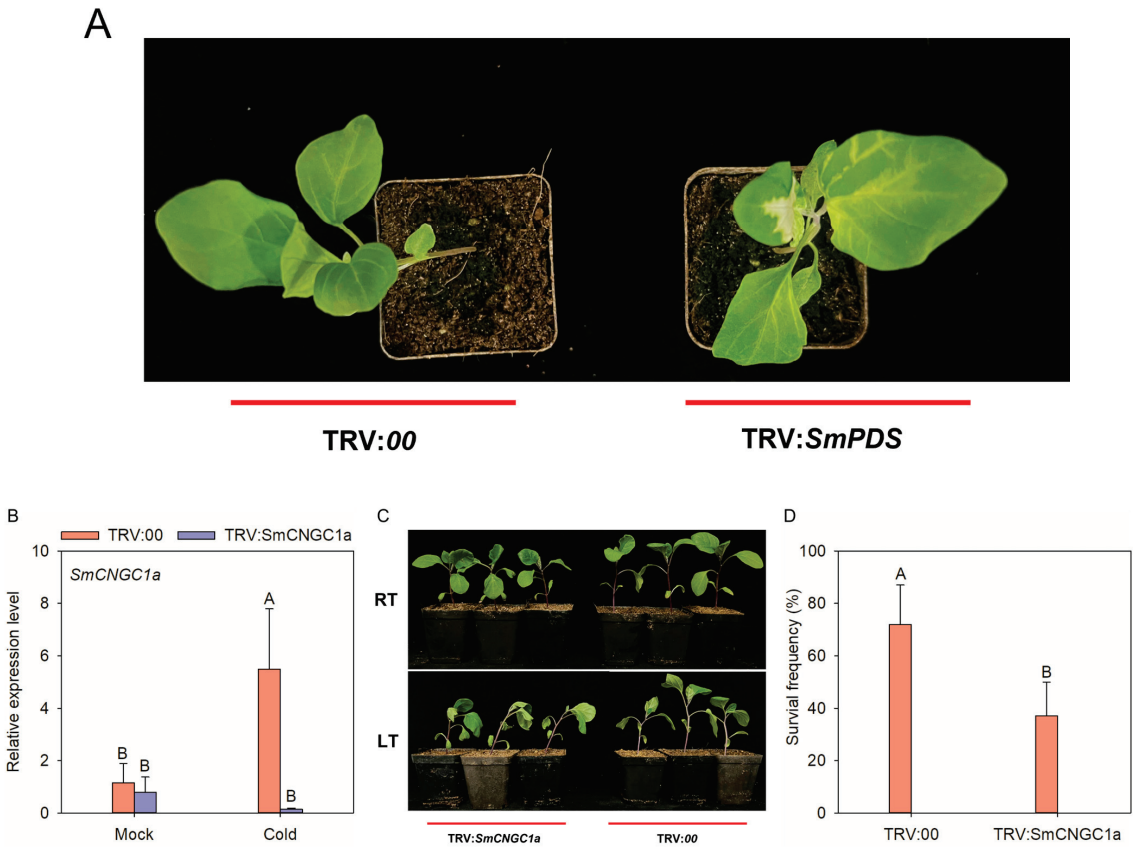


Figure 9. *Cont.*

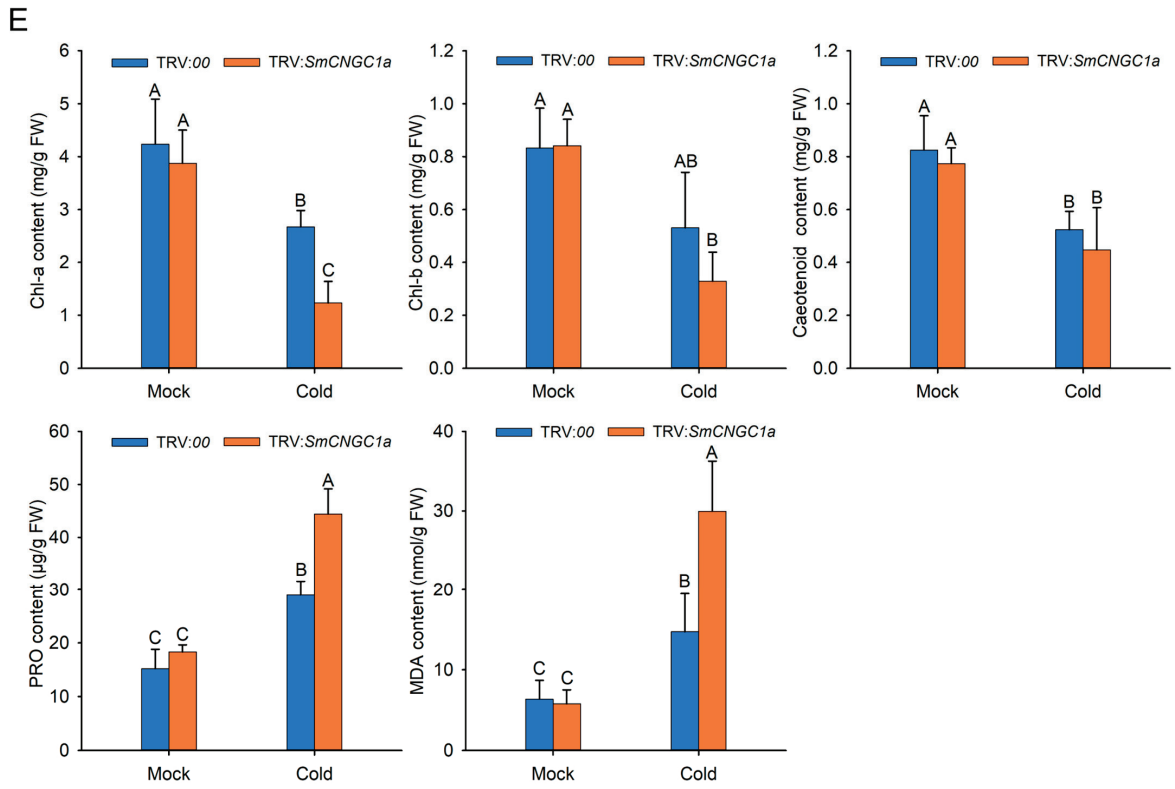


Figure 9. Effects of *SmCNGC1a* silencing on eggplant responses to cold stress (**A**) TRV: the albino phenotype of *SmPDS* indicates the success of the silencing system. (**B**) Silencing efficiency of *SmCNGC1a* in plants under cold stress based on a qRT-PCR assay. (**C**) Phenotype of *SmCNGC1a*-silenced and control plants challenged with cold stress at 2 h post-treatment. RT: 25 °C; LT: 4 °C. (**D**) Survival frequencies of *SmCNGC1a*-silenced and control plants subjected to cold stress at 2 h post-treatment. (**E**) Contents of chlorophyll a, chlorophyll b, carotenoids, proline, and malondialdehyde in *SmCNGC1a*-silenced and control plants after 2 h of cold stress. Chl-a: chlorophyll a; Chl-b: chlorophyll b; PRO: proline; MDA: malondialdehyde; FW: fresh weight. Different capital letters between samples denote significant differences according to one-way ANOVA and Tukey's test ($p < 0.01$).

2.10. Analysis Interaction Network of SmCNGC1a in Eggplant

To further investigate the function of *SmCNGC1a*, we utilized STEING v11.5 to predict and analyze its interaction network in eggplant. The results revealed that *SmCNGC1a* interacts with proteins related to defense response by callose deposition in the cell wall, regulation of anion channel activity, and receptor-mediated endocytosis (Smechr0201361.1, Smechr0201358.1, Smechr0201356.1, Smechr0201360.1). Additionally, it interacts with proteins associated with protein serine/threonine kinase activity, adenyl ribonucleotide binding, and ATP binding (Smechr0201224.1, Smechr0801839.1, Smechr0201360.1, Smechr0103284.1, Smechr0201226.1, Smechr0201225.1). The predicted interacting proteins' KEGG pathways indicate their potential involvement in the plant MAPK signaling pathway and plant–pathogen interaction. The results of STEING local network cluster analysis suggest that these proteins may participate in cGMP-binding defense response by cell wall thickening and plant MAPK signaling pathway (Figure 10).

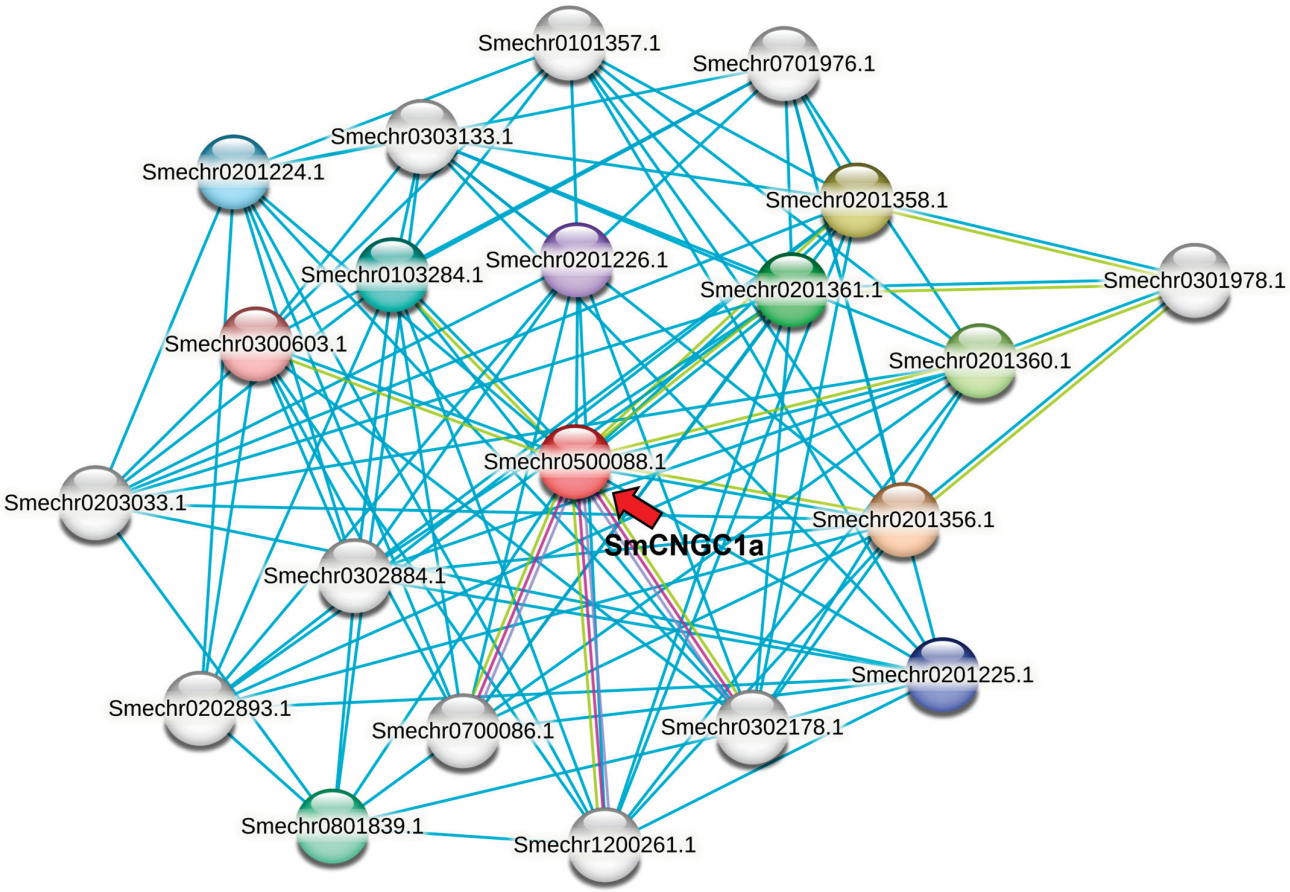


Figure 10. Interaction network analysis of *SmCNGC1a*. Nodes represent different proteins, with colored nodes indicating proteins that may have first shell of interactions with *SmCNGC1a*, and uncolored nodes indicating proteins that may have second shell of interactions with *SmCNGC1a*. Lines represent known or predicted protein–protein interactions, with blue lines indicating interactions predicted from homologous genes in other species, red lines indicating interactions that have been experimentally validated, green lines indicating interactions predicted using text mining methods, and purple lines indicating the presence of homology between the connected proteins. The "Smechr0500088.1" referred to by the red arrow is identified as SmCNGC1a.

3. Discussion

With the increasing exploration of reference genomes in various plant species, many gene families have been identified; however, the CNGC gene family in eggplant remains unknown. In this study, we identified the CNGC gene family in eggplant and conducted comprehensive bioinformatics analyses of its physicochemical properties, phylogenetic evolutionary relationships, chromosome localization, gene structure, cis-acting elements, and gene duplication events. Subsequently, we analyzed the expression patterns of the CNGC gene family under cold stress and identified a significantly upregulated gene, *SmCNGC1a*, in response to cold stress. We further analyzed its subcellular localization and tissue-specific expression patterns. Finally, we employed VIGS technology to validate the function of this gene and analyzed its protein interaction network. Our physicochemical property analysis revealed that most members of the *SmCNGC* family are hydrophilic proteins localized in the plasma membrane, which confirms their role as ion channel proteins. Phylogenetic evolutionary analysis showed that the *SmCNGC* family has relatively more members compared to Arabidopsis and tomato, with a distinct branch exhibiting a distant evolutionary relationship with other species. This indicates that CNGC proteins in eggplant may have undergone unique evolutionary processes, possibly due to differences in the environmental conditions for eggplant survival during natural evolution compared to Ara-

bidopsis and tomato, leading to adaptive evolution of the genes. Motif analysis identified a total of 10 motifs, among which the majority of *SmCNGC* proteins exhibit 8 conserved motifs characterized by similar sequences. However, a unique motif 6 is present in another branch, while motif 5 is absent, suggesting that this subset of *SmCNGC* proteins may have distinct preferences and potentially exert specialized functions during evolution.

According to reports, Ca^{2+} channels play a crucial role in plant responses to temperature stress [32]. Studies conducted on rice (*Oryza sativa* L.), cabbage (*Brassica oleracea* L.), tobacco (*Nicotiana benthamiana*), and mango (*Mangifera indica* L.) have shown that the expression of CNGC gene family undergoes changes under cold-stress conditions. For instance, 10 *OsCNGCs* in rice [33], 13 *BoCNGCs* in cabbage [34], 10 *NtabcCNGCs* in tobacco [35], and the genes *MiCNGC15* and *MiCNGC15II* in mango fruit peel [36] were found to be upregulated under cold stress. Additionally, a study in 2020 identified 15 *ZjCNGCs* in the jujube (*Ziziphus zizyphus*) genome, where *ZjCNGC2*, *8*, *10*, and *ZjCNGC15* showed downregulation within 24 h of cold-stress treatment, while the expression levels of *ZjCNGC4* and *ZjCNGC12* increased approximately four-fold and two-fold, respectively, after 1 h of cold treatment [37]. According to the research on temperature-stress treatment in Chinese cabbage (*Brassica rapa pekinensis*), it was found that the expression of *BrCNGC1*, *2*, *3*, *10*, *17*, *22*, *23*, *27*, and *BrCNGC29* was upregulated under low-temperature stress [38]. Apart from cold stress, CNGCs have also been found to be involved in regulating heat stress. Specifically, the genes *AtCNGC6* and *AtCNGC2* are involved in plant response to heat stress and are closely associated with plant thermotolerance. Under high-temperature induction, *AtCNGC6* triggers Ca^{2+} influx and the adaptive expression of heat-shock proteins. Mutants of *AtCNGC2*, namely *cngc2-1* and *cngc2-2*, exhibit enhanced tolerance to heat stress and accumulate more heat-responsive proteins [39,40]. Conversely, interference with *CNGCb* (a homologous gene of *AtCNGC2*) from the moss Physcomitrella patens leads to a super-thermosensitive phenotype in plants [41]. In rice, *OsCNGC14* and *OsCNGC16* also play important roles in plant thermotolerance. Mutants *cngc14* and *cngc16* show lower survival rates under high and low-temperature stress, and the extent of heat-stress induction and inhibition of certain genes is altered in the cngc16 mutant. Furthermore, the absence of *OsCNGC14* or *OsCNGC16* reduces or eliminates the cytosolic Ca^{2+} signals induced by temperature stress [42]. In this study, we found numerous light-responsive elements, hormone-responsive elements, and transcription factor binding sites in the promoter region of the *SmCNGC* gene, indicating that this gene family may respond to environmental signals and stress, potentially playing an important role in regulating eggplant's environmental adaptation. The expression pattern analysis of *SmCNGC* genes under cold stress also suggests that many genes in this family can respond to low-temperature signals. Specifically, we observed significant upregulation of *SmCNGC1a* under cold stress, and this finding was further validated through VIGS experiments, highlighting its importance in eggplant's response to low temperatures.

In addition, we predict complex interactions between *SmCNGC1a* and various proteins in eggplant, including cell wall defense response proteins, anion channel activity regulatory proteins, and serine/threonine-kinase-related proteins. Cell wall defense response proteins participate in the synthesis and deposition of pectin, enhancing the defensive capacity of plant cell walls. They are important mechanisms for plants to resist stress and adversity [43]. Anion channel activity regulatory proteins, similar to CNGCs, constitute a class of membrane protein channels that can regulate intracellular ion balance and signal transduction [44]. During stress and adversity, these proteins may interact synergistically with CNGCs to resist stress. Serine/threonine kinases are a class of enzyme proteins that phosphorylate serine (Ser) or threonine (Thr) residues in target proteins and participate in the regulation of various signal transduction pathways [45]. We speculate that *SmCNGC1a* may interact with the aforementioned proteins to transmit cold-stress signals and activate cold-stress-defense-related genes. However, this speculation requires further experimental validation. It is worth noting that although our study confirms the positive regulatory role of *SmCNGC1a* in cold tolerance in eggplant, there are still many unresolved questions that

require further investigation. For example, a deeper understanding of the process by which *SmCNGC1a* transmits cold-stress signals and clarification of its regulatory relationships with other genes, as well as the detailed mechanisms of the regulatory pathways, are needed. Our study provides a reference for further research on cold-tolerance genes in eggplant.

4. Materials and Methods

4.1. Identification of CNGC Gene Family in Eggplant Genome

The reference genome, coding sequences (CDS), and protein sequences of eggplant were obtained from the Eggplant Genome Database (Table 2). Hidden Markov model (HMM) profiles for the cyclic nucleotide-binding domain (cNMP, PF00027) and the ion transport protein domain (iTP, PF00520) used by the eggplant CNGC gene family were obtained from the InterPro protein family database (Table 2). The HMM profiles were searched using HMMER v3.0 software to identify genes containing both cNMP and iTP domains. These genes were then extracted, and further manual verification was performed using the NCBI CDD (Table 2) to ensure the inclusion of genes with both conserved domains and to remove pseudogenes. The final set of *SmCNGC* gene family members was obtained. The molecular weight, theoretical isoelectric point, instability index, aliphatic index, and average hydrophilicity of all the genes in the family were predicted using the Expasy ProtParam tool (Table 2). Additionally, the subcellular localization of the *SmCNGCs* was predicted using the WoLF PSORT subcellular localization prediction tool (Table 2).

Table 2. Online Analysis Websites and URLs.

Website	URL
Eggplant Genome Database	http://eggplant-hq.cn/Eggplant/home/index (accessed on 3 March 2023)
InterPro protein family database	https://www.ebi.ac.uk/interpro/ (accessed on 3 March 2023)
NCBI CDD	https://www.ncbi.nlm.nih.gov/cdd/ (accessed on 3 March 2023)
Expasy ProtParam tool	https://web.expasy.org/protparam/ (accessed on 7 April 2023)
WoLF PSORT subcellular localization prediction tool	https://wolfpsort.hgc.jp/ (accessed on 7 April 2023)
TAIR	https://www.arabidopsis.org/ (accessed on 16 March 2023)
Tomato genome database	https://solgenomics.net/ (accessed on 16 March 2023)
Evolview v3	https://www.evolgenius.info/evolview/ (accessed on 16 March 2023)
MEME	http://memesuite.org/tools/meme/ (accessed on 30 March 2023)
MapGene2Chromosome v2.1	http://mg2c.iask.in/mg2c_v2.1/ (accessed on 30 March 2023)
PlantCARE	http://bioinformatics.psb.ugent.be/webtools/plantcare/html/ (accessed on 30 March 2023)

4.2. Multiple Sequence Alignment and Phylogenetic Analysis

The protein sequence information of the Arabidopsis CNGC gene family was obtained from TAIR (Table 2). The protein sequences of the tomato CNGC gene family were referenced based on the identification results from previous studies conducted by other researchers [17] and obtained from the tomato genome database (Table 2). The amino acid sequences of CNGC gene families in Arabidopsis, tomato, and eggplant were aligned using MEGA v7.0. After removing non-conserved gaps, the alignment results were used to construct a phylogenetic tree using the maximum likelihood (ML) method with a bootstrap value set at 1000. The generated phylogenetic tree file was imported into Evolview v3 (Table 2) for visualization.

4.3. Gene Structures and Conserved Motifs Analysis

The genomic annotation file (.gff) for eggplant was obtained from the Eggplant Genome Database. The software TBtools v1.120 was used to extract gene length and positional information from the genomic annotation file of *SmCNGCs*. The extracted data were then visualized. The amino acid sequence motifs of *SmCNGCs* were analyzed using MEME (Table 2), with the search model set as "anr" and a minimum motif length of 6 and a maximum motif length of 50. Only the top 10 most reliable pieces of motif information were retained. The analysis results were visualized using MEME and Tbtools software [46].

4.4. Gene Distribution and Cis-Acting Elements Prediction

After extracting the chromosome position information of *SmCNGCs* using TBtools, the extracted results were imported into MapGene2Chromosome v2.1 (Table 2) to generate a chromosome location map of the gene family members. The promoter sequences of the *SmCNGCs* family were obtained using TBtools, specifically the 1500bp upstream of the gene's start codon. These promoter sequences were then analyzed using the PlantCARE website (Table 2) to identify and retain regulatory elements associated with environmental response, hormone response, and transcription factor binding sites that occur frequently. The retained elements were visualized using TBtools [46].

4.5. Collinearity Analysis

To uncover intraspecies microsynteny groups in eggplant, we utilized the MCScanX plugin integrated within the TBtools software to perform collinearity analysis. Subsequently, the Circos plugin in TBtools was employed to generate circos plots depicting the microsynteny groups. Additionally, we employed MCScanX to analyze the co-occurrence relationships between the CNGC gene family in eggplant and those in Arabidopsis thaliana and tomato. Visualization was carried out using TBtools [46].

4.6. Plant Materials and Treatments

The plant materials used in this experiment were eggplant variety "JS221" and *N. benthamiana*. Seeds were germinated by placing them on moist filter paper in Petri dishes, followed by cultivation in a growth chamber. The growth chamber was kept at a constant temperature of 25 °C, with a relative humidity of 60–70% and a photoperiod of 16 h of light (800 μmol m^{-2} s^{-1}) followed by 8 h of darkness. After germination, the seedlings were transplanted into seedling trays and subsequently transferred to small pots once the cotyledons had expanded, allowing for further growth. The plants were cultivated at a temperature of 25 °C and a photoperiod of 16 h of light (800 μmol m^{-2} s^{-1}) followed by 8 h of darkness until they reached the 4–6 leaf stage for subsequent experiments. Cold-stress treatment was conducted in the growth chamber at a temperature of 4 °C and heat-stress treatment was conducted in the growth chamber at 42 °C, while the other conditions remained unchanged. After the treatment, the eggplant leaves were washed with ddH_2O. Salt-stress treatment was conducted following the methods described in previous studies, with sampling focused on the root [47]. The samples were cryogenically preserved by immersing them in liquid nitrogen and then stored at a temperature of −80 °C to facilitate further analysis.

4.7. Subcellular Localization

SmCNGC1a was constructed into the pCambia1300-35S-EGFP plant overexpression vector to generate a fusion vector containing 35S: *SmCNGC1a*-GFP, which was then transformed into *Agrobacterium* strain *GV3101*. The composition of *Agrobacterium* suspension culture medium refers to previous studies in the field [47]. Healthy *N. benthamiana* were selected as the experimental material. Using a sterile syringe, the leaf surface of *N. benthamiana* was injected with OD_{600} = 0.8 suspension of pCambia1300-35S-EGFP empty vector and pCambia1300-*SmCNGC1a*-GFP separately, avoiding the leaf veins. This process was repeated three times. Subsequently, the injected *N. benthamiana* were cultivated in a growth chamber under appropriate lighting conditions for 48 h. The *N. benthamiana* leaves were then cut into suitable sizes, rinsed with sterile water, and mounted on glass slides. The fluorescence signals of the genes in the leaves were observed using a laser scanning confocal microscope (LSM 880NLO; Leica Microsystems, Wetzlar, Germany) to determine the cellular localization of *SmCNGC1a*.

4.8. Functional Analysis of SmCNGC1a Based on VIGS Method

The VIGS method referred to in previous studies [47] was employed. RNA was isolated from the leaves of pTRV2-*SmCNGC1a*-silenced plants and pTRV2 negative control

plants, followed by qRT-PCR analysis to assess the silencing efficiency of *SmCNGC1a*. Cold-stress treatment was conducted according to the procedure outlined in Section 2.6. After 2 h treatment, the phenotypes of eggplant seedlings were observed under different experimental conditions, and the survival rates were calculated.

4.9. Gene Expression Analysis by qPCR

To analyze the transcriptional expression levels of the target gene, qPCR was performed following the method established by previous studies [47]. Four biological replicates were used to detect the transcript expression levels of the target gene. The primers used in this study are listed in Table 3.

Table 3. Primers used in this study.

Gene Name	Gene ID	Forward Primer (5′ –> 3′)	Reverse Primer (5′ –> 3′)
SmCNGC1a	Smechr0500088.1	AACCAACGTTTAGCTCGTTGA	TAGAGGATGCATGCGAATTG
SmCNGC1b	Smechr0101357.1	AAAGCCACCAATCTGCTCAT	AGGAAAGGGATGCACATTGA
SmCNGC1c	Smechr0302178.1	CGGCAAATTTGGAGTGTTCT	TTTGGCCAGAAGGCAACTTA
SmCNGC2	Smechr0202893.1	CAACCTGATAACAGCGACGA	TCACAACTGGTGGAATGGAA
SmCNGC3	Smechr0400998.1	GGAAGTGAAATATTCATCATATGGTTT	CCACCTCTCTCACCGTACCT
SmCNGC4a	Smechr1000491.1	GGACAAGGATGTGGATGAGG	ACACGACCACGACCACTACA
SmCNGC4b	Smechr0400207.1	GCTCGAGTGATCTGATTGTTGA	TCCAAAATAAGTGATACCGATCC
SmCNGC5	Smechr1200261.1	TTGTTGATCTTTTGAGCTTTGC	TTTACACAATCGGTGTATATAAAACTC
SmCNGC6	Smechr1100117.1	TGGAGGTCGAGCAGAGTATG	TTTGCCGGCTAATTTTTCTC
SmCNGC7	Smechr0700086.1	GAGTCGAGTTTGAGGGCTTG	TCGCAGTCTTGCTGATGAAC
SmCNGC8	Smechr0303133.1	TTGGAGGGCAAAAAGAAAAG	TGGTTACATGCCCACCAGTA
SmCNGC9	Smechr0801568.1	CTGAAGGATCTGGATTCTTTGC	TCATCTTGACATCTTAACTTATGGA
SmCNGC10	Smechr0800622.1	GGACATGGAAAGCAAACCAA	CGTCCACAACTTTCACCTTC
SmCNGC11	Smechr0801458.1	ATTGCTTGTGGACAATGGTG	TCACCTCCATACCGGATGAT
SmCNGC12	Smechr0201396.1	GGGATTTGGAGGTTTTGGTT	TGTCCATCACCTTTCTCTTGC
SmCNGC13	Smechr0900702.1	TTTCAGAAATGTATCTGATTGACC	CTCAATGACTAGAATTCCGCTGT
SmCNGC14	Smechr0801673.1	AGCTGGCCAAAGAACTTTACA	TGTTGATCATCCTCGGGTTC
SmCNGC15a	Smechr0900248.1	CCTCGAGGAGGTCCTATAAACA	CCATGGGATAACTTGCATCC
SmCNGC15b	Smechr0600577.1	CAGTTGTAACTTGTAAGATAAGATGGA	ATGGCACAAAAGCTGCAGTA
SmCNGC15c	Smechr1201893.1	CAAATGTGGAAGGGTGTTTT	GTTTCTCTTCCCCCTCTTGC
SmCNGC16	Smechr0302884.1	CAGGGAAAGTCGTTTTGGAA	GGAAGCAGCAAAAACAGAGG
SmCNGC17	Smechr0700014.1	TGGGAGGAAAAGCAGACAGT	CCTTTTTAGGCCTCCCAAAC
SmCNGC18	Smechr0203033.1	GGTGGCGTCAGATTTTTGAT	TGACGAAAGGGACGAAGAAG
SmCNGC19	Smechr0302451.1	GCCAAAGAAGTTCAGGCAGA	AGTAATTCCGCAGCCATTTG
SmCNGC20	Smechr0302478.1	TTGGTCGAGAGCCTGAGAAT	TACGCCAACCATTTCGTTCT
SmCNGC21	Smechr0500154.1	AATCGTCGAGAAGCAGCAGT	GAGGCCATTGATGACGTTTT
SmCNGC22	Smechr0103753.1	AAAAACAGAGGAAACAAATATAATGAA	TGCTATCATGTTCATCTCATTACCA
SmCNGC23	Smechr1200076.1	TGGAGCAGCACAAGAAATTG	TTGCCGATCATAAGGTGAAA
SmCNGC24	Smechr0100701.1	TGCAAATGAGCCATTCATACA	TGCTACTCCCATGGCTATCA

4.10. Physiological Parameter Determination

The determination method of photosynthetic pigments is as follows: Firstly, 0.2 g of leaf samples were weighed into a mortar, and 12 mL of 95% ethanol was added. The samples were ground until the tissue turned white and left to stand for 3–5 min. Then, the mixture was filtered into a 25 mL brown volumetric flask using filter paper. The filter paper and residue were rinsed several times until no residues remained. Finally, the volume was adjusted to a fixed volume with ethanol and shaken well. The absorbance values (A_{665}, A_{649}, and A_{470}) of the chlorophyll extract were measured at wavelengths of 665 nm, 649 nm, and 470 nm, respectively. The concentrations (mg/L) of chlorophyll a, chlorophyll b, and carotenoids were calculated using the following formulas: Ca = 13.95 × A6656.88 × A649; Cb = 24.96 × A_{649} − 0.32 × A_{665}; Cx·c = (1000 × A − 2.05 × Ca − 114.8 × Cb)/245. Subsequently, the pigment content in the tissue (mg/g) was calculated using the following formula: Pigment content = Pigment concentration (calculated as above) × Extract volume × Dilution factor/FW (fresh weight) [48]. Proline content was determined using the Proline (Pro) Content Assay Kit (AKAM003C) provided by BoxBio (Beijing BoxBio Science & Technology Co., Ltd., Beijing, China). Malondialdehyde (MDA) content was determined using the Malondialdehyde (MDA) Content Assay Kit (AKFA013C) provided by BoxBio

(Beijing BoxBio Science & Technology Co., Ltd.). The experimental procedures for these assays were performed according to the respective kit instructions. All experiments were conducted with three biological replicates to ensure reproducibility.

4.11. Statistic Analysis

The statistical analysis data were analyzed utilizing Microsoft Office Excel 2019 and IBM SPSS Statistics 26. To assess the variations among the samples, one-way analysis of variance (ANOVA) was conducted, followed by Tukey's test ($p < 0.01$) for post hoc comparisons.

5. Conclusions

In general, we identified the CNGC gene family in eggplant and analyzed its expression patterns under cold stress. Finally, functional analysis was performed on the cold-responsive gene *SmCNGC1a*. These studies provide valuable information for the research on cold tolerance in eggplant. Further investigations can be based on these findings to explore the function of *SmCNGC* and the regulatory mechanisms of *SmCNGC1a* in cold tolerance in eggplant. This will provide a theoretical basis for the development of cold-tolerant eggplant varieties and breeding strategies aimed at enhancing cold tolerance in eggplant.

Author Contributions: All authors contributed experimental design oversight. X.Y. and L.S. conceived the experiments. Z.J., L.D., X.X. and L.Z. carried out the experiments with the help of J.H. and X.Y. contributed the plant materials and data analysis. Z.J. wrote the manuscript and J.H. edited the manuscript. All authors contributed to the article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The National Key Research and Development Program of China (2019YFD1000300).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Pebriana, E.; Maghfoer, M.D.; Widaryanto, E. Effect of grafting using wild eggplant as rootstock on growth and yield of four eggplant (*Solanum melongena* L.) cultivars. *Biosci. Res.* **2018**, *15*, 337–347.
2. Kadivec, M.; Kopjar, M.; Znidarcic, D.; Pozrl, T. Potential of Eggplant Peel as by-Product. *Acta Aliment.* **2015**, *44*, 126–131. [CrossRef]
3. Hanson, J.B. Impairment of Respiration, Ion Accumulation, and Ion Retention in Root Tissue Treated with Ribonuclease and Ethylenediamine Tetraacetic Acid. *Plant Physiol.* **1960**, *35*, 372–379. [CrossRef] [PubMed]
4. Dodd, A.N.; Kudla, J.; Sanders, D. The language of calcium signaling. *Annu. Rev. Plant Biol.* **2010**, *61*, 593–620. [CrossRef]
5. Kudla, J.; Batistic, O.; Hashimoto, K. Calcium signals: The lead currency of plant information processing. *Plant Cell* **2010**, *22*, 541–563. [CrossRef]
6. Guo, K.M.; Babourina, O.; Christopher, D.A.; Borsic, T.; Rengel, Z. The cyclic nucleotide-gated channel AtCNGC10 transports Ca^{2+} and Mg^{2+} in Arabidopsis. *Physiol. Plant* **2010**, *139*, 303–312. [CrossRef]
7. Maser, P.; Thomine, S.; Schroeder, J.I.; Ward, J.M.; Hirschi, K.; Sze, H.; Talke, I.N.; Amtmann, A.; Maathuis, F.J.; Sanders, D.; et al. Phylogenetic relationships within cation transporter families of Arabidopsis. *Plant Physiol.* **2001**, *126*, 1646–1667. [CrossRef]
8. Yuen, C.C.Y.; Christopher, D.A. The group IV-A cyclic nucleotide-gated channels, CNGC19 and CNGC20, localize to the vacuole membrane in Arabidopsis thaliana. *Aob Plants* **2013**, *5*, plt012. [CrossRef]
9. Talke, I.N.; Blaudez, D.; Maathuis, F.J.; Sanders, D. CNGCs: Prime targets of plant cyclic nucleotide signalling? *Trends Plant Sci.* **2003**, *8*, 286–293. [CrossRef]
10. Ma, W.; Berkowitz, G.A. Ca^{2+} conduction by plant cyclic nucleotide gated channels and associated signaling components in pathogen defense signal transduction cascades. *New. Phytol.* **2011**, *190*, 566–572. [CrossRef]
11. Spalding, E.P.; Harper, J.F. The ins and outs of cellular Ca^{2+} transport. *Curr. Opin. Plant Biol.* **2011**, *14*, 715–720. [CrossRef] [PubMed]
12. Zelman, A.K.; Dawe, A.; Gehring, C.; Berkowitz, G.A. Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels. *Front. Plant Sci.* **2012**, *3*, 95. [CrossRef] [PubMed]

13. Young, E.C.; Krougliak, N. Distinct structural determinants of efficacy and sensitivity in the ligand-binding domain of cyclic nucleotide-gated channels. *J. Biol. Chem.* **2004**, *279*, 3553–3562. [CrossRef]
14. Schuurink, R.C.; Shartzer, S.F.; Fath, A.; Jones, R.L. Characterization of a calmodulin-binding transporter from the plasma membrane of barley aleurone. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1944–1949. [CrossRef]
15. Arazi, T.; Kaplan, B.; Fromm, H. A high-affinity calmodulin-binding site in a tobacco plasma-membrane channel protein coincides with a characteristic element of cyclic nucleotide-binding domains. *Plant Mol. Biol.* **2000**, *42*, 591–601. [CrossRef]
16. Kohler, C.; Neuhaus, G. Characterisation of calmodulin binding to cyclic nucleotide-gated ion channels from Arabidopsis thaliana. *FEBS Lett.* **2000**, *471*, 133–136. [CrossRef] [PubMed]
17. Saand, M.A.; Xu, Y.P.; Li, W.; Wang, J.P.; Cai, X.Z. Cyclic nucleotide gated channel gene family in tomato: Genome-wide identification and functional analyses in disease resistance. *Front. Plant Sci.* **2015**, *6*, 303. [CrossRef]
18. Dietrich, P.; Moeder, W.; Yoshioka, K. Plant Cyclic Nucleotide-Gated Channels: New Insights on Their Functions and Regulation. *Plant Physiol.* **2020**, *184*, 27–38. [CrossRef]
19. Tian, W.; Hou, C.; Ren, Z.; Wang, C.; Zhao, F.; Dahlbeck, D.; Hu, S.; Zhang, L.; Niu, Q.; Li, L.; et al. A calmodulin-gated calcium channel links pathogen patterns to plant immunity. *Nature* **2019**, *572*, 131–135. [CrossRef]
20. Jarratt-Barnham, E.; Wang, L.; Ning, Y.; Davies, J.M. The Complex Story of Plant Cyclic Nucleotide-Gated Channels. *Int. J. Mol. Sci.* **2021**, *22*, 874. [CrossRef]
21. Zhang, S.; Pan, Y.; Tian, W.; Dong, M.; Zhu, H.; Luan, S.; Li, L. Arabidopsis CNGC14 Mediates Calcium Influx Required for Tip Growth in Root Hairs. *Mol. Plant* **2017**, *10*, 1004–1006. [CrossRef] [PubMed]
22. Brost, C.; Studtrucker, T.; Reimann, R.; Denninger, P.; Czekalla, J.; Krebs, M.; Fabry, B.; Schumacher, K.; Grossmann, G.; Dietrich, P. Multiple cyclic nucleotide-gated channels coordinate calcium oscillations and polar growth of root hairs. *Plant J.* **2019**, *99*, 910–923. [CrossRef] [PubMed]
23. Zeb, Q.; Wang, X.; Hou, C.; Zhang, X.; Dong, M.; Zhang, S.; Zhang, Q.; Ren, Z.; Tian, W.; Zhu, H.; et al. The interaction of CaM7 and CNGC14 regulates root hair growth in Arabidopsis. *J. Integr. Plant Biol.* **2020**, *62*, 887–896. [CrossRef] [PubMed]
24. Tan, Y.Q.; Yang, Y.; Zhang, A.; Fei, C.F.; Gu, L.L.; Sun, S.J.; Xu, W.; Wang, L.; Liu, H.; Wang, Y.F. Three CNGC Family Members, CNGC5, CNGC6, and CNGC9, Are Required for Constitutive Growth of Arabidopsis Root Hairs as Ca^{2+}-Permeable Channels. *Plant Commun.* **2020**, *1*, 100001. [CrossRef]
25. Moon, J.Y.; Belloeil, C.; Ianna, M.L.; Shin, R. Arabidopsis CNGC Family Members Contribute to Heavy Metal Ion Uptake in Plants. *Int. J. Mol. Sci.* **2019**, *20*, 413. [CrossRef]
26. Tunc-Ozdemir, M.; Tang, C.; Ishka, M.R.; Brown, E.; Groves, N.R.; Myers, C.T.; Rato, C.; Poulsen, L.R.; McDowell, S.; Miller, G.; et al. A cyclic nucleotide-gated channel (CNGC16) in pollen is critical for stress tolerance in pollen reproductive development. *Plant Physiol.* **2013**, *161*, 1010–1020. [CrossRef]
27. Tunc-Ozdemir, M.; Rato, C.; Brown, E.; Rogers, S.; Mooneyham, A.; Frietsch, S.; Myers, C.T.; Poulsen, L.R.; Malho, R.; Harper, J.F. Cyclic nucleotide gated channels 7 and 8 are essential for male reproductive fertility. *PLoS ONE* **2013**, *8*, e55277. [CrossRef]
28. Pan, Y.; Chai, X.; Gao, Q.; Zhou, L.; Zhang, S.; Li, L.; Luan, S. Dynamic Interactions of Plant CNGC Subunits and Calmodulins Drive Oscillatory Ca^{2+} Channel Activities. *Dev. Cell* **2019**, *48*, 710–725.e715. [CrossRef]
29. Gao, Q.F.; Fei, C.F.; Dong, J.Y.; Gu, L.L.; Wang, Y.F. Arabidopsis CNGC18 is a Ca^{2+}-permeable channel. *Mol. Plant* **2014**, *7*, 739–743. [CrossRef]
30. Gao, Q.F.; Gu, L.L.; Wang, H.Q.; Fei, C.F.; Fang, X.; Hussain, J.; Sun, S.J.; Dong, J.Y.; Liu, H.; Wang, Y.F. Cyclic nucleotide-gated channel 18 is an essential Ca^{2+} channel in pollen tube tips for pollen tube guidance to ovules in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 3096–3101. [CrossRef]
31. Shi, J.; Zuo, J.; Xu, D.; Gao, L.; Wang, Q. Effect of low-temperature conditioning combined with methyl jasmonate treatment on the chilling resistance of eggplant (*Solanum melongena* L.) fruit. *J. Food Sci. Technol.* **2019**, *56*, 4658–4666. [CrossRef] [PubMed]
32. Hong-Bo, S.; Li-Ye, C.; Ming-An, S.; Shi-Qing, L.; Ji-Cheng, Y. Bioengineering plant resistance to abiotic stresses by the global calcium signal system. *Biotechnol. Adv.* **2008**, *26*, 503–510. [CrossRef] [PubMed]
33. Nawaz, Z.; Kakar, K.U.; Saand, M.A.; Shu, Q.Y. Cyclic nucleotide-gated ion channel gene family in rice, identification, characterization and experimental analysis of expression response to plant hormones, biotic and abiotic stresses. *BMC Genom.* **2014**, *15*, 853. [CrossRef] [PubMed]
34. Kakar, K.U.; Nawaz, Z.; Kakar, K.; Ali, E.; Almoneafy, A.A.; Ullah, R.; Ren, X.L.; Shu, Q.Y. Comprehensive genomic analysis of the CNGC gene family in Brassica oleracea: Novel insights into synteny, structures, and transcript profiles. *BMC Genom.* **2017**, *18*, 869. [CrossRef]
35. Nawaz, Z.; Kakar, K.U.; Ullah, R.; Yu, S.; Zhang, J.; Shu, Q.Y.; Ren, X.L. Genome-wide identification, evolution and expression analysis of cyclic nucleotide-gated channels in tobacco (*Nicotiana tabacum* L.). *Genomics* **2019**, *111*, 142–158. [CrossRef]
36. Sivankalyani, V.; Sela, N.; Feygenberg, O.; Zemach, H.; Maurer, D.; Alkan, N. Transcriptome Dynamics in Mango Fruit Peel Reveals Mechanisms of Chilling Stress. *Front. Plant Sci.* **2016**, *7*, 1579. [CrossRef]
37. Wang, L.; Li, M.; Liu, Z.; Dai, L.; Zhang, M.; Wang, L.; Zhao, J.; Liu, M. Genome-wide identification of CNGC genes in Chinese jujube (*Ziziphus jujuba* Mill.) and ZjCNGC2 mediated signalling cascades in response to cold stress. *BMC Genom.* **2020**, *21*, 191. [CrossRef]
38. Li, Q.; Yang, S.; Ren, J.; Ye, X.; Jiang, X.; Liu, Z. Genome-wide identification and functional analysis of the cyclic nucleotide-gated channel gene family in Chinese cabbage. *3 Biotech* **2019**, *9*, 114. [CrossRef]

39. Gao, F.; Han, X.; Wu, J.; Zheng, S.; Shang, Z.; Sun, D.; Zhou, R.; Li, B. A heat-activated calcium-permeable channel—Arabidopsis cyclic nucleotide-gated ion channel 6—Is involved in heat shock responses. *Plant J.* **2012**, *70*, 1056–1069. [CrossRef]
40. Katano, K.; Kataoka, R.; Fujii, M.; Suzuki, N. Differences between seedlings and flowers in anti-ROS based heat responses of Arabidopsis plants deficient in cyclic nucleotide gated channel 2. *Plant Physiol. Biochem.* **2018**, *123*, 288–296. [CrossRef]
41. Finka, A.; Cuendet, A.F.; Maathuis, F.J.; Saidi, Y.; Goloubinoff, P. Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell* **2012**, *24*, 3333–3348. [CrossRef]
42. Cui, Y.; Lu, S.; Li, Z.; Cheng, J.; Hu, P.; Zhu, T.; Wang, X.; Jin, M.; Wang, X.; Li, L.; et al. Cyclic Nucleotide-Gated Ion Channels 14 and 16 Promote Tolerance to Heat and Chilling in Rice. *Plant Physiol.* **2020**, *183*, 1794–1808. [CrossRef]
43. Jin, L.; Mackey, D.M. Measuring Callose Deposition, an Indicator of Cell Wall Reinforcement, during Bacterial Infection in Arabidopsis. *Methods Mol. Biol.* **2017**, *1578*, 195–205. [CrossRef] [PubMed]
44. McCulloch, S.R.; Laver, D.R.; Walker, N.A. Anion channel activity in the Chara plasma membrane: Co-operative subunit phenomena and a model. *J. Exp. Bot.* **1997**, *48*, 383–397. [CrossRef] [PubMed]
45. Castano, A.; Silvestre, M.; Wells, C.I.; Sanderson, J.L.; Ferrer, C.A.; Ong, H.W.; Lang, Y.; Richardson, W.; Silvaroli, J.A.; Bashore, F.M.; et al. Discovery and characterization of a specific inhibitor of serine-threonine kinase cyclin dependent kinase-like 5 (CDKL5) demonstrates role in hippocampal CA1 physiology. *bioRxiv* **2023**. [CrossRef]
46. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]
47. Shen, L.; Zhao, E.; Liu, R.; Yang, X. Transcriptome Analysis of Eggplant under Salt Stress: AP2/ERF Transcription Factor SmERF1 Acts as a Positive Regulator of Salt Stress. *Plants* **2022**, *11*, 2205. [CrossRef] [PubMed]
48. Wang, X.; Fang, Z.; Zhao, D.; Tao, J. Effects of High-Temperature Stress on Photosynthetic Characteristics and Antioxidant Enzyme System of Paeonia ostii. *Phyton* **2022**, *91*, 599–615. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

Comprehensive Analysis of the Pectate Lyase Gene Family and the Role of *FaPL1* in Strawberry Softening

Yuanxiu Lin †, Hao He †, Yanling Wen, Shuaipeng Cao, Zisen Wang, Ziqing Sun, Yunting Zhang, Yan Wang, Wen He, Mengyao Li, Qing Chen, Yong Zhang, Ya Luo, Xiaorong Wang and Haoru Tang *

College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China; linyx@sicau.edu.cn (Y.L.)
* Correspondence: htang@sicau.edu.cn
† These authors contributed equally to this work.

Abstract: Fruit softening is a crucial factor that controls shelf life and commercial value. Pectate lyase (PL) has a major role in strawberry fruit softening. However, the *PL* gene family in strawberry has not been comprehensively analyzed. In this study, 65 *FaPL* genes were identified in the octoploid strawberry genome. Subcellular localization prediction indicated that FaPLs are mostly localized to the extracellular and cytoplasmic spaces. Duplication event analysis suggested that *FaPL* gene family expansion is mainly driven by whole genome or segmental duplication. The *FaPL* family members were classified into six groups according to the phylogenetic analysis. Among them, *FaPL1*, *3*, *5*, *20*, 25, *42*, and *57* had gradually increased expressions during strawberry fruit development and ripening and higher expression levels in the fruits with less firmness than that in firmer fruit. This result suggested that these members are involved in strawberry softening. Furthermore, overexpression of *FaPL1* significantly reduced the fruit firmness, ascorbic acid (AsA), and malondialdehyde (MDA) content but obviously increased the anthocyanins, soluble proteins, and titratable acidity (TA), while it had no apparent effects on flavonoids, phenolics, and soluble sugar content. These findings provide basic information on the *FaPL* gene family for further functional research and indicate that *FaPL1* plays a vital role in strawberry fruit softening.

Keywords: pectate lyase; strawberry; fruit softening; fruit firmness

1. Introduction

Due to its rich nutrients (such as vitamin C, vitamin A, anthocyanins, etc.), unique flavor, sweetness, and bright color, the strawberry (*Fragaria* × *ananassa* Duch.) is cultivated and consumed worldwide. However, the extreme soft texture of strawberry brings a very short shelf life and poor fruit quality, thus leading to enormous marketing and economic losses. Therefore, the mechanisms underlying fruit softening and its manipulation are of great interest for both consumers and breeders.

Softening, a typical characteristic of ripening in most fleshy fruits, is a crucial factor that controls shelf life and commercial value. In certain species and cultivar, some degree of softening is desirable to consumers. However, during fruit ripening, excessive softening often leads to postharvest damage or infection decay that results in diminished fruit quality and significant economic losses. Although softening has been shown to be regulated by endogenous phytohormones (e.g., ethylene and abscisic acid) [1,2] and degradation of starch [3], it is mainly caused by modification or remodeling of the cell wall (CW) [4]. The plant CW comprises the primary CW, secondary CW, and middle lamella (ML). The primary CW consists of of cellulose, hemicellulose, and pectin; however, the weakening and disassembly of the CW are predominantly due to the solubilization and depolymerization of hemicellulose and pectin.

Pectin is the most complex polysaccharide and is also an important component of plant primary CW and ML. It plays an important role in intercellular adhesion, maintaining the

stability, strength, and integrity of the CW. Accompanied with pectin degradation, the CW dissociates, and the fruit softens. The degradation of pectin is the result of the joint action of several metabolic enzymes, including pectin methylesterase (PME), polygalacturonase (PG), and pectate lyase (PL). PG is one of the most studied enzymes, which can hydrolyze homogalacturonan (HG), leading to the degradation of pectin. However, silencing of the ripening-related PG genes has a relatively small effect on slowing down the rate of fruit softening, and PG enzyme activity is relatively low in strawberries [5]. Hence, it is generally believed that the role of PG genes in strawberry fruit softening is relatively small. Recent studies have shown that there are a total of 82 PG genes in the strawberry genome of which *FaPG1* and *FaPG2* are highly expressed in the fruit. Silencing *FaPG1* and *FaPG2* significantly increases the firmness of strawberry fruit [6–8], indicating that these two genes play an important role in strawberry fruit softening. Similarly, PME has also been screened and identified in strawberry. In total, 54 PME genes have been identified in the strawberry genome of which *FvPME38* and *FvPME39* are involved in fruit softening and are regulated by MYB transcription factors [9,10]. However, the screening of the PL gene family in strawberry and the identification of its members related to fruit softening have not been reported yet.

Pectate lyase (PL) belongs to polysaccharide lyase family 1 (PL1), can randomly breaks the β-1,4 glycosidic bond, producing unsaturated C4-C5 bond oligomeric galactose uronic acid, thereby degrading pectin. PL exists in many plant species and has been identified in tomato [11], *Arabidopsis* [12], peach [13], and grape [14]. Its important role in fruit softening has also been confirmed in several species, such as banana [15], tomato [16], mango [17], and grape [18]. Previous studies have shown that silencing of the PL gene in tomatoes reduces the content of water-soluble pectin, which significantly improves fruit firmness and prolongs the storage period of the fruit [19]. In strawberry, it has been previously reported that the expression levels of three members of the PL family (*FaPLa*, *FaPLb*, and *FaPLc*) increased along with fruit maturation, indicating these genes are associated with fruit ripening and softening [20,21]. Meanwhile, by antisense inhibition of *FaPLc* gene expression, the fruit firmness was significantly increased [22,23], confirming the involvement of the PL gene in strawberry fruit softening.

Although a few softening-related PL genes in strawberry have been reported, the genome-wide systematic examination is still missing. In this study, we identified all the PL family members in strawberry; the basic information and expression profiles were obtained during fruit ripening. In addition, transient overexpression was used to clarify the function of *FaPL1* in strawberry fruit softening. The findings provide a foundation for further investigation of the function of PL family members in strawberry fruit softening, aimed to better reveal the molecular mechanism underlying strawberry fruit softening.

2. Results

2.1. Identification of PL Genes in Strawberry

A total of 65 *FaPL* genes were identified by searching and confirming the conserved PL domains (PF00544) in the genome of cultivated strawberry. According to their distribution order on chromosomes, all the 65 *FaPL* genes were renamed as *FaPL1* to *FaPL65* (Figure 1). The sixty-five *FaPL* genes were unevenly distributed across the twenty-eight chromosomes in the four subgenomes of cultivated strawberry, with an apparent concentration on the chromosome 6. A maximum of seven *FaPL* genes were located on chromosome 6 from the first, second, and third subgenomes (Fvb6-1, Fvb6-2, Fvb6-3), while the minimum number was 1 on chromosomes Fvb3-1, Fvb2-2, Fvb3-2, Fvb2-4, and Fvb7-4. However, there were no *FaPL* genes on chromosomes 1 and 2 from the first subgenome (Fvb1-1, Fvb2-1), chromosome 2 from the second subgenome (Fvb2-2), chromosomes 1, 2, and 3 from the third genome (Fvb1-3, Fvb2-3 and Fvb3-3), and chromosomes 1 and 3 from the fourth subgenome (Fvb1-4 and Fvb3-4), which is not shown.

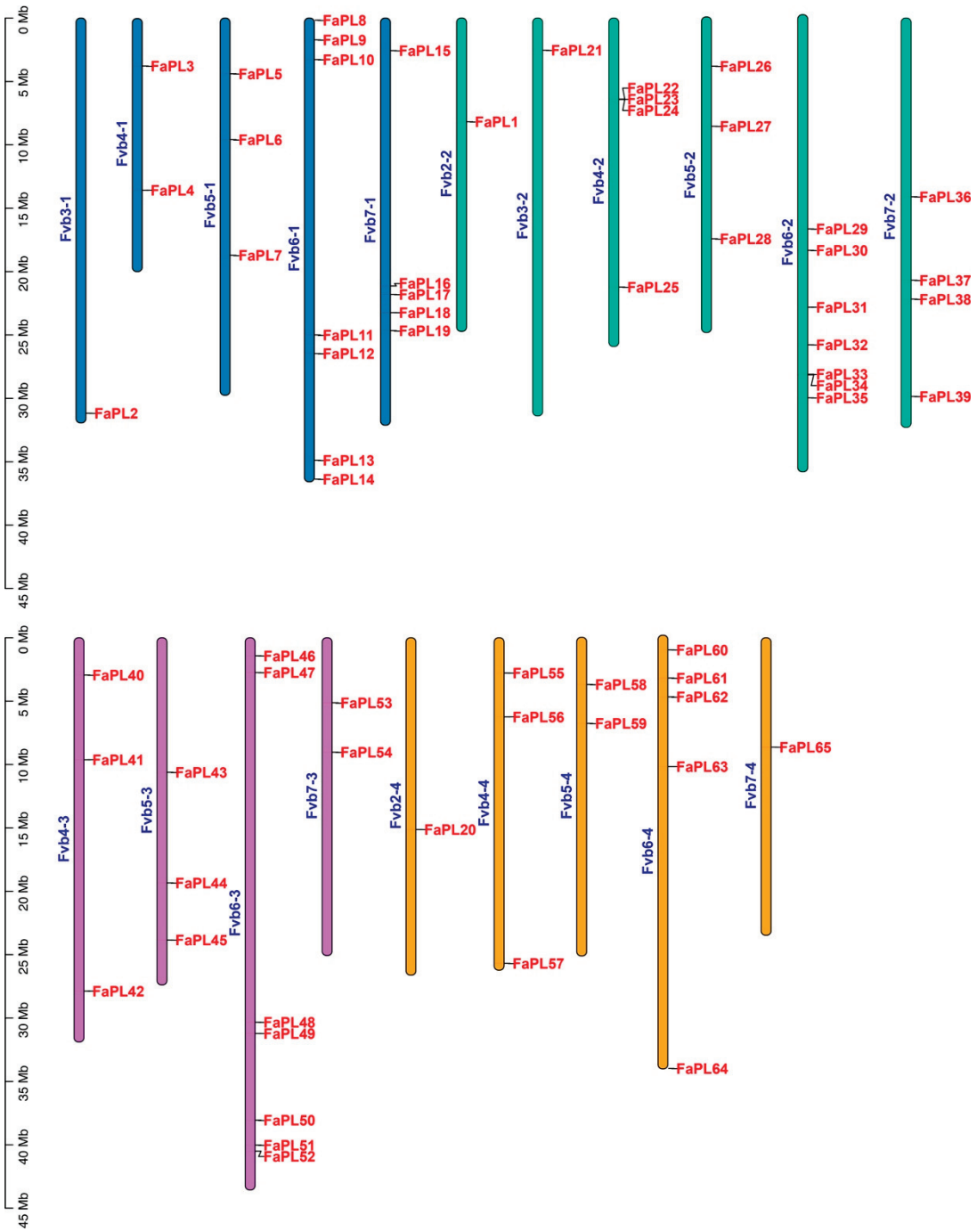


Figure 1. Chromosomal distribution and location of *FaPLs* in strawberry. Different colors indicate the chromosomes from different subgenomes of cultivated strawberry.

The characteristics and physicochemical properties of the deduced 65 FaPL proteins is shown in Supplementary Table S1. The number of amino acids varied from 91 to 500 aa, most of them (46) were concentrated from 400 to 500 aa. There were only two FaPL proteins comprising amino acids below 100 aa. The molecular weights (MW) were from 10.282 to 53.933 KDa. Only 21 FaPL proteins had isoelectric points (pI) below 7, while the others were all above or equal to 7 and three of which had pI above 10. Furthermore, based on the subcellular localization prediction results (Supplementary Table S1), most FaPL proteins were predicted to be located in the extracellular space (21), suggesting they might be secreted proteins. Several FaPL proteins were located in the cytoplasm

(13), plasma membrane (9), and vacuolar (8); some FaPLs were localized in mitochondria, chloroplast, and peroxisome. Interestingly, a few FaPLs were predicted to be dual-localized, as examples, FaPL24 was located in both chloroplast and nuclear, while FaPL25 and FaPL57 were located in chloroplast or vacuolar (Supplementary Table S1). Subsequently, the origins of duplication events of *FaPLs* in strawberry were detected using MCScanX package. As a result, three types of duplication events were found, including whole genome duplication or segmental (WGD/segmental), dispersed, and proximal (Supplementary Table S1). Most *FaPLs* were duplicated by WGD/segmental, only seven and two *FaPLs* were duplicated from dispersed and proximal duplication events separately.

2.2. Phylogenetic and Gene Structure Analysis for FaPL Genes

According to the result (Figure 2), all the sixty-five *FaPL* genes were classified into six clusters. Among them, group I is the largest group containing 16 members, followed by groups II and III, which had 14 and 12 members of *FaPL* genes, respectively. Both groups IV and V had ten *FaPL* members, whereas there were only two *FaPLs* included in group VI.

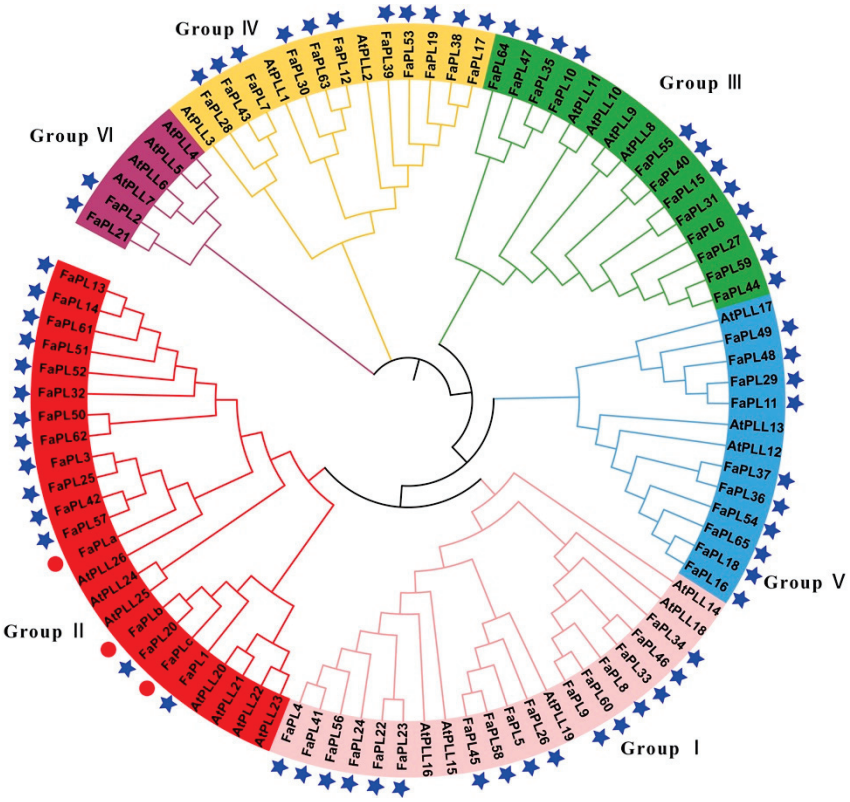


Figure 2. Phylogenetic tree of *FaPLs* from strawberry and *Arabidopsis*. Different branch colors represent the different groups. PL family members from strawberry identified in this study are marked with blue stars. The red dots indicate the previously reported *FaPLs*.

To better elucidate the structural characteristics of the *FaPL* genes, their intron/exon distributions were analyzed and visualized (Figure 3). Overall, six *FaPLs*, including in Group II, had no intron, and other members displayed discontinuous sequences due to the distribution of different number of introns. The exons numbers ranged from 1 to 7. Specifically, all members belonging to group V had four exons. It was noted that six *FaPLs* members (*FaPL1, 20, 3, 25, 57, and 42*) in Group II contained the most exons, while the other eight members of *FaPLs* only had one or two exons. The exons ranged from three to five of *FaPLs* divided into other groups. Meanwhile, the conserved motifs of FaPL proteins were analyzed using MEME Suite online software (version 5.5.3). The motifs number and distribution order were similar in FaPL members, most of which contained 10 motifs.

However, FaPL15, 31, and 39 only contained two conserved motifs, FaPL50 had five motifs, and most members classed into group IV had six motifs (Figure 3). As shown (Figure 3), all the FaPL proteins contained the core motifs 1, 2, 6, or 10, which were annotated as PL domains. Motif 3 encoded a zinc finger domain, while the others were unknown.

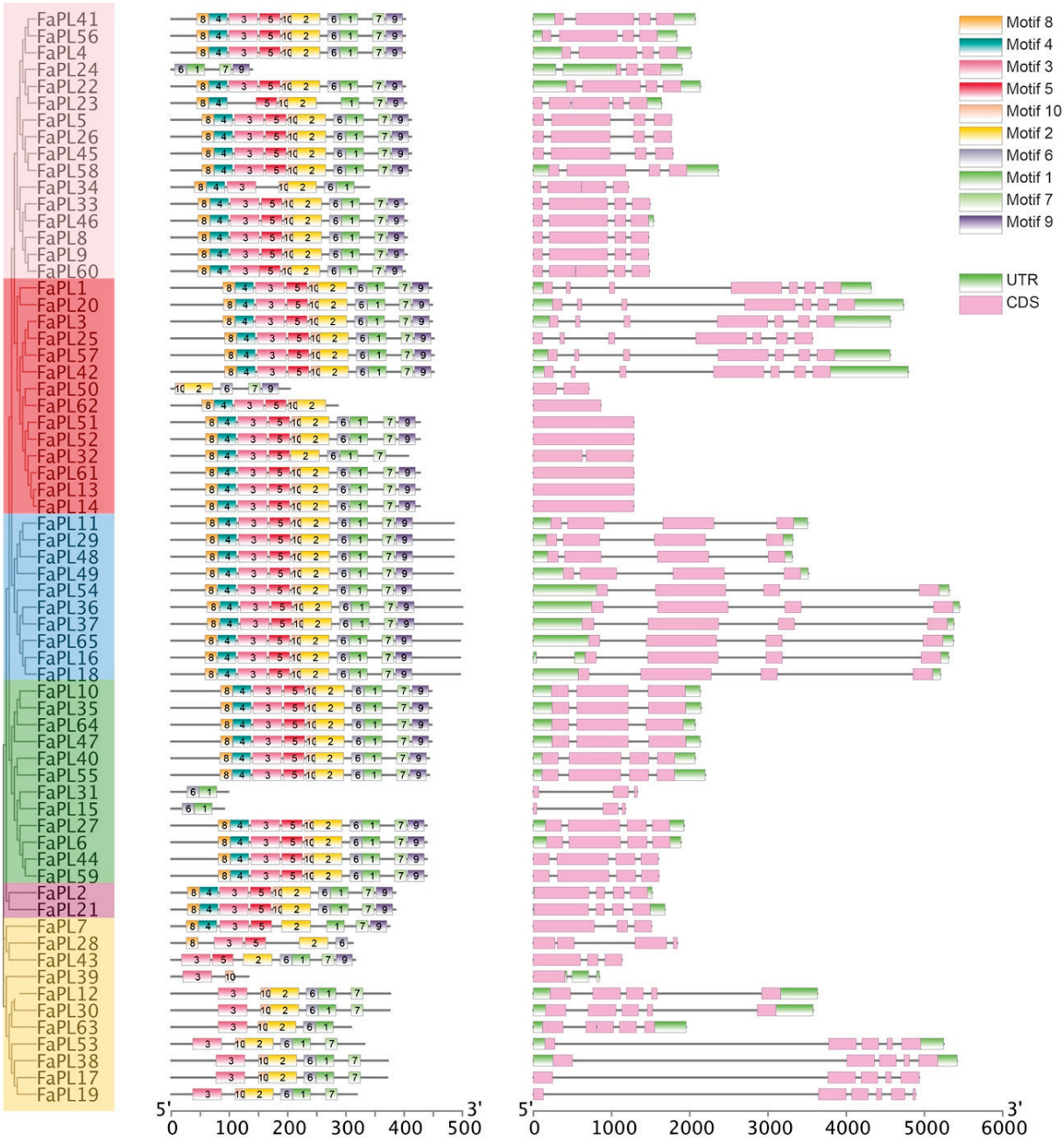


Figure 3. Conserved motifs and gene structure analysis of *FaPLs*. Left part indicated an unroot tree of strawberry *FaPLs*, middle part displayed the distribution of conserved motifs on each FaPL protein, and the right part showed the exon–intron distribution of *FaPLs*.

2.3. Collinearity Analysis

The collinearity analysis among *Arabidopsis*, woodland strawberry (*Fragaria vesca*), and cultivated strawberry (*Fragaria* × *ananassa*) was carried out to explore the evolutionary relationship of *FaPLs*. According to the result, *57 FaPLs*, 18 *AtPLs*, and 16 *FvPLs* were involved to form 129 collinear pairs (Supplementary Table S2). In particular, 52 pairs

between *Arabidopsis* and cultivated strawberry and 57 between woodland strawberry and cultivated strawberry are highlighted in Figure 4.

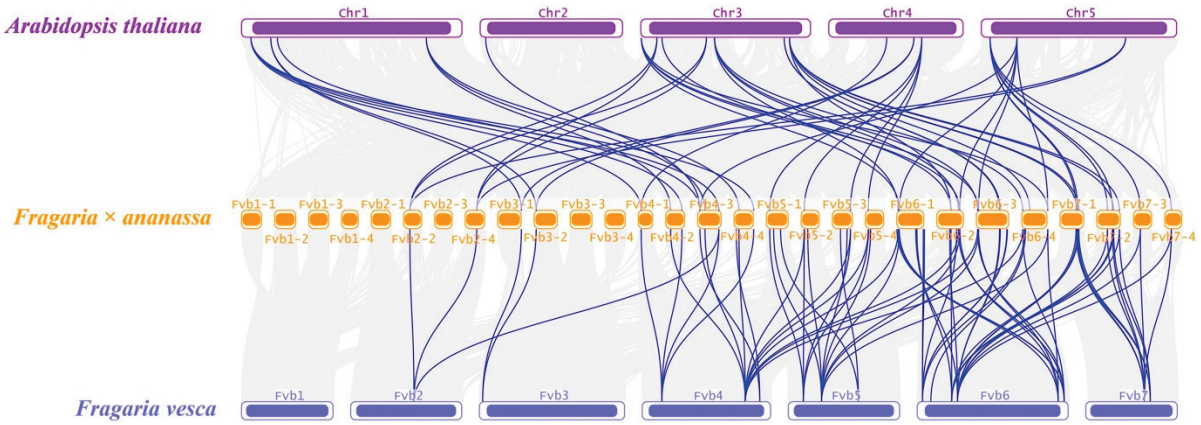


Figure 4. Collinearity analysis of *PL* genes among *Arabidopsis thaliana*, *Fragaria* × *ananassa*, and *Fragaria vesca* genomes. Grey lines indicate collinear blocks within the three genomes, while the blue lines represent collinear *PL* gene pairs. The purple, yellow, and blue columns indicate the chromosomes from *Arabidopsis thaliana*, *Fragaria* × *ananassa*, and *Fragaria vesca* genomes, respectively. Chromosome numbers are displayed at the side of chromosomes.

2.4. Expression Profiles of FaPLs during Fruit Development and Ripening

To identify the *FaPLs* related to strawberry fruit ripening, their expression patterns during fruit development were examined based on transcriptome data. As shown in Figure 5A, 52 out of 65 *FaPLs* distinctly expressed during the fruit development. Interestingly, most *FaPL* genes were highly expressed in the large green (LG) stage, while barely expressed in the partial red (PR) and full red (FR) stages. On the contrary, there were seven *FaPLs* (*FaPL1*, *FaPL3*, *FaPL5*, *FaPL20*, *FaPL25*, *FaPL42*, *FaPL57*) that exhibited lower expressions in the LG stage and gradually increased as the fruit ripened, indicating that they may be associated with strawberry fruit ripening. Subsequently, the expressions of these seven *FaPLs* were further assessed in fruit in comparison with firmness. As a result (Figure 5B), all of them had much higher expression levels in fruit with weak firmness than that in fruit with strong firmness, confirming their potential role in strawberry fruit softening. The transcriptome FPKM values are listed in Supplementary Table S3.

2.5. Functional Analysis of FaPL1 in Strawberry Fruit Softening

Among all the seven potential softening related *FaPLs*, *FaPL3*, *FaPL25*, *FaPL42*, and *FaPL57* were close to the previously reported *FaPLa*, *FaPL20* was close to the previously reported *FaPLb* (Figure 2), while *FaPL5* had the lowest expression (Figure 5B and Table S3). Therefore, the *FaPL1* was selected for further expression and functional analysis. The temporal and spatial expression analysis result (Figure 6A) revealed that *FaPL1* expressed in various tissues, with the lowest level in functional leaves (the fully expanded leaf) and the highest level in fruit. In addition, the expression of *FaPL1* gradually increased along with fruit development and ripening, which showed a complete negative correlation with fruit firmness (Figure 6B). These findings suggested that the *FaPL1* may play a vital role in strawberry fruit softening. Furthermore, *FaPL1* was transiently overexpressed in strawberry fruit to validate its function in softening. The phenotypic result showed that overexpression of *FaPL1* did not apparently affect the fruit skin color (Figure 6C). qRT-PCR analysis of *FaPL1* expression indicated a seven times higher level in the overexpressed sample compared to the control (Figure 6D), suggesting *FaPL1* was successfully overexpressed. Moreover, the fruit firmness was significantly decreased by overexpression of *FaPL1* (Figure 6E), confirming the important role of *FaPL1* in strawberry softening.

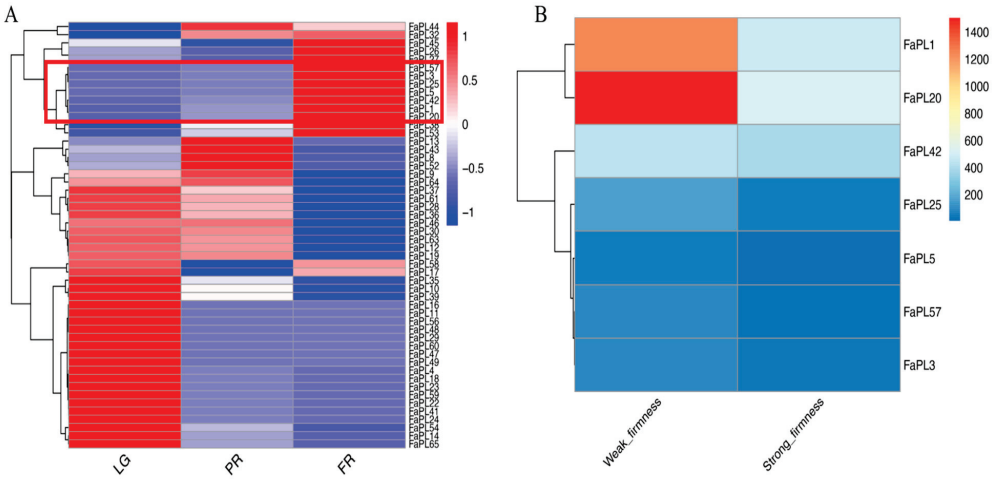


Figure 5. Heat maps showing the transcriptome-based expression of *FaPLs*. (**A**) The RNAseq retrieved expression during fruit development and ripening. A three-color scale was used with blue, white, and red indicating lowly, intermediately, and highly expressed genes, respectively. (**B**) Transcript abundance of *FaPL* genes in fruit with contrasting firmness. The color scale indicated the expression levels in FPKM value. LG, large green; PR, partial red; FR, full red. Red box indicates the seven *FaPL* members comprising gradual increases of expression during fruit development and ripening.

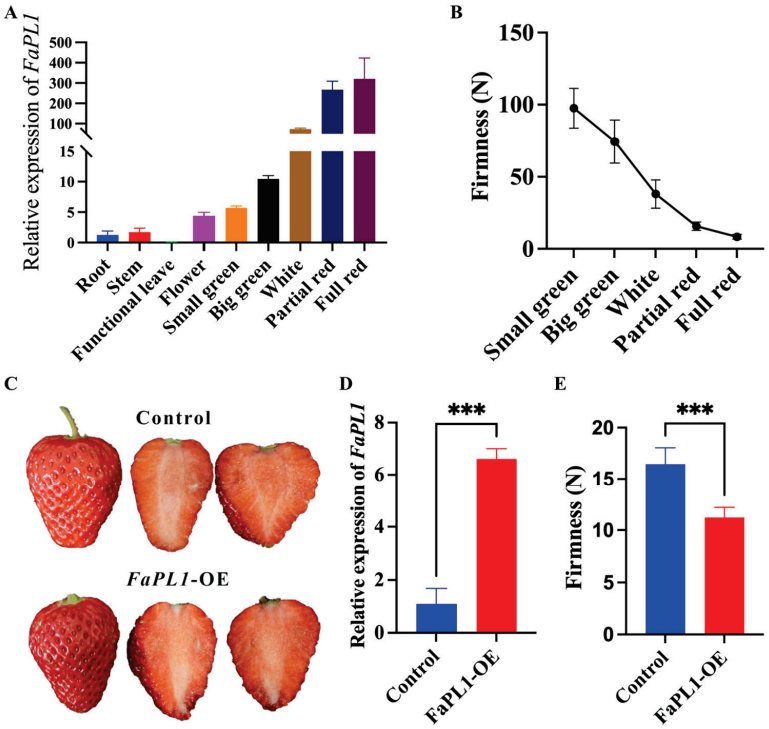


Figure 6. qRT-PCR based expression analysis and overexpression of *FaPL1* in strawberry. (**A**) The expression patterns of *FaPL1* in different tissues and during fruit development and ripening. (**B**) The change of fruit firmness of strawberry. (**C**) The phenotype of strawberry injected with empty (control) and *FaPL1*-overexpressing recombinant plasmid. (**D**) The relative expression of *FaPL1* in the full red *FaPL1*-overexpressed fruit and control fruit. (**E**) The firmness of full red *FaPL1*-overexpressed fruit and control fruit. OE, overexpressing. Triple asterisk indicated statistical difference at $p \leq 0.001$ level.

2.6. The Effects of FaPL1 Overexpression on Fruit-Ripening-Related Traits

According to the result, the content of total anthocyanins, titratable acidity (TA), and soluble protein was remarkable higher in *FaPL1*-overexpressed fruit than that in the control fruit (Figure 7A,C,E). By contrast, AsA and malondialdeehyde (MDA) contents exhibited

significantly lower levels in the *FaPL1*-overexpressed fruit compared to the control fruit (Figure 7D,H). However, the contents of soluble sugar, total flavonoid, and phenolic were similar in *FaPL1*-overexpressed fruit and the control fruit, which showed no obvious differences (Figure 7B,F,G).

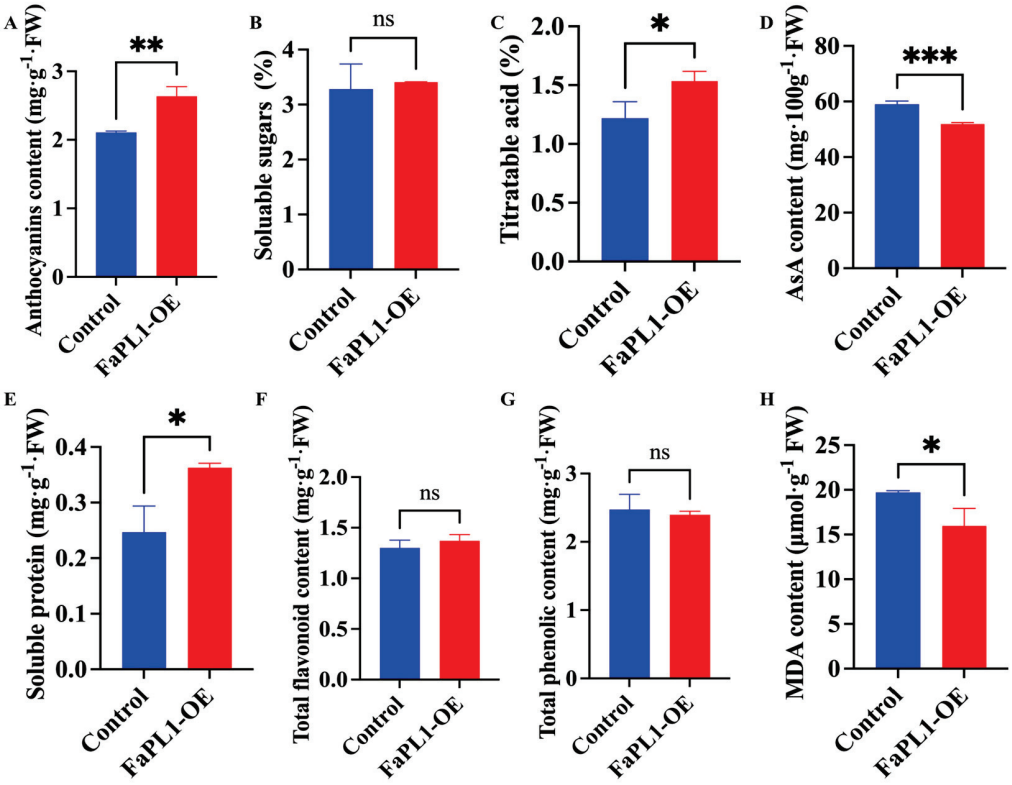


Figure 7. The effects of *FaPL1* overexpressing on the ripening-related traits. (**A**–**H**) indicate total anthocyanins content, soluble sugar, titratable acidity, AsA, soluble protein, total flavonoid, phenolic, and MDA content, respectively. Single, double, and triple asterisks indicate statistical differences at $p \leq 0.05$, 0.01, and 0.001 levels. ns, no significant difference was found.

3. Discussion

Due to their important roles in a broad range of physiological processes associated with pectin degradation, such as plant growth, development, fruit softening and ripening, PL genes have been identified in various plant species. It has been reported that a total of 26, 20, 12, 46, 22, and 16 *PL* family members were identified in *Arabidopsis* [24], peach [13], rice [25], *Brassica rapa* [26], tomato [19], and grape [14], respectively. In strawberry, several *FaPL* genes were obtained from multiple varieties, including *FaPLa*, *FaPLb*, and *FaPLc*, from 'Chandler', *FaSCPL* from 'Sweet Charlie', *FaTPL* from 'Toyonoka', and five subtype sequences from 'Elsanta' [20,21,27]. However, to our knowledge, the genome-wide analysis of this family remains limited. In this study, 65 *FaPL* genes were identified in strawberry based on a genome-wide investigation, which is more than the numbers in the above-mentioned species. It is possibly because the cultivated strawberry has undergone a whole genome duplication during the evolutionary process [28]. Segmental and tandem are two main duplication events driving the expansion of gene families [29]. For example, *PpePL5*, *6*, *7*, and *8* have been regarded as arising from tandem repeats [13], while the *GhPELs* in cotton seemed likely to be driven by segmental duplication [30]. Here, we have found that most (55 out of 65) *FaPL* genes were duplicated from segmental events (Table S1), which may contribute to the gene family expansion and their diverse structures and functions. These results indicated different expansion mechanisms of the *PL* gene family among different species. Moreover, the unrooted tree separated the *FaPL* genes into six different groups (Figure 3), which is different from the five groups from peach [13], tomato, [19] and

cotton [30]. This may be caused by the larger number of *PL* genes in strawberry. Except for Group II, the *FaPL* genes involved in the same group have similar exon–intron structures. Whereas, in Group II, a multiplicity of exon numbers was found (Figure 3), suggesting their probable functional differentiation.

Fruit ripening is a complex process that involves substantive alterations in gene expression resulting in changes in flavor, aroma, and texture. Being one of the important cell wall modification genes, the *PL* expression has been found to be related to fruit ripening in various species. For instance, the PL gene is mainly expressed in ripe fruit but not the unripe fruit of banana [31], the expression of five peach *PpePL* genes, and three strawberry *FaPL* genes accumulated during fruit ripening [13,20,22]. Consistent with the previous studies, we found that there were seven *FaPL* genes (*FaPL1, 3, 5, 20, 25, 42, and 57*) that have gradually increased expression patterns during fruit development and ripening (Figure 5). The predominant and high expression of *FaPL1* in fruit during ripening was also confirmed by a qRT-PCR experiment (Figure 6A), revealing these genes are associated with strawberry ripening. Additionally, it has been well documented that *PL* genes play a central role in fruit softening [13,14,19]. *FaPLa*, *FaPLb*, and *FaPLc* have been suggested to participate in strawberry softening. Silencing *FaPLc* resulted in 30% firmer fruit than the control [22]. Here, according to the phylogenetic tree (Figure 2), it was found that among the seven ripening related *FaPLs, FaPL3, 25, 42*, and *57* were classified into the same clade with *FaPLa, FaPL20* was clustered with *FaPLb, and FaPL1* was close to *FaPLc*. This result demonstrated these genes may have similar functions in strawberry softening. Furthermore, we found that all of the seven ripening related *FaPL* genes apparently had higher expression levels in the fruit with weak firmness, compared to the fruit with strong firmness (Figure 5B). Transient overexpression of *FaPL1* significantly decreased the strawberry firmness (Figure 6E), confirming its key role in strawberry softening. Notably, *FaPL5* was clustered with *AtPLL19* in Group I, which is different from the other six ripening related *FaPLs* (Figure 2). *AtPLL19* was identified by its xylem-specific expression in *Arabidopsis* [32]. Combined with the fact that *FaPL5* had the lowest expression among the seven ripening related *FaPLs* (Figure 5B), it can be speculated that *FaPL5* may mainly function in xylem vascular development rather than fruit softening, which needs further research.

Developing methods without influencing the edible and appealing aspects of fruit, including color, aroma, or nutritional value, has currently become the major goal for controlling softening [4]. It has been suggested that the antisense expression of the *FaPLc* gene in strawberry did not affect the fruit color [22]; while overexpression of *VvPL15* in tomato accelerated the fruit ripening and coloring [14]. In the present study, we found that overexpression of *FaPL1* significantly increased anthocyanin content (Figure 7A). The possible explanation is that *FaPL1* may have a similar function with *VvPL15* in promoting fruit ripening and coloring. This may also explain the decrease of AsA in *FaPL1*-overexpressed fruit (Figure 7D). Because it has been suggested that the AsA decreased during strawberry storage and senescence [33], the facilitation of ripening by overexpression of *FaPL1* may lead to fruit senescence faster than the control and thus contained a lower AsA content. In addition, we have also found that *FaPL1* overexpression increased that soluble protein content in strawberry (Figure 7E). This is probably because the overexpression of *FaPL1* caused the degradation of the cell wall, resulting in the release of proteins. Moreover, it has been reported that acidic pH can cause cell wall loosening by inducing *PL* expression [34]; however, how the overexpression of the *PL* gene increases the content of titratable acidity (Figure 7C) is still to be studied in the future.

4. Materials and Methods

4.1. Identification and Comprehensive Analysis of FaPL Genes

The genome of cultivated strawberry was downloaded from the GDR (Genome Database for Rosaceae) (https://www.rosaceae.org, accessed on 23 August 2023) [35]. The specific Hidden Markov Model (HMM) file for the PL conserved domain (PF00544) was downloaded from the Pfam database (https://www.ebi.ac.uk/interpro/, accessed

on 23 August 2023) [36] and used as a query to search the candidate *FaPL* genes by the HMMsearch program. The e value was set to 10^{-5}, and the other parameters were set as default. The sequences with complete PL domain were further confirmed by searching the NCBI conserved domain database [37]. The deduced amino acid number, molecular weight (MW), and isoelectric point (pI) of putative proteins were obtained using a perl script. The chromosome locations of *FaPLs* were retrieved from the genome annotation file; the conserved motifs were analyzed using the MEME suite online program (version 5.5.3) and visualized together with the gene structure using TBtoos software (version 2.001). The subcellular localization prediction was performed by WOLF PSORT program (https://wolfpsort.hgc.jp, accessed on 23 August 2023).

4.2. Phylogenetic and Evolutionary Analysis of FaPL Genes in Strawberry

Based on the multiple alignment of FaPL proteins obtained by the MUSCLE program, a phylogenetic tree was constructed by MEGA X software (version 10.1.8) using the maximum likelihood method [38]. The beautification of the tree was subsequently carried out using the iTol online tool (https://itol.embl.de/about.cgi, accessed on 23 August 2023) [39]. Duplication events and the collinear gene pairs were determined using MCScanX software (http://chibba.pgml.uga.edu/mcscan2/, accessed on 23 August 2023). All the analysis was conducted using the default parameters of specific software according to the user instructions.

4.3. Expression Analysis

The RNAseq-based expression profiles of *FaPLs* in different fruit developmental stages and in strawberry fruits with contrasting firmness were retrieved from the previously published transcriptome data PRJNA838938 (https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA838938, accessed on 23 August 2023) and PRJNA662854 (https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA662854, accessed on 23 August 2023), respectively. The expression level was represented by the FPKM values. The heatmap was created using "pheatmap" package of R software (version 4.2.2) with a normalization in row.

RT-qPCR-based expression analysis were carried out using SYBR Green Premix Ex Taq^{TM} (Takara, Tokyo, Japan) on a CFX96 RT-qPCR system (Bio-Rad, Hercules, CA, USA). The total RNA was extracted from the plant sample using the improved cetyltrimethylammonium bromide (CTAB) method. The strawberry tissues and fruit at different developmental stages were collected in the previous study [40]. The first strand cDNA was synthesized following the operating manual of PrimeScriptTM RT reagent Kit with gDNA Eraser (Takara, Tokyo, Japan). The relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method [41]. The 26-18S interspacer RNA sequence [42] was used as the internal reference. Expression data was reflected by mean± standard deviation (SD) of three independent biological replicates. Specific primers used for RT-qPCR were designed using the NCBI online tools. All the primer sequences are listed in Supplementary Table S4.

4.4. Transient Overexpression of FaPL1 Gene

The full length of *FaPL1* CDS was amplified using the primers *PL1*-TF and *PL1*-TR (Supplementary Table S4) according to the sequence retrieved from the strawberry genome and was substantially homologous recombined into a modified overexpression vector pCAMBIA1301 [42]. The recombinant plasmid was transformed into the strawberry cv. 'Benihoppe' fruit at the white stage using the previous agrobacterium-mediated transformation method [42]. The agrobacterium GV3101 strain was cultured at 28 °C until the OD_{600} reached 0.8. Each fruit was injected with 500 µL of bacterial solution and placed into a cultivation incubator. The fruit injected with empty vector was used as the control. At least 20 fruits were injected for overexpression and the control group separately. The injection fruit side was samples after 7 days for further measurement.

4.5. Determination of Fruit Firmness, Soluble Sugar and TA

Fruit firmness was determined two times on each injection part side of the fruit by a Texture Analyzer TA XT2i (Stable Micro systems, Godalming, Surrey, UK) with a 5 mm diameter cylinder needle and a penetration depth of 10 mm. Firmness was expressed as newton (N). Soluble sugar content was measured by the previously described colorimetric method [43]. Around 0.1 g of frozen stored fruit was completely extracted in 1 mL distilled water. After that, the 250 µL of the extract was diluted into 750 µL distilled water and 250 µL 2% (*w*/*v*) anthrone-ethyl acetate. The mixed solution was subsequently added to 2.5 mL concentrated sulfuric acid and put in a boiled water bath for 1 min. After cooling it down to room temperature, the absorbance of the extraction solution was recorded at 620 nm using a spectrophotometer, and the soluble sugars content was quantified by comparison to an external standard. The TA content was estimated by titrating the fruit extract against 0.1 N sodium hydroxide (NaOH) to the end point of pH 8.2 (faint pink) and represented as citric acid percentage.

4.6. MDA and Soluble Proteins

The MDA was assayed according to the formerly described procedure with slight modification [43]. Briefly, 0.5 g of frozen fruit sample was completely homogenized with 10% trichloroacetic acid. After a 10 min centrifugation at 4 °C, the clear solution was mixed with 0.67% 2-thiobarbituric acid. The mixture was then placed into a water bath at 100 °C for 10 min and immediately cooled on ice. The absorption values at 450 nm, 523 nm, and 600 nm were read separately. The results were represented as µmol per g FW.

The soluble protein content was measured according to the previous study. In brief, 0.5 g of fruit sample was homogenized in 5 mL of distilled water. The upper phase was collected and added with CBBG. After centrifugation, the absorbance of the mixture was tested at 595 nm. The content of soluble protein was quantified by a standard curve constructed using bovine serum albumin (BSA) protein.

4.7. Total Flavonoid, Phenolic, Anthocyanin, and AsA Content

Based on the previously described procedure [44], approximately 3 g of fruit were extracted in 5 mL of 80% acetone for 1 h at room temperature. After centrifugation for 10 min at 4500 rpm, the supernatant was collected for total flavonoids and phenolic content measurement. The photographic densities of 415 nm and 650 nm were read for the calculations of total flavonoids and phenolic content, respectively. The quercetin and gallic acid were used as external standards to construct the calibration curves separately. The total flavonoid content was presented as mg quercetin per kg of FW, and the total phenolic content was expressed as g gallic acid per kg of FW.

The determination of total anthocyanins was performed by pH differential method [45]. As previously demonstrated, the fruit sample was extracted in an acetic acid: water: acetone: methanol (1:2:4:4) solution. The mixture was incubated at room temperature for 30 min and then placed into a 40 °C water bath for 4 h. The clear extract was added with KCl (0.025 M, pH 1.0) and sodium acetate and then detected by recording the absorption values of 496 and 700 nm. The content of total anthocyanins was expressed as g pelargonidin 3-glucoside (Pg3G) per kg of FW.

AsA content was detected following the procedure described by Jiang et al. [43]. The content of AsA was calculated using the photographic density of the fruit extract at 534 nm and expressed as g AsA per kg of FW.

4.8. Statistical Analysis

All experiments were carried out in three independent biological replicates. Experimental data were expressed as mean values ± SD. The statistical differences were analyzed using Prism 9 software. The differences between the overexpression and control groups were determined using *t*-test. The results with p value below 0.05 were considered as statistically significantly different.

5. Conclusions

To summarize, 65 *FaPLs* gene family members were identified in strawberry and characterized. Among those, *FaPL1*, *3*, *5*, *20*, *25*, *42*, and *57* are likely to function in strawberry softening due to their increasing expression during fruit development and ripening and higher expression in weak firmness fruit. Transient overexpression of *FaPL1* significantly reduced the fruit firmness, confirming its role in strawberry softening. This work provides a basis for better understanding the function of the *FaPL* gene family in fruit ripening and softening.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms241713217/s1.

Author Contributions: Conceptualization, Y.L. (Yuanxiu Lin) and H.T.; methodology, H.H., Y.W. (Yanling Wen) and Z.W.; software, Q.C., M.L. and W.H.; validation, S.C. and Z.S.; formal analysis, Y.Z. (Yunting Zhang); resources, H.T. and Y.L. (Ya Luo); writing—original draft preparation, Y.L. (Yuanxiu Lin); writing—review and editing, Y.W. (Yan Wang), Y.Z. (Yong Zhang) and X.W.; supervision, H.T.; funding acquisition, Y.L. (Yuanxiu Lin). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Natural Science Foundation of Sichuan Province (2023NSFSC1239), and National Students Innovation Training Project of Sichuan Agricultural University (202210626049).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Li, B.J.; Grierson, D.; Shi, Y.; Chen, K.S. Roles of abscisic acid in regulating ripening and quality of strawberry, a model non-climacteric fruit. *Hortic. Res.* **2022**, *9*, uhac089. [CrossRef] [PubMed]
2. Li, S.; Chen, K.; Grierson, D. Molecular and Hormonal Mechanisms Regulating Fleshy Fruit Ripening. *Cells* **2021**, *10*, 1136. [CrossRef] [PubMed]
3. Shiga, T.M.; Soares, C.A.; Nascimento, J.R.; Purgatto, E.; Lajolo, F.M.; Cordenunsi, B.R. Ripening-associated changes in the amounts of starch and non-starch polysaccharides and their contributions to fruit softening in three banana cultivars. *J. Sci. Food Agric.* **2011**, *91*, 1511–1516. [CrossRef]
4. Wang, D.; Yeats, T.H.; Uluisik, S.; Rose, J.K.C.; Seymour, G.B. Fruit Softening: Revisiting the Role of Pectin. *Trends Plant Sci.* **2018**, *23*, 302–310. [CrossRef] [PubMed]
5. Mercado, J.A.; Pliego-Alfaro, F.; Quesada, M.A. Fruit Shelf Life and Potential for Its Genetic Improvement. In *Breeding for Fruit Quanlity*; Matthew, A., Bebeli, P.J., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2011; pp. 81–104.
6. Paniagua, C.; Ric-Varas, P.; García-Gago, J.A.; López-Casado, G.; Blanco-Portales, R.; Muñoz-Blanco, J.; Schückel, J.; Knox, J.P.; Matas, A.J.; Quesada, M.A.; et al. Elucidating the role of polygalacturonase genes in strawberry fruit softening. *J. Exp. Bot.* **2020**, *71*, 7103–7117. [CrossRef] [PubMed]
7. García-Gago, J.A.; Posé, S.; Muñoz-Blanco, J.; Quesada, M.A.; Mercado, J.A. The polygalacturonase *FaPG1* gene plays a key role in strawberry fruit softening. *Plant Signal Behav.* **2009**, *4*, 766–768. [CrossRef]
8. Quesada, M.A.; Blanco-Portales, R.; Posé, S.; García-Gago, J.A.; Jiménez-Bermúdez, S.; Muñoz-Serrano, A.; Caballero, J.L.; Pliego-Alfaro, F.; Mercado, J.A.; Muñoz-Blanco, J. Antisense down-regulation of the *FaPG1* gene reveals an unexpected central role for polygalacturonase in strawberry fruit softening. *Plant Physiol.* **2009**, *150*, 1022–1032. [CrossRef]
9. Xue, C.; Guan, S.C.; Chen, J.Q.; Wen, C.J.; Cai, J.F.; Chen, X. Genome wide identification and functional characterization of strawberry pectin methylesterases related to fruit softening. *BMC Plant Biol.* **2020**, *20*, 13. [CrossRef]
10. Cai, J.; Mo, X.; Wen, C.; Gao, Z.; Chen, X.; Xue, C. FvMYB79 Positively Regulates Strawberry Fruit Softening via Transcriptional Activation of FvPME38. *Int. J. Mol. Sci.* **2021**, *23*, 101. [CrossRef]
11. Ke, X.; Wang, H.; Li, Y.; Zhu, B.; Zang, Y.; He, Y.; Cao, J.; Zhu, Z.; Yu, Y. Genome-Wide Identification and Analysis of Polygalacturonase Genes in Solanum lycopersicum. *Int. J. Mol. Sci.* **2018**, *19*, 2290. [CrossRef]
12. Cao, J. The pectin lyases in Arabidopsis thaliana: Evolution, selection and expression profiles. *PLoS ONE* **2012**, *7*, e46944. [CrossRef] [PubMed]

13. Xu, Z.; Dai, J.; Kang, T.; Shah, K.; Li, Q.; Liu, K.; Xing, L.; Ma, J.; Zhang, D.; Zhao, C. PpePL1 and PpePL15 Are the Core Members of the Pectate Lyase Gene Family Involved in Peach Fruit Ripening and Softening. *Front. Plant Sci.* **2022**, *13*, 844055. [CrossRef] [PubMed]
14. Ma, Y.; Wang, C.; Gao, Z.; Yao, Y.; Kang, H.; Du, Y. VvPL15 Is the Core Member of the Pectate Lyase Gene Family Involved in Grape Berries Ripening and Softening. *Int. J. Mol. Sci.* **2023**, *24*, 9318. [CrossRef] [PubMed]
15. Pua, E.C.; Ong, C.K.; Liu, P.; Liu, J.Z. Isolation and expression of two pectate lyase genes during fruit ripening of banana (*Musa acuminata*). *Physiol. Plant* **2001**, *113*, 92–99. [CrossRef]
16. Uluisik, S.; Chapman, N.H.; Smith, R.; Poole, M.; Adams, G.; Gillis, R.B.; Besong, T.M.D.; Sheldon, J.; Stiegelmeyer, S.; Perez, L.; et al. Genetic improvement of tomato by targeted control of fruit softening. *Nat. Biotechnol.* **2016**, *34*, 950–952. [CrossRef] [PubMed]
17. Chourasia, A.; Sane, V.A.; Nath, P. Differential expression of pectate lyase during ethylene-induced postharvest softening of mango (*Mangifera indica* var. Dashehari). *Physiol. Plant* **2006**, *128*, 546–555. [CrossRef]
18. Nunan, K.J.; Davies, C.; Robinson, S.P.; Fincher, G.B. Expression patterns of cell wall-modifying enzymes during grape berry development. *Planta* **2001**, *214*, 257–264. [CrossRef]
19. Yang, L.; Huang, W.; Xiong, F.; Xian, Z.; Su, D.; Ren, M.; Li, Z. Silencing of SlPL, which encodes a pectate lyase in tomato, confers enhanced fruit firmness, prolonged shelf-life and reduced susceptibility to grey mould. *Plant Biotechnol. J.* **2017**, *15*, 1544–1555. [CrossRef]
20. Benítez-Burraco, A.; Blanco-Portales, R.; Redondo-Nevado, J.; Bellido, M.L.; Moyano, E.; Caballero, J.L.; Muñoz-Blanco, J. Cloning and characterization of two ripening-related strawberry (*Fragaria* × *ananassa* cv. Chandler) pectate lyase genes. *J. Exp. Bot.* **2003**, *54*, 633–645. [CrossRef]
21. Severo, J.; Tiecher, A.; Chaves, F.C.; Silva, J.A.; Rombaldi, C.V. Gene transcript accumulation associated with physiological and chemical changes during developmental stages of strawberry cv. Camarosa. *Food Chem.* **2011**, *126*, 995–1000. [CrossRef]
22. Jiménez-Bermúdez, S.; Redondo-Nevado, J.; Muñoz-Blanco, J.; Caballero, J.L.; López-Aranda, J.M.; Valpuesta, V.; Pliego-Alfaro, F.; Quesada, M.A.; Mercado, J.A. Manipulation of strawberry fruit softening by antisense expression of a pectate lyase gene. *Plant Physiol.* **2002**, *128*, 751–759. [CrossRef] [PubMed]
23. Santiago-Doménech, N.; Jiménez-Bemúdez, S.; Matas, A.J.; Rose, J.K.C.; Muñoz-Blanco, J.; Mercado, J.A.; Quesada, M.A. Antisense inhibition of a pectate lyase gene supports a role for pectin depolymerization in strawberry fruit softening. *J. Exp. Bot.* **2008**, *59*, 2769–2779. [CrossRef] [PubMed]
24. Palusa, S.G.; Golovkin, M.; Shin, S.B.; Richardson, D.N.; Reddy, A.S.N. Organ-specific, developmental, hormonal and stress regulation of expression of putative pectate lyase genes in Arabidopsis. *New Phytol.* **2007**, *174*, 537–550. [CrossRef] [PubMed]
25. Zheng, Y.; Yan, J.; Wang, S.; Xu, M.; Huang, K.; Chen, G.; Ding, Y. Genome-wide identification of the pectate lyase-like (PLL) gene family and functional analysis of two PLL genes in rice. *Mol. Genet. Genomics* **2018**, *293*, 1317–1331. [CrossRef] [PubMed]
26. Jiang, J.; Yao, L.; Miao, Y.; Cao, J. Genome-wide characterization of the Pectate Lyase-like (PLL) genes in Brassica rapa. *Mol. Genet. Genom.* **2013**, *288*, 601–614. [CrossRef]
27. Zhou, H.; Li, G.; Zhao, X. Comparative analysis of pectate lyase in relation to softening in strawberry fruits. *Can. J. Plant Sci.* **2016**, *96*, 604–612. [CrossRef]
28. Tennessen, J.A.; Govindarajulu, R.; Ashman, T.-L.; Liston, A. Evolutionary Origins and Dynamics of Octoploid Strawberry Subgenomes Revealed by Dense Targeted Capture Linkage Maps. *Genome Biol. Evol.* **2014**, *6*, 3295–3313. [CrossRef]
29. Cannon, S.B.; Mitra, A.; Baumgarten, A.; Young, N.D.; May, G. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. *BMC Plant Biol.* **2004**, *4*, 10. [CrossRef]
30. Sun, H.; Hao, P.; Ma, Q.; Zhang, M.; Qin, Y.; Wei, H.; Su, J.; Wang, H.; Gu, L.; Wang, N.; et al. Genome-wide identification and expression analyses of the pectate lyase (PEL) gene family in cotton (*Gossypium hirsutum* L.). *BMC Genom.* **2018**, *19*, 661. [CrossRef]
31. Marín-Rodríguez, M.C.; Smith, D.L.; Manning, K.; Orchard, J.; Seymour, G.B. Pectate lyase gene expression and enzyme activity in ripening banana fruit. *Plant Mol. Biol.* **2003**, *51*, 851–857. [CrossRef]
32. Bai, Y.; Tian, D.; Chen, P.; Wu, D.; Du, K.; Zheng, B.; Shi, X. A Pectate Lyase Gene Plays a Critical Role in Xylem Vascular Development in *Arabidopsis*. *Int. J. Mol. Sci.* **2023**, *24*, 10883. [CrossRef] [PubMed]
33. Mishra, R.; Kar, A. Effect of storage on the physicochemical and flavour attributes of two cultivars of strawberry cultivated in Northern India. *Sci. World JL* **2014**, *2014*, 794926. [CrossRef] [PubMed]
34. Phyo, P.; Gu, Y.; Hong, M. Impact of acidic pH on plant cell wall polysaccharide structure and dynamics: Insights into the mechanism of acid growth in plants from solid-state NMR. *Cellulose* **2019**, *26*, 291–304. [CrossRef]
35. Jung, S.; Lee, T.; Cheng, C.H.; Buble, K.; Zheng, P.; Yu, J.; Humann, J.; Ficklin, S.P.; Gasic, K.; Scott, K.; et al. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. *Nucleic Acids Res.* **2019**, *41*, D1137–D1145. [CrossRef] [PubMed]
36. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangradorvegas, A. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [CrossRef] [PubMed]
37. Aron, M.B.; Shennan, L.; Anderson, J.B.; Farideh, C.; Derbyshire, M.K.; Carol, D.W.S.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R. CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **2011**, *39*, 225–229.

38. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
39. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]
40. Yue, M.; Jiang, L.; Zhang, N.; Zhang, L.; Liu, Y.; Wang, Y.; Li, M.; Lin, Y.; Zhang, Y.; Zhang, Y.; et al. Importance of FaWRKY71 in Strawberry (*Fragaria ananassa*) Fruit Ripening. *Int. J. Mol. Sci.* **2022**, *23*, 12483. [CrossRef]
41. 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]
42. Jiang, L.; Yue, M.; Liu, Y.; Zhang, N.; Lin, Y.; Zhang, Y.; Wang, Y.; Li, M.; Luo, Y.; Zhang, Y.; et al. A novel R2R3-MYB transcription factor FaMYB5 positively regulates anthocyanin and proanthocyanidin biosynthesis in cultivated strawberries (*Fragaria* × *ananassa*). *Plant Biotechnol. J.* **2023**, *21*, 1140–1158. [CrossRef] [PubMed]
43. Jiang, L.; Chen, X.; Gu, X.; Deng, M.; Li, X.; Zhou, A.; Suo, M.; Gao, W.; Lin, Y.; Wang, Y.; et al. Light Quality and Sucrose-Regulated Detached Ripening of Strawberry with Possible Involvement of Abscisic Acid and Auxin Signaling. *Int. J. Mol. Sci.* **2023**, *24*, 5681. [CrossRef] [PubMed]
44. Zhang, Y.; Li, S.; Deng, M.; Gui, R.; Liu, Y.; Chen, X.; Lin, Y.; Li, M.; Wang, Y.; He, W.; et al. Blue light combined with salicylic acid treatment maintained the postharvest quality of strawberry fruit during refrigerated storage. *Food Chem. X* **2022**, *15*, 100384. [CrossRef]
45. Zhang, Y.; Hu, W.; Peng, X.; Sun, B.; Wang, X.; Tang, H. Characterization of anthocyanin and proanthocyanidin biosynthesis in two strawberry genotypes during fruit development in response to different light qualities. *J. Photochem. Photobiol. B* **2018**, *186*, 225–231. [CrossRef] [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Genome-Wide Identification of Catalase Gene Family and the Function of *SmCAT4* in Eggplant Response to Salt Stress

Lei Shen †, Xin Xia †, Longhao Zhang, Shixin Yang and Xu Yang *

College of Horticulture and Landscape Architecture, Yangzhou University, Yangzhou 225009, China; shenlei07@yzu.edu.cn (L.S.); mz120221425@stu.yzu.edu.cn (X.X.); mz120231508@stu.yzu.edu.cn (L.Z.); mz120221438@stu.yzu.edu.cn (S.Y.)

* Correspondence: yangxu@yzu.edu.cn

† These authors contributed equally to this work.

Abstract: Salinity is an important abiotic stress, damaging plant tissues by causing a burst of reactive oxygen species (ROS). Catalase (CAT) enzyme coded by *Catalase (CAT)* genes are potent in reducing harmful ROS and hydrogen peroxide (H_2O_2) produced. Herein, we performed bioinformatics and functional characterization of four *SmCAT* genes, retrieved from the eggplant genome database. Evolutionary analysis *CAT* genes revealed that they are divided into subgroups I and II. The RT-qPCR analysis of *SmCAT* displayed a differential expression pattern in response to abiotic stresses. All the CAT proteins of eggplant were localized in the peroxisome, except for SmCAT4, which localized in the cytomembrane and nucleus. Silencing of *SmCAT4* compromised the tolerance of eggplant to salt stress. Suppressed expression levels of salt stress defense related genes *SmTAS14* and *SmDHN1*, as well as increase of H_2O_2 content and decrease of CAT enzyme activity was observed in the *SmCAT4* silenced eggplants. Our data provided insightful knowledge of *CAT* gene family in eggplant. Positive regulation of eggplant response to salinity by *SmCAT4* provides resource for future breeding programs.

Keywords: eggplant (*Solanum melongena*); catalase; *SmCAT4*; salt stress

1. Introduction

Due to their sessile nature, plants must endure various kinds of stresses such as heat [1,2], cold [3–5], salt [6,7], drought [8], heavy metal [9], insect infestation [10,11], and pathogens [12,13]. Reactive oxygen species (ROS) including hydrogen peroxide (H_2O_2), oxygen free radical or superoxide (O^{2-}), hydrogen free radical (·OH), and non-radical singlet oxygen ($^1O_2$) act as signaling molecules that modulate plant homeostatic mechanism [14,15]. However, excessive accumulation of ROS can damage cells, eventually affect growth, development, or, in severe cases, plant death. Plant have evolved sophisticated mechanisms such as enzymatic or non-enzymatic detoxification systems to alleviate the damage of ROS. The antioxidant enzyme machinery including peroxidase (POD), catalase (CAT), superoxide dismutase, ascorbate peroxidase (APX), glutathione S transferase (GST), glutathione peroxide (GPX), dehydroascorbate reductase (DHAR), and glutathione reductase (GR) [16–19]. The POD, CAT, APX, and SOD act as a protector by scavenging H_2O_2, and enhance plant immunity against abiotic and biotic stresses [16].

CAT converging with other components of antioxidant machinery is key in plant stress biology [20]. CAT can effectively scavenge H_2O_2, further relieving the damaging effects of oxidative stress. CAT is a tetrameric heme protein composed of four subunits. They are mainly distributed in the peroxisome, glyoxylate circulator, and cytoplasm Apart from that, a small amount of distribution in mitochondria and chloroplast was also recorded [21]. CATs are encoded by multiple genes with numbers varying across different species. For instance, 3 CATs in *Arabidopsis thaliana* [22], 3 CATs in rice (*Oryza sativa*) [23], 7 CATs in cotton (*Gossypium hirsutum*) [24], 10 CATs in wheat (*Triticum aestivum*) [20], and 14 CATs in

Brassica napus [25] were mapped. CATs regulated all aspects of plant growth and development by maintaining intracellular redox homeostasis [26]. ABA INSENSITIVE 5 (ABI5) activated *AtCAT1* expression by the directly binding to the promoter of *AtCAT1* to affect the ROS homeostasis, and then promoted the seed germination in Arabidopsis [27]. The hot pepper (*Capsicum annuum*) *CaCAT1* and *CaCAT2* were involved in the modulation of circadian rhythms and their expression was tightly organ-specific [28]. In maize, the P31 protein of chlorotic mottle virus (MCMV) interacts with ZmCAT1 and the P31-ZmCAT1 cascade further suppresses the expression of *PR1* gene thus enhancing the accumulation of viral load. Similarly, the interaction of *Phytophthora sojae* effector PsAvh113 with transcription factor GmDPB decreases the expression level of *GmCAT1*, thereby decreasing plant resistance to *Phytophthora* [29]. In rice, E3 ligase APIP6 degraded catalase OsCATC to negatively regulate rice innate immunity against blast fungus *Magnaporthe oryzae* [30]. These findings indicate that CATs play important roles in plants against various pathogenic microorganism. CAT facilitation of plant immunity against abiotic stresses are well studied. Induced expression of *MeCu/ZnSOD* and *MeCAT1* enhanced the tolerance of cassava against cold and drought stresses [31]. WD40 protein TaWD40-4b.1C interacts with TaCAT1A and TaCAT3A, promote their oligomerization and enzyme activities. This interaction amplifies the wheat response to drought stress by reducing H_2O_2 level [32]. Although many studies reported that CATs play an important role in plant response to abiotic stresses, the functions and regulatory mechanism of CATs remain unclear.

Eggplant (*Solanum melongena*) is a popular *Solanaceae* vegetable cultivated all around the world [33]. The growth, development, and yield of eggplant were vulnerable to various environmental stresses such as salt, drought, high temperature, and cold stress. The continuous environmental shift encourages the breeding of new eggplant varieties with high tolerance to multiple abiotic stresses. In the present study, a comprehensive genome-wide analysis of eggplant *CAT* gene family was performed. A total of four *CAT* genes were isolated from the eggplant genome database. A detailed transcript expression analysis of *CATs* in eggplant was carried out. Virus induced gene silencing (VIGS) assay showed that *SmCAT4* plays a positive role in eggplant response to salt stress.

2. Results

2.1. Identification and Physicochemical Property Analysis of Catalase Proteins in Eggplant

According to the amino acid sequences of Arabidopsis *CAT* gene family members, we identified four *CAT* genes (*SmCAT1*, *SmCAT2*, *SmCAT3*, and *SmCAT4*) in the eggplant genome by using TBtools software (version 2.019) and the NCBI website. We further analyzed the physicochemical properties of SmCATs (Table 1). *SmCAT1-3* encoded 492 amino acids with 1479 bp, while *SmCAT4* encoded 491 amino acids with 1476 bp. The molecular weight, instability index, aliphatic index, and average hydropathicity of SmCAT1-4 are 56,545~56,996 Da, 38.18~40.25, 68.94~71.91, and −0.582~−0.515, respectively. We also found that the theoretical pI and grand average of hydropathicity of SmCAT1-4 proteins are 6.80~7.31 and −0.582~−0.515, respectively, implying that SmCAT1-4 proteins are hydrophilic non-transmembrane proteins. In addition, the prediction results of subcellular localization showed that SmCAT1-4 proteins are located in the peroxisome.

Table 1. Physicochemical properties of CAT proteins of *Solanum melongena*.

Gene Name	Base Number	Amino Acid Number	Molecular Weight	Instability Index	Aliphatic Index	Grand Average of Hydropathicity	Average of Hydropathicity	Theoretical pI	Subcellular Localization
SmCAT1	1479	492	56,594.07	39.78	71.91	−0.515	−0.515	6.86	Peroxisome
SmCAT2	1479	492	56,996.23	40.25	68.94	−0.582	−0.582	6.88	Peroxisome
SmCAT3	1479	492	56,934.43	38.18	70.35	−0.552	−0.552	6.80	Peroxisome
SmCAT4	1476	491	56,545.00	38.37	71.87	−0.575	−0.575	7.31	Peroxisome

2.2. Phylogenic Analysis of Catalase Proteins

We constructed the phylogenetic tree by using MEGA 7.0 software (version 7.0.26) to investigate the phylogenetic relationships of the *CAT* gene family members in eggplant with their homologs from Arabidopsis, tomato (*Solanum lycopersicum*), and soybean (*Glycine max*). All *CAT* genes in the phylogenetic tree are divided into subgroups I and II (Figure 1). The subgroup I consists of four *CATs* (*OsCATA-D*) of rice, *AtCAT3*, and three *SmCATs* (*SmCAT1*, *SmCAT2*, *SmCAT4*). The subgroup II consists of four *CATs* (*GmCAT1/2-5*) of soybean, *SmCAT3*, and *AtCAT1* and *AtCAT*2. In subgroup I, four soybean *CATs* were grouped into the same subgroup as the two members *AtCAT1* and *AtCAT*2 of Arabidopsis, implying that their protein sequences have high similarity. Similarly, four *OsCATs* were divided into subgroup II with three *SmCATs*, suggesting that the protein sequences of *OsCATs* have high similarity with the *SmCATs*.

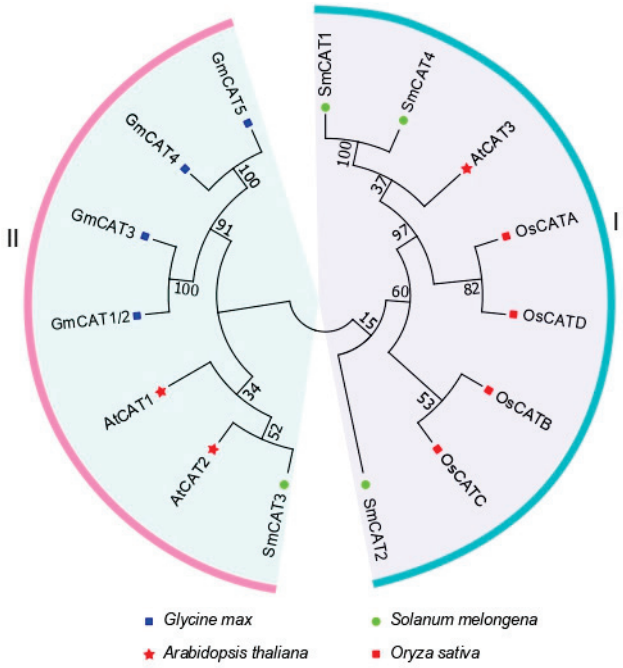


Figure 1. Analysis of phylogenetic relationship of *CAT* gene family members. Phylogenetic tree was generated to analyze the phylogenetic relationship between eggplant *CAT* gene family members and the *CAT* members from *Arabidopsis thaliana*, rice (*Oryza sativa*), and soybean (*Glycine max*) by MEGA 7.0 software.

2.3. Sequence Analysis of CAT Genes in Eggplant

The visualization of the chromosome location of eggplant *CATs* was performed using TBtools software (version 2.019) [34]. The *SmCAT* genes are scattered over chromosomes 02, 04, and 05. We also found that *SmCAT1* and *SmCAT4* were located very close to each other on chromosome 5 (Figure 2a). We next investigated the gene structures of *CATs* in eggplant by identifying their exon-intron orientation. *SmCAT1-3* has eight exons, one unique 5′-UTR and 3′-UTR regions, while *SmCAT4* has seven exons but no 5′-UTR and 3′-UTR regions. The size of *SmCTAs* is about 4~5 kb except for *SmCAT4*, which is close to 2 kb (Figure 2b). The conserved domain of SmCATs was predicted by the SMART website and visualized by DOG 2.0 software. All SmCAT proteins have one Catalase and Catalase-rel (catalase-related immune response) domain, and the locations of these two domains within four SmCAT proteins are very close to each other (Figure 2c). The motifs in SmCAT proteins were analyzed by searching the MEME website. The result showed that all SmCAT proteins contain 10 identical motifs. The majority of them are at least 50 amino acids in length, except for motifs 9 and 10, which have lower sizes (Figure 2d). We further predicted the *cis*-elements within the promoters of *SmCATs* by searching the PlantCARE website and

visualizing by TBtools software. The observed *cis*-elements include light responsiveness or unknown functions. Multiple phytohormone responsive *cis*-elements, such as ethylene (ET) response element (ERE), salicylic acid (SA) response element (TCA-element), jasmonic acid methyl ester (MeJA) response element (TGACG-motif), and ABA response element (ABRE) were identified. Important transcription factor binding sites such as WRKY transcription factor binding site (W-box), MYB binding site (MYB), MYC binding site (MYC), and bZIP binding site (A-box and G-box) as well as stress responsive *cis*-elements including low-temperature response element (LTR) and anaerobic induction element (ARE) were harbored in the promoter of these four *SmCATs* (Figure 2e). In addition, the tertiary structures of four SmCAT proteins were analyzed by searching the SWISS-MODEL website. We found similarities between SmCAT1, SmCAT3, SmCAT4, and A0A5J5B3V7.1.A catalase model of 80.08, 90.85, and 77%, respectively. The similarity between SmCAT2 and B9S6U0.1.A catalase model is 90.2% (Figure 2f).

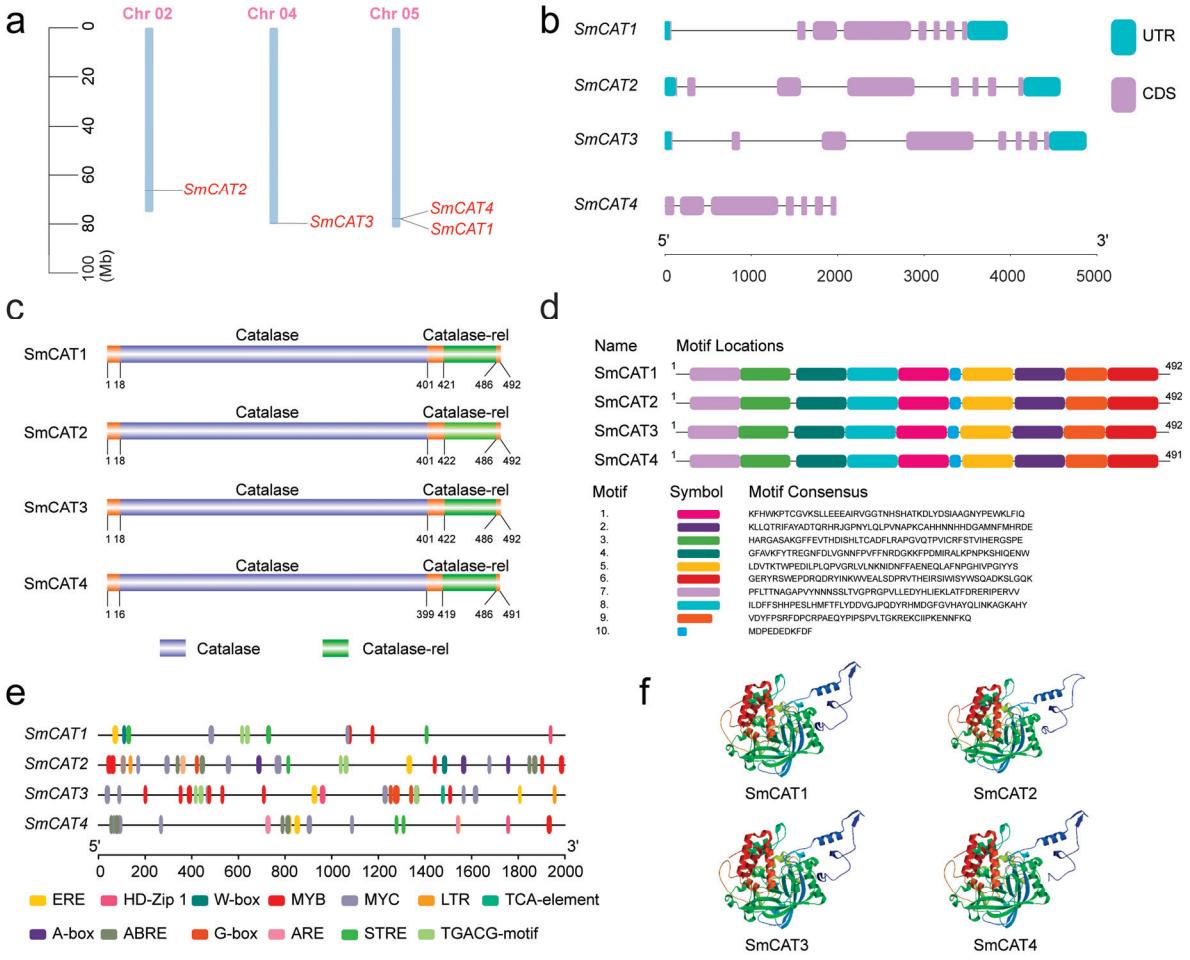


Figure 2. Analysis of chromosomal assignment, gene structures, conserved domain and motifs, *cis*-elements, and tertiary structure of *catalase* gene family members in eggplant. (**a**) Distribution of eggplant *CAT* genes on chromosomes. (**b**) The gene structure of eggplant *CAT* genes predicted by TBtools software. UTR, untranslated region; CDS, coding sequence. (**c**) The conserved domains of eggplant CAT proteins predicted by SMART website. (**d**) Analysis of conserved motifs of eggplant CAT proteins predicted by MEME website. (**e**) Analysis of *cis* elements within the promoters of eggplant *CTA* genes by searching PlantCARE website. (**f**) The tertiary structure of eggplant CAT proteins including SmCAT1, SmCAT2, SmCAT3, and SmCAT4 by searching the SWISS-MODEL website.

2.4. Collinearity Analysis of Eggplant CAT Genes and Their Homologs from Arabidopsis and Tomato

Tandem duplications and segmental duplications significantly contributed to the expansion of gene families [35]. So, we analyzed the gene collinearity relationship of *SmCAT* genes and their homologs from Arabidopsis and tomato by using TBtools software. Two duplicated segments (*SmCAT4* and *SmCAT3*, *SmCAT3*, and *SmCAT2*) were identified in the eggplant genome (Figure 3a). Additionally, we found that tandem duplication events of *CAT* gene family members did not occur in the eggplant genome. To further explore the evolutionary relationship of the *CAT* genes between eggplant, Arabidopsis, and tomato, we investigated the collinearity of *SmCATs* with that of Arabidopsis and tomato. We found that *SmCAT2* showed a synteny relationship with *AtCAT1* of Arabidopsis and *SlCAT2* and *SlCAT3* of tomato, respectively. Meanwhile, both *SmCAT3* and *SmCAT4* exhibited a synteny relationship with *AtCAT1* of Arabidopsis and *SlCAT1*, *SlCAT2*, and *SlCAT3* of tomato, respectively (Figure 3b).

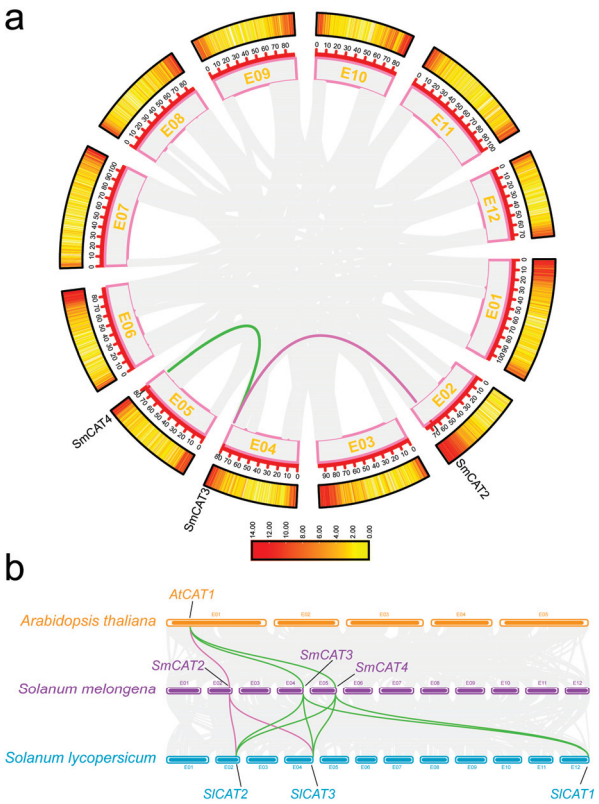


Figure 3. The collinearity analysis among the eggplant *CAT* genes (**a**) and of eggplant CAT genes with *Arabidopsis thaliana CAT* genes and tomato *CAT* genes (**b**). In (**a**), the gray lines represent all co-linear gene pairs in eggplant, and the green and pinkish red lines respectively indicate co-linear gene pairs *SmCAT3* and *SmCAT4* as well as *SmCAT2* and *SmCAT3*. In (**b**), the gray lines represent all co-linear gene pairs in eggplant, *Arabidopsis thaliana*, and tomato, and the pinkish red, green, and red lines respectively represent the collinear gene pairs of *SmCAT2*, *SmCAT3*, and *SmCAT4* in eggplant, *Arabidopsis thaliana*, and tomato, respectively.

2.5. Expression Analysis of Eggplant CATs under Abiotic Stress Treatment and in Different Tissues

We analyzed the expression of *SmCATs* under various abiotic stresses, including dehydration, salt, high temperature (HT), low temperature (LT), H_2O_2, and ABA by real-time quantitative PCR (RT-qPCR) assay (Figure 4). In response to dehydration stress, the expression levels of four eggplant *CATs* were significantly up-regulated (Figure 4a). Under salt stress, the expression levels of *SmCAT3* and *SmCAT4* were significantly up-regulated, whereas *SmCAT1* expression level was significantly down-regulated. On the other hand, the expression levels of *SmCAT2* displayed no obvious change (Figure 4b).

To HT treatment, the expression levels of *SmCAT1* and *SmCAT2* were significantly down-regulated while *SmCAT4* were significantly up-regulated (Figure 4c). The expression levels of *SmCAT1*, *SmCAT2*, and *SmCAT4* showed a trend of significant up-regulation and then down, while the *SmCAT3* were significantly down-regulated in response to LT treatment (Figure 4d). Under H_2O_2 treatment, *SmCAT1*, *SmCAT2*, and *SmCAT3* showed a significant up-regulation expression trajectory, but *SmCAT4* were down-regulated (Figure 4e). ABA treatment could induce the expression levels of *SmCAT1*, *SmCAT2*, and *SmCAT3*—but not *SmCAT4* (Figure 4f).

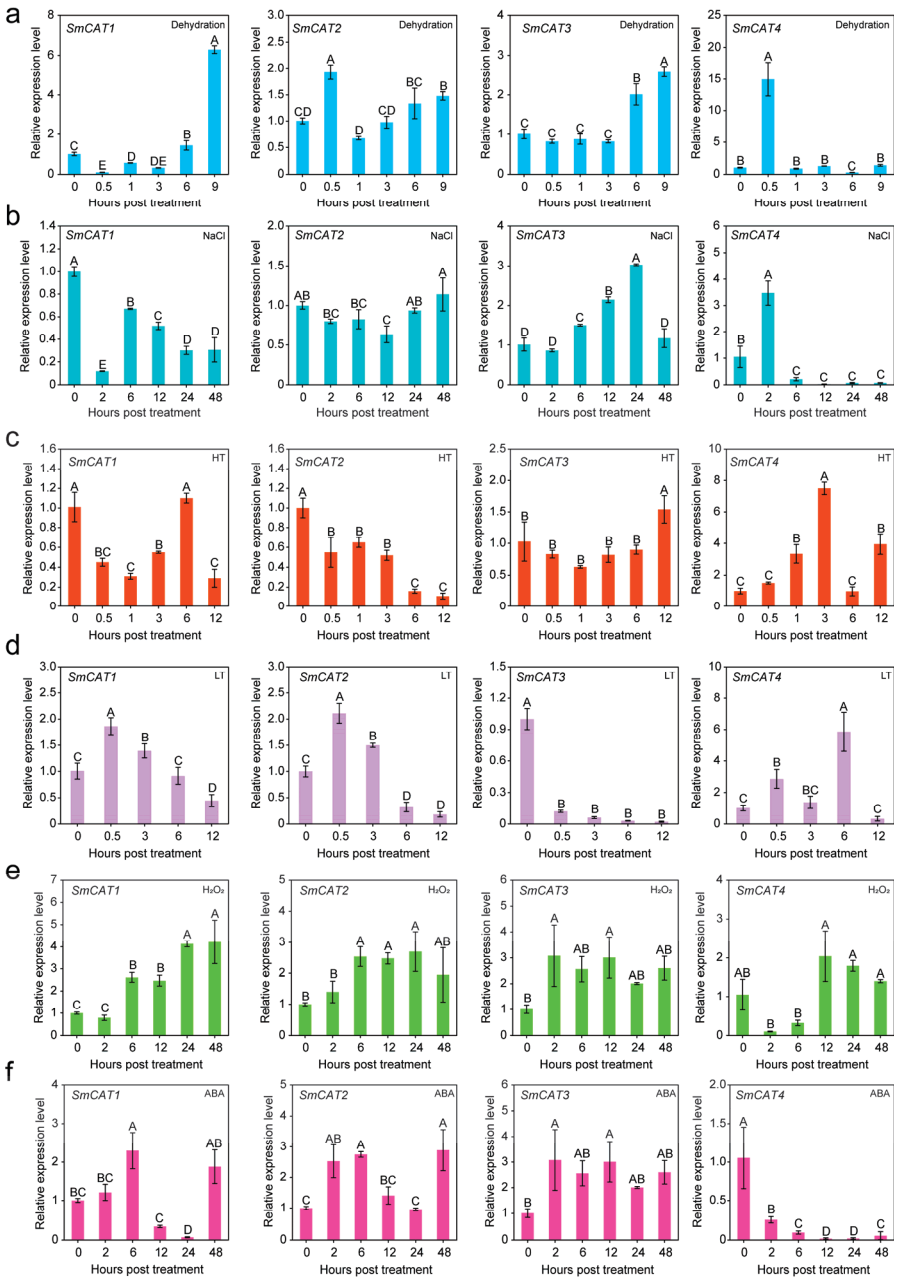


Figure 4. Analysis of the relative transcript expression levels of *SmCAT1*, *SmCAT2*, *SmCAT3*, and *SmCAT4* under the condition of dehydration (**a**), 200 mM NaCl (**b**), high temperature (HT, 43 °C) (**c**), low temperature (LT, 4 °C) (**d**), H_2O_2 (**e**), and 100 μM abscisic acid (ABA) (**f**) treatment. Data indicate the means ± SD from three biological repeats. Different upper letters indicate significant differences, as determined by Fisher's protected LSD test ($p < 0.01$).

We further investigated the expression levels of eggplant *CATs* in the different tissues of seedlings and mature plants (Figure 5). The expression levels of *SmCAT1*, *SmCAT2*, *SmCAT3*, and *SmCAT4* were highest in the young leaf (YL), root (RT), and stem (ST) of seedlings. All four *SmCATs* showed high expressions in the sepal (SE), fully expanded leaf (FEL), flower (FL), and ST of mature eggplant.

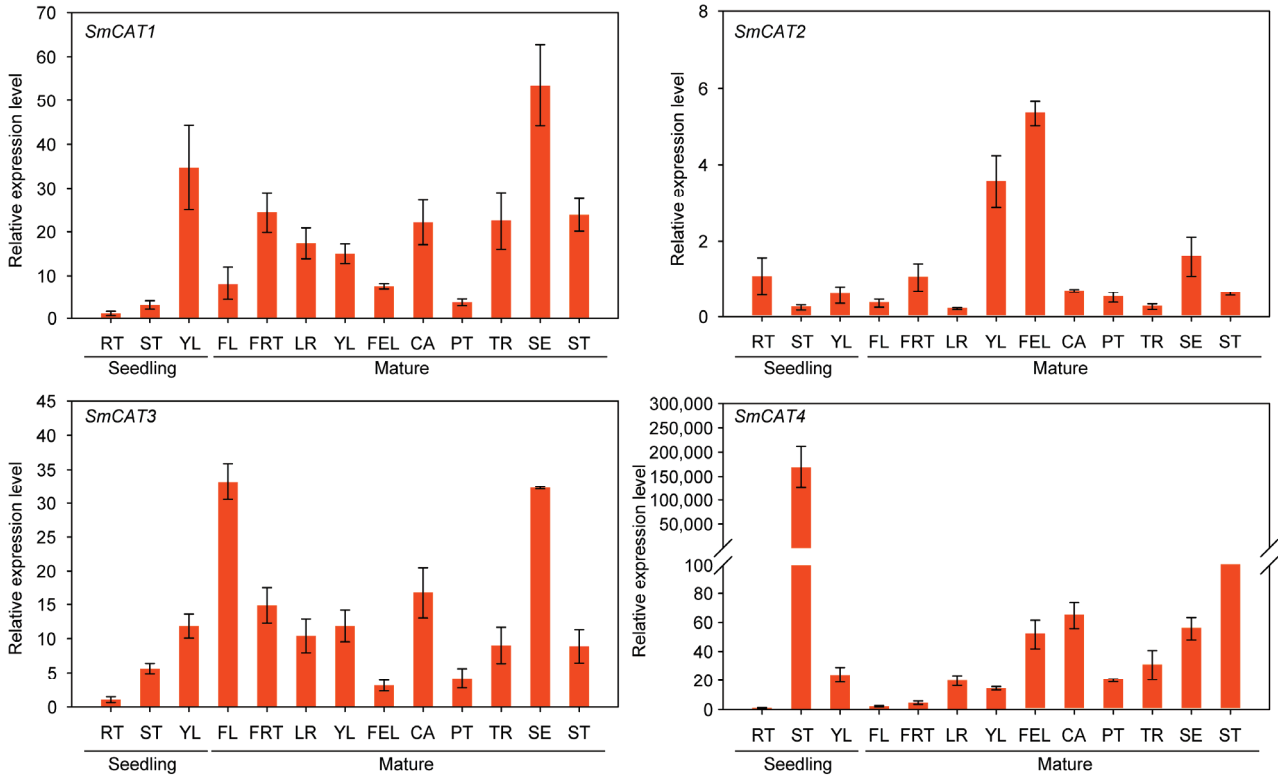


Figure 5. Analysis of the relative transcript expression levels of four eggplant *CAT* genes in various organs of eggplant seedlings and mature plants. RT, root; ST, stem; YL, young leaf; FL, flower; FRT, fruit; LR, lateral root; FEL, fully expanded leaf; CA, carpopodium; PT, petiole; TR, tap root; SE, sepal. Data indicate the means ± SD from three biological repeats.

2.6. Subcellular Localization of Eggplant CAT Proteins

The subcellular location was predicted using an online server and it revealed that all SmCATs are resided in the peroxisomes (Table 1). To test this result, we performed the transient expression of *SmCATs* in the leaves of *Nicotiana benthamiana* via Agrobacterium-mediated transformation. The full-length ORF of eggplant *SmCATs* were cloned into plant overexpression vector pBinGFP2 (Figure 6a). After 48 h infiltration, we observed the fluorescence signal by using a laser scanning confocal microscope. We found that the green fluorescence signal of all eggplant CAT proteins appears in the peroxisome, except for SmCAT4-GFP whose green fluorescence signal appeared in the cytomembrane and nucleus, while the green fluorescence signal of GFP exists in the whole cells (Figure 6b).

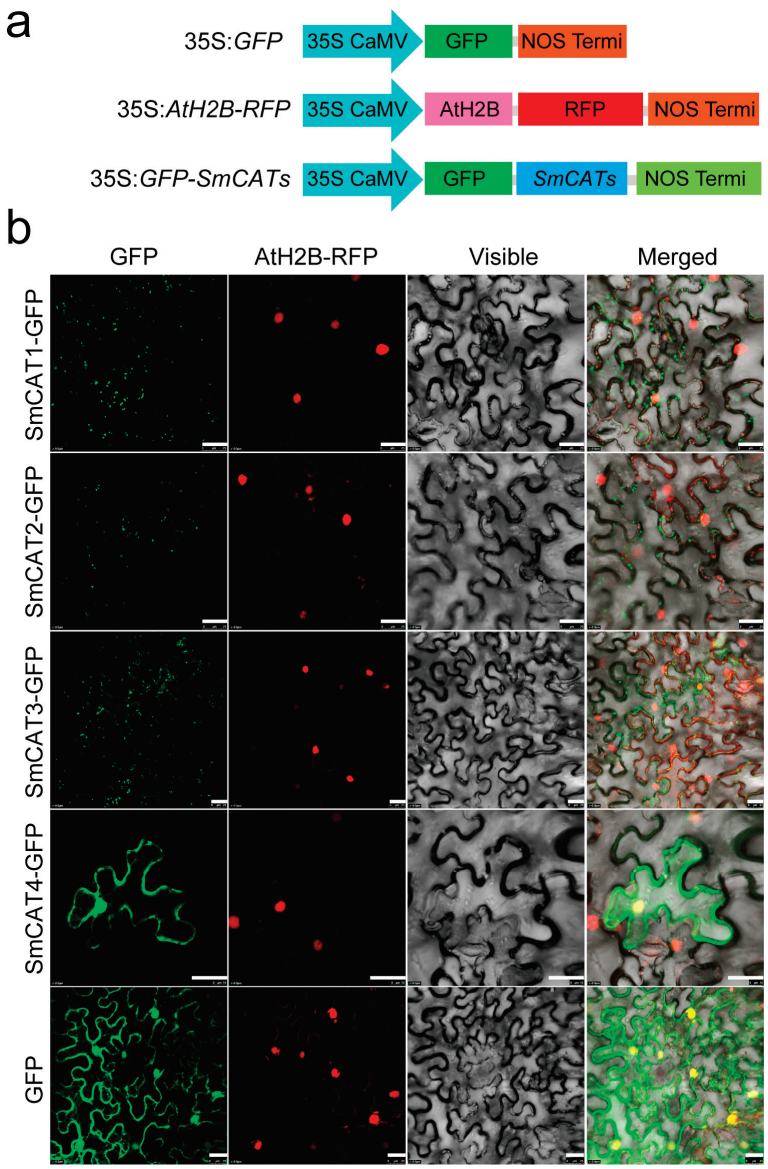


Figure 6. Analysis of subcellular localization of eggplant CAT proteins. (**a**) The schematic diagrams of 35S:*GFP* (empty vector), 35S:*AtH2B-RFP* (a subcellular localization marker of nucleus), and 35S:*GFP-SmCATs* structures. (**b**) Subcellular localization of eggplant CAT proteins in the epidermic cells of *Nicotiana benthamiana* leaves. Bar = 25 μm.

2.7. Recombinant SmCAT4 Enzyme Activity Analysis

We found that the expression of *SmCAT4* was up-regulated by dehydration, salt, HT, and LT treatment (Figure 4), indicating that *SmCAT4* may play an essential role in eggplant response to the above abiotic stresses. Hence, we selected *SmCAT4* to explore its functions in eggplant response to salt stress. We first analyzed *SmCAT4* enzyme activity by recombinant CAT enzyme assay in vitro. The recombinant SmCAT4-GST proteins or GST proteins were expressed by prokaryotic expression assay and purified by GST magnetic beads and then verified by SDS-PAGE assay (Figure 7a). To detect the CAT enzyme activity of SmCAT4, we tested the capacity of recombinant SmCAT4-GST proteins to increase H_2O_2 scavenging. Compared to the control (GST proteins), SmCAT4-GST exhibited distinctly H_2O_2 scavenging capacity (Figure 7b). These data indicate that SmCAT4 exhibits CAT enzyme activity to scavenge H_2O_2.

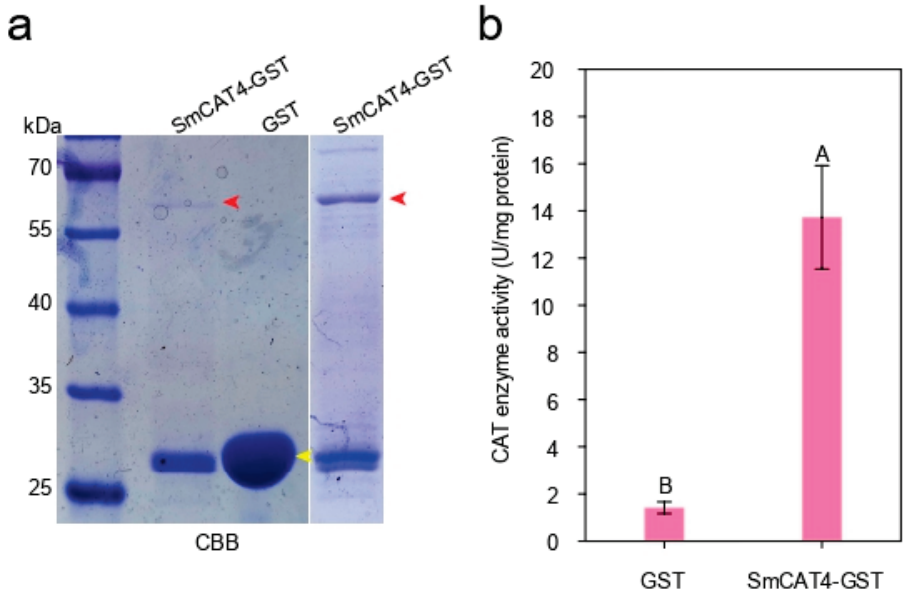


Figure 7. Analysis of CAT enzyme activity of SmCAT4-GST in vitro. (**a**) Analysis of purification of recombinant protein SmCAT4-GST and GST. The recombinant protein SmCAT4-GST and GST were separated by SDS-PAGE assay and the gel was then stained by Coomassie brilliant blue (CBB) solution and de-stained by a destainer. Red arrows indicate SmCAT4-GST protein, and yellow arrow represent GST protein. (**b**) Detection of CAT enzyme activity of SmCAT4-GST and GST protein in vitro. Data indicate the means ± SD from three biological repeats. Different upper letters indicate significant differences, as determined by Student's *t*-test ($p < 0.01$).

2.8. Silencing of SmCAT4 Enhanced Susceptibility of Eggplant to Salt Stress

To investigate the function of *SmCAT4* in eggplant, we assessed the effect of *SmCAT4* silencing on the tolerance of eggplant to salt stress using VIGS assay. We first detected the silencing efficiency of *SmCAT4* by RT-qPCR assay. Compared to the control plants (TRV:*00*), the expression level of *SmCAT4* in the roots of *SmCAT4*-silenced (TRV:*SmCAT4*) plants was significantly reduced by approximately 70% under salt stress treatment (Figure 8a). Silencing of *SmCAT4* enhanced the susceptibility of eggplant to salt stress and exhibiting a lower survival rate compared to the control plants at 48 h post salt stress treatment (Figure 8b,c). In addition, we also found that silencing of *SmCAT4* significantly down-regulated the expression levels of salt stress defense-related marker genes *SmATS14* and *SmDHN1* (Figure 8d), accompanied by an increase of H_2O_2 content and a decrease of CAT enzyme activity (Figure 8e). These data imply that *SmCAT4* positively functions in eggplant response to salt stress.

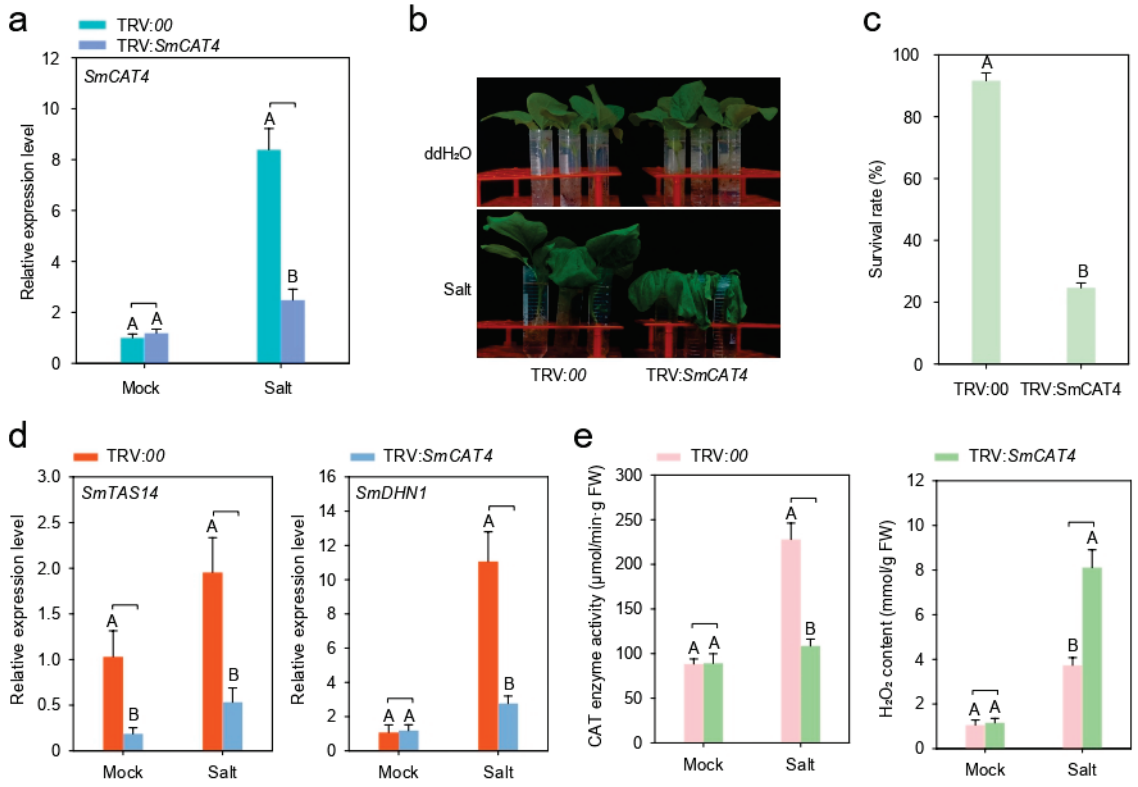


Figure 8. Silencing of *SmCAT4* enhanced susceptibility of eggplant against salt stress. (**a**) Analysis of silencing efficiency of *SmCAT4* by performing RT-qPCR assay. (**b**) Enhanced salt sensitivity of *SmCAT4*-silenced eggplants. The roots of the control and *SmCAT4*-silenced plants were soaked into 200 mM NaCl solution or ddH_2O for 2 days. (**c**) Calculation of survival rate of the control and *SmCAT4*-silenced plants treated with 200 mM NaCl solution at 48 h post treatment. (**d**) Analysis of transcript expression levels of salt stress defense related marker genes *SmTAS14* and *SmDHN1* in the roots of *SmCAT4*-silenced or control plants at 24 h post salt stress treatment. (**e**) Measurement of CAT enzyme activity and H_2O_2 content in the roots of *SmCAT4*-silenced or control plants at 24 h post salt stress treatment. In (**a**,**c**–**e**), the data indicate the means ± SD from three biological repeats. Different upper letters indicate significant differences, as determined by Student's *t*-test ($p < 0.01$).

3. Discussion

Excessive accumulation of ROS caused serious damage to the stability of the plasma membrane in plant cells. The ROS scavenging system, comprising enzymatic and non-enzymatic components [36], is key to plant survival under adverse conditions. CATs, one of the most important ROS scavenging enzymes, has been involved in regulating plant response to abiotic stresses such as salt [37,38], drought [31,39–41], cold [31], heat [42,43], osmotic stress [37], and biotic stresses including pathogenic microorganisms attack [30,44–50], and insect invasion [51]. In addition, growth and development [27,52] and senescence [53] are also regulated by CATs in conjunction with growth-related hormones. However, the underlying mechanisms of CATs in plants involved in the regulation of salt stress resistance largely remain unclear. In this study, we identified four *CAT* genes in the eggplant genome, analyzed their sequences and structures and explored their expressions under NaCl, dehydration, LT, HT, H_2O_2, and ABA treatment. We further provided evidence that *SmCAT4* positively functions in eggplant response to salt stress.

ROS has dual functions, and the causal link between ROS production and stress tolerance is not as straightforward as one may expect [54]. Upon normal circumstances, ROS acts as a second messenger to integrate signaling pathways involved in the regulation of growth, development, gravitropism, phytohormone signal transduction, defense response, and many other physiological processes [55–59]. When ROS production is excessive, un-

controlled oxidation ultimately leads to cellular damage and even cell death. To avoid excessive cell damage, excessive accumulation of ROS should be scavenged, and the antioxidant defenses must keep ROS under control. Therefore, analyzing ROS scavenging is beneficial for a better understanding of the functions and mechanisms of ROS in various physiological processes in plants. CATs act as the essential ROS scavenging enzyme and play an important role in plant response to salt stress [37,38]. However, to date, the information and functions of *CAT* genes in eggplant remain indistinct. Herein, we identified four *CAT* genes in the eggplant genome, which exhibited similar physicochemical properties (Table 1). By generating a phylogenetic tree, the *CAT* genes of eggplant mainly clustered with four rice *CAT* genes and *AtCAT3* apart from *SmCAT3*, while *SmCAT3* exhibited the closest phylogenetic relationship with *AtCAT1*, *AtCAT2*, and four soybean *CAT* genes (Figure 1), implying that there is likely a functional differentiation between *SmCAT3* and the three other *CAT* genes in eggplant. We analyzed sequences and structures of *CAT* genes in eggplant and found that four eggplant *CAT* genes were located in chromosomes 02, 04, and 05 (Figure 2a). Notably, *SmCAT1* and *SmCAT4* were located very close to each other in chromosome 05, which was similar to *AtCAT1* and *AtCAT3* [14], suggesting that *SmCAT1* and *SmCAT4* most likely evolved from a common ancestral gene and had similar functions. We analyzed the gene structures of eggplant *CAT* genes and found that *SmCAT1-3* has eight exons, while *SmCAT4* has seven exons (Figure 2b). Each of the four eggplant CAT proteins has conserved Catalase and Catalase-rel domain as well as 10 identical conservative motifs (Figure 2c,d) but did not have a PP-Binding domain such as Arabidopsis CAT proteins [14]. Apart from the *cis*-elements related to light response, we counted the *cis*-elements related to phytohormone response, transcription factors binding, and stress response within the promoters of four eggplant *CAT* genes by searching the PlantCARE website and found the phytohormone response-related *cis*-elements including SA response-related element CAT-element, ET response-related element ERE, ABA response-related element ABRE, and MeJA response-related element TGACG-motif, and transcription factors binding elements such as HD-Zip I, W-box, MYB, MYC, G-box, and A-box, as well as stress response-related elements LTR, ARE, and STRE within the promoters of these four *CAT* genes (Figure 2e). Abundant *cis*-elements within the eggplant *CAT* gene promoters implied that *CAT* genes may be involved in multiple physiological processes. Gene duplication events act as a primary contributor to gene expansion, leading to the evolution and diversification of genes in plants [60]. In general, the functional relationships of homologs within the same species or in interspecific species reveal their conserved functions and the co-evolutionary origins of members within subgroups [61]. In eggplant, there are two tandem gene clusters, including *SmCAT3* and *SmCAT2* as well as *SmCAT3* and *SmCAT4*, but segmentally duplicated gene pairs do not exist (Figure 3a), indicating that *CAT* genes of eggplant have not been duplicated in large fragments during the evolutionary process, and this may be the reason that the eggplant *CAT* family has fewer members. The collinear analysis displayed that three *CAT* genes (*SmCAT2-4*) from eggplant were orthologous to Arabidopsis (*AtCAT1*) and tomato (*SlCAT2-4*), respectively (Figure 3b), suggesting that these genes may be functionally similar.

Accumulating pieces of evidence demonstrated that *CAT* genes have functions in plant response to abiotic stresses [31,38,39]. We detected the transcript expression levels of four eggplant *CAT* genes under stress, dehydration, LT, HT, H_2O_2, and ABA treatment and found that the transcript expression levels of the four *CAT* genes from eggplant were up- or down-regulated to different extents, and some *CAT* genes expression did not have obvious change (Figure 4). Notably, transcript expression levels of *SmCAT4* were significantly up-regulated by salt, dehydration, LT, HT, and H_2O_2 treatment, suggesting that *SmCAT4* may play an important role in eggplant against salt, dehydration, LT, and HT stresses. In addition, we tested the transcript expression levels of *SmCAT1-4* in the different tissues of seedlings or mature plants, and in seedlings, the expression of *SmCAT1* and *SmCAT3* was higher than that of in stem and young leaf (YL), while *SmCAT2* and *SmCAT4* were highly expressed in stem (ST) and root (RT), respectively. In mature plants,

SmCAT1-4, respectively, were highly expressed in sepal (SE), fully expanded leaf (FEL), flower (FL), and stem (ST) (Figure 5). Previous studies showed that CAT proteins are mainly located in the peroxisome, glyoxylate circulator and cytoplasm [21,62]. Although we predicted the subcellular localization of these four CAT proteins, which localized in the peroxisome, we provided evidence that eggplant CAT1-3 proteins located in the peroxisome and SmCAT4 localized in the cytomembrane and nucleus (Figure 6). Similar studies revealed that *Pinellia ternata* catalase protein PtCAT2 localized in the cytoplasm and membrane [63]. Sugarcane catalase protein ScCAT2 is distributed in the nucleus, plasma membrane, and cytoplasm in the epidermal cells of *Nicotiana benthamiana* leaves [47]. These reports imply that not all CAT proteins function in peroxisomes. To explore the functions of *CAT* genes in eggplant's response to abiotic stresses, we selected SmCAT4 to further analyze its function in eggplant's response to salt stress. Silencing of *SmCAT4* decreased the tolerance of eggplant to salt stress (Figure 8b) and significantly down-regulated the transcript expression levels of salt stress defense-related genes *SmATS14* and *SmDNH1* (Figure 8d), accompanied by an increase of H_2O_2 content and a decrease of the enzyme activities of CAT (Figure 8e). In addition, we tested the CAT enzyme activity of SmCAT4 in vitro and found that SmCAT4 could scavenge H_2O_2 in vitro, indicating that SmCAT4 exhibits CAT enzyme activity. Likewise, the OsCAT3 protein has an obvious ability to remove H_2O_2 in vitro [23]. These data indicate that *SmCAT4* positively functions in eggplant response to salt stress via scavenging H_2O_2. Previous studies showed that Arabidopsis CAT2 promotes LAP2 hydrolysis activity with leucine-4-methylcoumaryl-7-amides as a substrate by interacting with LAP2 in vivo and in vitro to confer Arabidopsis salt and osmotic stress tolerance [37].

In conclusion, we identified four *CAT* genes in the eggplant genome, analyzed their sequences, structures and expressions, and provided evidence that SmCAT4 has an obvious ability to remove H_2O_2 in vitro and positively functions in eggplant response to salt stress.

4. Material and Methods

4.1. Identification of CAT Genes in Solanum melongena Genome

The amino acid sequences of three Arabidopsis CAT obtained from The Arabidopsis Information Resource database (https://www.arabidopsis.org/, accessed on 10 September 2023) were acted as reference sequences [22], and the members of CAT genes in eggplant genome were identified by combining TBtools software [34], National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/, accessed on 10 September 2023) website, and Eggplant Genome Database (http://eggplant-hq.cn/Eggplant/home/index, accessed on 10 September 2023).

4.2. Physicochemical Properties Analysis of CAT Proteins in Eggplant

The physicochemical properties of eggplant CATs protein amino acid sequences, including total number of atoms, instability index, aliphatic index, average of hydropathicity, and theoretical pI, were analyzed by searching the ProtParam Expasy (https://web.expasy.org/protparam/, accessed on 11 September 2023) website [64]. The subcellular localization prediction of eggplant CATs was carried out by searching the Plant-mPLoc website (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/, accessed on 11 September 2023) with their amino acid sequences.

4.3. Multiple Sequence Alignment and Phylogenetic Analysis

The amino acid sequences of CAT members in soybean (*Glycine max*) and rice were respectively downloaded from SoyBase (https://soybase.org/, accessed on 10 September 2023) database [29] and The Rice Annotation Project (RAP) (https://rapdb.dna.affrc.go.jp/, accessed on 10 September 2023) database [14]. Multiple sequence alignment and phylogenetic analysis were carried out by using MEGA 7 (version. 7.0.26) software. The evolutionary tree was constructed by MEGA 7 using the neighbor-joining (NJ) method with 1000 bootstrap replications [65].

4.4. Analysis of Chromosomal Location, Gene Structure, Conserved Domain and Motif, cis-Elements, and Tertiary Structure of Eggplant CATs

The chromosomal locations and gene structures of eggplant CATs were visualized by using TBtools software with Gene Location Visualize from the GTF/GFF functional module and Visualize Gene Structure (Basic), respectively. The conserved domain of eggplant CATs was predicted by searching the SMART (http://smart.embl-heidelberg.de/, accessed on 11 September 2023) website and visualized by DOG 2.0 software [66]. The conserved motifs of eggplant CATs were predicted by searching Multiple Em for Motif Elicitation (MEME, Version 5.5.3) website (https://meme-suite.org/meme/tools/meme, accessed on 11 September 2023) with eggplant CATs amino acid sequences. The prediction of the *cis* elements within the promoters of eggplant *CAT* genes was performed by searching PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 11 September 2023) website and then visualized by using TBtools software. The tertiary structure of eggplant CATs was predicted by searching the SWISS-MODEL website (https://swissmodel.expasy.org/interactive, accessed on 11 September 2023) with the build model function.

4.5. Collinearity Analysis

The collinearity relationships of eggplant CATs were analyzed and visualized by using TBtools software with the Multiple Collinearity Scan toolkit (MCScanX) functional modules [67]. The duplicated genes of eggplant CTAs with their homologs in tomato and Arabidopsis were performed by using TBtools software with the Advanced Circos functional module.

4.6. Plant Materials and Growth Conditions

The seeds of cultivated eggplant ML41 were packaged with clean gauze and incubated in the 55 °C water bath for 15 min, and then were socked in the ddH_2O at room temperature overnight. The seeds were sowed in the nutrient soil [peat moss: perlite, 2:1 (v/v)] and placed in the illumination incubator for germination. The eggplant seedlings were transmitted into plastic pots (7 cm × 7 cm) with nutrient soil. The *Nicotiana benthamiana* seeds were sowed on wet filter paper and placed in the illumination incubator for germination. The seedlings of *Nicotiana benthamiana* were transmitted in the plastic pot (7 cm × 7 cm) with nutrient soil. Both the seedlings of eggplant and *Nicotiana benthamiana* were grown in the illumination incubator with 25 °C, 60% relative humidity, 16 h light/8 h dark photoperiod.

4.7. Salt Stress, Dehydration, High Temperature, Low Temperature, ABA, and H_2O_2 Treatment

The 4-to-6 leaf-stage eggplants were pulled out gently from the soil, and the eggplant roots were washed with clean tap water. The eggplant roots were soaked into the Hoagland nutrient solution for 2 days, and then the roots were respectively soaked in 200 mM NaCl, 100 μM ABA, or 1 mM H_2O_2 solution. The treated roots were harvested at the time points of 0, 2, 6, 12, 24, and 48 h post-treatment and frozen in liquid nitrogen. For dehydration stress treatment, the cleaned eggplant roots were dried with filter paper and placed on the bench. The roots were harvested at the time points of 0, 0.5, 1, 3, 6, and 9 h post treatment. For HT or LT stress treatment, the eggplants were respectively placed in the illumination incubator with 43 °C or 4 °C. The treated leaves were harvested in the 2 mL centrifuge tube with 3 small steel balls at the time points of 0, 0.5, 1, 3, 6, and 12 h post-treatment.

4.8. Plant Total RNA Extraction, cDNA Synthesis and RT-qPCR Analysis

The treated roots or leaves were ground into powder by a plant tissue crusher under low-temperature conditions, and the total RNA was extracted by using FastPure Universal Plant Total RNA Isolation Kit (RC411-01, Vazyme, Nanjing city, China). The mRNA was reversely transcribed into cDNA by using HiScript III RT SuperMix for qPCR (+gDNA wiper) (R323-01, Vazyme, China). For real-time quantitative PCR (RT-qPCR) analysis, the ChamQ Universal SYBR qPCR Master Mix (Q711-02, Vazyme, China) and the specific

primer pairs listed in Table S1 were used to carry out the RT-qPCR assay to detect the relative transcript expression levels of target genes. *SmActin* (Smechr1100649) was used as the reference gene to normalize the transcript expression levels of target genes. Three biological replications were used, and the relative transcript expression levels of target genes were analyzed by the $2^{-\Delta\Delta CT}$ method [68].

4.9. Vector Construction

For subcellular localization, the full-length open reading frames (ORF) of *SmCAT1*, *SmCAT2*, *SmCAT3*, and *SmCAT4* were respectively amplified by PCR assay with the specific primer pairs and then cloned into the plant overexpression vector pBinGFP2 linearized by DNA restriction endonuclease *Sma* I using ClonExpress II One Step Cloning Kit (C112-01, Vazyme, China). For prokaryotic expression, the full-length ORF of *SmCAT4* was amplified by PCR with the specific primer pairs and then cloned into the multiple clone site *Sma* I in the pGEX-6P-1 vector by using ClonExpress II One Step Cloning Kit (C112-01, Vazyme, Nanjing city, China). For the VIGS assay, the 300 bp specific DNA fragment of *SmCAT4* was amplified by PCR and then cloned into entry pDONR207 vector by BP reaction, and then transmitted into the destination vector pTRV2 by LR reaction. The primer pairs used for vector construction in this study were listed in Table S1.

4.10. Agrobacterium tumefaciens Cultivation and Infiltration and Subcellular Localization Analysis

Agrobacterium tumefaciens strain GV3101 containing 35S:*SmCAT1-GFP*, 35S:*SmCAT2-GFP*, 35S:*SmCAT3-GFP*, or 35S:*SmCAT4-GFP* constructs were cultivated in the liquid Luria-Bertani (LB) medium with 50 μg/mL kanamycin and 50 μg/mL rifampicin antibiotics at the conditions of 28 °C, 200 rpm, overnight. The bacteria solution was centrifuged at 6000 rpm for 5 min at room temperature, and then the bacterial cells were resuspended by infiltration buffer (10 mM MES, 10 mM $MgCl_2$, 200 mM acetosyringone, pH = 5.4) to adjust OD_{600} to 0.8. The bacterial solution was infiltrated into the leaves of *Nicotiana benthamiana* using a disposable sterilized syringe without a needle. The infiltrated plants were cultivated in the illumination incubator. After 48 h, the signaling in the epidemic cells of *Nicotiana benthamiana* leaves was observed by laser scanning confocal microscope (TCS SP8, Leica Microsystems, Weztlar, Germany).

4.11. VIGS Assay

The *Agrobacterium tumefaciens* GV3101 cells harboring pTRV1, pTRV2:*00*, pTRV2:*SmPDS*, or pTRV2:*SmCAT4* constructs were cultivated in liquid LB medium overnight and then adjusted OD_{600} into 0.8 by infiltration buffer. The bacterial cells containing the pTRV1 vector were mixed with the GV3101 cells carrying pTRV2:*00*, pTRV2:*SmPDS*, or pTRV2:*SmCAT4* constructs at 1:1 ratio, and the mixtures were tenderly incubated in the 28 °C shaker at 60 rpm for 3 h. The Agrobacterium solution was infiltrated into the cotyledon of 2-to-3 leaf-stage eggplants. The infiltrated eggplant seedlings were placed in the illumination incubator without light at 20 °C for 48 h. The treated seedlings grew in the illumination incubator for 3 weeks until the leaves of TRV:*SmPDS* plants turned white.

4.12. Prokaryotic Expression and Purification of Recombinant Protein

The *Escherichia coli* strain BL21 (DE3) harboring pGEX-6P-1:*SmCAT4* (containing a GST protein tag) constructs or empty vector pGEX-6P-1 were cultivated in the liquid LB medium with 100 μg/mL ampicillin at 37 °C, 200 rpm until the OD_{600} of the bacterial solution to 0.6, and then added 2 mM isopropyl-β-d-thiogalactoside into the bacterial solution. The bacterium was cultivated at 18 °C, 200 rpm overnight. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie bright blue staining assays were performed to detect the target proteins, whether they were expressed or not. For the purification of recombinant proteins, the bacterium cells were broken by an ultrasonic cell crusher. The supernatant was incubated with the magnetic beads of BeaverBeads™ GSH kit (70601-5, Suzhou Beaver Biomedical Engineering Co., LTD,

Suzhou city, China) at 4 °C for 1 h, and the target proteins on the magnetic beads were eluted by elution buffer after washing.

4.13. Recombinant CAT Enzyme Assay

The detection of the purified recombinant SmCAT4-GST or GST proteins was performed following a previously described study [69].

4.14. Physiological Index Measurement

The eggplant roots of TRV:*00* or TRV:*SmCAT4* plants treated with salt stress were harvested, and the physiological index measurement of CAT enzyme activity and H_2O_2 content were performed following previously described studies [69,70].

5. Conclusions

In this study, we identified four *CAT* genes in the eggplant genome and further analyzed their sequences, structures, and expressions in eggplant. Phylogenetic classification divided CAT proteins into subgroups I and II. Further sequences and structure analysis of *CAT* genes revealed a high degree of conservation among the *CAT* gene family members in eggplant. Based on the result of different gene expressions of *CAT* genes under abiotic stress such as salt, drought, cold, and high temperature stress, we found that *CAT* genes may play an important role in eggplant response to these abiotic stresses. Moreover, eggplant SmCAT1-3 proteins were localized in the peroxisome, and SmCAT4 was localized in the cytomembrane and nucleus. *SmCAT4* expression was induced by salt, dehydration, LT, and HT treatment. VIGS assay revealed that silencing of *SmCAT4* could compromise the tolerance level of eggplant to salt stress. Our data provide new insight into a comprehensive understanding of the *CAT* gene family and help further explore the functions of *CAT* family members in eggplant response to various abiotic stresses.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms242316979/s1.

Author Contributions: L.S. and X.Y. conceived the idea. L.S. designed the experiments and analyzed the results. L.S., X.X., L.Z. and S.Y. performed the experiments. L.S. wrote the manuscript. L.S. and X.Y. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data will be made available on request.

Acknowledgments: We are thankful to Rahat Sharif (College of Horticulture and Landscape architecture, Yangzhou University) for revising the language of manuscript.

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

References

1. Li, Y.; Chen, M.; Khan, A.H.; Ma, Y.; He, X.; Yang, J.; Zhang, R.; Ma, H.; Zuo, C.; Li, Y.; et al. Histone H3 lysine 27 trimethylation suppresses jasmonate biosynthesis and signaling to affect male fertility under high temperature in cotton. *Plant Commun.* **2023**, *4*, 100660. [CrossRef] [PubMed]
2. Wei, H.; Song, Z.; Xie, Y.; Cheng, H.; Yan, H.; Sun, F.; Liu, H.; Shen, J.; Li, L.; He, X.; et al. High temperature inhibits vascular development via the PIF4-miR166-HB15 module in Arabidopsis. *Curr. Biol.* **2023**, *33*, 3203–3214.e4. [CrossRef]
3. Gu, S.; Zhang, Z.; Li, J.; Sun, J.; Cui, Z.; Li, F.; Zhuang, J.; Chen, W.; Su, C.; Wu, L.; et al. Natural variation in OsSEC13 HOMOLOG 1 modulates redox homeostasis to confer cold tolerance in rice. *Plant Physiol.* **2023**, *193*, 2180–2196. [CrossRef]
4. An, J.P.; Liu, Z.Y.; Zhang, X.W.; Wang, D.R.; Zeng, F.; You, C.X.; Han, Y. Brassinosteroid signaling regulator BIM1 integrates brassinolide and jasmonic acid signaling during cold tolerance in apple. *Plant Physiol.* **2023**, *193*, 1652–1674. [CrossRef] [PubMed]
5. Mei, C.; Yang, J.; Mei, Q.; Jia, D.; Yan, P.; Feng, B.; Mamat, A.; Gong, X.; Guan, Q.; Mao, K.; et al. MdNAC104 positively regulates apple cold tolerance via CBF-dependent and CBF-independent pathways. *Plant Biotechnol. J.* **2023**, *21*, 2057–2073. [CrossRef]
6. Hu, Y.; Zeng, L.; Lv, X.; Guo, J.; Li, X.; Zhang, X.; Wang, D.; Wang, J.; Bi, J.; Julkowska, M.M.; et al. NIGT1.4 maintains primary root elongation in response to salt stress through induction of ERF1 in Arabidopsis. *Plant J.* **2023**, *116*, 173–186. [CrossRef] [PubMed]
7. Yang, J.; Qiu, L.; Mei, Q.; Sun, Y.; Li, N.; Gong, X.; Ma, F.; Mao, K. MdHB7-like positively modulates apple salt tolerance by promoting autophagic activity and Na(+) efflux. *Plant J.* **2023**, *116*, 669–689. [CrossRef]

8. Zhang, T.; Zhang, W.; Ding, C.; Hu, Z.; Fan, C.; Zhang, J.; Li, Z.; Diao, S.; Shen, L.; Zhang, B.; et al. A breeding strategy for improving drought and salt tolerance of poplar based on CRISPR/Cas9. *Plant Biotechnol. J.* **2023**, *21*, 2160–2162. [CrossRef]
9. Gautam, N.; Tiwari, M.; Kidwai, M.; Dutta, P.; Chakrabarty, D. Functional characterization of rice metallothionein OsMT-I-Id: Insights into metal binding and heavy metal tolerance mechanisms. *J. Hazard. Mater.* **2023**, *458*, 131815. [CrossRef]
10. Xia, J.; Guo, Z.; Yang, Z.; Han, H.; Wang, S.; Xu, H.; Yang, X.; Yang, F.; Wu, Q.; Xie, W.; et al. Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell* **2021**, *184*, 1693–1705.e17. [CrossRef]
11. Zhang, Y.; Fu, Y.; Liu, X.; Francis, F.; Fan, J.; Liu, H.; Wang, Q.; Sun, Y.; Zhang, Y.; Chen, J. SmCSP4 from aphid saliva stimulates salicylic acid-mediated defence responses in wheat by interacting with transcription factor TaWKRY76. *Plant Biotechnol. J.* **2023**, *21*, 2389–2407. [CrossRef] [PubMed]
12. Wang, G.; Wang, X.; Song, J.; Wang, H.; Ruan, C.; Zhang, W.; Guo, Z.; Li, W.; Guo, W. Cotton peroxisome-localized lysophospholipase counteracts the toxic effects of Verticillium dahliae NLP1 and confers wilt resistance. *Plant J.* **2023**, *115*, 452–469. [CrossRef] [PubMed]
13. Yang, S.; Cai, W.; Wu, R.; Huang, Y.; Lu, Q.; Hui, W.; Huang, X.; Zhang, Y.; Wu, Q.; Cheng, X.; et al. Differential CaKAN3-CaHSF8 associations underlie distinct immune and heat responses under high temperature and high humidity conditions. *Nat. Commun.* **2023**, *14*, 4477. [CrossRef] [PubMed]
14. Alam, N.B.; Ghosh, A. Comprehensive analysis and transcript profiling of *Arabidopsis thaliana* and *Oryza sativa* catalase gene family suggests their specific roles in development and stress responses. *Plant Physiol. Biochem.* **2018**, *123*, 54–64. [CrossRef] [PubMed]
15. Mhamdi, A.; Queval, G.; Chaouch, S.; Vanderauwera, S.; Van Breusegem, F.; Noctor, G. Catalase function in plants: A focus on Arabidopsis mutants as stress-mimic models. *J. Exp. Bot.* **2010**, *61*, 4197–4220. [CrossRef] [PubMed]
16. Ren, L.; Wang, M.R.; Wang, Q.C. ROS-induced oxidative stress in plant cryopreservation: Occurrence and alleviation. *Planta* **2021**, *254*, 124. [CrossRef] [PubMed]
17. Gill, S.S.; Tuteja, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* **2010**, *48*, 909–930. [CrossRef]
18. Nawaz, M.A.; Huang, Y.; Bie, Z.; Ahmed, W.; Reiter, R.J.; Niu, M.; Hameed, S. Melatonin: Current status and future perspectives in plant science. *Front. Plant Sci.* **2015**, *6*, 1230. [CrossRef]
19. Sah, S.K.; Reddy, K.R.; Li, J. Abscisic acid and abiotic stress tolerance in crop plants. *Front. Plant Sci.* **2016**, *7*, 571. [CrossRef]
20. Zhang, Y.; Zheng, L.; Yun, L.; Ji, L.; Li, G.; Ji, M.; Shi, Y.; Zheng, X. Catalase (*CAT*) gene family in wheat (*Triticum aestivum* L.): Evolution, expression pattern and function analysis. *Int. J. Mol. Sci.* **2022**, *23*, 524. [CrossRef]
21. Hu, L.; Yang, Y.; Jiang, L.; Liu, S. The catalase gene family in cucumber: Genome-wide identification and organization. *Genet. Mol. Biol.* **2016**, *39*, 408–415. [CrossRef] [PubMed]
22. Du, Y.Y.; Wang, P.C.; Chen, J.; Song, C.P. Comprehensive functional analysis of the catalase gene family in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* **2008**, *50*, 1318–1326. [CrossRef] [PubMed]
23. Jiang, W.; Ye, Q.; Wu, Z.; Zhang, Q.; Wang, L.; Liu, J.; Hu, X.; Guo, D.; Wang, X.; Zhang, Z.; et al. Analysis of *CAT* gene family and functional identification of *OsCAT3* in rice. *Genes* **2023**, *14*, 138. [CrossRef] [PubMed]
24. Wang, W.; Cheng, Y.; Chen, D.; Liu, D.; Hu, M.; Dong, J.; Zhang, X.; Song, L.; Shen, F. The catalase gene family in cotton: Genome-wide characterization and bioinformatics analysis. *Cells* **2019**, *8*, 86. [CrossRef]
25. Raza, A.; Su, W.; Gao, A.; Mehmood, S.S.; Hussain, M.A.; Nie, W.; Lv, Y.; Zou, X.; Zhang, X. Catalase (*CAT*) gene family in rapeseed (*Brassica napus* L.): Genome-wide analysis, identification, and expression pattern in response to multiple hormones and abiotic stress conditions. *Int. J. Mol. Sci.* **2021**, *22*, 4281. [CrossRef]
26. Sofo, A.; Scopa, A.; Nuzzaci, M.; Vitti, A. Ascorbate peroxidase and catalase activities and their genetic regulation in plants subjected to drought and salinity stresses. *Int. J. Mol. Sci.* **2015**, *16*, 13561–13578. [CrossRef] [PubMed]
27. Bi, C.; Ma, Y.; Wu, Z.; Yu, Y.T.; Liang, S.; Lu, K.; Wang, X.F. Arabidopsis ABI5 plays a role in regulating ROS homeostasis by activating CATALASE 1 transcription in seed germination. *Plant Mol. Biol.* **2017**, *94*, 197–213. [CrossRef]
28. Lee, S.H.; An, C.S. Differential expression of three catalase genes in hot pepper (*Capsicum annuum* L.). *Mol. Cells* **2005**, *20*, 247–255.
29. Zhu, X.; Guo, L.; Zhu, R.; Zhou, X.; Zhang, J.; Li, D.; He, S.; Qiao, Y. *Phytophthora sojae* effector PsAvh113 associates with the soybean transcription factor GmDPB to inhibit catalase-mediated immunity. *Plant Biotechnol. J.* **2023**, *21*, 1393–1407. [CrossRef]
30. You, X.; Zhang, F.; Liu, Z.; Wang, M.; Xu, X.; He, F.; Wang, D.; Wang, R.; Wang, Y.; Wang, G.; et al. Rice catalase OsCATC is degraded by E3 ligase APIP6 to negatively regulate immunity. *Plant Physiol.* **2022**, *190*, 1095–1099. [CrossRef]
31. Xu, J.; Duan, X.; Yang, J.; Beeching, J.R.; Zhang, P. Coupled expression of Cu/Zn-superoxide dismutase and catalase in cassava improves tolerance against cold and drought stresses. *Plant Signal Behav.* **2013**, *8*, e24525. [CrossRef] [PubMed]
32. Tian, G.; Wang, S.; Wu, J.; Wang, Y.; Wang, X.; Liu, S.; Han, D.; Xia, G.; Wang, M. Allelic variation of TaWD40-4B.1 contributes to drought tolerance by modulating catalase activity in wheat. *Nat. Commun.* **2023**, *14*, 1200. [CrossRef]
33. Wei, Q.; Wang, J.; Wang, W.; Hu, T.; Hu, H.; Bao, C. A high-quality chromosome-level genome assembly reveals genetics for important traits in eggplant. *Hortic. Res.* **2020**, *7*, 153. [CrossRef]
34. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y.; et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* **2023**, *16*, 1733–1742. [CrossRef] [PubMed]
35. Cao, Y.; Han, Y.; Meng, D.; Li, D.; Jiao, C.; Jin, Q.; Lin, Y.; Cai, Y. B-BOX genes: Genome-wide identification, evolution and their contribution to pollen growth in pear (*Pyrus bretschneideri* Rehd.). *BMC Plant Biol.* **2017**, *17*, 156. [CrossRef] [PubMed]

36. Verma, D.; Kaushal, N.; Balhara, R.; Singh, K. Genome-wide analysis of Catalase gene family reveal insights into abiotic stress response mechanism in *Brassica juncea* and *B. rapa*. *Plant Sci.* **2023**, *330*, 111620. [CrossRef] [PubMed]
37. Zhang, Y.; Wang, L.F.; Li, T.T.; Liu, W.C. Mutual promotion of LAP2 and CAT2 synergistically regulates plant salt and osmotic stress tolerance. *Front. Plant Sci.* **2021**, *12*, 672672. [CrossRef]
38. Luo, X.; Wu, J.; Li, Y.; Nan, Z.; Guo, X.; Wang, Y.; Zhang, A.; Wang, Z.; Xia, G.; Tian, Y. Synergistic effects of *GhSOD1* and *GhCAT1* overexpression in cotton chloroplasts on enhancing tolerance to methyl viologen and salt stresses. *PLoS ONE* **2013**, *8*, e54002. [CrossRef]
39. Wang, L.; Zhou, Y.; Ding, Y.; Chen, C.; Chen, X.; Su, N.; Zhang, X.; Pan, Y.; Li, J. Novel flavin-containing monooxygenase protein FMO1 interacts with CAT2 to negatively regulate drought tolerance through ROS homeostasis and ABA signaling pathway in tomato. *Hortic. Res.* **2023**, *10*, uhad037. [CrossRef]
40. Zhao, Q.; Hu, R.S.; Liu, D.; Liu, X.; Wang, J.; Xiang, X.H.; Li, Y.Y. The AP2 transcription factor NtERF172 confers drought resistance by modifying NtCAT. *Plant Biotechnol. J.* **2020**, *18*, 2444–2455. [CrossRef]
41. Zou, J.J.; Li, X.D.; Ratnasekera, D.; Wang, C.; Liu, W.X.; Song, L.F.; Zhang, W.Z.; Wu, W.H. Arabidopsis CALCIUM-DEPENDENT PROTEIN KINASE8 and CATALASE3 function in abscisic acid-mediated signaling and H_2O_2 homeostasis in stomatal guard cells under drought stress. *Plant Cell* **2015**, *27*, 1445–1460. [CrossRef]
42. Ono, M.; Isono, K.; Sakata, Y.; Taji, T. CATALASE2 plays a crucial role in long-term heat tolerance of *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* **2021**, *534*, 747–751. [CrossRef]
43. Wang, J.; Xu, J.; Wang, L.; Zhou, M.; Nian, J.; Chen, M.; Lu, X.; Liu, X.; Wang, Z.; Cen, J.; et al. SEMI-ROLLED LEAF 10 stabilizes catalase isozyme B to regulate leaf morphology and thermotolerance in rice (*Oryza sativa* L.). *Plant Biotechnol. J.* **2023**, *21*, 819–838. [CrossRef]
44. Wu, Q.; Chen, Y.; Zou, W.; Pan, Y.B.; Lin, P.; Xu, L.; Grisham, M.P.; Ding, Q.; Su, Y.; Que, Y. Genome-wide characterization of sugarcane catalase gene family identifies a *ScCAT1* gene associated disease resistance. *Int. J. Biol. Macromol.* **2023**, *232*, 123398. [CrossRef]
45. Liao, Y.; Ali, A.; Xue, Z.; Zhou, X.; Ye, W.; Guo, D.; Liao, Y.; Jiang, P.; Wu, T.; Zhang, H.; et al. Disruption of LLM9428/OsCATC represses starch metabolism and confers enhanced blast resistance in rice. *Int. J. Mol. Sci.* **2022**, *23*, 3827. [CrossRef]
46. Zhang, Y.; Song, R.F.; Yuan, H.M.; Li, T.T.; Wang, L.F.; Lu, K.K.; Guo, J.X.; Liu, W.C. Overexpressing the N-terminus of CATALASE2 enhances plant jasmonic acid biosynthesis and resistance to necrotrophic pathogen Botrytis cinerea B05.10. *Mol. Plant Pathol.* **2021**, *22*, 1226–1238. [CrossRef]
47. Sun, T.; Liu, F.; Wang, W.; Wang, L.; Wang, Z.; Li, J.; Que, Y.; Xu, L.; Su, Y. The role of sugarcane catalase gene ScCAT2 in the defense response to pathogen challenge and adversity stress. *Int. J. Mol. Sci.* **2018**, *19*, 2686. [CrossRef] [PubMed]
48. Giri, M.K.; Singh, N.; Banday, Z.Z.; Singh, V.; Ram, H.; Singh, D.; Chattopadhyay, S.; Nandi, A.K. GBF1 differentially regulates *CAT2* and *PAD4* transcription to promote pathogen defense in Arabidopsis thaliana. *Plant J.* **2017**, *91*, 802–815. [CrossRef]
49. Yuan, H.; Jin, C.; Pei, H.; Zhao, L.; Li, X.; Li, J.; Huang, W.; Fan, R.; Liu, W.; Shen, Q.H. The powdery mildew effector CSEP0027 interacts with barley catalase to regulate host immunity. *Front. Plant Sci.* **2021**, *12*, 733237. [CrossRef] [PubMed]
50. Jiao, Z.; Tian, Y.; Cao, Y.; Wang, J.; Zhan, B.; Zhao, Z.; Sun, B.; Guo, C.; Ma, W.; Liao, Z.; et al. A novel pathogenicity determinant hijacks maize catalase 1 to enhance viral multiplication and infection. *New Phytol.* **2021**, *230*, 1126–1141. [CrossRef] [PubMed]
51. Rasool, B.; Karpinska, B.; Pascual, J.; Kangasjarvi, S.; Foyer, C.H. Catalase, glutathione, and protein phosphatase 2A-dependent organellar redox signalling regulate aphid fecundity under moderate and high irradiance. *Plant Cell Environ.* **2020**, *43*, 209–222. [CrossRef] [PubMed]
52. Yang, Z.; Mhamdi, A.; Noctor, G. Analysis of catalase mutants underscores the essential role of CATALASE2 for plant growth and day length-dependent oxidative signalling. *Plant Cell Environ.* **2019**, *42*, 688–700. [CrossRef]
53. Chen, H.J.; Wu, S.D.; Huang, G.J.; Shen, C.Y.; Afiyanti, M.; Li, W.J.; Lin, Y.H. Expression of a cloned sweet potato catalase SPCAT1 alleviates ethephon-mediated leaf senescence and H_2O_2 elevation. *J. Plant Physiol.* **2012**, *169*, 86–97. [CrossRef]
54. Bose, J.; Rodrigo-Moreno, A.; Shabala, S. ROS homeostasis in halophytes in the context of salinity stress tolerance. *J. Exp. Bot.* **2014**, *65*, 1241–1257. [CrossRef]
55. Mittler, R. Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci.* **2002**, *7*, 405–410. [CrossRef]
56. Apel, K.; Hirt, H. Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* **2004**, *55*, 373–399. [CrossRef]
57. Mittler, R.; Vanderauwera, S.; Gollery, M.; Van Breusegem, F. Reactive oxygen gene network of plants. *Trends Plant Sci.* **2004**, *9*, 490–498. [CrossRef] [PubMed]
58. Miller, G.; Shulaev, V.; Mittler, R. Reactive oxygen signaling and abiotic stress. *Physiol. Plant* **2008**, *133*, 481–489. [CrossRef]
59. FOYER, C.H.; NOCTOR, G. Oxidant and antioxidant signalling in plants: A re-evaluation of the concept of oxidative stress in a physiological context. *Plant Cell Environ.* **2005**, *28*, 1056–1071. [CrossRef]
60. Hanada, K.; Zou, C.; Lehti-Shiu, M.D.; Shinozaki, K.; Shiu, S.H. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol.* **2008**, *148*, 993–1003. [CrossRef]
61. Li, C.; Zhang, H.; Qi, Y.; Zhao, Y.; Duan, C.; Wang, Y.; Meng, Z.; Zhang, Q. Genome-wide identification of PYL/PYR-PP2C (A)-SnRK2 genes in Eutrema and their co-expression analysis in response to ABA and abiotic stresses. *Int. J. Biol. Macromol.* **2023**, *253*, 126701. [CrossRef]

62. Hackenberg, T.; Juul, T.; Auzina, A.; Gwizdz, S.; Malolepszy, A.; Van Der Kelen, K.; Dam, S.; Bressendorff, S.; Lorentzen, A.; Roepstorff, P.; et al. Catalase and NO CATALASE ACTIVITY1 promote autophagy-dependent cell death in Arabidopsis. *Plant Cell* **2013**, *25*, 4616–4626. [CrossRef] [PubMed]
63. Xu, J.; Du, N.; Dong, T.; Zhang, H.; Xue, T.; Zhao, F.; Zhao, F.; Duan, Y.; Xue, J. A novel Pinellia ternata catalase gene PtCAT2 regulates drought tolerance in Arabidopsis by modulating ROS balance. *Front. Plant Sci.* **2023**, *14*, 1206798. [CrossRef] [PubMed]
64. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999**, *112*, 531–552. [PubMed]
65. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]
66. Ren, J.; Wen, L.; Gao, X.; Jin, C.; Xue, Y.; Yao, X. DOG 1.0: Illustrator of protein domain structures. *Cell Res.* **2009**, *19*, 271–273. [CrossRef] [PubMed]
67. Du, Y.; Zhang, Z.; Gu, Y.; Li, W.; Wang, W.; Yuan, X.; Zhang, Y.; Yuan, M.; Du, J.; Zhao, Q. Genome-wide identification of the soybean cytokinin oxidase/dehydrogenase gene family and its diverse roles in response to multiple abiotic stress. *Front. Plant Sci.* **2023**, *14*, 1163219. [CrossRef]
68. 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]
69. Yang, T.; Peng, Q.; Lin, H.; Xi, D. Alpha-momorcharin preserves catalase activity to inhibit viral infection by disrupting the 2b-CAT interaction in Solanum lycopersicum. *Mol. Plant Pathol.* **2023**, *24*, 107–122. [CrossRef]
70. Li, Y.; Feng, Z.; Wei, H.; Cheng, S.; Hao, P.; Yu, S.; Wang, H. Silencing of *GhKEA4* and *GhKEA12* revealed their potential functions under salt and potassium stresses in upland cotton. *Front. Plant Sci.* **2021**, *12*, 789775. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

Comprehensive Analysis of Highbush Blueberry Plants Propagated In Vitro and Conventionally

Marzena Mazurek [1,*], Aleksandra Siekierzyńska [1], Tomasz Piechowiak [2], Anna Spinardi [3] and Wojciech Litwińczuk [1]

[1] Department of Physiology and Plant Biotechnology, Institute of Agricultural Sciences, Environment Management and Protection University of Rzeszow, Ćwiklińskiej 2, 35-601 Rzeszow, Poland; asiekierzynska@ur.edu.pl (A.S.)
[2] Department of Chemistry and Food Toxicology, Institute of Food Technology and Nutrition, University of Rzeszow, St. Cwiklinskiej 1a, 35-601 Rzeszow, Poland; tpiechowiak@ur.edu.pl
[3] Department of Agricultural and Environmental Sciences—Production, Landscape, Agroenergy, Università degli Studi di Milano, 20133 Milan, Italy
* Correspondence: mmazurek@ur.edu.pl

Abstract: In vitro culture allows the production of numerous plants with both desirable and undesirable traits. To investigate the impact of the propagation method on highbush blueberry plants, an analysis was performed on four groups of differentially propagated plants: in vitro with axillary (TC-Ax) or adventitious shoots (TC-Ad), conventionally (SC) and using a mixed method (TC/SC). The analysis included plant features (shoot length and branching, chlorophyll and fluorescence and DNA methylation) and fruit properties (antioxidant compounds). The data obtained indicated significant differences between plants propagated conventionally and in vitro, as well as variations among plants derived from in vitro cultures with different types of explants. SC plants generally exhibited the lowest values of morphological and physiological parameters but produced fruits richest in antioxidant compounds. TC/SC plants were dominant in length, branching and fluorescence. Conversely, TC-Ax plants produced fruits with the lowest levels of antioxidant compounds. The methylation-sensitive amplified polymorphism (MSAP) technique was employed to detect molecular differences. TC-Ad plants showed the highest methylation level, whereas SC plants had the lowest. The overall methylation level varied among differentially propagated plants. It can be speculated that the differences among the analysed plants may be attributed to variations in DNA methylation.

Keywords: plant in vitro culture; somaclonal variation; *Vaccinium* sp.; DNA methylation; antioxidant

1. Introduction

Blueberries are referred to as 'superfruits' due to their health-promoting effects. Fruits of highbush blueberries contain high levels of water (84%) and carbohydrates (9.7%) and low concentrations of proteins (0.6%) and fats (0.4%). Moreover, they are also characterised by high vitamin C content at approx. 10 mg/100 g fresh weight, corresponding to one-third of the recommended daily intake [1]. *Vaccinium corymbosum* L. is one of the fruits with the highest antioxidant potential due to its high level of polyphenols [2]. The main group of phenolic compounds in blueberries is flavonoids. This group includes anthocyanins, proanthocyanidins, as well as flavonols (mainly quercetin derivatives). The present phenolic acids mainly include chlorogenic, coumaric and ellagic acids [3,4]. The content of polyphenols in berry plants ranges from 48 to 304 mg per 100 g fresh weight. It strictly depends on the variety, growing conditions and fruit ripening. The most common anthocyanins are malvidin, delphinidin, petunidin, cyanidin and peonidin, and they occur in combination with glucose, galactose and arabinose sugar molecules. The content of anthocyanins in the fruits of *Vaccinium* plants ranges from 25 to 495 mg/100 g of fresh fruit [1]. Metabolites contained in the fruits of *Vaccinium* sp. exhibit health-promoting [5–8], anti-inflammatory

and anticancer effects [5,9–11]. Regular consumption of blueberries is known to contribute to disease prevention [1,7,9,12].

Due to the taste and attractiveness of berries, the highbush blueberry quickly became an object of widespread interest among consumers, growers and fruit nurserymen [13]. Propagation of blueberries by conventional means can be performed by generative or vegetative methods. Generative reproduction has not been widely used in the cultivation of *Vaccinium* plants due to low seed yield, poor germination and low quality of seedlings [14,15] Vegetative propagation of blueberry plants relies on the use of semi-woody shoots collected from the donor plant, which are subsequently rooted under high-humidity conditions [16]. Technological progress and the development of in vitro culture techniques resulted in the production of berry plants, including highbush blueberries, under laboratory conditions [15,16]. The first in vitro cultures of highbush blueberry were initiated in 1979–1980 by Cohen and Elliot [17] and Zimmerman and Broome [18]. Currently, the in vitro culture method is the most widely used technique in plant biotechnology [19].

Both micropropagation and conventional plant propagation exhibit advantages and disadvantages. The main advantage of conventional methods of vegetative propagation of blueberry plants is the possibility of obtaining plant seedlings that are homogeneous with the mother plants. However, this method is slow and labour-intensive [15,19]. Other drawbacks include dependence on weather and a limited supply of seedlings dependent on the number of mother plants. These disadvantages render the conventional method insufficient to meet the growing demand for highbush blueberry plants [20]. The solution to the limitations of conventional highbush blueberry seedling production lies in the in vitro technique. In vitro propagation allows for the efficient multiplication of plants, enabling the generation of a substantial number of seedlings in a relatively short time. In vitro cultures carried out in laboratory conditions are independent of weather conditions and seasons, facilitating seedling production throughout the year. The main goal of micropropagation is to produce multiple plant clones with the same characteristics as the donor plants [19,21–23] Micropropagation, despite many advantages, is associated with obtaining plants that are heterogeneous compared to the mother plants. This manifested itself in various phenotypes and (epi)genetic properties [24–28], a phenomenon known as somaclonal variation [25–27,29–31]. Somaclonal variation occurring during in vitro cultures may be caused by pre-existing variability (explant-specific variability) or variability induced by in vitro culture [32–34]. The in vitro culture environment may therefore have a mutagenic effect. Plants obtained from in vitro culture of callus and protoplasts and through somatic embryogenesis may also show phenotypic changes caused by alterations at the DNA level [29]. Moreover, genetic stability during in vitro culture depends on numerous factors, including the type of medium [35], type and concentration of growth regulators [36,37], culture conditions (temperature, light, etc.) [28,38], culture duration and the number of passages [37,39]. The origin of the explants also influences somaclonal variation [40,41]. According to De Klerk [42], plants developed from adventitious shoots by indirect organogenesis are the most common carriers of somaclonal variation. It is worth noting that the appearance of adventitious shoots during in vitro cultures of many plants, including highbush blueberries, is a common, difficult-to-control phenomenon [18,43,44]. Adventitious shoots develop concurrently with axillary shoots and are often in close proximity. Some of them reach the size of axillary shoots at the end of the passage, making them difficult to distinguish [44]. According to Litwińczuk [44], tolerating adventitious shoots may lead to the uncontrolled selection of mutated or epigenetically altered cultures. The effect of somaclonal variation occurring during micropropagation of blueberry plants leads to both undesirable and desirable traits. Scientific reports [15,44,45] suggest that micropropagated *Vaccinium* plants may exhibit delayed entry into the fruiting period, lower yield in the first years, smaller fruit size and reduced fruit mass. Other studies indicated a smaller number of flowers in the inflorescences and fewer inflorescences themselves, but these features may also depend on species and variety [15,46–48]. On the other hand, other studies [15,44] indicated that blueberry plants originating from in vitro cultures exhibited

increased root production, more vigorous growth and enhanced branching compared to traditionally propagated plants.

To identify true-to-type plants obtained after in vitro propagation, it is essential to use methods for evaluating somaclonal variation [27]. A growing number of studies indicate that the variability observed in plants regenerated through in vitro cultures is not solely attributed to genetic alteration but also to epigenetic changes. Additionally, not all genetic variations in somaclones are phenotypically expressed. These modifications can either occur in non-coding sequences or may not significantly alter the gene product [27]. To detect somaclonal variation at the epigenetic level, characterised by alterations in DNA methylation, the methylation-sensitive amplification length polymorphism technique (MSAP) is most commonly applied [27,49–53]. This method is based on the different sensitivities of restriction enzymes to cytosine methylation at their cleavage sites [49,54]. It permits the comparison of the DNA methylation status of different organisms based on their differential digestion patterns. A frequently employed method involves HpaII and MspI isoschizomers, both recognising the same 5′-CCGG sequence. Although HpaII cleaves hemimethylated sequences (only one DNA strand is methylated), it is most sensitive when one or both cytosines are fully methylated (both strands are methylated). In contrast, MspI cleaves at the C5mCGG site, whether hemimethylated or with both strands methylated, but does not cleave at the 5mC5mCGG or 5mCCGG sites [53]

Previous studies on the impact of micropropagation on plants of the genus *Vaccinium* have mainly focused on lowbush blueberries, with fewer investigations conducted on highbush blueberries. These studies focused on the comparison of conventionally propagated plants with plants derived from in vitro cultures, without taking into account the type of culture [15,45–47]. Furthermore, existing research did not analyse comprehensively the impact of propagation methods on plants and fruits, focusing on morphological [15,44–47,55] differences mainly, or less, on physiological [56], biochemical [15,46,48] or epigenetic differences [57]. Moreover, there are no scientific reports presenting analyses conducted on a wide group of differentially propagated plants. Therefore, the present study aimed to demonstrate the differences between highbush blueberry plants ('Brigitta' cultivar) and their fruits, propagated conventionally, in vitro (with different types of explants) and using a combination of these methods (in vitro and conventional). Differences between plants were determined at the morphological, physiological and epigenetic levels, whereas differences between fruits were determined based on analyses of bioactive compounds.

2. Results

The analysis demonstrated that the propagation method (in vitro or conventional) of the highbush blueberry cultivar 'Brigitta Blue' (*Vaccinium* × *corymbosum* L.) significantly influenced both the plants and the fruits. Differences were detected at the morphological, physiological, epigenetic and biochemical levels.

2.1. Plant Analysis

Morphological Measurements

The studied group of blueberry plants displayed significant differences in growth strength and branching. Variations were observed in the number of shoots (both main and lateral), the average length of main shoots and the maximum length of main shoots. Plants propagated in vitro (TC-Ax and TC-Ad) and those propagated in vitro first and then conventionally (TC/SC) showed higher values of the examined parameters compared to plants propagated only conventionally (SC) (Table 1).

The highest values of the analysed features were mainly recorded in TC/SC plants, including the maximum length of shoots, the number of main and lateral shoots and the total number of shoots. In contrast, SC plants consistently demonstrated the lowest values (Table 1), whereas plants obtained directly from in vitro culture (TC-Ax and TC-Ad) showed a similar number of shoots. However, TC-Ax plants surpassed TC-Ad and other groups in terms of the average length of main shoots (Table 1).

Table 1. Shoot measurements of highbush blueberry plants propagated by in vitro and conventional methods.

Parameters	TC-Ax	TC-Ad	TC/SC	SC
Average length of main shoots (cm)	39.4 ^{a}	33.6 ^{c}	37.8 ^{ab}	34.5 ^{bc}
Maximum length of main shoots (cm)	61.5 ^{b}	57.2 ^{bc}	67.3 ^{a}	53.3 ^{c}
Number of main shoots	5.5 ^{a}	6.1 ^{a}	6.1 ^{a}	3.8 ^{b}
Number of lateral shoots	11.9 ^{b}	11.5 ^{b}	16.0 ^{a}	11.0 ^{b}
Total number of shoots	17.5 ^{b}	17.6 ^{b}	22.1 ^{a}	14.8 ^{c}

The mean value marked with different lowercase letters differs significantly at $p \leq 0.05$.

2.2. Chlorophyll Content and Fluorescence

The physiological measurements of chlorophyll (a and b) content and the efficiency of the light phase of photosynthesis revealed differences between individual groups of highbush blueberry plants (Figures 1 and 2). Fluorometric analyses confirmed different efficiencies of the light phase of photosynthesis between the plant groups studied. Statistically significant differences were observed in minimum (F0), maximum (Fm) and variable fluorescence (Fv), as well as in the efficiency of the PSII photosystem (Fv/Fm) (Figure 1).

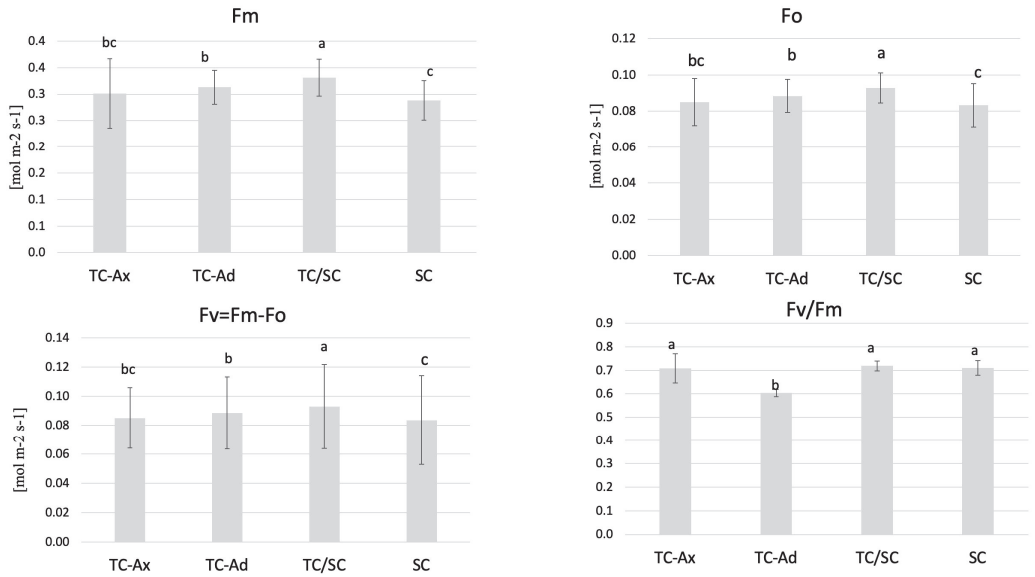


Figure 1. Fluorescence parameters. The mean value marked with different lowercase letters differs significantly at $p \leq 0.05$.

The collected data indicated that TC/SC highbush blueberry plants exhibited the highest fluorescence values: minimum (F0), maximum (Fm) and variable (Fv) fluorescence. Conventionally propagated blueberry plants (SC) displayed significantly lower fluorescence parameters, except for Fv/Fm. However, blueberry plants derived directly from in vitro (TC-Ax and TC-Ad) cultures demonstrated similar values of these parameters (F0, Fm and Fv). Nonetheless, TC-Ad plants indicated the lowest efficiency of the PSII photosystem (Fv/Fm) compared to TC-Ax plants, as well as other groups (TC/SC and SC). Measurements of chlorophyll a and b content revealed statistically significant differences between plants propagated using different methods (Figure 2). Plants directly derived from in vitro cultures (TC-Ax and TC-Ad) and those propagated using combined methods (TC/SC) showed similar levels of chlorophyll a, b and a + b. SC plants, on the other hand, had significantly lower chlorophyll a and a + b contents (Figure 2).

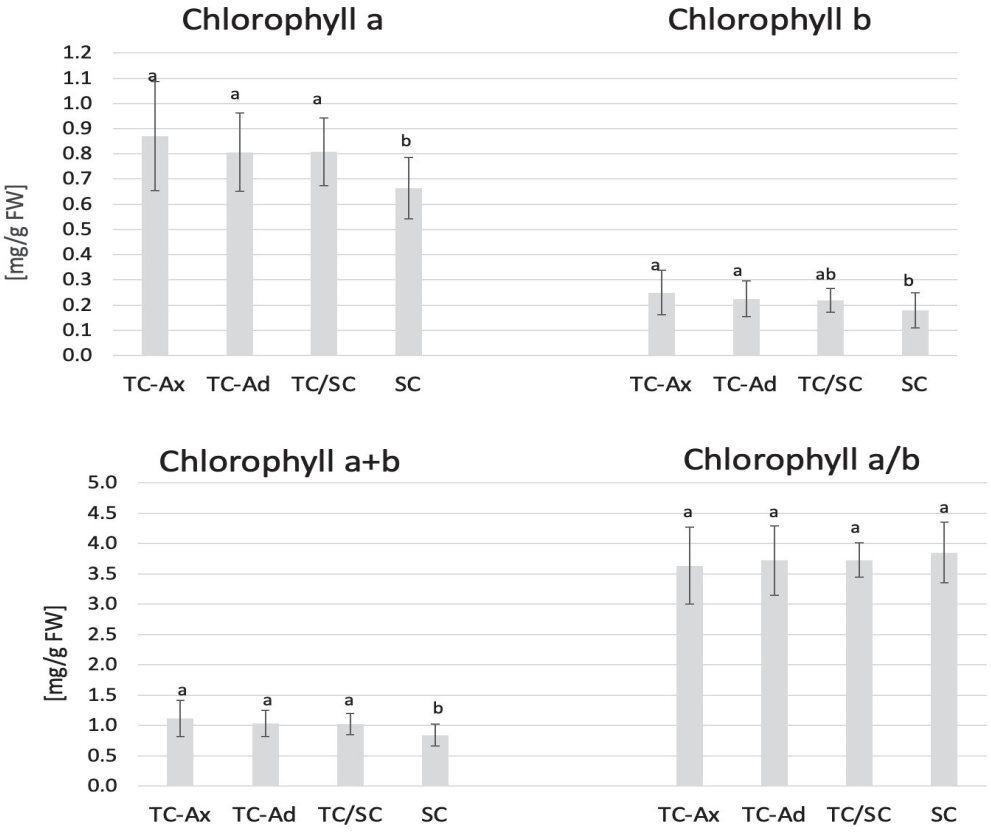


Figure 2. Chlorophyll content. The mean value of chlorophyll marked with different lowercase letters differs significantly at $p \leq 0.05$.

2.3. Fruit Analysis

Fruits were picked up from propagated in vitro and conventionally 5-year-old blueberry plants and stored at −80 °C. Homogenised fruits were used for analyses of their antioxidant potential. The result indicated differences in some parameters between in vitro and conventionally propagated blueberry plants.

2.3.1. Antioxidant Activity

The value of the antioxidant activity varied between blueberry plants propagated using different methods. The highest DPPH value was observed in SC (18, 57) and TC-Ad plants (18, 35), whereas the ABTS value was higher in TC-Ad and TC/SC plants. However, a lower level of antioxidant activity (DPPH and ABTS) was recorded for TC-Ax plants (Figure 3).

2.3.2. Polyphenols and Anthocyanins

The determination of polyphenol and anthocyanin contents also showed that SC plants had the highest concentrations of these compounds, while TC-Ax plants (as well as TC/SC) had their lowest concentrations (Figure 4).

2.3.3. Ascorbic Acid

The ascorbic acid content varied in blueberry fruits from plants obtained directly from in vitro culture, propagated with different types of explants (axillary or adventitious shoots). TC-Ad plants had the highest level of ascorbic acid, while TC-Ax had the lowest level. These values differed significantly from each other as well as from other groups of plants (TC/SC and SC). The analysis of the ascorbic acid levels indicated similar values of this parameter for plants propagated conventionally (SC) or using combined methodologies (in vitro propagation followed by conventional cultures) (TC/SC) (Figure 5).

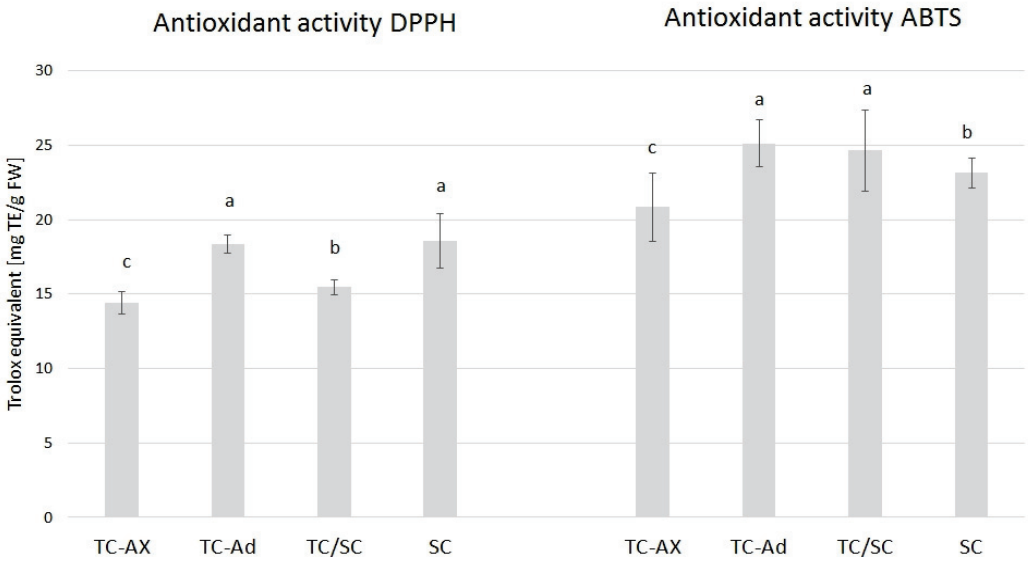


Figure 3. Results of antioxidant activity. The mean value marked with different lowercase letters differs significantly at $p \leq 0.05$.

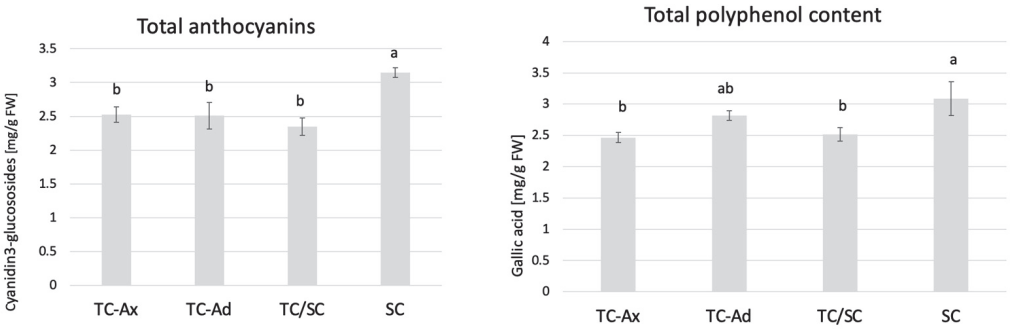


Figure 4. Total anthocyanins and total polyphenol contents. The mean value marked with different lowercase letters differs significantly at $p \leq 0.05$.

2.3.4. DNA Methylation

Molecular analyses were performed using MSAP markers. This method utilises *Eco*RI/*Hpa*II and *Eco*RI/*Msp*I restriction enzymes with different sensitivity to cytosine methylation within the recognised 5′CCGG3′ DNA sequence. MSAP analysis produced polymorphic DNA fragments, distinguishing individual plant groups from each other (Figure 6). Notably, polymorphisms in DNA band patterns were also detected between plants derived from in vitro culture propagated by axillary (TC-Ax) or adventitious shoots (TC-Ad) (Figure 6B,C).

DNA fragments characterising M-type (fully/symmetric methylation of the 5′CCmGG3′ sequence) and H-type (hemimethylation of the 5′CmCGG3′ sequence) methylation events were identified according to the procedure described by Xiong et al. [54] and Walder et al. [58]. For all analysed groups of blueberry plants, MSAP analysis yielded a significantly higher number of amplification products, indicating M-type methylation sites compared to H-type methylation events (Table 2).

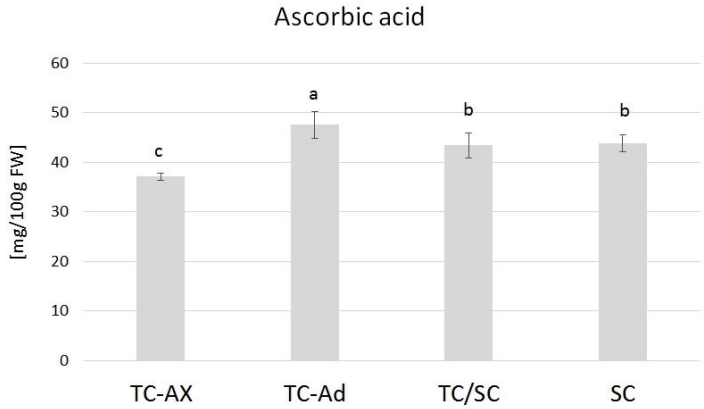


Figure 5. Ascorbic acid content. The mean value marked with different lowercase letters differs significantly at $p \leq 0.05$.

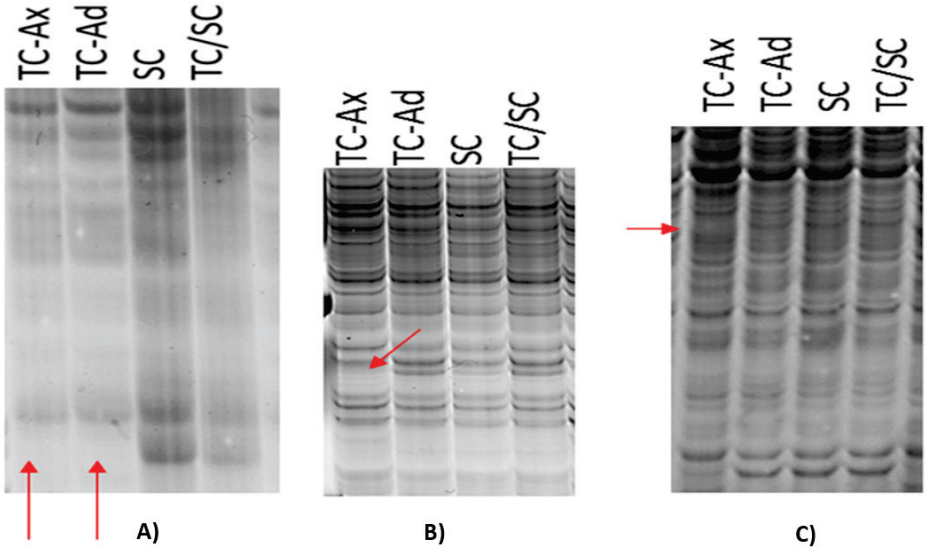


Figure 6. Differences in DNA band patterns—products of selective amplification as part of MSAP analysis. Red arrows indicate the lack of specific DNA bands for in vitro-derived plants (TC-Ax and TC-Ad) (**A**) or for TC-Ax plants only (**B**,**C**).

Table 2. Methylation frequency based on MSAP analysis.

Type of Propagation	TC-Ax	TC-Ad	TC/SC	SC
Types of Methylation	**Methylation Frequency (%)**			
M	13.2	14.3	14.0	13.0
H	10.6	10.5	10.5	10.6
M + H	23.8	24.8	24.5	23.6

Methylation frequency analysis showed different levels of methylation between blueberry plants propagated conventionally or in vitro, as well as using a combination of these methods. The highest methylation frequency was recorded for TC-Ad plants, while TC/SC plants had slightly lower levels of methylation. On the other hand, the lowest occurrence of methylation events was found in traditionally propagated SC plants and in vitro-derived plants (TC-Ax) (Table 2). Interestingly, plants from in vitro culture, propagated through axillary or adventitious shoots, demonstrated different methylation frequencies (Table 2). These findings underscore the presence of somaclonal variation during in vitro culture.

3. Discussion

Historically, the primary objective of micropropagation was to obtain true-to-type plants faithfully replicating the characteristics of the donor plant in a relatively short time. In scientific terms, in vitro culture methods for plants were generally regarded as a means of cloning a specific genotype [23,28,57]. Consequently, in breeding programmes, plants with desirable agronomic traits were selected for micropropagation to quickly obtain clones with the same properties (i.e., plants morphologically and genetically identical to the mother plant) [22,23,59].

Compared to the conventional propagation method, micropropagation is a more effective, reproducible and season-independent approach [21,23]. As a result, micropropagation has rapidly become, if not the primary technique, a prominent method for propagating large numbers of plants, including those belonging to the genus *Vaccinium* sp. [43,57].

For most nurserymen and cultivators, using the vegetative propagation method is important to preserve key agronomic features that characterise an elite variety [22,59]. However, the occurrence of somaclonal variation makes it difficult to meet this requirement. Uncontrolled changes during in vitro cultures can result in different phenotypes with both undesirable and improved features [30]. Nevertheless, when in vitro cultures are used to clone true-to-type plants with specific and selected characteristics, any changes at the morphological and/or molecular level are not desirable [29]. Regarding micropropagated blueberry plants, it has been proven so far that emerging undesirable traits include lower yield, delayed fruiting, smaller fruit size and fewer flowers on the inflorescence [15,45].

The present study has confirmed the hypothesis that the method of plant propagation affects both the plants and the fruits of highbush blueberries. The results indicated differences among plants propagated conventionally (by semi-woody cuttings; SC plants) and those propagated through the in vitro method using various explants (axillary shoots, TC-Ax, or adventitious shoots, TC-Ad), as well as plants propagated through a combined approaches (in vitro followed by conventional propagation; TC/SC plants). Variations between the analysed plant groups were detected at the morphological, physiological, epigenetic and biochemical levels. Specifically, SC plants exhibited notably lower values in terms of the number of main and total shoots (Table 1) and chlorophyll a and a + b contents (Figure 2) compared to other groups. Conversely, TC/SC plants displayed the highest values for parameters such as the maximum length of main shoots and the number of total and lateral shoots (Table 1). TC-Ax plants demonstrated the statistically significantly highest value for the average length of main shoots. Highbush blueberry plants directly originating from in vitro cultures (propagated by axillary or adventitious shoots) generally displayed intermediate values of the analysed parameters when compared to SC and TC/SC plants. Notably, differences between plants propagated through in vitro culture and those propagated conventionally were also previously observed by Goali et al. [46] and Litwińczuk [44]. Goyali et al. [46] highlighted a significant interaction between the propagation method, plant branching and chlorophyll content in lowbush blueberries. Similar observations were reported by Litwińczuk et al. [55]. The latter authors noted that highbush blueberry plants derived from softwood cuttings exhibited slower growth, produced significantly fewer and shorter shoots and showed greater variability compared to micropropagated plants. Such differences between cutting-derived and micropropagated blueberry plants were also observed earlier by Grout et al. [60] and El-Shiekh et al. [61]. The aforementioned differences could be the result of the juvenile characteristics of micropropagated plants [43,57], which may persist in the long term and may facilitate rapid establishment of micropropagated highbush blueberry plants in a new planting area. According to El-Shiekh et al. [61], the enhanced branching and spreading characteristics of tissue culture-derived blueberry plants can persist for up to 10 years. Scientific reports indicate that plants obtained by micropropagation manifested juvenile features, including high rooting capacity of shoots and intensive vegetative growth [43,57]. Research performed on *Vaccinium* sp. plants has validated these phenomena in highbush and lowbush blueberries [15,43–47,55], lingonberries [48,62,63] and cranberries [64,65]. However, most

authors of the cited studies did not consider the origin of the cultures (derived from axillary or adventitious shoots) and focused solely on comparing plants from in vitro cultures to conventionally propagated ones.

In the present study, we demonstrated variations in chlorophyll content and fluorescence between plants derived from in vitro cultures and those propagated conventionally. Differences were also observed in the reaction to stress conditions of highbush blueberry plants subjected to diverse propagation methods [56]. Many authors [66–68] have indicated that in vitro culture conditions may affect photosynthetic processes. Some researchers have observed that in vitro conditions induce alterations in the shape of chloroplasts, starch accumulation or irregular orientation of the thylakoid system [68,69]. Therefore, such abnormalities are believed to be responsible for the reduced photosynthetic efficiency of regenerated plants [70]. The diverse fluorescence values obtained in this study for micropropagated plants and traditionally propagated plants may be due to the application of growth regulators during in vitro cultures [57]. This notion is supported by the studies of Stefanova et al. [68] and Dobránszki and Drienyovszki [66], who demonstrated that growth regulators could modify the morphology and anatomy as well as the functioning of the photosynthetic apparatus.

Analysis of antioxidant parameters of fruits, such as antioxidant activity and levels of ascorbic acid, total polyphenols and anthocyanins, also revealed differences between highbush blueberry plants propagated conventionally and by using in vitro methods (Figures 3–5). Interestingly, conventionally propagated blueberry plants (SC) frequently demonstrated superior values of the analysed parameters (total polyphenolics and anthocyanins and antioxidant activity—DPPH activity) (Figures 3 and 4) compared to other groups of plants. Equally high (total anthocyanins and antioxidant activity—DPPH activity) or even higher values (ascorbic acid level—ABTS activity) were recorded for blueberry plants derived from in vitro cultures using adventitious shoots as explants (TC-Ad) (Figures 3 and 4). It is noteworthy that plants propagated in vitro by axillary shoots (TC-Ax) as explants demonstrated the lowest values in almost all analysed parameters (antioxidant activity, total polyphenol and anthocyanin contents, as well as ascorbic acid levels) (Figures 3–5). Goyali and colleagues published two reports in 2013 [46] and 2015 [15], providing evidence for a strong and positive correlation between total phenolic and anthocyanin contents and antioxidant activity in blueberries.

There are numerous literature reports highlighting the influence of various factors on the quality of blueberry fruit. For *Vaccinium* plants, the content of bioactive compounds, such as vitamin C, sugars and other components like minerals, depends on geographical location, soil, climatic and habitat conditions, crop care and harvest date [16]. However, there is limited scientific literature addressing the impact of the propagation methods of blueberry plants on the content of bioactive compounds. Goyali et al. [46] indicated that the propagation method exerted a significant effect on total polyphenols in lowbush blueberries. However, the next detailed analyses performed by the same group [15] demonstrated that micropropagated lowbush blueberry plants were characterised by a higher content of polyphenols and flavonoids compared to traditionally propagated plants [28,71]. The latter findings contradict the results obtained in the presented research for highbush blueberry plants, which can be attributed to different combinations and concentrations of plant growth regulators (PGR) applied during in vitro cultures. Various types and concentrations of PGRs have been reported to exert diverse regulatory effects on developmental processes and concentrations of secondary metabolites in plants [72,73]. Hormone concentration is the major factor in secondary product accumulation, such as phenolics and flavonoids [74–76]. Plant hormones are chemical compounds and a group of key signal molecules that are actively involved in the synthesis of plant secondary metabolites and also in regulating development and plant growth [74–76]. According to Baskaran et al. [73], a combination of glutamine and N6-benzyladenine significantly increased the accumulation of phenolics and flavonoids in vitro compared to the separate application of these compounds. In our study, only cytokinins, specifically 2iP (6-γ,γ-dimethylallylamine) at a

concentration of 10 mg/L, were used in the medium. Research performed by Al-Khayri et al. [74] indicated that cell suspension cultures of date palm (*Phoenix dactylifera* L.), containing 2,4-D and 2iP, yielded the maximum accumulation of phenolics only in the case when 2iP is combined with 2,5-D at a concentration of 2.5 mg/L (2iP) and 5 mg/L (2,4-D). Whereas the cell suspension culture medium supplemented with a higher concentration of auxin/cytokinin (10 mg/L 2,4-D + 5 mg/L 2iP) led to the accumulation of the least concentration of the total phenolic content and flavonoids of date palm. On the other hand, Bairu et al. [29] and Vitamvas et al. [77] have documented a clear association between culture type, in vitro morphogenesis and the incidence of somaclonal variation. Leva et al. [31] categorised cultures based on genetic stability, ranking those established by explants with pre-formed apical and lateral meristems, that is, growth tips and nodal shoot sections, as the most stable. Following these were adventitious shoot cultures started from de novo meristems, particularly through direct organogenesis, somatic embryo cultures developed by direct embryogenesis, organ cultures regenerated from callus (direct morphogenesis), suspension cultures (of cells) and protoplast cultures. The appearance of callus and adventitious shoots during blueberry in vitro cultures is usually considered a potential source of somaclonal variation [44]. Comparing TC-Ax plants to TC-Ad ones, we showed statistically significant differences for parameters such as PSII photosystem efficiency (Fv/Fm) and the average length of main shoots. These parameters reached higher values for TC-Ax plants compared to TC-Ad and SC plants (average main shoot length).

The conducted research revealed that plants derived from tissue cultures and propagated by axillary shoots (TC-Ax plants) exhibited more significant differences from conventionally propagated plants (SC) than those propagated by adventitious shoots (TC-Ad). This observation is intriguing, as adventitious shoot cultures are commonly considered to be a source of somaclonal variation. Perhaps the researchers' focus on finding evidence of somaclonal variability in plants obtained from adventitious shoots, generally considered as a source of somaclonal variation, was misplaced. Debnath [78], in his molecular analyses of plants obtained from adventitious shoots, did not include plants obtained from axillary shoots (commonly regarded as being genetically stable). Soneji et al. [79], on the other hand, observed phenotypic differences in pineapple micropropagation using axillary buds, particularly in fruit colour and thorn generation [79]. According to the aforementioned authors, cultures derived from axillary shoots are the source of somaclonal variation. The reason for differences between plants propagated in in vitro cultures may be attributed to changes occurring at the genetic and/or epigenetic level.

Molecular analyses conducted by researchers focusing on determining the clonal identity of plants of the genus *Vaccinium* confirmed the absence of differences between clones of adventitious origin of lowbush blueberry [78], as well as between cultures and mother plants of highbush blueberry [80]. However, it is crucial to note that these analyses applied EST-PCR [78] and RAPD [80] molecular markers. These techniques are applied to detect alterations in DNA sequence; epigenetic analyses, on the other hand, are designed to detect changes in DNA structure, encompassing mechanisms such as gene silencing or activation, often influenced by changes in DNA methylation patterns [28,49,81]. Therefore, the methylation process plays a pivotal role in gene expression in eukaryotes, influencing various aspects of plant growth and development as well as responses to stress [28,50,81]. In vitro cultures create conditions conducive to changes in gene expression. Hence, it can be inferred that obtaining micropropagated plants different from the initial plant may not necessarily involve mutations but could result from a modified DNA methylation profile [57].

Epigenetic analyses based on the determination of DNA methylation using MSAP markers showed different methylation levels among the analysed groups of differentially propagated highbush blueberry plants (Table 2). A prevalence of symmetric or fully methylated 5′CmCGG 3′ DNA sequence (Type M) over hemimethylated 5′mCCGG 3′ DNA sequence (Type H) was found in all plant groups (Table 2). This predominance of total (symmetric) cytosine methylation over hemimethylation has been previously confirmed in

other scientific reports [51–53]. Goyali et al. [51] showed a higher frequency of symmetric methylation events compared to hemimethylated events in both conventionally and in vitro propagated lowbush blueberry plants. However, the highest degree of methylation (frequency) was found in TC-Ad plants, slightly lower in TC/SC plants and lowest in SC plants and TC-Ax (Table 2).

DNA methylation has been recognised as a major regulatory epigenetic mechanism, closely associated with diverse gene functions [51]. Epigenetic control of gene expression plays a fundamental role in cell division and differentiation, influencing various functions [51,82]. This regulatory system of gene expression is related to developmental stage, tissue type, environmental conditions and vigour [83]. Changes in methylation patterns are therefore fundamental to plant development processes [84]. Consequently, it can be hypothesised that the observed differences between the plants propagated conventionally (SC), using in vitro techniques (TC-Ad and TC-Ax), as well as by a combination of these approaches (TC/SC), at the morphological, physiological and biochemical levels, may be attributed to variations in DNA methylation levels.

4. Materials and Methods

The experiments included four groups of nursery plants of the highbush blueberry (*Vaccinium* × *corymbosum* L.) cv. 'Brigitta Blue'. Two groups were obtained from in vitro cultures propagated through axillary (TC-Ax) or adventitious shoots (TC-Ad). The control group of plants (SC) was obtained by rooting semi-woody cuttings derived from continuously conventionally propagated plants. An additional fourth group (TC/SC) was also included, consisting of plants obtained by rooting semi-woody cuttings from conventionally propagated plants originally derived from in vitro cultures. All groups contained approximately 40 plants.

4.1. In Vitro Conditions

For micropropagation, a modified ZB medium (Z-2) [18] was used with the addition of L-cysteine (5 mg/L), vitamin C (100 mg/L), sucrose (25 g/L) and fructose (5 g/L). The medium was supplemented with the following growth regulators: cytokinins (10 mg/L 2iP; 6-γ,γ-dimethylallylamine), adenine sulphate (80 mg/L) and auxin (4 mg/L IAA; indolyl-3-acetic acid). Stabilised highbush blueberry cultures, maintained in in vitro culture for 3 years, were used as the starting material for TC-Ax and TC-AD in vitro cultures. The donor plant material for the initiation of the in vitro stabilised culture consisted of one-year-old nursery pot plants grown in a glasshouse. They were obtained conventionally by rooting cuttings collected from mother plants in the previous year. The mother plants were propagated in the same manner for at least 2 generations. The explants for in vitro culture initiation were prepared from new shoots/accretions taken from mother plants grown in a glasshouse. Two types of in vitro cultures were established from 2-node shoot explants of different origins. For the TC-Ax in vitro culture, a 2-node axillary shoot fragment, strongly connected with the donor culture (by vascular bundles), was used as an explant. Axillary shoots were developed from a bud formed in the leaf axils. For the TC-Ad in vitro culture, a 2-node adventitious shoot fragment that had developed from a callus or was weakly connected with the donor culture (without vascular bundles) was used as an explant. Examples of explants and in vitro cultures are included in supplementary material. The highbush blueberry cultures were maintained for 10–12 weeks and then passaged to fresh ZB medium. Four passages were performed. Plant cultures were maintained under a 16-h photoperiod, with light provided by cool-white fluorescent lamps (photosynthetic photon flux density (PPF) of approximately 12 μmol m^{-2} s^{-1}) at 24 ± 2 °C.

4.2. In Vivo and Field Conditions

All plant groups were obtained by in vivo rooting of 3-node shoot fragments derived from donor plants (SC and TC/SC) or in vitro cultures (TC-Ax and TC-Ad). Rooting was

carried out in 'mini-greenhouses' under high relative air humidity at 26 °C. The substrate used for rooting was a mixture of peat and perlite in a ratio of 3:1 (*v*/*v*).

The rooted shoots were placed in pots (9 x9 cm, initially with a capacity of 0.5 L and 1.5 L afterwards). After three years, the plants were transplanted into the ground. The blueberry plants grew in a substrate prepared from a mixture of high peat with a pH of 3.5–4.0 and sand in a ratio of 2:1 (*v*/*v*). The plants were watered at 1–3 day intervals (depending on climatic conditions) and fertilised from April to September with a 0.3% solution of 'Florovit' (IncoVeritas, Góra Kalwaria, Poland) containing the macronutrients (N, K and P (4:6:6)) and micronutrients(B, Cu, Fe, Mn and Zn). The plants grew under field conditions in a sunny location protected from the wind in south-eastern Poland, in Karolówka, Podkarpackie Voivodeship.

The analyses were performed in the following periods: 2014—obtaining two types of in vitro culture of highbush blueberries, TC-Ax and TC-AD, and passaging to fresh medium; 2015—rooting of blueberry shoots derived from donor plants; 2018—morphological, physiological and epigenetic analyses; 2020—fruit collection; 2023—analysis of bioactive compounds.

4.3. Plant Analysis

4.3.1. Morphological Measurements

Morphological analysis was performed on 3-year-old seedlings before transplanting them from pots to the ground. Measurements included the length of each shoot and the total number of shoots, with only those exceeding a minimum length of 3 cm taken into account.

4.3.2. Chlorophyll Content and Fluorescence

An IMAGING-PAM M-Series Chlorophyll Fluorometer (MINI version) (Waltz GmbH Eichenring, 6 91090 Effeltrich, Germany) was used to record chlorophyll fluorescence parameters. The leaves were kept in the dark for at least 20 min before measurements. The analyses were performed in 4 replicates for 20 leaves from randomly selected plants from each group. The following parameters were determined: minimum fluorescence (F0) and maximum fluorescence (Fm). Based on these measurements, the variable fluorescence (Fv = Fm − F0) and PSII photosystem efficiency (Fv/Fm) were calculated.

The content of photosynthetic pigments (chlorophyll a, chlorophyll b and total a + b) was determined on 20 leaves sampled from randomly selected plants of different origins. Pigments were extracted from 50 mg of leaf tissues with 2.5 mL of DMSO (dimethyl sulfoxide, Toruń, Poland). Absorbance measurements were performed at 665 and 649 nm using a Specol Aquamate VIS device (Thermo Scientific, Waltham, MA, USA). The content of individual pigments was calculated according to the formulas given below and described earlier by [85]:

- chlorophyll a = $[(12.9 \cdot A663) - (3.45 \cdot A649)]$.
- Chlorophyll b = $[(21.91 \cdot A649) - (5.32 \cdot A663)]$.

where A is absorbance at a specific wavelength.

4.4. DNA Methylation

DNA methylation was determined in young, fully developed leaves of highbush blueberry plants collected from 40 plants in each analysed group. DNA was isolated from pooled samples according to the protocol of Doyle and Doyle [86]. The methylation-sensitive amplification polymorphism (MSAP) technique, based on the protocols described in [49,54], was used. The set of primers utilised for MSAP analysis is listed in Table 3.

DNA (50 ng) from each sample was digested with EcoRI/MspI (Thermo Scientific, Waltham, MA, USA) and EcoRI/HpaII (Thermo Scientific, Waltham, MA, USA) restriction enzymes and ligated with EcoRI- and MspI- or HpaII-specific adapters (Genomed, Warsaw, Poland) at 37 °C for six hours. The ligated DNA was diluted 10 times and pre-amplified using EcoRI and MspI or HpaII primers (Genomed, Warsaw, Poland). The PCR conditions were as follows: 94 °C—1 min, 65 °C—1 min and 72 °C—1 min for 35 cycles, with a

final extension at 72 °C for 7 min. The pre-amplified product was diluted ten times with Tris-EDTA (TE) buffer and selectively amplified with different combinations of EcoRI and MspI or HpaII MSAP primers (Genomed, Warsaw, Poland) (Table 3), each with two to three selective nucleotides.

Table 3. Sequences of primers and adapters used in MSAP analysis.

MSAP Stage	Primer/Adapter	Sequence
Ligation	EcoRI-Adapter	5′CTCGTAGACTGCGTACC 3′ 3′CATCTGACGCATGGTTAA 5′
	MspI-HpaII-Adapter	5′CGACTCAGGACTCAT3′ 3′TGAGTCCTGAGTAGCAG5′
Preamplification	Pre-EcoRI	5′GACTGCGTACCAATTC 3′
	Pre-MspI-HpaII	5′GATGAGTCCTGAGTCGG 3′
Selective amplification	EcoRI-ACT	5′GACTGCGTACCAATTCACT 3′
	EcoRI-AG	5′GACTGCGTACCAATTCAG3′
	EcoRI-AC	5′GACTGCGTACCAATTCAC 3′
	EcoRI-AT	5′GACTGCGTACCAATTCAT 3′
	MspI/HpaII-ATG	5′GATGAGTCCTGAGTCGGATG3′
	MspI/HpaII-CTA	5′GATGAGTCCTGAGTCGGCTA3′
	MspI/HpaII-CTC	5′GATGAGTCCTGAGTCGGCTC3′
	MspI/HpaII-CAT	5′GATGAGTCCTGAGTCGGCAT3′
	MspI/HpaII-CT	5′GATGAGTCCTGAGTCGGCT3′
	MspI/HpaII-GT	5′GATGAGTCCTGAGTCGGGT3′
	MspI/HpaII-CA	5′GATGAGTCCTGAGTCGGCA3′

For selective amplification, PCR conditions were as follows: 94 °C—1 min, 65 °C—1 min (temperature was reduced by 0.7 °C in each successive cycle) and 72 °C—1 min for 11 cycles, and 94 °C—1 min, 56 °C—1 min and 72 °C—1 min for 23 cycles, with a final extension at 72 °C for 7 min. An equal volume of formamide dye was added to the PCR products, which were subjected to electrophoretic separation on a 6% denaturing polyacrylamide gel. Gels were stained with silver nitrate (Avantor Performance Materials Poland S.A., Gliwice, Poland (according to the method described by Bassam and Gresshoff [87]) and subsequently scanned for data recording.

Methylation Analysis

Methylation analysis was performed according to the method outlined by Walder [58] and Xiong et al. [54]. The presence of bands from the EcoRI + MspI (M) primer mixture and their simultaneous absence in the EcoR I + HpaII (H) reaction indicated DNA methylation. The internal cytosine of the 5′CCGG 3′ sequence was methylated (5′CmCGG 3′). This is referred to as 'symmetric or full methylation'. The absence of bands from the EcoRI + MspI (M) reaction and their simultaneous presence in the EcoR I + HpaII (H) reaction primer mixture indicated DNA methylation, specifically the methylation of the external cytosine of one DNA strand (5′mCCGG 3′). This phenomenon is referred to as the 'hemimethylated state'. Methylation frequency was calculated following the method described by Li et al. [53] as follows:

$$\text{Methylation (\%)} = (\text{number of methylated bands})/\text{total number of bands}) \times 100.$$

4.5. Fruit Analysis

Fruits were harvested from each group of 5-year-old blueberry plants and stored at −80 °C until analysis. The analyses were conducted on ten fruits with similar shapes and weights in 3 replicates. Homogenised fruits were used for analyses of basic quality parameters and antioxidant potential.

4.5.1. Antioxidant Activity

The sample for determining antiradical activity (AA) was prepared by homogenising the tissue (5 g) with 15 mL of 75% methanol solution and centrifuging the homogenate at 10,000× *g* for 30 min.

The antiradical activity (AA) of fruits was assayed using synthetic ABTS●+ and DPPH● radicals, following the method described by Piechowiak et al. [3]. The results of AA were expressed as Trolox equivalents (mg TE/g FW).

Ascorbic acid content was analysed using the 2,6-dichlorophenolindophenol/spectrophotometric method as outlined by Piechowiak et al. [4]. Ascorbic acid from the fruits was extracted by homogenising the fruit tissue (5 g) with 15 mL of 2% oxalic acid and centrifuging at 10,000× *g* for 30 min. The supernatant was used to detect the ascorbic acid content. The absorbance of the reaction mixture was measured at 500 nm, and the results were presented as ascorbic acid per 100 g fresh weight (FW) of fruit tissue.

4.5.2. Total Polyphenol and Anthocyanin Contents

Approximately 8 g of frozen berries of equal size were homogenised in an ice-cooled mortar with 20 mL of cold extraction solution EtOH/formic acid/H_2O (25/2/73). The homogenate was placed in an ultrasonic bath for 15 min. The suspension was centrifuged at 10,000× *g* for 20 min at 4 °C, and the supernatant was collected. The pellet was resuspended in 10 mL of extraction solution and treated as described above. Then, the supernatants were combined, and the volume was adjusted to 40 mL using the extraction solution. All extracts were stored at −80 °C before spectrophotometric analysis. For each group of plants, 3 pools of berries were extracted, yielding 3 biological replicates.

The total polyphenol content was determined following the Folin-Ciocalteau method [88]. To 0.05 mL of the extract in a 10 mL volumetric flask, 0.05 mL of Folin-Ciocalteau reagent, 4.45 mL of distilled water, and 2 mL of 10% Na_2CO_3 were added and immediately diluted with distilled water. The optical density was measured after 90 min at 700 nm using a 10S UV–vis spectrophotometer (Thermoscientific, Waltham, MA, USA). Results were expressed as milligrams of gallic acid per 1 g fresh weight (FW).

The total anthocyanin content was determined using a 10S UV–vis spectrophotometer (Thermoscientific, Waltham, MA, USA) by measuring the absorption peak of anthocyanin pigments at 530 nm. The total anthocyanin content was expressed as cyanidin 3-glucoside equivalent using a molar extinction coefficient of 26.9 L mol^{-1} cm^{-1} and reported as milligrams per 1 g FW.

4.6. Statistical Analysis

Statistical analysis was performed using Statistica (Stat Soft, Krakow, Poland, 13.1 version). An ANOVA test was used to identify significant differences between groups. Tukey's HSD post hoc test was performed to determine and verify differences at a significance level of $p \leq 0.05$.

5. Conclusions

In summary, the result indicated that the method of blueberry plant propagation (conventional and in vitro) influenced plant seedlings, including shoot length and branching, chlorophyll content and fluorescence, as well as DNA methylation. Additionally, differences were observed in the concentrations of fruit antioxidant compounds, including total polyphenols and anthocyanins, ascorbic acid and antioxidant activity. These differences were primarily observed between conventionally propagated highbush blueberry plants and those propagated in vitro. However, some differences were also recorded between plants derived from in vitro cultures with axillary (Tc-Ax) or adventitious shoots (TC-Ad) used as explants. The differences between the analysed groups of plants may be the result of somaclonal variation occurring during the micropropagation process. The underlying factor for somaclonal variation could be attributed to changes in DNA methylation levels detected in the analysed plant groups. It can be hypothesised that the detected differences

between conventionally and in vitro propagated highbush blueberry plants may be the result of differences in DNA methylation levels. Nevertheless, further analyses should be performed to confirm the present results for other varieties of highbush blueberry plants.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25010544/s1.

Author Contributions: Conceptualisation, M.M. and W.L.; methodology: M.M., T.P., A.S. (Anna Spinardi) and A.S. (Aleksandra Siekierzynska); software, A.S. (Aleksandra Siekierzynska); validation, M.M., T.P. and A.S. (Anna Spinardi); formal analysis, M.M., T.P., A.S. (Aleksandra Siekierzynska) and A.S. (Anna Spinardi); investigation, M.M. and T.P.; resources, M.M.; data curation, M.M.; writing—original draft preparation, M.M.; writing—review and editing, M.M. and A.S. (Aleksandra Siekierzynska); visualisation, A.S. (Aleksandra Siekierzynska); supervision, W.L.; project administration, A.S. (Aleksandra Siekierzynska); funding acquisition, W.L. and A.S. (Aleksandra Siekierzynska) All authors have read and agreed to the published version of the manuscript.

Funding: The project is financed by the program of the Minister of Education and Science named "Regional Initiative of Excellence" in the years 2019–2023, project number 026/RID/2018/19, the amount of financing PLN 9 542 500.00.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to a lot of data.

Acknowledgments: The authors would like to express their gratitude to Massimo Fronte Maranzano of the Department of Agricultural and Environmental Sciences—Production, Landscape, Agroenergy, for providing technical support in the analyses performed in the laboratory of the Università degli Studi di Milano.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript or in the decision to publish the result.

References

1. Michalska, A.; Łysiak, G. Bioactive Compounds of Blueberries: Post-Harvest Factors Influencing the Nutritional Value of Products. *Int. J. Mol. Sci.* **2015**, *16*, 18642–18663. [CrossRef] [PubMed]
2. Cocetta, G.; Cavenago, B.; Bulgari, R.; Spinardi, A. Benzothiadiazole Enhances Ascorbate Recycling and Polyphenols Accumulation in Blueberry in a Cultivar-Dependent Manner. *Front. Plant Sci.* **2022**, *13*, 1032133. [CrossRef] [PubMed]
3. Piechowiak, T.; Skóra, B.; Grzelak-Błaszczyk, K.; Sójka, M. Extraction of Antioxidant Compounds from Blueberry Fruit Waste and Evaluation of Their In Vitro Biological Activity in Human Keratinocytes (HaCaT). *Food Anal. Methods* **2021**, *14*, 2317–2327. [CrossRef]
4. Piechowiak, T.; Antos, P.; Józefczyk, R.; Kosowski, P.; Skrobacz, K.; Balawejder, M. Impact of Ozonation Process on the Microbiological Contamination and Antioxidant Capacity of Highbush Blueberry (*Vaccinum corymbosum* L.) Fruit during Cold Storage. *Ozone Sci. Eng.* **2019**, *41*, 376–385. [CrossRef]
5. Ma, L.; Sun, Z.; Zeng, Y.; Luo, M.; Yang, J. Molecular Mechanism and Health Role of Functional Ingredients in Blueberry for Chronic Disease in Human Beings. *Int. J. Mol. Sci.* **2018**, *19*, 2785. [CrossRef] [PubMed]
6. Ren, T.; Huang, C.; Cheng, M. Dietary Blueberry and Bifidobacteria Attenuate Nonalcoholic Fatty Liver Disease in Rats by Affecting SIRT1-Mediated Signaling Pathway. *Oxid. Med. Cell Longev.* **2014**, *2014*, 469059. [CrossRef] [PubMed]
7. Shukitt-Hale, B.; Lau, F.C.; Josep, J.A. Berry Fruit Supplementation and the Aging Brain. *J. Agric. Food Chem.* **2008**, *56*, 636–641. [CrossRef]
8. Papandreou, M.A.; Dimakopoulou, A.; Linardaki, Z.I.; Cordopatis, P.; Klimis-Zacas, D.; Margarity, M.; Lamari, F.N. Effect of a Polyphenol-Rich Wild Blueberry Extract on Cognitive Performance of Mice, Brain Antioxidant Markers and Acetylcholinesterase Activity. *Behav. Brain Res.* **2009**, *198*, 352–358. [CrossRef]
9. Cheatham, C.L.; Vazquez-Vidal, I.; Medlin, A.; Voruganti, V.S. Blueberry Consumption Affects Serum Uric Acid Concentrations in Older Adults in a Sex-Specific Manner. *Antioxidants* **2016**, *5*, 43. [CrossRef]
10. Kalt, W.; Cassidy, A.; Howard, L.R.; Krikorian, R.; Stull, A.J.; Tremblay, F.; Zamora-Ros, R. Recent Research on the Health Benefits of Blueberries and Their Anthocyanins. *Adv. Nutr.* **2020**, *11*, 224–236. [CrossRef]
11. Kausar, H.; Jeyabalan, J.; Aqil, F.; Chabba, D.; Sidana, J.; Singh, I.P.; Gupta, R.C. Berry Anthocyanidins Synergistically Suppress Growth and Invasive Potential of Human Non-Small-Cell Lung Cancer Cells. *Cancer Lett.* **2012**, *325*, 54–62. [CrossRef] [PubMed]

12. Wu, T.; Tang, Q.; Gao, Z.; Yu, Z.; Song, H.; Zheng, X.; Chen, W. Blueberry and Mulberry Juice Prevent Obesity Development in C57BL/6 Mice. *PLoS ONE* **2013**, *8*, e77585. [CrossRef] [PubMed]
13. Duan, Y.; Tarafdar, A.; Chaurasia, D.; Singh, A.; Bhargava, P.C.; Yang, J.; Li, Z.; Ni, X.; Tian, Y.; Li, H.; et al. Blueberry Fruit Valorization and Valuable Constituents: A Review. *Int. J. Food Microbiol.* **2022**, *381*, 109890. [CrossRef] [PubMed]
14. Song, G.Q.; Hancock, J.F. Vaccinium. In *Wild Crop Relatives: Genomic and Breeding Resources*; Kole, C., Ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 197–222.
15. Goyali, J.C.; Igamberdiev, A.U.; Debnath, S.C. Propagation Methods Affect Fruit Morphology and Antioxidant Properties but Maintain Clonal Fidelity in Lowbush Blueberry. *HortScience* **2015**, *50*, 888–896. [CrossRef]
16. Pliszka, K. *Borówka Wysoka*; Państwowe Wydawnictwo Rolnicze I Leśne: Warszawa, Poland, 2002.
17. Cohen, D.; Eliot, D. Micropropagation Methods for Blueberries and Tamarillos. *Int. Plant Propagators' Soc.* **1979**, *29*, 144–146.
18. Zimmerman, R.H.; Broome, O.C. Blueberry Micropropagatio. In *Proceedings of the Conference on Nursery Production of Fruit Plants through Tissue Culture-Applications and Feasibility*; Agricultural Research Results ARR-NE; Zimmerman, Ed.; USDASEA: Tucson, AZ, USA, 1980; pp. 44–47.
19. Debnath, S.C. Clonal Fidelity and Morphological and Chemical Variations in Micropropagated *Vaccinium* Plants. *Acta Hortic.* **2017**, *1180*, 111–116. [CrossRef]
20. Mohamed, G.; Khusnetdinova, L.; Timofeeva, O. Elaboration of Micropropagation Protocol for *Vaccinium* Corymbosum Cv. "Sunt Blue Giant". *Asian J. Plant Sci. Res.* **2018**, *8*, 1–11.
21. Hussain, A.; Ahmed, I.; Nazir, H.; Ullah, I. Plant Tissue Culture: Current Status and Opportunities. In *Recent Advances in Plant In Vitro Culture*; Delve Publishing: Burlington, ON, Canada, 2012.
22. Kumar, R.; Raturi, P.; Jat, M.J. Plant Propagation: Importance and Scope. In *Plant Propagation Andnursery Management for Fruit Crops*; NIPA, Genx Electronic Resources & Solutions P. LTD: New Delhi, India, 2023; pp. 1–15.
23. Sharma, K. Plant Tissue Culture and Micropropagation. In *Plant Propagation Andnursery Management for Fruit Crops*; Genx Electronic Resources & Solutions P. LTD: New Delhi, India, 2023; pp. 83–95.
24. Larkin, P.J.; Scowcroft, W.R. Somaclonal Variation—A Novel Source of Variability from Cell Cultures for Plant Improvement. *Theor. Appl. Genet.* **1981**, *60*, 197–214. [CrossRef]
25. Kaeppler, S.M.; Kaeppler, H.F.; Rhee, Y. Epigenetic Aspects of Somaclonal Variation in Plants. *Plant Mol. Biol.* **2000**, *43*, 179–188. [CrossRef]
26. Kaeppler, S.M.; Phillips, R.L.; Olhoft, P. Molecular Basis of Heritable Tissue Culture-Induced Variation in Plants. *Somaclonal Var. Induc. Mutat. Crop Improv.* **1998**, 465–484.
27. Duta-Cornescu, G.; Constantin, N.; Pojoga, D.M.; Nicuta, D.; Simon-Gruita, A. Somaclonal Variation—Advantage or Disadvantage in Micropropagation of the Medicinal Plants. *Int. J. Mol. Sci.* **2023**, *24*, 838. [CrossRef] [PubMed]
28. Ghosh, A.; Igamberdiev, A.U.; Debnath, S.C. Tissue Culture-Induced DNA Methylation in Crop Plants: A Review. *Mol. Biol. Rep.* **2021**, *48*, 823–841. [CrossRef] [PubMed]
29. Bairu, M.W.; Aremu, A.O.; van Staden, J. Somaclonal Variation in Plants: Causes and Detection Methods. *Plant Growth Regul.* **2011**, *63*, 147–173. [CrossRef]
30. Jain, S.M. Tissue Culture-Derived Variation in Crop Improvement. *Euphytica* **2001**, *118*, 153–166. [CrossRef]
31. Leva, A.R.; Petruccelli, R.; Rinaldi, L.M.R. Somaclonal Variation in Tissue Culture: A Case Study with Olive. In *Recent Advances in Plant In Vitro Culture*; INTECH Open Access Publisher: Croatia, Balkans, 2012.
32. Bhatia, R.; Singh, K.P.; Jhang, T.; Sharma, T.R. Assessment of Clonal Fidelity of Micropropagated Gerbera Plants by ISSR Markers. *Sci. Hortic.* **2009**, *119*, 208–211. [CrossRef]
33. Kawiak, A.; Łojkowska, E. Application of RAPD in the Determination of Genetic Fidelity in Micropropagated Drosera Plantlets. *Vitr. Cell. Dev. Biol.-Plant* **2004**, *40*, 592–595. [CrossRef]
34. Skirvin, R.M.; McPheeters, K.D.; Norton, M. Sources and Frequency of Somaclonal Variation. *HortScience* **2019**, *29*, 1232–1237. [CrossRef]
35. Bouman, H.; De Klerk, G.J. Measurement of the Extent of Somaclonal Variation in Begonia Plants Regenerated under Various Conditions. Comparison of Three Assays. *Theor. Appl. Genet.* **2001**, *102*, 111–117. [CrossRef]
36. Aremu, A.O.; Bairu, M.W.; Szüčová, L.; Doležal, K.; Finnie, J.F.; Van Staden, J. Genetic Fidelity in Tissue-Cultured "Williams" Bananas—The Effect of High Concentration of Topolins and Benzyladenine. *Sci. Hortic.* **2013**, *161*, 324–327. [CrossRef]
37. Razani, M.; Kayat, F.; Mohamed Re, R.; Susanto, D. Effect of Somaclonal Variation in Musa Acuminata Cv. Berangan through Micropropagation Using RAPD. *Biotechnology* **2018**, *18*, 9–14. [CrossRef]
38. Bednarek, P.T.; Orłowska, R. Plant Tissue Culture Environment as a Switch-Key of (Epi)Genetic Changes. *Plant Cell Tissue Organ. Cult.* **2020**, *140*, 245–257. [CrossRef]
39. Gimenez, M.D.; Yañez-Santos, A.M.; Paz, R.C.; Quiroga, M.P.; Marfil, C.F.; Conci, V.C.; García-Lampasona, S.C. Assessment of Genetic and Epigenetic Changes in Virus-Free Garlic (*Allium Sativum* L.) Plants Obtained by Meristem Culture Followed by in Vitro Propagation. *Plant Cell Rep.* **2016**, *35*, 129–141. [CrossRef] [PubMed]
40. Guo, W.; Li, Y.; Gong, L.; Li, F.; Dong, Y.; Liu, B. Efficient Micropropagation of Robinia Ambigua Var. Idahoensis (Idaho Locust) and Detection of Genomic Variation by ISSR Markers. *Plant Cell Tissue Organ. Cult.* **2006**, *84*, 343–351. [CrossRef]

41. Guo, W.L.; Gong, L.; Ding, Z.F.; Li, Y.D.; Li, F.X.; Zhao, S.P.; Liu, B. Genomic Instability in Phenotypically Normal Regenerants of Medicinal Plant Codonopsis Lanceolata Benth. et Hook. f., as Revealed by ISSR and RAPD Markers. *Plant Cell Rep.* **2006**, *25*, 896–906. [CrossRef] [PubMed]
42. De Klerk, G.-J. How to Measure Somaclonal Variation. *Acta Bot. Neerl.* **1990**, *39*, 129–144. [CrossRef]
43. Debnath, S.C. Strategies to Propagate *Vaccinium* Nuclear Stocks for the Canadian Berry Industry. *Can. J. Plant Sci.* **2007**, *87*, 911–922. [CrossRef]
44. Litwińczuk, W. *Rozmnażanie Borówki Wysokiej (Vaccinium corymbosum L.) w Kulturach In Vitro. Wpływ Mikrorozmnażania Na Wzrost i Owocowanie Krzewów*; Wydawnictwo Uniwersytetu Rzeszowskiego: Rzeszów, Poland, 2007.
45. Jamieson, A.R.; Nickerson, N.L. Field Performance of the Lowbush Blueberry Propagated by Seed, Stem Cuttings and Micropropagation. *Proc. Acta Hortic.* **2003**, *626*, 423–428. [CrossRef]
46. Goyali, J.C.; Igamberdiev, A.U.; Debnath, S.C. Morphology, Phenolic Content and Antioxidant Capacity of Lowbush Blueberry (*Vaccinium angustifolium* Ait.) Plants as Affected by in Vitro and Ex Vitro Propagation Methods. *Can. J. Plant Sci.* **2013**, *93*, 1001–1008. [CrossRef]
47. Morrison, S.; Smagula, J.M.; Litten, W. Morphology, Growth, and Rhizome Development of *Vaccinium Angustifolium* Ait. Seedlings, Rooted Softwood Cuttings, and Micropropagated Plantlets. *HortScience* **2000**, *35*, 738–741. [CrossRef]
48. Foley, S.L.; Debnath, S.C. Influence of in Vitro and *Ex. Vitro* Propagation on Anthocyanin Content and Anti-Oxidant Activity of Lingonberries. *J. Hortic. Sci. Biotechnol.* **2007**, *82*, 114–118. [CrossRef]
49. Peraza-Echeverria, S.; Herrera-Valencia, V.A.; Kay, A.J. Detection of DNA Methylation Changes in Micropropagated Banana Plants Using Methylation-Sensitive Amplification Polymorphism (MSAP). *Plant Sci.* **2001**, *161*, 359–367. [CrossRef] [PubMed]
50. Zhang, M.; Kimatu, J.N.; Xu, K.; Liu, B. DNA Cytosine Methylation in Plant Development. *J. Genet. Genom.* **2010**, *37*, 1–12. [CrossRef] [PubMed]
51. Goyali, J.C.; Igamberdiev, A.U.; Debnath, S.C. DNA Methylation in Lowbush Blueberry (*Vaccinium Angustifolium* Ait.) Propagated by Softwood Cutting and Tissue Culture. *Can. J. Plant Sci.* **2018**, *98*, 1035–1044. [CrossRef]
52. Xu, M.; Li, X.; Korban, S.S. AFLP-Based Detection of DNA Methylation. *Plant Mol. Biol. Report.* **2000**, *18*, 361–368. [CrossRef]
53. Li, X.; Xu, M.; Korban, S.S. DNA Methylation Profiles Differ between Fiel and in Vitro Grown Leaves of Apple. *J. Plant Physiol.* **2002**, *159*, 1229–1234. [CrossRef]
54. Xiong, L.Z.; Xu, C.G.; Saghai Maroof, M.A.; Zhang, Q. Patterns of Cytosine Methylation in an Elite Rice Hybrid and Its Parental Lines, Detected by a Methylation-Sensitive Amplification Polymorphism Technique. *Mol. General. Genet.* **1999**, *261*, 439–446. [CrossRef]
55. Litwińczuk, W.; Szczerba, G.; Wrona, D. Field Performance of Highbush Blueberries (*Vaccinium*×*corymbosum* L.) Cv. 'Herbert' Propagated by Cuttings and Tissue Culture. *Sci. Hortic.* **2005**, *106*, 162–169. [CrossRef]
56. Mazurek, M.; Siekierzyńska, A.; Jacek, B.; Litwińczuk, W. Differences in Response to Drought Stress among Highbush Blueberry Plants Propagated Conventionally and by Tissue Culture. *Plant Biosyst.* **2021**, *155*, 172–178. [CrossRef]
57. Debnath, S.C.; Ghosh, A. Phenotypic Variation and Epigenetic Insight into Tissue Culture Berry Crops. *Front. Plant Sci.* **2022**, *13*, 1042726. [CrossRef]
58. Walder, R.Y.; Langtimm, C.J.; Chaterjee, R.; Walder, J.A. Cloning of the *Msp*I modification enzyme. The site of modification andits effect on cleavage by *Msp*I and *HpaII*. *J. Biol. Chem.* **1982**, *258*, 1235–1241. [CrossRef]
59. Krishna, H.; Alizadeh, M.; Singh, D.; Singh, U.; Chauhan, N.; Eftekhari, M.; Sadh, R.K. Somaclonal Variations and Their Applications in Horticultural Crops Improvement. *Biotech* **2016**, *6*, 54. [CrossRef] [PubMed]
60. Grout, J.M.; Read, P.E.; Wildung, D.K. Influence of Tissue Culture and Leaf-Bud Propagation on the Growth Habit of 'Northblue' Blueberry. *J. Am. Soc. Hortic. Sci.* **2022**, *111*, 372–375. [CrossRef]
61. El-Shiekh, A.; Wildung, D.K.; Luby, J.J.; Sargent, K.L.; Read, P.E. Long-Term Effects of Propagation by Tissue Culture or Softwood Single-Node Cuttings on Growth Habit, Yield, and Berry Weight of "northblue" Blueberry. *J. Am. Soc. Hortic. Sci.* **1996**, *121*, 339–342. [CrossRef]
62. Debnath, S.C. Effects of Carbon Source and Concentration on Development of Lingonberry (Vaccinium *Vitis-Idaea* L.) Shoots Cultivated in Vitro from Nodal Explants. *In Vitro Cell. Dev. Biol.-Plant* **2005**, *41*, 145–150. [CrossRef]
63. Gustavsson, B.A.; Stanys, V. Field Performance of "Sanna" Lingonberry Derived by Micropropagation vs. Stem Cuttings. *HortScience* **2000**, *35*, 742–744. [CrossRef]
64. Debnath, S.C. Zeatin-Induced One-Step in Vitro Cloning Affects the Vegetative Growth of Cranberry (*Vaccinium* Macrocarpon Ait.) Micropropagules over Stem Cuttings. *Plant Cell Tissue Organ. Cult.* **2008**, *93*, 231–240. [CrossRef]
65. Serres, R.; McCown, B. Rapid Flowering of Microcultured Cranberry Plants. *HortScience* **1994**, *29*, 159–161. [CrossRef]
66. Dobránszki, J.; Mendler-Drienyovszki, N. Cytokinin-Induced Changes in the Chlorophyll Content and Fluorescence of in Vitro Apple Leaves. *J. Plant Physiol.* **2014**, *171*, 1472–1478. [CrossRef]
67. Martins, J.P.R.; Rodrigues, L.C.A.; Santos, E.R.; Batista, B.G.; Gontijo, A.B.P.L.; Falqueto, A.R. Anatomy and Photosystem II Activity of in Vitro Grown Aechmea Blanchetiana as Affected by 1-Naphthaleneacetic Acid. *Biol. Plant* **2018**, *62*, 211–221. [CrossRef]
68. Stefanova, M.; Koleva, D.P.; Ganeva, T.G.; Dimitrova, M. Effect of Plant Growth Regulators on the Regeneration of in Vitro Propagated *Lamium album* L. Plants. *Bulg. J. Agric. Sci.* **2013**, *19*, 1208–1212.

69. Majada, J.P.; Fal, M.A.; Tadeo, F.; Sánchez-Tamés, R. Effects of Natural Ventilation on Leaf Ultrastructure of *Dianthus caryophyllus* L. Cultured in Vitro. *Vitr. Cell. Dev. Biol.-Plant* **2002**, *38*, 272–278. [CrossRef]
70. Hazarika, B.N. Morpho-Physiological Disorders in in Vitro Culture of Plants. *Sci. Hortic.* **2006**, *108*, 105–120. [CrossRef]
71. Debnath, S.C.; Goyali, J.C. In Vitro Propagation and Variation of Antioxidant Properties in Micropropagated *Vaccinium* Berry Plants—A Review. *Molecules* **2020**, *25*, 788. [CrossRef] [PubMed]
72. Palacio, L.; Baeza, M.C.; Cantero, J.J.; Cusidó, R.; Goleniowski, M.E. In Vitro Propagation of "Jarilla" (*Larrea divaricata* Cav.) and Secondary Metabolite Production. *Biol. Pharm. Bull.* **2008**, *31*, 2321–2325. [CrossRef] [PubMed]
73. Baskaran, P.; Ncube, B.; van Staden, J. In Vitro Propagation and Secondary Product Production by *Merwilla plumbea* (Lindl.) Speta. *Plant Growth Regul.* **2012**, *67*, 235–245. [CrossRef]
74. Al-Khayri, J.M.; Naik, P.M. Influence of 2iP and 2,4-D Concentrations on Accumulation of Biomass, Phenolics, Flavonoids and Radical Scavenging Activity in Date Palm (*Phoenix dactylifera* L.) Cell Suspension Culture. *Horticulturae* **2022**, *8*, 683. [CrossRef]
75. Bienaimé, C.; Melin, A.; Bensaddek, L.; Attoumbré, J.; Nava-Saucedo, E.; Baltora-Rosset, S. Effects of plant growth regulators on cell growth and alkaloids production by cell cultures of *Lycopodiella inundata*. *Plant Cell Tissue Organ. Cult.* **2015**, *123*, 523–533. [CrossRef]
76. Zhao, J.; Davis, L.C.; Verpoorte, R. Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnol. Adv.* **2005**, *23*, 283–333. [CrossRef]
77. Vitamvas, J.; Viehmannova, I.; Cepkova, P.H.; Mrhalova, H.; Eliasova, K. Assessment of Somaclonal Variation in Indirect Morphogenesis-Derived Plants of Arracacia Xanthorrhiza. *Pesqui. Agropecu. Bras.* **2019**, *54*. [CrossRef]
78. Debnath, S.C. Adventitious Shoot Regeneration in a Bioreactor System and EST-PCR Based Clonal Fidelity in Lowbush Blueberry (*Vaccinium angustifolium* Ait.). *Sci. Hortic.* **2011**, *128*, 124–130. [CrossRef]
79. Soneji, J.R.; Rao, P.S.; Mhatre, M. Somaclonal Variation in Micropropagated Dormant Axillary Buds of Pineapple (*Ananas comosus* L., Merr.). *J. Hortic. Sci. Biotechnol.* **2002**, *77*, 28–32. [CrossRef]
80. Gajdošová, A.; Ostrolucká, M.G.; Libiaková, G.; Ondrušková, E.; Šimala, D. Microclonal Propagation of *Vaccinium* sp. and *Rubus* sp. and Detection of Genetic Variability in Culture in vitro. *J. Fruit. Ornam. Plant Res.* **2006**, *14*, 1.
81. Orłowska, R.; Machczyńska, J.; Oleszczuk, S.; Zimny, J.; Bednarek, P.T. DNA Methylation Changes and TE Activity Induced in Tissue Cultures of Barley (*Hordeum vulgare* L.). *J. Biol. Res.* **2016**, *23*, 19. [CrossRef] [PubMed]
82. Wnuk, M.; Lewinska, A.; Gurgul, A.; Zabek, T.; Potocki, L.; Oklejewicz, B.; Bugno-Poniewierska, M.; Wegrzyn, M.; Slota, E. Changes in DNA Methylation Patterns and Repetitive Sequences in Blood Lymphocytes of Aged Horses. *Age* **2014**, *36*, 31–48. [CrossRef] [PubMed]
83. Stephens, K.E.; Miaskowski, C.A.; Levine, J.D.; Pullinger, C.R.; Aouizerat, B.E. Epigenetic Regulation and Measurement of Epigenetic Changes. *Biol. Res. Nurs.* **2013**, *15*, 373–381. [CrossRef] [PubMed]
84. Takatsuka, H.; Umeda, M. Epigenetic Control of Cell Division and Cell Differentiation in the Root Apex. *Front. Plant Sci.* **2015**, *6*, 1178. [CrossRef]
85. Wellburn, A.R. The Spectral Determination of Chlorophylls a and b, as Well as Total Carotenoids, Using Various Solvents with Spectrophotometers of Different Resolution. *J. Plant Physiol.* **1994**, *144*, 307–313. [CrossRef]
86. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **1990**, *12*, 13–15.
87. Bassam, B.J.; Gresshoff, P.M. Silver Staining Dna in Polyacrylamide Gels. *Nat. Protoc.* **2007**, *2*, 2649–2654. [CrossRef]
88. Waterhouse, A.L. Determination of Total Phenolics. In *Current Protocols in Food Analytical Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003.

International Journal of
Molecular Sciences

MDPI

Article

Differential Response of MYB Transcription Factor Gene Transcripts to Circadian Rhythm in Tea Plants (*Camellia sinensis*)

Zhihang Hu [1,2], Nan Zhang [2], Zhiyuan Qin [1], Jinwen Li [1], Ni Yang [1], Yi Chen [1], Jieyu Kong [1], Wei Luo [1], Aisheng Xiong [2,*] and Jing Zhuang [1,*]

[1] Tea Science Research Institute, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China; 2022204007@stu.njau.edu.cn (Z.H.); 2022804179@stu.njau.edu.cn (Z.Q.); 2020204040@stu.njau.edu.cn (J.L.); 2021204040@stu.njau.edu.cn (N.Y.); 2019204042@njau.edu.cn (Y.C.); 2021104085@stu.njau.edu.cn (J.K.); 2022104084@stu.njau.edu.cn (W.L.)
[2] State Key Laboratory of Crop Genetics & Germplasm Enhancement and Utilization, Nanjing Agricultural University, Nanjing 210095, China; 2022104054@stu.njau.edu.cn
* Correspondence: xiongaisheng@njau.edu.cn (A.X.); zhuangjing@njau.edu.cn (J.Z.)

Abstract: The circadian clock refers to the formation of a certain rule in the long-term evolution of an organism, which is an invisible 'clock' in the body of an organism. As one of the largest TF families in higher plants, the MYB transcription factor is involved in plant growth and development. MYB is also inextricably correlated with the circadian rhythm. In this study, the transcriptome data of the tea plant 'Baiyeyihao' were measured at a photoperiod interval of 4 h (24 h). A total of 25,306 unigenes were obtained, including 14,615 unigenes that were annotated across 20 functional categories within the GO classification. Additionally, 10,443 single-gene clusters were annotated to 11 sublevels of metabolic pathways using KEGG. Based on the results of gene annotation and differential gene transcript analysis, 22 genes encoding MYB transcription factors were identified. The G10 group in the phylogenetic tree had 13 members, of which 5 were related to the circadian rhythm, accounting for 39%. The G1, G2, G8, G9, G15, G16, G18, G19, G20, G21 and G23 groups had no members associated with the circadian rhythm. Among the 22 differentially expressed MYB transcription factors, 3 members of LHY, RVE1 and RVE8 were core circadian rhythm genes belonging to the G10, G12 and G10 groups, respectively. Real-time fluorescence quantitative PCR was used to detect and validate the expression of the gene transcripts encoding MYB transcription factors associated with the circadian rhythm.

Keywords: *Camellia sinensis*; MYB; circadian rhythm; transcription factor; transcriptome analysis

1. Introduction

The circadian rhythm is an endogenous timing mechanism that operates on a 24 h cycle, reflecting the relative length of day and night [1]. The input, central oscillator and output pathways are the three components that constitute the circadian rhythm system in higher plants. Among them, the central oscillator, which is the central element of the biological clock system, is essential for controlling plant growth [2,3]. Previous studies have identified core clock genes, such as the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*), *PSEUDO RESPONSE REGULATOR* (*PRR*) family, *TIMING OF CAB EXPRESSION 1* (*TOC1*), *GIGANTEA* (*GI*), *REVEILLEs* (*RVEs*), *NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED* (*LNK*) and *EARLY FLOWERING* (*ELF*) genes [4]. At present, research related to the circadian rhythm mainly focuses on *Arabidopsis*, rice, tomato, etc. [5–7]. Huang et al. investigated photosynthetic efficiency through the circadian clock network under different light intensities in tomatoes [7]. The appearance of mutations in EID1 and LNK2 during domestication delayed both the phase and period of its circadian rhythm [6]. Most clock genes

in *Arabidopsis* have homologous genes in rice [8]. *OsELF3* promotes flowering by repressing *Ghd7* at the transcriptional level and maintains circadian accumulation by interacting with *HAF1* under long-day conditions [9]. The photoperiodic control of flowering time and photoregulation of growth are two examples of features that can be altered by clock gene mutations [10,11]. When comparing the phenotypes of wild-type and mutant plants with clock-changing functions, it can be found that different physiological characteristics (total CO_2 assimilation rate, sugar status and photosynthetic utilization efficiency) are affected by the circadian rhythm [12,13]. The expression profiles of genes related to the circadian rhythm show different patterns in the photoperiods of plants.

Transcription factors (TFs) are a group of proteins that specifically bind to upstream sequences at the 5′ end of the genes, thereby ensuring targeted gene expression with precise spatiotemporal control. MYBs are an important family of transcription factors in higher plants [14,15]. In higher plants, there are four main types of MYB transcription factors: 1R-MYB, R2R3-MYB, 3R-MYB and 4R-MYB [16–19]. Genome-wide descriptions have been relatively detailed for model plants such as *Arabidopsis*, rice and carrot [20,21]. MYB transcription factors exhibit diverse functions in higher plants, including the regulation of various subbiological metabolic synthesis and abiotic stress responses [22–25]. MYB transcription factors have also been reported in tea plants [26,27]. *CsMYB4a* is potentially involved in the regulation of gene expression related to flavonoid biosynthesis [28]. *CsMYB2* and *CsMYB26* are two other regulatory genes involved in the biosynthesis of tea flavonoids [27].

In higher plants, MYB transcription factors are associated with the circadian rhythm [29–33]. The MYB transcription factors CCA1 and LHY (LATE ELONGATED HYPOCOTYL) play crucial roles in maintaining the diurnal cycle of the central oscillator in the biological rhythm system and in regulating the circadian rhythm in *Arabidopsis* [30,34]. REVEILLE1 (RVE1), a transcription factor that is similar to MYB, is homologous to LHY and CCA1. As a node in the auxin network that connects elements of the circadian rhythm, RVE1 provides a mechanistic link between two important signaling pathways [35]. Another member of the RVE family, RVE8, is an MYB transcription factor. RVE8 interacts with the transcriptional coactivators LNK1 and LNK2 to promote the expression of clock genes during nighttime [36]. Duan et al. identified a novel MYB transcription factor in rice, CMYB1, through transcriptome analysis and observed significant rhythmic expression profiles, suggesting that CMYB1 plays a crucial role in cold tolerance via circadian regulation [32]. MYB-like transcription factors LUX (LUX ARRHYTHMO), ELF3 (EARLY FLOWERING 3) and ELF4 constitute the Evening Complex (EC) [37]. In the morning, the EC was inhibited by CCA1 and LHY, and in the evening, it was inhibited by TOC1. MYB112 can directly bind to the promoter of the *LUX* gene (the core component of the central oscillator of the biological clock) and inhibit its transcription, weakening the inhibitory effect of LUX on the transcription of PIF4 and thus promoting its transcription of the *PIF4* gene [38]. Circadian rhythm gene expression and various regulatory relationships within the core oscillator are critical for maintaining the circadian rhythm. The circadian rhythm core oscillator forms a highly linked regulatory network with other circadian rhythm components, affecting a wide range of plant signal transduction and metabolic pathways [39].

To date, studies of MYB transcription factors have mainly focused on anthocyanin biosynthesis regulation [40,41], lignin biosynthesis regulation [42] and the abiotic stress response [43]. To the best of our knowledge, there are few reports on the circadian rhythm of tea plants. The response of MYB transcription factors to the circadian rhythm in tea plants has also not been reported. Understanding the mechanism of MYB transcription factors in regulating the circadian rhythm can enhance the adaptability and stress resistance of tea plants, change the regional restriction of tea plants and improve tea quality.

The tea plant (*Camellia sinensis* (L.) O. Kuntze) has a history of over 5000 years and boasts rich germplasm resources. As a natural, nonalcoholic beverage popular worldwide, tea plays a significant role in the prevention of cancer and cardiovascular diseases [44,45]. Tea is derived from the leaves of tea plants. As the primary organs responsible for pho-

tosynthesis and transpiration, plant leaves provide both the material and energy necessary for growth and development [46]. With the development of the tea genome and sequencing technology, omics studies of tea plants have been reported [47–49]. Here, the high-throughput sequencing technology DNBSEQ was used to sequence the transcriptome of the tea plant 'Baiyeyihao' at intervals of 4 h within one photoperiod (24 h). We identified prospective genes involved in the circadian rhythm, and differentially expressed gene transcripts with MYB transcription factor annotation information were further screened. The selected candidate genes encoding MYB transcription factors were evaluated using a heatmap to understand the response of the circadian rhythm in tea plants. Real-time fluorescence quantitative PCR was used to identify and confirm the expression profiles of circadian-rhythm-related MYB genes in tea plants. This is the first transcriptome analysis of circadian rhythm transcription factors in tea plants. These results also provide a reference for further research on the roles of MYB transcription factors in the circadian rhythm of higher plants.

2. Results

2.1. Sequencing Quality Analysis

Through transcriptome sequencing analysis of tea leaves at a photoperiod interval of 4 h (0 h, 4 h, 8 h, 12 h, 16 h, 20 h), 10.42~10.45 Gb of effective data was obtained. The values of Q20 (error rate less than 1%) were above 96.21%, and the values of Q30 (error rate less than 1%) were above 91.00%. The percentages of GC contents were 45.046%, 44.611%, 44.659%, 44.157%, 44.306% and 44.529%, respectively (Table 1).

Table 1. Tea plant sequencing data volume statistics in a photoperiod.

Sample	Clean Reads (n)	Clean Bases (bp)	Q20 Rate/%	Q30 Rate/%	GC Content/%
0 h	70,014,514	10,426,133,452	97.201	91.627	45.046
4 h	70,039,110	10,436,107,844	96.278	91.847	44.611
8 h	70,054,394	10,441,015,266	96.206	91.676	44.659
12 h	70,005,768	10,434,845,980	97.420	92.273	44.157
16 h	70,049,578	10,445,672,732	97.170	91.545	44.306
20 h	70,017,578	10,420,595,674	96.997	91.001	44.529

2.2. Analysis of Gene Transcripts

Adaptor sequences, duplicated sequences and low-quality reads were excluded. The filtered data were qualitatively controlled using fastqc, and the StringTie software (Version: 2.1.4; Parameter: default) was used to assemble the transcripts. In total, 25,306 transcripts were obtained. Among them, 'o' is other overlapped parts of the same chain with reference exons, with a total of 694; 'j' is for multiple exons with at least one match, of which there are 14,060; 'x' is exon overlap on the antichain; 'i' is completely contained in the intron of the reference transcript; and 'u' is 9415 unknown new transcripts. TransDecoder was used to predict the encoding region of a new transcript. The results showed that the length of the unigenes was mainly between 100 and 1500 bp, and the length of the N50 was 1392 bp (Figure 1).

2.3. Functional Annotation and Categorization

To obtain functional information on the transcripts, functional annotations were drawn up for transcripts from seven databases: Nr (NCBI non-redundant) [50], Pfam [51], Uniprot [52], KEGG (Kyoto Encyclopedia of Genes and Genomes) [53], GO (Gene Ontology) [54], COG (Clusters of Orthologous Groups) [55] and PATHWAY [53]. A total of 19,872 unigenes were annotated in the Nr database, accounting for 78.53%. A total of 19,641 unigenes were annotated in the Uniport database, accounting for 77.61%. A total of 48 unigenes were annotated in the KOG database, accounting for 0.19%. A total of 14,615 unigenes were annotated in the GO database, accounting for 57.75%. A total of

5779 unigenes were annotated in the KEGG database, accounting for 22.84% (Supplementary Table S1).

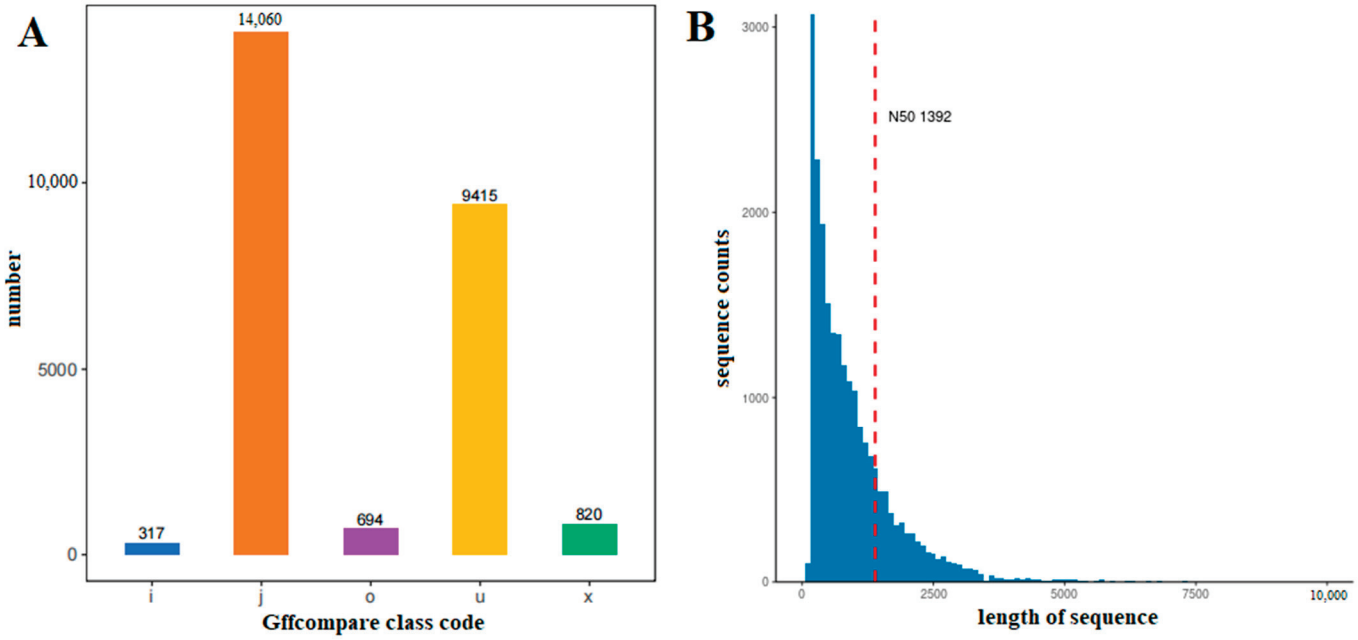


Figure 1. Quantity and differential gene transcript (DET) length distribution of new transcript types in *C. sinensis* within a photoperiod. (**A**): Quantity distribution of new transcript types in the sample; 'o': other overlapped parts of the same chain with reference exons; 'j': multiple exons with at least one match; 'x': exon overlap on the antichain; 'i': completely contained in the intron of the reference transcript; 'u': unknown new transcripts. (**B**): Differential gene transcript length distribution of new transcripts. The abscissa indicates the length of reads. The ordinate indicates the number of reads in that length range, where the red dotted line indicates the length of N50.

The 25,306 assembled unigenes were compared with the related databases. Among these, 19,872 unigenes were annotated in the Nonredundant Collection of Nucleotide Sequences (nr). According to the comparative annotation results of the Nr database with *Camellia sinensis* as a reference, the top 10 species with the most comparisons were determined to be as follows: *Camellia sinensis* (17,547), *C. sinensis* var. *sinensis* (1558), *Aactinidia chinensis* var. *chinensis* (108), *Rhododendron griersonianum* (41), *Actinidia rufa* (41), *Vaccinium darrowii* (37), *Vitis vinifera* (31), *Nyssa sinensis* (31), *Rhodoendron simsii* (22) and *Carya illinoinensis* (15) (Figure 2).

2.4. GO Analysis

Gene Ontology (GO) is an international standard classification system of gene function. The GO database was used to classify the annotated tea transcriptome data. A total of 14,615 unigenes of *C. sinensis* had Gene Ontology (GO) annotations. GO describes gene function from three aspects: cellular component (CC), molecular function (MF) and biological process (BP). The top 20 most common GOslim secondary categories in each category were selected for plotting (Figure 3). Among them, transcription regulation & DNA—templated (351), defense response (273) and translation (241) involved the largest number of differentially expressed gene transcripts (DETs). Among the cell components, the integral components of the membrane (3857), nucleus (1395) and cytoplasm (697) were involved in the most DETs. The most common terms in molecular functions were ATP binding (2173), metal ion binding (1161) and RNA binding (727).

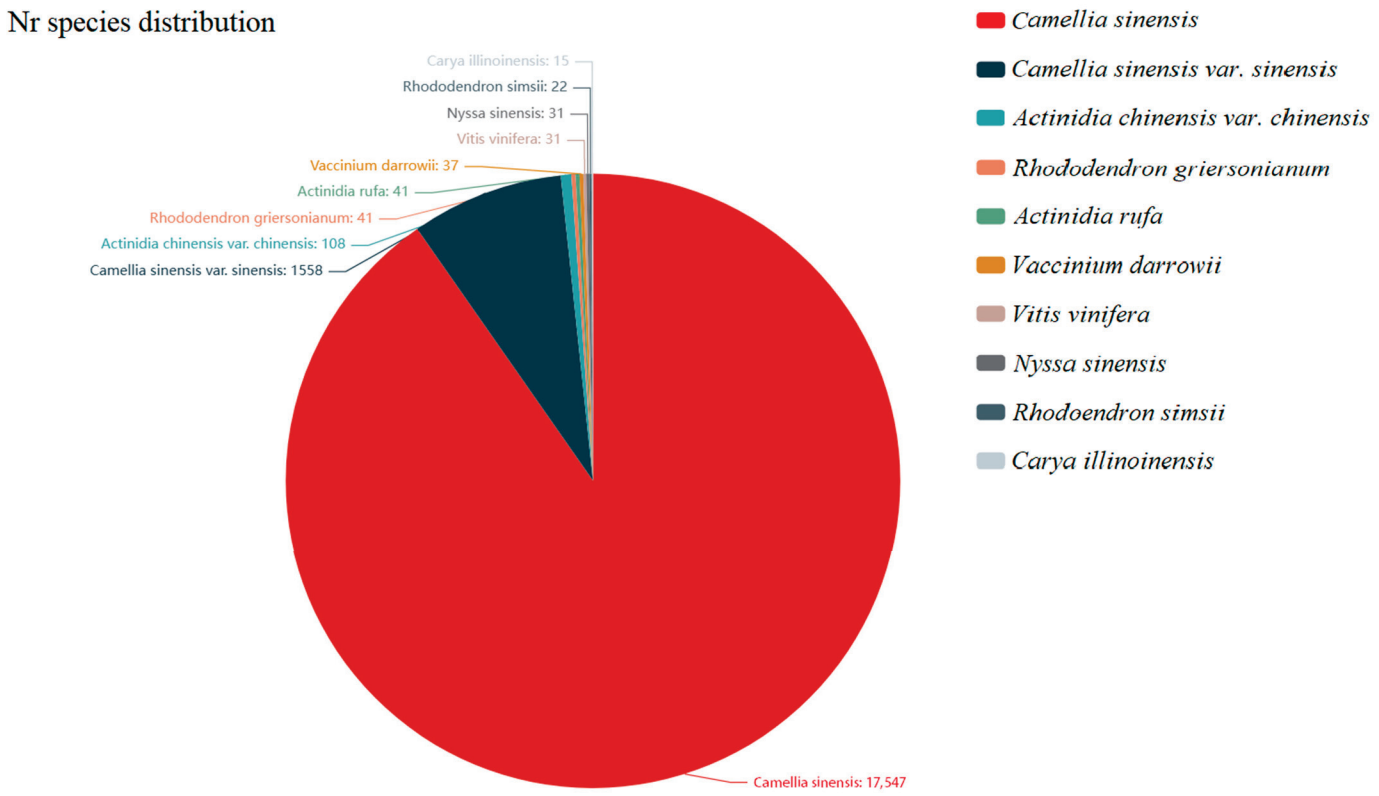


Figure 2. Nr annotated species distribution map in *C. sinensis*. The top 10 species with the most comparisons were counted, and the numbers represent the numbers of different species in the Nr database, while different colors represent different species. The species distribution of BLAST hits for each unigene was determined with a cut-off of 1×10^{-5}.

2.5. KEGG Classification

After KEGG annotation of transcript sequences, they were classified according to the KEGG metabolic pathways in which they participated, and the results are shown in Figure 4. A total of 10,443 DETs were annotated. This classification is divided into five levels: cellular processes, environmental information processing, genetic information processing, metabolism and organismal systems. The largest level of these was metabolism, which had 11 sublevels. Among the 11 sublevel metabolic pathways were global and overview maps (2279), carbohydrate metabolism (657) and amino acid metabolism (471). In the processing of genetic information, translation (n = 722) and folding, sorting and degradation (n = 609) occupied large proportions.

2.6. Differential Gene Transcript Analysis

In our differential expression analysis, the expression levels of the gene transcripts in each sample were obtained through expression quantitation of read count data. A total of 772, 2411, 1731, 1434 and 393 differentially expressed gene transcripts were identified in "0 h_vs._4 h", "0 h_vs._8 h", "0 h_vs._12 h", "0 h_vs._16 h" and "0 h_vs._20 h." In "0 h_vs._4 h", a total of 469 differentially expressed gene transcripts were upregulated and 303 were downregulated. In "0 h_vs._8 h", a total of 1215 differentially expressed gene transcripts were up- and 1196 downregulated. In "0 h_vs._12 h", a total of 697 differentially expressed gene transcripts were up- and 1034 downregulated. In "0 h_vs._16 h", a total of 532 differentially expressed gene transcripts were up- and 902 downregulated. In "0 h_vs._20 h", a total of 235 differentially expressed gene transcripts were up- and 158 downregulated (Figure 5).

Figure 3. GO annotation of functional genes in *C. sinensis*. We selected the top 20 most annotated GOslim under each classification, and each unigene was classified into at least one GO term. All unigenes were grouped into three categories: molecular function, cellular component and biological process.

2.7. Differential Expression of Gene Transcripts Encoding Transcription Factors

DETs encoding transcription factors were predicted using PlantTFDB [56]. The top 20 transcription factor families with differential expression were selected for plotting, and the proportion statistics are shown in Figure 6. A total of 894 DETs encoding transcription factors were detected, the most abundant of which were the B3 (71) and Nin-like (54) transcription factor families. Twenty-three MYB transcription factors and thirty-three MYB-related transcription factors were identified, accounting for 3.35% and 4.80% of the total differentially expressed transcription factors (894), respectively. According to the results of transcription factor annotation, DETs with transcription factor annotation information were further screened, and duplicate transcription factors were removed. A total of 22 transcripts of the MYB transcription factors were obtained for subsequent analysis. We identified one upregulated and four downregulated unigenes in "0 h_vs._4 h"; two upregulated and twelve downregulated unigenes in "0 h_vs._8 h"; one upregulated and eight downregulated unigenes in "0 h_vs._12 h"; one upregulated and eight downregulated

unigenes in "0 h_vs._16 h"; and one upregulated and zero downregulated unigenes in "0 h_vs._20 h."

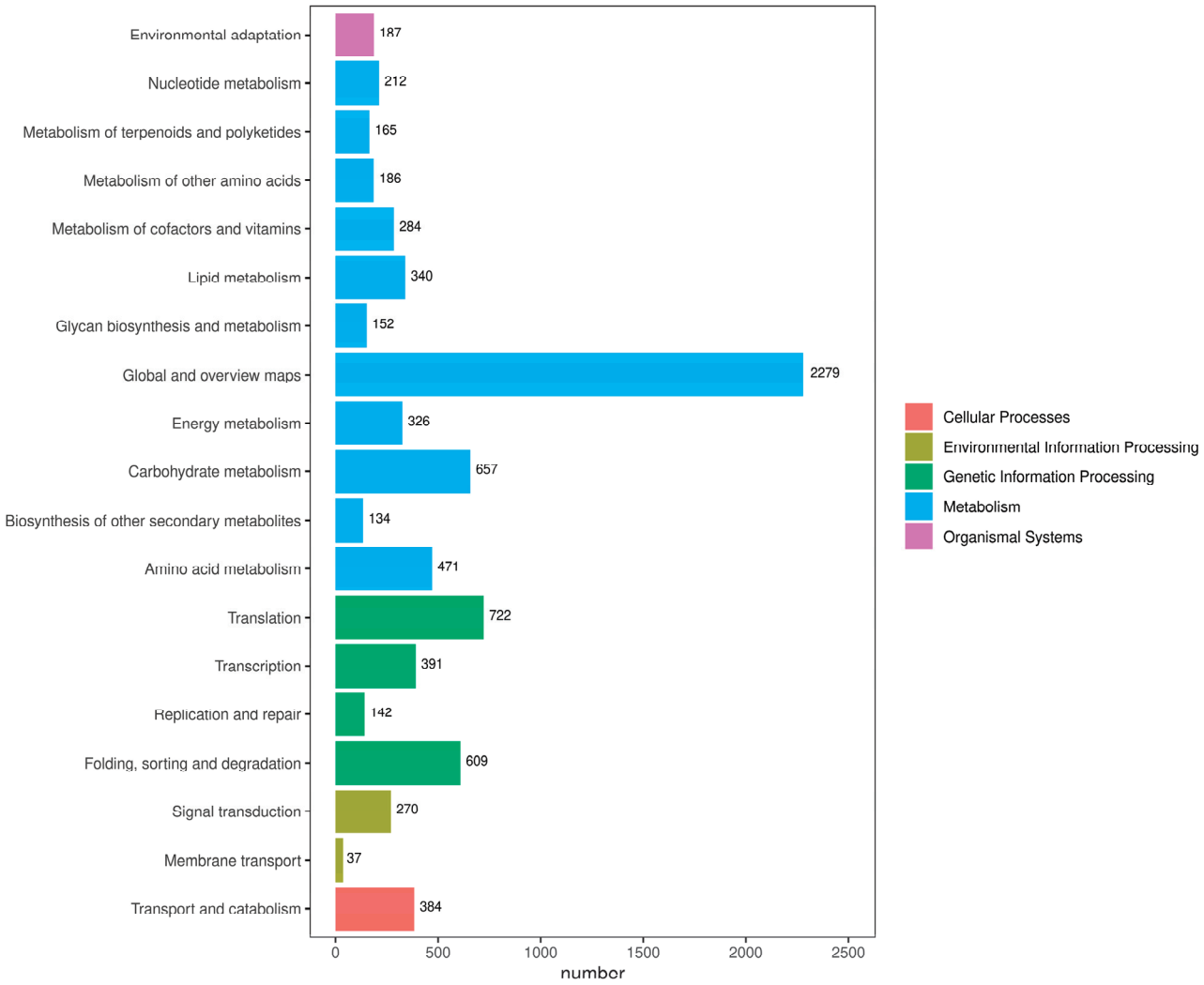


Figure 4. KEGG metabolism pathway categories in *C. sinensis*. Different colors represent the five classes involved in the KEGG metabolic pathway, and numbers represent the numbers of DETs in different classes.

2.8. Distribution of the DETs Encoding MYB Transcription Factors

A total of 222 MYB transcription factors were retrieved from the TPIA database. Multiple sequence alignment was performed on those, and an NJ phylogenetic tree was generated based on the alignment results (Figure 7). The results showed that tea plant MYB members can be divided into 23 subfamilies, named Groups 1 to 23. The 22 selected DETs encoding MYB and MYB-related transcription factors related to the circadian rhythm of tea plants were distributed among the different subfamilies. The numbers of DET MYB transcription factors belonging to the G10, G11 and G12 groups were higher than those in the other groups. Among them, the G10 group had the highest number of transcripts (a total of five) associated with the circadian rhythm. In the G1, G2, G8, G9, G15, G16, G18, G19, G20, G21 and G23 groups, no member was associated with the circadian rhythm.

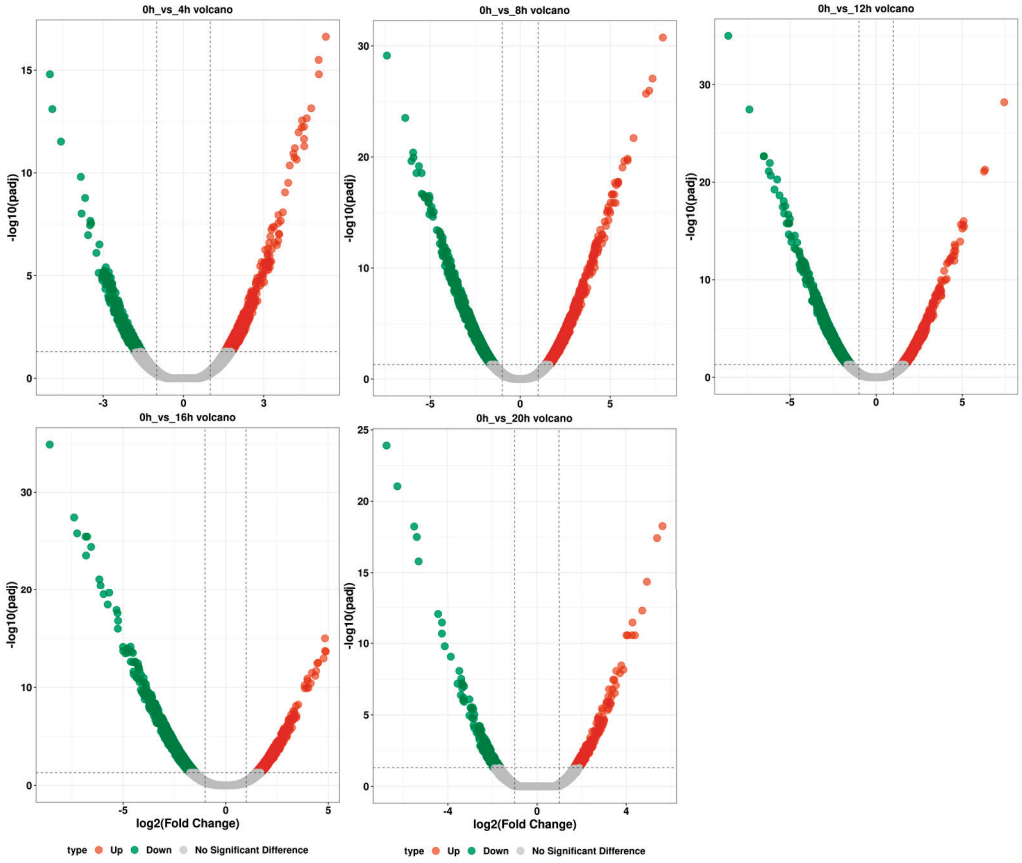


Figure 5. Volcano map of differentially expressed gene transcripts. Each point represents a transcript, and the horizontal coordinate is logFC, which represents the base −2 pair value of the expression difference multiple of a certain transcript in the two samples. The ordinate represents the base 10 pair value of PValue (too few differential transcripts) or error detection rate FDR multiplied by −1.

2.9. Expression Analysis of MYB Transcription Factors in the Circadian Rhythm of Tea Plants

In studying the expression patterns of the 22 MYB transcription factors in the circadian rhythm of tea plants, heatmap analysis showed the expression levels of the gene transcripts within a photoperiod (Figure 8). The heatmap showed that the expression levels of *MYB3*, *MYB159* and *MYB171* were high during the day and low at night, and the process from 0 h treatment showed an obvious rise to fall. The expression levels of *MYB15*, *MYB58*, *MYB140* and *MYB168* were high at night and low during the day, showing a process of decline to rise. The expression levels of *MYB84*, *MYB85*, *MYB86* and *MYB212* were the highest at 0 h and then began to decline, rise again, fall again and rise again at night, then reached a peak, showing a zigzag pattern. The expression levels of *MYB7*, *MYB95*, *MYB106* and *MYB107* were the highest at 0 h, then began to decline and then began to increase after 20 h treatment. The daytime period is 0 h–12 h (9:00–21:00) on the horizontal coordinate, and the nighttime period is 12 h–20 h (21:00–5:00) on the horizontal coordinate.

The 22 MYB gene transcripts showed differential responses in terms of tea plant circadian rhythm based on the results of the RPKM method (Table S3). The expression levels of *MYB3*, *MYB7*, *MYB15*, *MYB40*, *MYB55*, *MYB58*, *MYB84*, *MYB85*, *MYB86*, *MYB95* (*RVE8*), *MYB106* (*LHY*), *MYB107*, *MYB110*, *MYB111*, *MYB119*, *MYB140*, *MYB141*, *MYB159*, *MYB163* (*RVE1*), *MYB168*, *MYB171* and *MYB212* were measured and analyzed using RT-qPCR. The results showed that the gene expression profiles of the 22 MYB transcription factors can be divided into four patterns according to photoperiod: low expression level in a photoperiod, high expression level during the day and low expression level at night, low expression level during the day and high expression level at night, and irregularity. In general,

the expression levels of the 22 DETs encoding MYB transcription factors in the tea plant circadian rhythm were consistent with the trend of the transcriptome sequencing results.

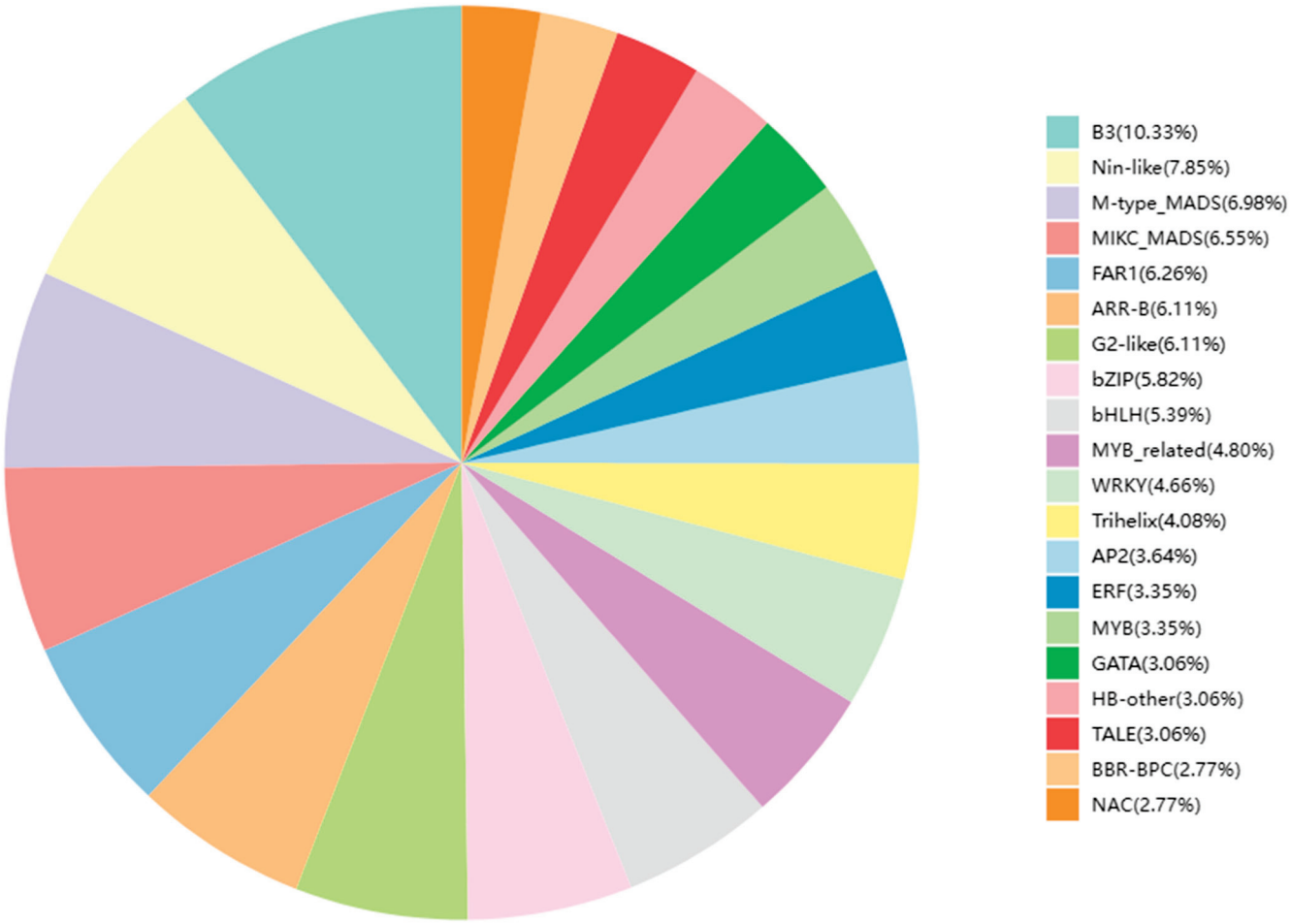


Figure 6. Analysis of transcription factor families in *C. sinensis*. Different colors represent different families of transcription factors, and the numbers in brackets indicate the proportions to the total differential transcription factors.

2.10. The Expression Profiles of the MYB Gene Transcripts

2.10.1. High Expression Level in the Day and Low Expression Level at Night

When studying the RPKM (Figure 9A) and quantitative RT-PCR (Figure 9B) results, it was found that the expression levels of *MYB3*, *MYB159* and *MYB171* gene transcripts were high in the day and low at night. The *MYB3* and *MYB159* gene transcripts reached a peak value at 8 h and the lowest value at 20 h. The expression levels of *MYB3* and *MYB159* at the peak value were 2.78 and 5.17 times higher than that of the lowest value, respectively. The *MYB171* transcript reached its peak value at 4 h and reached its lowest value at 20 h, which was 6.71 times higher. The expression level of *MYB171* at the peak value was 6.71 times higher than that at the lowest value. The transcript levels of *MYB3*, *MYB159* and *MYB171* first increased and then decreased, which had a certain regularity (Figure 10A).

2.10.2. Low Expression Level in the Day and High Expression Level at Night

The expression levels of *MYB7*, *MYB15*, *MYB40*, *MYB55*, *MYB106* (*LHY*), *MYB107*, *MYB119*, *MYB140*, *MYB163* (*RVE1*) and *MYB168* were high at night and low in the day based on the results of the RPKM method (Figure 11A). *MYB106* (*LHY*) and *MYB163* (*RVE1*) are the core genes of the circadian rhythm. The expression level of the LHY transcript was the highest at 0 h and 20 h, when the expression level was up to 207.57 and 63.57 times that

at 8 h, respectively. The expression level of the RVE1 transcript increased gradually from 12 h to 20 h, while the expression level was low during the day. The expression level of the RVE1 transcript decreased gradually from 0 h to 8 h. The peak expression levels were more than 200 times higher than the trough levels (Figure 11B). The expression levels of these MYB gene transcripts showed a trend of first decreasing and then increasing (Figure 10B).

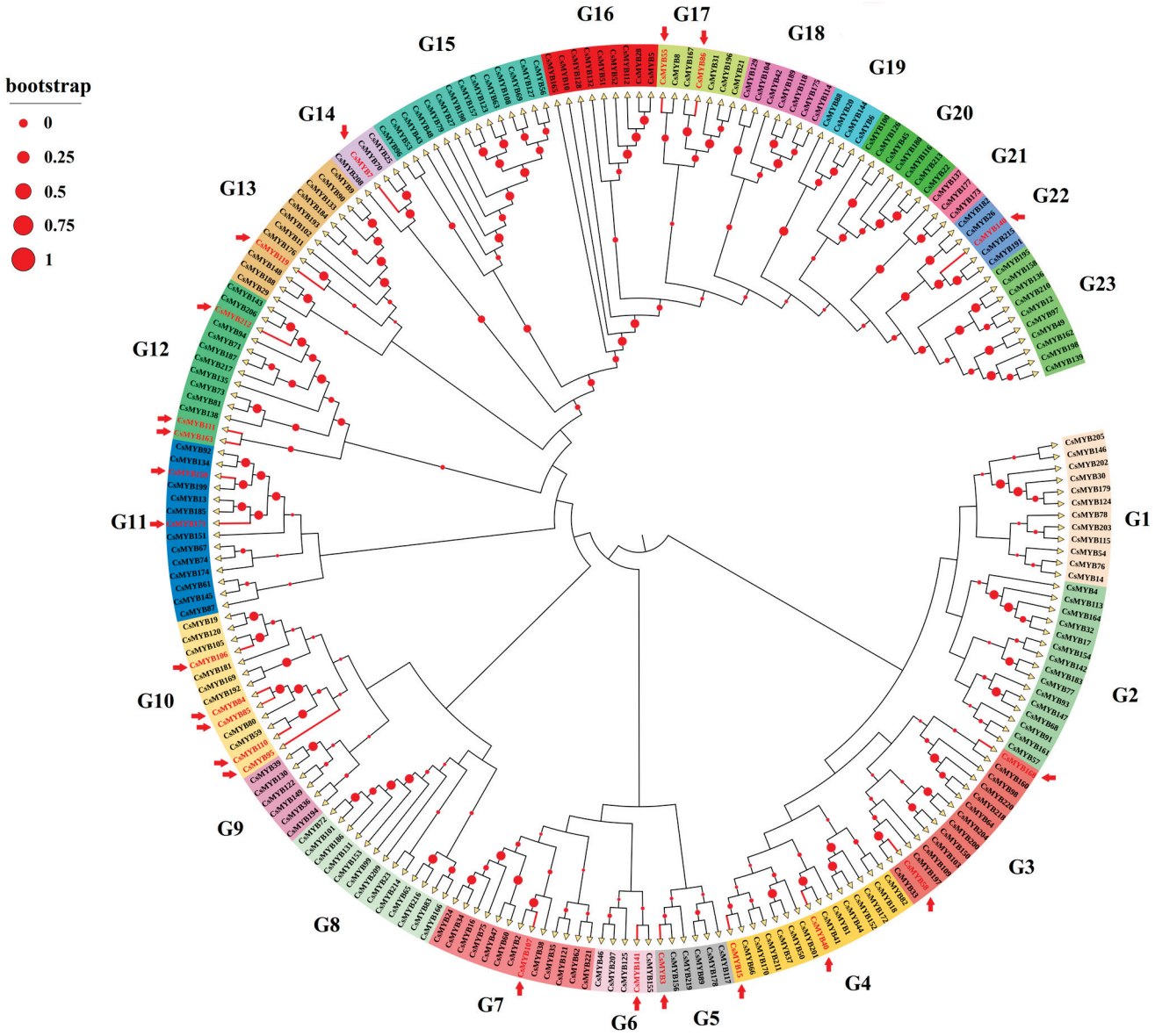


Figure 7. Evolutionary tree analysis of MYB transcription factor family in *C. sinensis*. Arrows represent 22 differentially expressed MYB transcription factors associated with circadian rhythm. Different colors indicate different subfamilies, and red circles indicate similarity.

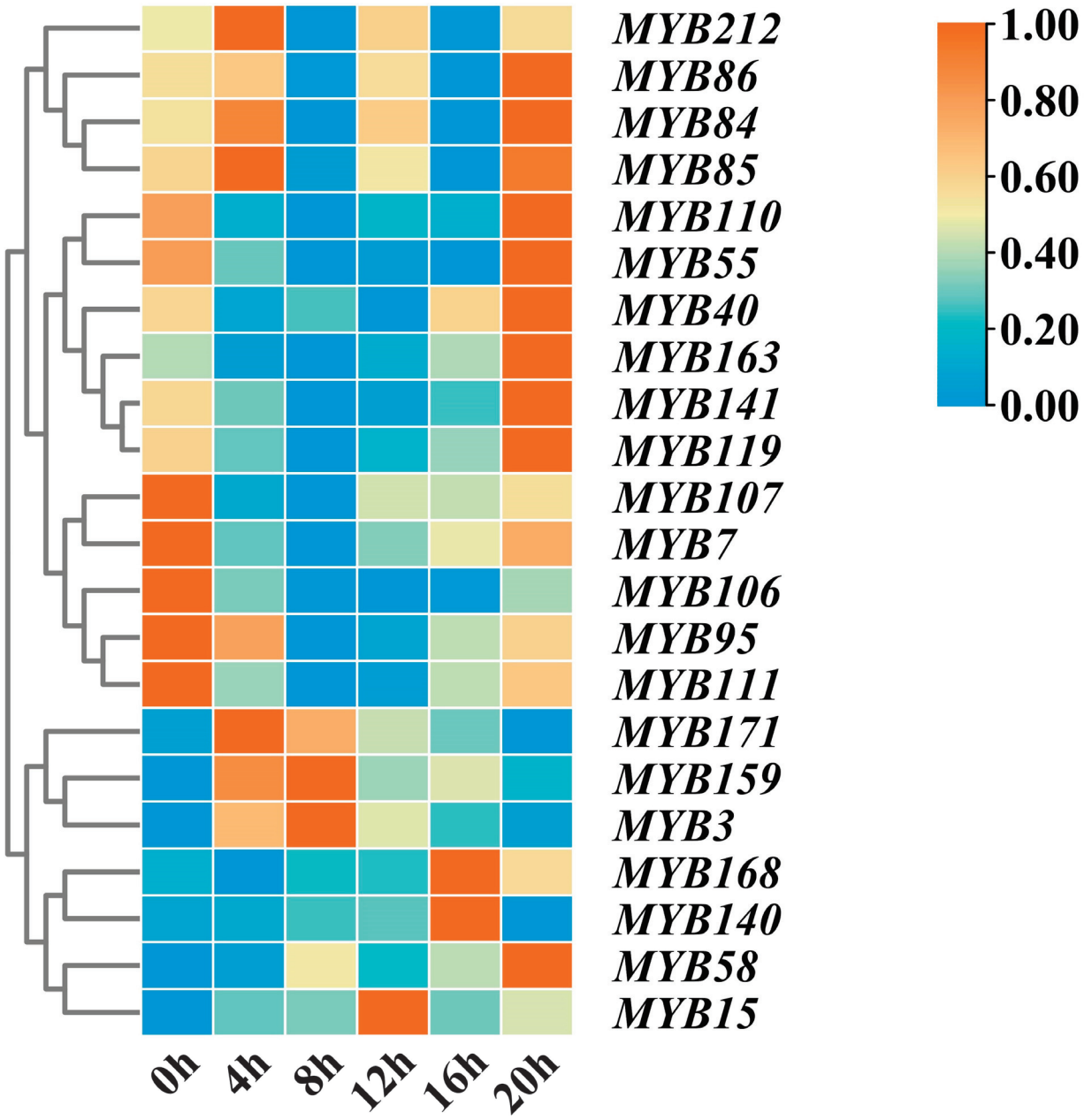


Figure 8. Heatmap of expression profiles of MYB transcription factors within a photoperiod. Color scores were normalized by the counts of RPKM values. Red represents high expression. Blue represents low expression.

2.10.3. The Expression Level Irregularity and Little Change

A total of 222 MYB transcription factor transcripts were found in the tea plant database. Through transcriptome identification and screening, it was estimated that the expression levels of the other 200 MYB gene transcripts showed little change during the photoperiod (24 h) (Figure 10C). The genes of *MYB84*, *MYB85*, *MYB86* and *MYB141* were highly expressed during both the day and night (Figure 12A,B).

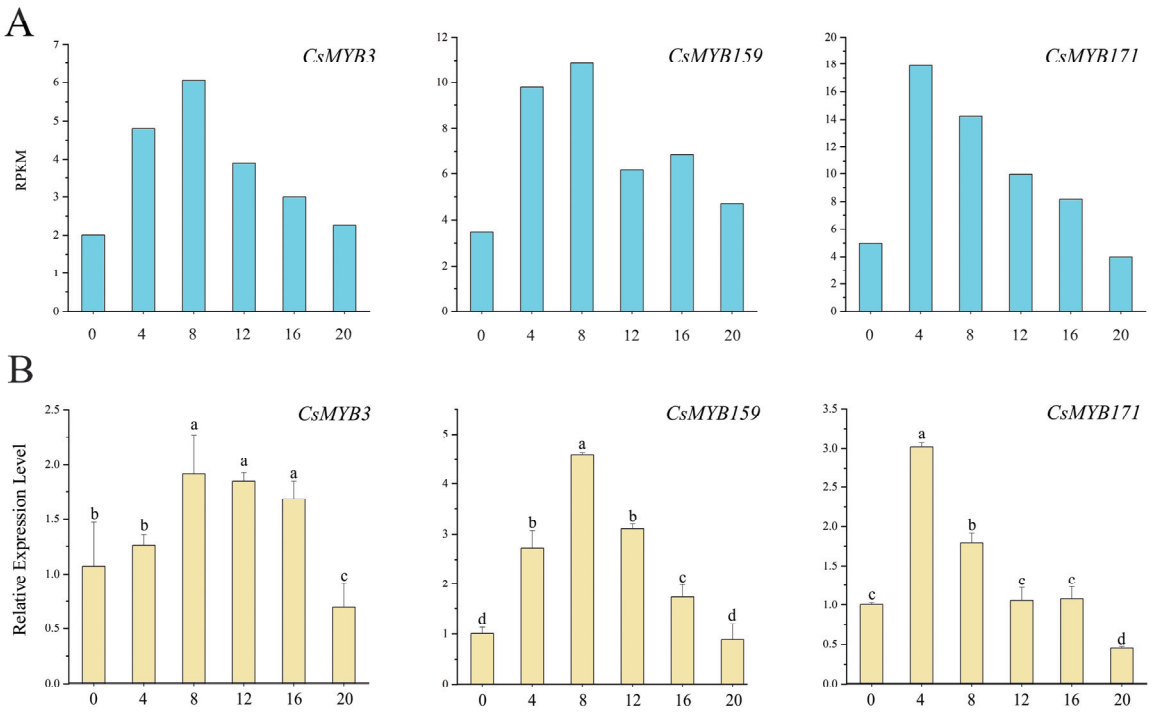


Figure 9. The expression levels of CsMYB gene transcripts in *C. sinensis* within a photoperiod, with profiles that are high in the day and low at night. (**A**): Reads per kilobase per million mapped reads. (**B**): Real-time PCR. The standard deviation (SD) is represented by the error bars. Different lowercase letters indicate significant differences at the 0.05 level.

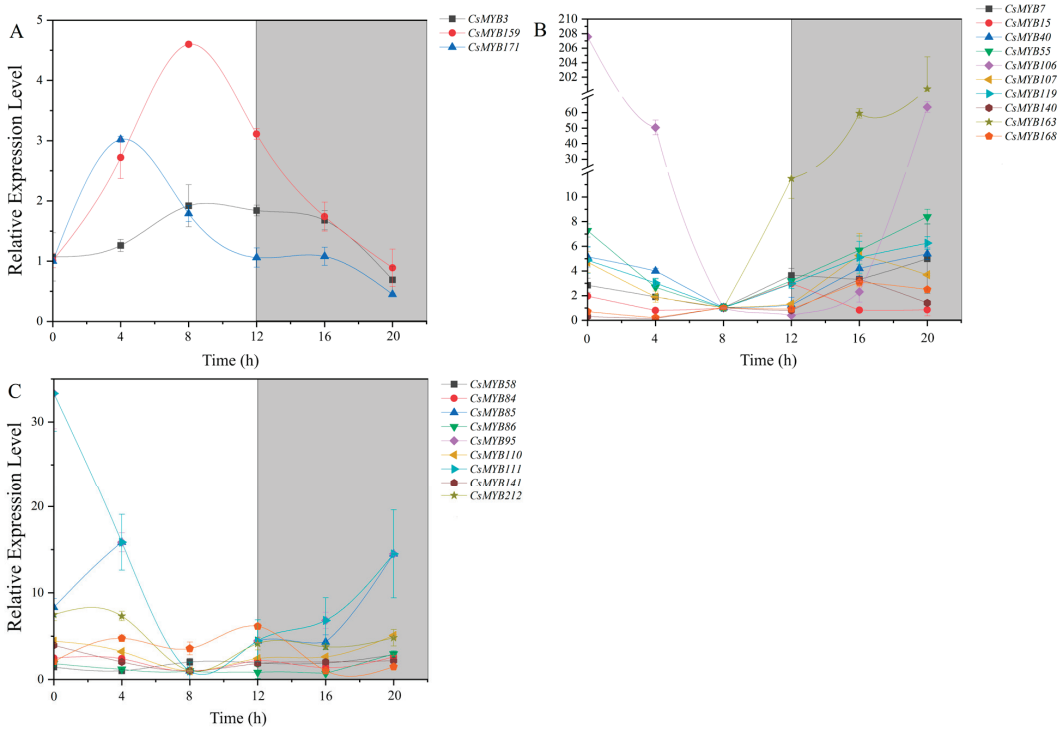


Figure 10. Dynamic behavior of MYB transcription factor in *C. sinensis* within a photoperiod. (**A**). Dynamic behavior of high expression level during the day and low expression level at night. (**B**). Dynamic behavior of low expression level during the day and high expression level at night. (**C**). Dynamic behavior of the expression level irregularity and little change. The gray bands denote dark, and the white bands represent light.

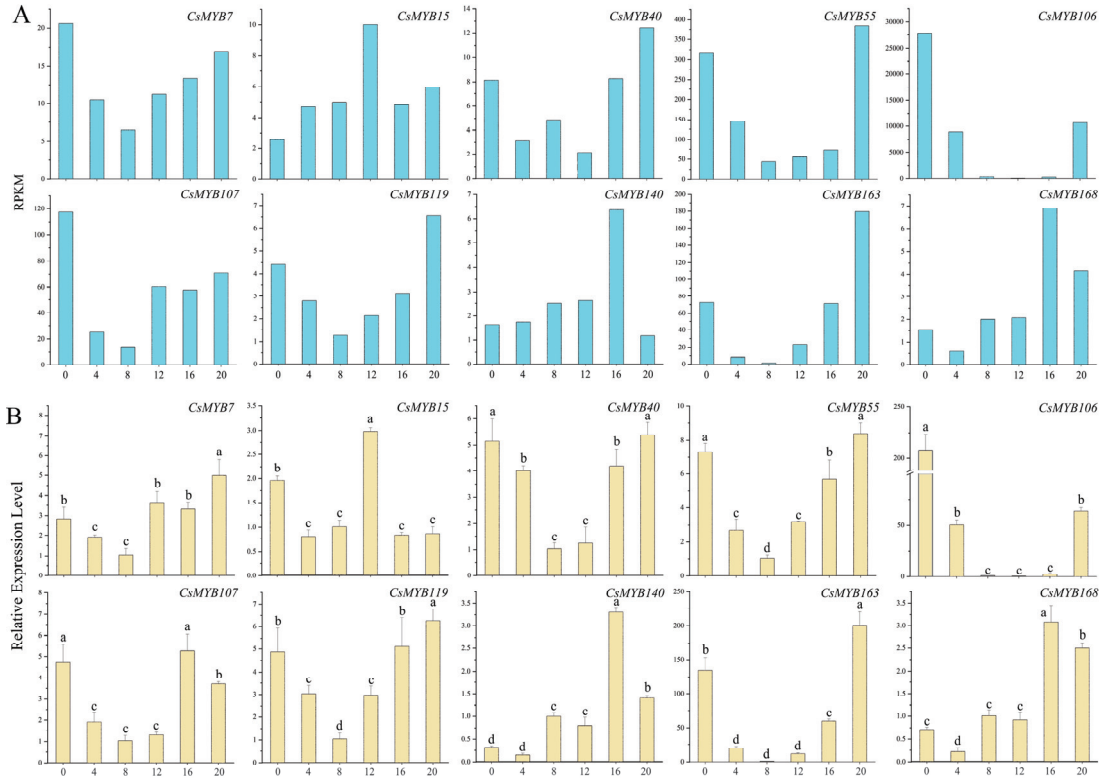


Figure 11. The expression levels of CsMYB gene transcripts in *C. sinensis* within a photoperiod, with profiles that are low in the day and high at night. (**A**): Reads per kilobase per million mapped reads. (**B**): Real-time PCR. The standard deviation (SD) is represented by the error bars; different lowercase letters indicate significant differences at the 0.05 level.

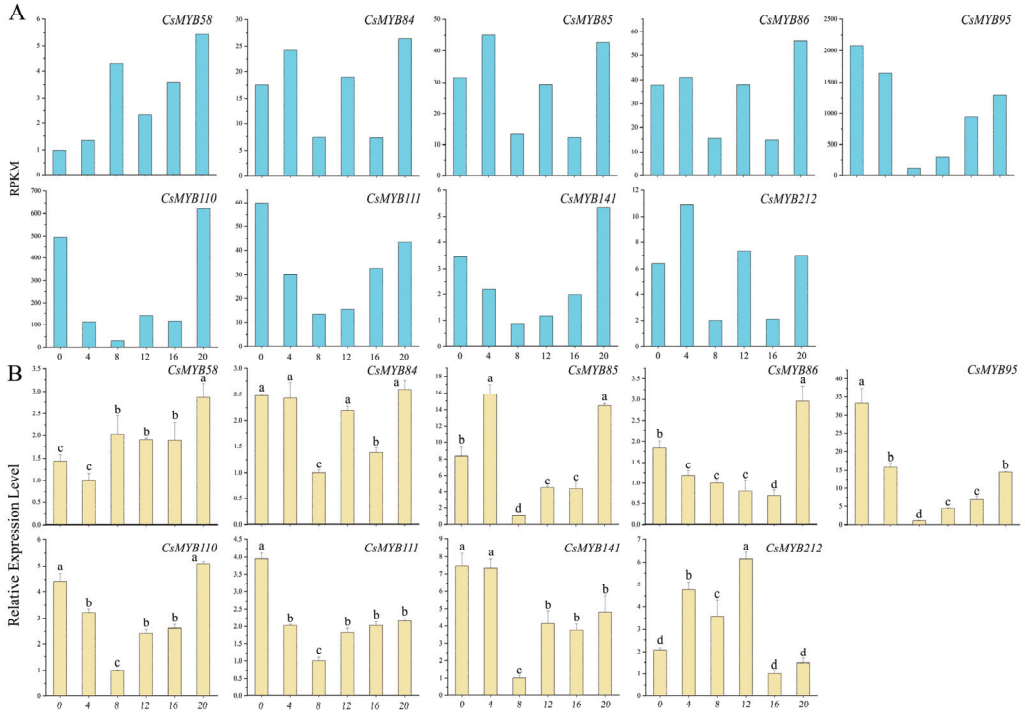


Figure 12. The expression levels of CsMYB gene transcripts with irregular expression profiles in *C. sinensis* within a photoperiod. (**A**): Reads per kilobase per million mapped reads. (**B**): Real-time PCR. The standard deviation (SD) is represented by the error bars. Different lowercase letters indicate significant differences at the 0.05 level.

3. Discussion

Tea plants are typically woody crops with rich genetic diversity during long-term natural growth [57]. In tea plants, the accumulation of secondary metabolites is inconsistent during the day and night and is affected by various factors, including the circadian rhythm [58]. At present, research on the circadian rhythm is mainly concentrated on *Arabidopsis* model plants, which are also the most in-depth plants in terms of photoperiod regulatory mechanisms [5,6]. Research of the circadian rhythm of horticultural plants has just started, mainly involving tomatoes [6], garlic [59], cabbage [60], etc.

At present, 58 transcription factor families have been identified in plants, among which AP2/ERF, MYB, GRAS, WRKY and NAC are the transcription factor families studied in-depth [61,62]. The MYB family of genes exhibits differential expression patterns in plant tissues and performs specific physiological regulatory functions [63]. The MYB family is one of the largest families of transcription factors in plants and plays a crucial role in the regulation of plant biochemical and physiological processes. Syntenic analysis demonstrated that tandem duplication contributed to its evolution. It has also been shown that the main drivers of peanut *R2R3-MYB* gene amplification are polyploidy and tandem and single repetition [64,65]. Chen et al. identified a total of 122 CsR2R3-MYB genes based on the chromosome-level genome of tea plants. *CsMYB12*, *CsMYB17*, *CsMYB25* and *CsMYB47* genes exhibited high expression in the tea plant leaves under PEG stress, but low expression under cold stress. At the same time, they found that *CsMYB17* enhanced anthocyanin accumulation in tea plants [43]. *CsMYB42* played a crucial role in activating the expression of CsGS1c and might be involved in the biosynthesis of theanine in albino tea leaves [66]. Li et al. predicted that *CsMYB8* and *CsMYB99* were involved in the biosynthesis of flavonoids (including catechins, anthocyanins and flavanols), while *CsMYB9* and *CsMYB49* were associated with theanine synthesis, as well as the MYB TFs that were likely involved in hormone-signaling-mediated environmental stress and defense responses [58]. A total of 222 transcripts of the MYB transcription factor are found in tea plants [57]. MYB transcription factors also play key roles in the regulation of the growth and development of tea plants [16]. In this study, information about MYB transcription factors was based on data provided by Li et al. [58]. According to the TPIA database, MYB transcription factors were systematically named CsMYB1–CsMYB221. The corresponding MYB transcription factor names were obtained through BLAST comparison (Table S4). RT-qPCR primers were designed using Primer Premier 6.0 software.

Nearly one-third of *Arabidopsis* genes are regulated by the circadian rhythm, with their expression peaking at different times of the photoperiod [67,68]. In plants, the proteins encoded by *CCA1* and *LHY* are important components in maintaining the diurnal cycle in the central oscillator of the biological clock system of higher plants and regulating the circadian rhythm [30,34]. The *LCL5* (*RVE8*) and *CCA1* gene sequences were homologous. However, these two transcription factors have opposite regulatory effects on the circadian rhythm expression of the *Arabidopsis TOC1* gene [69]. *CCA1* and its closely related homolog *LHY* encode a single MYB transcription factor that regulates the circadian rhythm of genes expressed in the morning and evening. The morning cycle is composed of *CCA1*, *LHY*, *PRR7* (*PSEUDO-RESPONSE REGULATOR 7*) and *PRR9* (*PSEUDO-RESPONSE REGULATOR 9*), in which *CCA1* and *LHY* can promote the transcription and expression of *PRR7* and *PRR9*. However, *PRR7* and *PRR9* can reverse the inhibition of *CCA1* and *LHY* [70]. The abundance of *CCA1* and *LHY* mRNA transcripts oscillates rhythmically, with peak expression at dawn and trough expression at dusk [71,72]. Similar to the expression patterns of *CCA1* and *LHY*, the expression of *RVE1*, *RVE2* and *RVE7* is clock-regulated, with peak transcriptional abundance near subjective dawn, suggesting that these genes may play an important role in clock regulation during the morning phase [73], which is also consistent with the results of this study. In addition, CCA1 and LHY play crucial roles in flowering. Flowering marks the transition from vegetative to reproductive growth, and proper flowering timing is critical for reproduction [74]. CCA1 and LHY participate in the regulation of photoperiodic flowering by regulating the GI-CO-FT pathway [75].

The MYB transcription factor is closely related to the circadian rhythm. By screening the DETs with transcription factor annotation information, a total of 22 MYB and MYB-related transcription factors related to the circadian rhythm of tea plants were obtained, including the typical circadian rhythm genes *LHY* (*MYB106*), *RVE1* (*MYB163*) and *RVE8* (*MYB95*). *REVEILLE1* (*RVE1*), which is homologous to *LHY* and *CCA1*, is an MYB-like transcription factor. Another member of the RVE family, *RVE8*, is an MYB transcription factor. RVE8 interacts with the transcriptional coactivators LNK1 and LNK2 to promote the expression of night clock genes [36]. RVE1 regulates auxin biosynthetic genes to connect the circadian clock and auxin pathways [35]. RVE8 regulates anthocyanin biosynthesis by promoting the expression of anthocyanin genes [76]. *GmMYB133* is a soybean *RVE* homologous gene involved in the biosynthesis of soybean isoflavones [77]. Transient overexpression of *PbRVE1b* can increase the anthocyanin content in birch peels [78]. These results support the importance of the RVE protein in plant development. All *PbRVE* genes in *Pyrus bretschneideri* leaves under a light/dark cycle exhibited a circadian rhythm. The four *PbRVE* genes also exhibited robust rhythms under constant light conditions [79]. In plants, *Arabidopsis* and *Oenanthe javanica* grow under various photoperiodic conditions, and the expression of the clock gene *Lhcb* always decreases from peak levels before dusk and increases from trough levels before dawn. This also suggests that the significance of circadian clock adaptation may be partly due to its regulation of light responses to photosynthesis and photoprotection mechanisms [80–82]. The circadian oscillator may contain multiple circuits that mediate positive and negative feedback, respectively. Interlocked feedback loops have been described in the circadian rhythm in Drosophila, tomatoes, fungi, etc. [7,83,84].

PbRVE expression patterns differed in different tissues of *P. bretschneideri*, suggesting that *PbRVE*s may have a potential function in plant development. *PbLHY* and *PbRVE8* are highly expressed in leaves, which is consistent with the previously reported cognates of poplar [85] and rice [86]. In addition, transcripts of *RVE* genes, such as *PbRVE3b* and *PbREVE6b*, show higher levels in the root of *P. bretschneideri*. Recently, transcriptome analysis of walnuts and alfalfa revealed that *RVE* genes may be involved in nitrogen metabolic pathways.

We measured expression patterns in different tissues of tea plants and further explored the function of MYB transcription factors in the following directions. Since the circadian rhythm in tea plants has not yet been reported, we are still in the preliminary stage of research. Changing the circadian rhythm of the tea plants and determining whether the growth state of the tea plant can be changed is also a good research direction.

4. Materials and Methods

4.1. Plant Materials

Cutting seedlings of two-year-old 'Baiyeyihao' (*C. sinensis* cv. Baiyeyihao) were planted in the growth chamber at the Tea Science Research Institute of Nanjing Agricultural University (Nanjing, China, 188.84° E, 32.04° N). Tea seedlings were planted in sandy loam with good drainage, an organic matter content over 1–2% and a pH value of 6.0. Tea seedlings with healthy growth were selected and cultured in a light incubator (temperature 25 ± 2 °C, photocycle 12 h light/12 h dark, light intensity 240 μmol·m^{-2}·s^{-1}, humidity 70 ± 10%) for one week. Plants were adjusted daily to ensure well-rounded lighting and watered daily early and late in the day. According to literature reports, the circadian rhythm time scheme was experimentally assessed with 4 h intervals [87]. A photocycle of 12 h/12 h is the normal growth time of the plant. Therefore, in order to better understand the influence of the circadian rhythm on the tea plants, we divided 24 h into 6 segments. At 9:00, the illumination began (the initial time was denoted as 0 h) and the first sample was taken. Then, healthy tea seedlings were selected at 4, 8, 12, 16 and 20 h, respectively. Horizontal coordinates 0 h, 4 h and 8 h belonged to the daytime period (9:00~17:00), while 12 h, 16 h and 20 h belonged to the night period (21:00~5:00). An individual sample containing one bud and two leaves was picked from one tea plant,

quickly frozen with liquid nitrogen and stored at −80 °C. Three biological replicates were performed for each sample.

4.2. RNA Extraction

Extraction of total RNA from tea leaves was performed using a reference RNA extraction kit (RNA simple total RNA Kit, Beijing Tiangen Company, Beijing, China). The concentration of RNA samples was determined using micro ultraviolet detector Nanodrop ND-1000 (Shanghai Spectrometer, Shanghai, China). The RNA quality was detected via 1.2% agarose gel electrophoresis.

4.3. cDNA Library Construction and High-Throughput Sequencing

Six groups of RNA samples were sent to Bena Technology Co., Ltd. (Wuhan, China) for construction and sequencing of the cDNA library. mRNA was enriched using oligo (dT) magnetic beads. Interrupt reagent was added to the enriched mRNA to fragment the mRNA. The interrupted mRNA was synthesized into one-strand cDNA and two-strand cDNA according to the corresponding procedure. After terminal repair, adding an A tail, adding a sequencing joint, purification, PCR amplification and product cyclization were among the steps were followed to complete the preparation of the library. We did so using the MGIEasy RNA Library Prep Kit (Huada intelligent manufacturing Technology Co., Ltd., Shenzhen, China). After qualified library detection, the high-throughput sequencing platform DNBSEQ was used for sequencing.

4.4. RNA-Seq Data and Enrichment Analysis of Differentially Expressed Transcripts

The transcription factors were predicted through the database, and the top 20 transcription factors were selected for analysis. In this study, we focused on MYB transcription factors involved in the response to the circadian rhythm. According to the transcription factor annotation results, after removing duplicate transcription factors, 22 transcripts of the MYB transcription factor were obtained. Since the original sequencing data may contain low-quality sequences, joint sequences, etc., in order to ensure the reliability of the information analysis results, the original sequencing data should be filtered to obtain clean reads. After quality control in FastQC (version: 0.11.9; Parameter: default) [88], the transcriptome samples were trimmed using Trim Galore [89]. The sequenced reads were mapped to the reference genome of tea plants using HISAT2 [90]. The reads for each gene were counted using Subread-featureCounts with default parameters [91]. The number of reads from each sample to each transcript was obtained using RSEM (RNA-Seq by Expectation Maximization) and converted using FPKM (Fragments Per Kilobase per Million bases) [92]. Paired-end reads from the same fragment were counted as one fragment to obtain the gene and transcript expression levels. Furthermore, differentially expressed gene transcripts (DETs) were identified using the edgeR package with FDR < 0.05 and $|\log FC| > 1.5$. If the number of significantly different transcripts is too small, the screening threshold is p-value < 0.05 and $|\log FC| > 1.5$. FDR values were used for significant difference filtering in this analysis [93]. This code is provided in Supplementary File S2.

4.5. Functional Annotation

In order to obtain comprehensive functional information of the new transcripts, the obtained Unigene sequences were BLAST analyzed against Nr, SwissProt, GO, KEGG, COG and other databases for protein functional annotation and classification of tea plant genes. KEGG annotations were associated with KEGG ORTHOLOGY and PATHWAY using KOBAS (version: 3.0). The Uniprot [52] database records the correspondence between each protein family and functional nodes in Gene Ontology [54]. This system predicts the biological functions of the protein sequence encoded by a transcript.

4.6. Differential Genes Were Verified Using qRT-PCR

Real-time quantitative PCR (qRT-PCR) was used to detect the expression levels of the 22 selected MYB gene transcripts related to the circadian rhythm in tea plants. Detection primers were designed using Primer Premier 5.0 software. The SYBR Premix Ex Taq kit (TaKaRa, Dalian, China) was used for qRT-PCR on the Bio-Rad IQ5 fluorescence quantitative PCR platform. The total RNA extracted from tea leaves was reverse-transcribed into cDNA using a reverse transcription kit (TaKaRa Biotech Co., Ltd., Dalian, China). The cDNA was diluted 18-fold with nuclease-free deionized water, and 2 μL of each sample was extracted as a template added to the reaction mixture (20 μL). Each cDNA required 1 μg of RNA. In this study, the *CsGAPDH* gene was selected as the internal reference gene in qRT-PCR analysis, and the amplification primers were CsGAPDH-F and CsGAPDH-R [48]. It has been reported that the stability of the *GAPDH* gene is high in *Coffea arabica* [94]. Wu et al. showed that *CsGAPDH* has a reasonable expression abundance ($19 < Cq < 30$) and can be used in the calculation of gene standardization in tea plants [48]. The cDNA of 'Baiyeyihao' was selected at intervals of 4 h within one photoperiod (24 h). The systems used for amplification were 20 μL: 10 μL SYBR Green I mix, 0.4 μL forward and reverse fluorescent quantitative primers, 2.0 μL cDNA, 7.2 μL ddH_2O [95]. The amplification program was set at 95 °C for 5 min, denatured at 95 °C for 10 s, annealed at 54 °C for 30 s and extended at 65 °C for 15 s, for a total of 40 cycles. The final concentration of primers in the reaction mix was 0.2 μM. Three biological replicates were performed, and $2^{-\Delta\Delta CT}$ was used to calculate relative gene expression levels [96]. Primers for RT-qPCR of MYB gene transcripts and the internal reference gene are listed in Table S2.

4.7. Data Processing and Analysis

Microsoft Excel 2019 software was used for data sorting. The significance of differences was analyzed using IBM SPSS Statistics 25.0, and Origin 8.0 software was used to complete the graph production. Heat maps were perpared using TBtools software (Version: 1.045) [97].

5. Conclusions

In this study, to explore the differential response of MYB transcription factors to the circadian rhythm of tea plants, the transcriptome of the tea plant 'Baiyeyihao' was sequenced at intervals of 4 h within one photoperiod (24 h). Then, 22 DETs encoding MYB and MYB-related transcription factors were identified, including three core oscillator genes (*LHY*, *RVE1* and *RVE8*). There are four types of circadian expression patterns of MYB transcription factors within a photoperiod in tea plants. Among them, the expression levels of the core oscillator genes *LHY* and *RVE1* were high at night and low during the day. The regulatory mechanism of the tea plant circadian rhythm is complex and is affected by various environmental factors. This study provides potential basic data for further research on the relevant regulatory mechanism of MYB gene transcripts in response to the circadian rhythm in tea plants.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25010657/s1.

Author Contributions: Conceptualization, J.Z., Z.H. and A.X.; methodology, Z.H., N.Z. and Z.Q.; investigation, Z.H., N.Z., Z.Q., J.L., N.Y., Y.C., J.K. and W.L.; data curation, Z.H.; writing—original draft preparation, Z.H.; writing—review and editing, Z.H., J.Z. and A.X. 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 (31870681), the Provincial Policy Guidance Program North Jiangsu Science and Technology Special Project (SZ-LYG202126), the Collection and Creation of Horticultural Crop Germplasm Resources of Jiangsu (JSFEM-202212) and the Priority Academic Program Development of Jiangsu Higher Education Institutions Project (PAPD).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw RNA-Seq reads can be found in the Genome Sequence Archive database (https://bigd.big.ac.cn/gsa/browse/CRA013441, accessed on 14 November 2023) with Bio Project ID CRA013441; the expression data (FPKM) are on Dryad (https://datadryad.org/stash/share/IJVErepIxRijiJis_wFRaBaIsS8BtDaT1oM5wy3yQX0, accessed on 22 December 2023).

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

References

1. Mori, T.; Johnson, C.H. Circadian programming in cyanobacteria. *Semin. Cell Dev. Biol.* **2001**, *12*, 271–278. [CrossRef] [PubMed]
2. Chow, B.Y.; Kay, S.A. Global approaches for telling time: Omics and the *Arabidopsis* circadian clock. *Semin. Cell Dev. Biol.* **2013**, *24*, 383–392. [CrossRef] [PubMed]
3. Wang, W.; Barnaby, J.Y.; Tada, Y.; Li, H.; Tör, M.; Caldelari, D.; Lee, D.U.; Fu, X.D.; Dong, X. Timing of plant immune responses by a central circadian regulator. *Nature* **2011**, *470*, 110–114. [CrossRef] [PubMed]
4. Inoue, K.; Araki, T.; Endo, M. Circadian clock during plant development. *J. Plant Res.* **2018**, *131*, 59–66. [CrossRef] [PubMed]
5. Song, Y.H.; Ito, S.; Imaizumi, T. Flowering time regulation: Photoperiod-and temperature-sensing in leaves. *Trends Plant Sci.* **2013**, *18*, 575–583. [CrossRef] [PubMed]
6. Xiang, Y.; Sapir, T.; Rouillard, P.; Ferrand, M.; Jiménez-Gómez, J.M. Interaction between photoperiod and variation in circadian rhythms in tomato. *BMC Plant Biol.* **2022**, *22*, 187. [CrossRef] [PubMed]
7. Huang, T.; Liu, H.; Tao, J.P.; Zhang, J.Q.; Zhao, T.M.; Hou, X.L.; Xiong, A.S.; You, X. Low light intensity elongates period and defers peak time of photosynthesis: A computational approach to circadian-clock-controlled photosynthesis in tomato. *Hortic. Res.* **2023**, *25*, 077. [CrossRef] [PubMed]
8. Song, Y.H.; Ito, S.; Imaizumi, T. Similarities in the circadian clock and photoperiodism in plants. *Curr. Opin. Plant Biol.* **2010**, *13*, 594–603. [CrossRef]
9. Zhu, C.; Peng, Q.; Fu, D.; Zhuang, D.; Yu, Y.; Duan, M.; Xie, W.; Cai, Y.; Ouyang, Y.; Lian, X.; et al. The E3 Ubiquitin Ligase HAF1 Modulates Circadian Accumulation of EARLY FLOWERING3 to Control Heading Date in Rice under Long-Day Conditions. *Plant Cell* **2018**, *30*, 2352–2367. [CrossRef]
10. Nozue, K.; Maloof, J.N. Diurnal regulation of plant growth. *Plant Cell Environ.* **2006**, *29*, 396–408. [CrossRef]
11. Yakir, E.; Hilman, D.; Harir, Y.; Green, R.M. Regulation of output from the plant circadian. *FEBS J.* **2007**, *274*, 335–345. [CrossRef] [PubMed]
12. Dodd, A.N.; Salathia, N.; Hall, A.; Kévei, E.; Tóth, R.; Nagy, F.; Hibberd, J.M.; Millar, A.J.; Webb, A.A. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **2005**, *309*, 630–633. [CrossRef] [PubMed]
13. Litthauer, S.; Battle, M.W.; Lawson, T.; Jones, M.A. Phototropins maintain robust circadian oscillation of PSII operating efficiency under blue light. *Plant J.* **2015**, *83*, 1034–1045. [CrossRef] [PubMed]
14. Du, H.; Zhang, L.; Liu, L.; Tang, X.F.; Yang, W.J.; Wu, Y.M.; Huang, Y.B.; Tang, Y.X. Biochemical and molecular characterization of plant MYB transcription factor family. *Biochemistry* **2009**, *74*, 1–11. [CrossRef] [PubMed]
15. Lippold, F.; Sanchez, D.H.; Musialak, M.; Schlereth, A.; Scheible, W.R.; Hincha, D.K.; Udvardi, M.K. AtMyb41 regulates transcriptional and metabolic responses to osmotic stress in *Arabidopsis*. *Plant Physiol.* **2009**, *149*, 1761–1772. [CrossRef] [PubMed]
16. Stracke, R.; Werber, M.; Weisshaar, B. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **2001**, *4*, 447–456. [CrossRef] [PubMed]
17. Dubos, C.; Stracke, R.; Grotewold, E.; Weisshaar, B.; Martin, C.; Lepiniec, L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **2010**, *15*, 573–581. [CrossRef]
18. Xu, Z.S.; Yang, Q.Q.; Feng, K.; Xiong, A.S. Changing Carrot Color: Insertions in *DcMYB7* Alter the Regulation of Anthocyanin Biosynthesis and Modification. *Plant Physiol.* **2019**, *181*, 195–207. [CrossRef]
19. Duan, A.Q.; Deng, Y.J.; Tan, S.S.; Xu, Z.S.; Xiong, A.S. A MYB activator, DcMYB11c, regulates carrot anthocyanins accumulation in petiole but not taproot. *Plant Cell Environ.* **2023**, *46*, 2794–2809. [CrossRef]
20. Duan, A.Q.; Tan, S.S.; Deng, Y.J.; Xu, Z.S.; Xiong, A.S. Genome-Wide Identification and Evolution Analysis of *R2R3-MYB* Gene Family Reveals S6 Subfamily R2R3-MYB Transcription Factors Involved in Anthocyanin Biosynthesis in Carrot. *Int. J. Mol. Sci.* **2022**, *23*, 11859. [CrossRef]
21. Li, M.Y.; Xu, Z.S.; Huang, Y.; Tian, C.; Wang, F.; Xiong, A.S. Genome-wide analysis of AP2/ERF transcription factors in carrot (*Daucus carota* L.) reveals evolution and expression profiles under abiotic stress. *Mol. Genet. Genom.* **2015**, *290*, 2049–2061. [CrossRef]
22. Wang, X.; Niu, Y.; Zheng, Y. Multiple Functions of MYB Transcription Factors in Abiotic Stress Responses. *Int. J. Mol. Sci.* **2021**, *22*, 6125. [CrossRef] [PubMed]
23. Baldoni, E.; Genga, A.; Cominelli, E. Plant MYB Transcription Factors: Their Role in Drought Response Mechanisms. *Int. J. Mol. Sci.* **2015**, *6*, 15811–15851. [CrossRef] [PubMed]

24. Feng, K.; Xu, Z.S.; Que, F.; Liu, J.X.; Wang, F.; Xiong, A.S. An R2R3-MYB transcription factor, OjMYB1, functions in anthocyanin biosynthesis in *Oenanthe javanica*. *Planta* **2018**, *247*, 301–315. [CrossRef]
25. Wu, W.; Zhu, S.; Zhu, L.; Wang, D.; Liu, Y.; Liu, S.; Zhang, J.; Hao, Z.; Lu, Y.; Cheng, T.; et al. Characterization of the Liriodendron Chinense *MYB* Gene Family and Its Role in Abiotic Stress Response. *Front. Plant Sci.* **2021**, *12*, 641280. [CrossRef] [PubMed]
26. Wang, W.L.; Wang, Y.X.; Li, H.; Liu, Z.W.; Cui, X.; Zhuang, J. Two MYB transcription factors (CsMYB2 and CsMYB26) are involved in flavonoid biosynthesis in tea plant [*Camellia sinensis* (L.) O. Kuntze]. *BMC Plant Biol.* **2018**, *18*, 288. [CrossRef]
27. Wang, W.L.; Cui, X.; Wang, Y.X.; Li, Z.W.; Zhuang, J. Members of R2R3-type MYB transcription factors from subgroups 20 and 22 are involved in abiotic stress response in tea plants. *Biotechnol. Biotechnol. Equip.* **2018**, *32*, 1141–1153. [CrossRef]
28. Li, M.; Li, Y.; Guo, L.; Gong, N.; Pang, Y.; Jiang, W.; Liu, Y.; Jiang, X.; Zhao, L.; Wang, Y.; et al. Functional Characterization of Tea (*Camellia sinensis*) MYB4a Transcription Factor Using an Integrative Approach. *Front. Plant Sci.* **2017**, *8*, 943. [CrossRef]
29. Katiyar, A.; Smita, S.; Lenka, S.K.; Rajwanshi, R.; Chinnusamy, V.; Bansal, K.C. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genom.* **2012**, *13*, 544. [CrossRef]
30. Carré, I.A.; Kim, J.Y. MYB transcription factors in the *Arabidopsis* circadian clock. *J. Exp. Bot.* **2002**, *53*, 1551–1557. [CrossRef]
31. Ramsay, N.A.; Glover, B.J. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* **2005**, *10*, 63–70. [CrossRef] [PubMed]
32. Duan, M.; Huang, P.; Yuan, X.; Chen, H.; Huang, J.; Zhang, H. CMYB1 encoding a MYB transcriptional activator is involved in abiotic stress and circadian rhythm in rice. *Sci. World J.* **2014**, *2014*, 178038. [CrossRef] [PubMed]
33. Farinas, B.; Mas, P. Histone acetylation and the circadian clock: A role for the MYB transcription factor RVE8/LCL5. *Plant Signal Behav.* **2011**, *6*, 541–543. [CrossRef] [PubMed]
34. Nakamichi, N.; Kiba, T.; Henriques, R.; Mizuno, T.; Chua, N.H.; Sakakibara, H. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the *Arabidopsis* circadian clock. *Plant Cell* **2010**, *22*, 594–605. [CrossRef]
35. Rawat, R.; Schwartz, J.; Jones, M.A.; Sairanen, I.; Cheng, Y.; Andersson, C.R.; Zhao, Y.; Ljung, K.; Harmer, S.L. REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16883–16888. [CrossRef]
36. Xing, H.; Wang, P.; Cui, X.; Zhang, C.; Wang, L.; Liu, X.; Yuan, L.; Li, Y.; Xie, Q.; Xu, X. LNK1 and LNK2 recruitment to the evening element require morning expressed circadian related MYB-like transcription factors. *Plant Signal Behav.* **2015**, *10*, e1010888. [CrossRef]
37. Nusinow, D.A.; Helfer, A.; Hamilton, E.E.; King, J.J.; Imaizumi, T.; Schultz, T.F.; Farré, E.M.; Kay, S.A. The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **2011**, *475*, 398–402. [CrossRef]
38. Cai, Y.; Liu, Y.; Fan, Y.; Li, X.; Yang, M.; Xu, D.; Wang, H.; Deng, X.W.; Li, J. MYB112 connects light and circadian clock signals to promote hypocotyl elongation in *Arabidopsis*. *Plant Cell* **2023**, *35*, 3485–3503. [CrossRef]
39. Bendix, C.; Marshall, C.M.; Harmon, F.G. Circadian Clock Genes Universally Control Key Agricultural Traits. *Mol. Plant* **2015**, *8*, 1135–1152. [CrossRef]
40. Zhao, X.; Li, P.; Zuo, H.; Peng, A.; Lin, J.; Li, P.; Wang, K.; Tang, Q.; Tadege, M.; Liu, Z.; et al. CsMYBL2 homologs modulate the light and temperature stress-regulated anthocyanin and catechins biosynthesis in tea plants (*Camellia sinensis*). *Plant J.* **2023**, *115*, 1051–1070. [CrossRef]
41. Jiao, T.; Huang, Y.; Wu, Y.L.; Jiang, T.; Li, T.; Liu, Y.; Liu, Y.; Han, Y.; Liu, Y.; Jiang, X.; et al. Functional diversity of subgroup 5 R2R3-MYBs promoting proanthocyanidin biosynthesis and their key residues and motifs in tea plant. *Hortic. Res.* **2023**, *10*, uhad135. [CrossRef] [PubMed]
42. Han, M.H.; Yang, N.; Wan, Q.W.; Teng, R.M.; Duan, A.Q.; Wang, Y.H.; Zhuang, J. Exogenous melatonin positively regulates lignin biosynthesis in *Camellia sinensis*. *Int. J. Biol. Macromol.* **2021**, *179*, 485–499. [CrossRef]
43. Chen, X.; Wang, P.; Gu, M.; Lin, X.; Hou, B.; Zheng, Y.; Sun, Y.; Jin, S.; Ye, N. R2R3-MYB transcription factor family in tea plant (*Camellia sinensis*): Genome-wide characterization, phylogeny, chromosome location, structure and expression patterns. *Genomics* **2021**, *113*, 1565–1578. [CrossRef] [PubMed]
44. Bordoni, A.; Hrelia, S.; Angeloni, C.; Giordano, E.; Guarnieri, C.; Caldarera, C.M.; Biagi, P.L. Green tea protection of hypoxia/reoxygenation injury in cultured cardiac cells. *J. Nutr. Biochem.* **2002**, *13*, 103–111. [CrossRef] [PubMed]
45. Lambert, J.D.; Hong, J.; Yang, G.Y.; Liao, J.; Yang, C.S. Inhibition of carcinogenesis by polyphenols: Evidence from laboratory investigations. *Am. J. Clin. Nutr.* **2005**, *81*, 284–291. [CrossRef]
46. Evans, J.R.; Loreto, F. Acquisition and diffusion of CO2 in higher plant leaves. *Photosynth. Physiol. Metab.* **2000**, *9*, 321–351.
47. Zhang, W.; Zhang, Y.; Qiu, H.; Guo, Y.; Wan, H.; Zhang, X.; Scossa, F.; Alseekh, S.; Zhang, Q.; Wang, P.; et al. Genome assembly of wild tea tree DASZ reveals pedigree and selection history of tea varieties. *Nat. Commun.* **2020**, *11*, 3719. [CrossRef] [PubMed]
48. Wu, Z.J.; Tian, C.; Jiang, Q.; Li, X.H.; Zhuang, J. Selection of suitable reference genes for qRT-PCR normalization during leaf development and hormonal stimuli in tea plant (*Camellia sinensis*). *Sci. Rep.* **2016**, *6*, 19748. [CrossRef]
49. Xia, E.H.; Li, F.D.; Tong, W.; Li, P.H.; Wu, Q.; Zhao, H.J.; Ge, R.H.; Li, R.P.; Li, Y.Y.; Zhang, Z.Z.; et al. Tea Plant Information Archive: A comprehensive genomics and bioinformatics platform for tea plant. *Plant Biotechnol. J.* **2019**, *17*, 1938–1953. [CrossRef]
50. Deng, Y.; Li, J.; Wu, S.; Zhu, Y.; Chen, Y.; He, F. Integrated nr database in protein annotation system and its localization. *Comput. Eng.* **2006**, *32*, 71–72.
51. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [CrossRef] [PubMed]

52. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; et al. UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115–D119. [CrossRef]
53. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [CrossRef]
54. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]
55. 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]
56. Jin, J.; Tian, F.; Yang, D.C.; Meng, Y.Q.; 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] [PubMed]
57. Liu, J.; Hua, W.; Yang, H.L.; Zhan, G.M.; Li, R.J.; Deng, L.B.; Wang, X.F.; Liu, G.H.; Wang, H.Z. The *BnGRF2* gene (*GRF2-like* gene from *Brassica napus*) enhances seed oil production through regulating cell number and plant photosynthesis. *J. Exp. Bot.* **2012**, *63*, 3727–3740. [CrossRef]
58. Li, P.; Xia, E.; Fu, J.; Xu, Y.; Zhao, X.; Tong, W.; Tang, Q.; Tadege, M.; Fernie, A.R.; Zhao, J. Diverse roles of MYB transcription factors in regulating secondary metabolite biosynthesis, shoot development, and stress responses in tea plants (*Camellia sinensis*). *Plant J.* **2022**, *110*, 1144–1165. [CrossRef]
59. Bian, S.C.; Lu, Y.N.; Xu, W.J.; Chen, B.Q.; Wang, G.L.; Xiong, A.S. Garlic Circadian Clock Genes *AsRVE1* and *AsRVE2* and Their Expression Analysis Under Osmotic Stress. *Acta Hortic. Sin.* **2021**, *48*, 1706–1716.
60. Kim, N.S.; Kim, S.J.; Jo, J.S.; Lee, J.G.; Lee, S.I.; Kim, D.H.; Kim, J.A. The *BrGI* Circadian Clock Gene Is Involved in the Regulation of Glucosinolates in Chinese Cabbage. *Genes* **2021**, *12*, 1664. [CrossRef]
61. Vannozzi, A.; Wong, D.C.J.; Höll, J.; Hmmam, I.; Matus, J.T.; Bogs, J.; Ziegler, T.; Dry, I.; Barcaccia, G.; Lucchin, M. Combinatorial Regulation of Stilbene Synthase Genes by WRKY and MYB Transcription Factors in Grapevine (*Vitis vinifera* L.). *Plant Cell Physiol.* **2018**, *59*, 1043–1059. [CrossRef] [PubMed]
62. Cai, H.; Chen, Y.; Zhang, M.; Cai, R.; Cheng, B.; Ma, Q.; Zhao, Y. A novel GRAS transcription factor, *ZmGRAS20*, regulates starch biosynthesis in rice endosperm. *Physiol. Mol. Biol. Plants* **2017**, *23*, 143–154. [CrossRef] [PubMed]
63. Xie, N.; Huang, X.; Zhou, J.; Song, X.; Lin, J.; Yan, M.; Zhu, M.; Li, J.; Wang, K. The R2R3-MYB transcription factor CsMYB42 regulates theanine biosynthesis in albino tea leaves. *Plant Sci.* **2023**, *336*, 111850. [CrossRef] [PubMed]
64. Allan, A.C.; Hellens, R.P.; Laing, W.A. MYB transcription factors that colour our fruit. *Trends Plant Sci.* **2008**, *13*, 99–102. [CrossRef]
65. Wang, S.; Xu, Z.; Yang, Y.; Ren, W.; Fang, J.; Wan, L. Genome-wide analysis of *R2R3-MYB* genes in cultivated peanut (*Arachis hypogaea* L.): Gene duplications, functional conservation, and diversification. *Front. Plant Sci.* **2023**, *14*, 1102174. [CrossRef]
66. Zeng, Y.; Li, Z.; Chen, Y.; Li, W.; Wang, H.B.; Shen, Y. Global dissection of R2R3-MYB in *Pogostemon cablin* uncovers a species-specific R2R3-MYB clade. *Genomics* **2023**, *115*, 110643. [CrossRef]
67. Covington, M.F.; Maloof, J.N.; Straume, M.; Kay, S.A.; Harmer, S.L. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biol.* **2008**, *9*, R130. [CrossRef]
68. Hazen, S.P.; Naef, F.; Quisel, T.; Gendron, J.M.; Chen, H.; Ecker, J.R.; Borevitz, J.O.; Kay, S.A. Exploring the transcriptional landscape of plant circadian rhythms using genome tiling arrays. *Genome Biol.* **2009**, *10*, R17. [CrossRef]
69. Farinas, B.; Mas, P. Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *Plant J.* **2011**, *66*, 318–329. [CrossRef]
70. Harmer, S.L. The circadian system in higher plants. *Annu. Rev. Plant Biol.* **2009**, *60*, 357–377. [CrossRef]
71. Schaffer, R.; Ramsay, N.; Samach, A.; Corden, S.; Putterill, J.; Carré, I.A.; Coupland, G. The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **1998**, *93*, 1219–1229. [CrossRef] [PubMed]
72. Wang, Z.Y.; Tobin, E.M. Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **1998**, *93*, 1207–1217. [CrossRef] [PubMed]
73. Wang, Z.Y.; Kenigsbuch, D.; Sun, L.; Harel, E.; Ong, M.S.; Tobin, E.M. A Myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhcb* gene. *Plant Cell* **1997**, *9*, 491–507.
74. Jung, C.; Muller, A.E. Flowering time control and applications in plant breeding. *Trends Plant Sci.* **2009**, *14*, 563–573. [CrossRef] [PubMed]
75. Shim, J.S.; Kubota, A.; Imaizumi, T. Circadian clock and photoperiodic fowering in *Arabidopsis*: CONSTANS is a hub for signal integration. *Plant Physiol.* **2017**, *173*, 5–15. [CrossRef]
76. Pérez-García, P.; Ma, Y.; Yanovsky, M.J.; Mas, P. Time-dependent sequestration of RVE8 by LNK proteins shapes the diurnal oscillation of anthocyanin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5249–5253. [CrossRef] [PubMed]
77. Bian, S.; Li, R.; Xia, S.; Liu, Y.; Jin, D.; Xie, X.; Dhaubhadel, S.; Zhai, L.; Wang, J.; Li, X. Soybean CCA1-like MYB transcription factor GmMYB133 modulates isofavonoid biosynthesis. *Biochem. Biophys. Res. Commun.* **2018**, *507*, 324–329. [CrossRef] [PubMed]
78. Li, X.; Wu, T.; Liu, H.; Zhai, R.; Wen, Y.; Shi, Q.; Yang, C.; Wang, Z.; Ma, F.; Xu, L. REVEILLE transcription factors contribute to the nighttime accumulation of anthocyanins in 'Red Zaosu' (*Pyrus bretschneideri* Rehd.) pear fruit skin. *Int. J. Mol. Sci.* **2020**, *21*, 1634. [CrossRef]

79. Liu, Z.; Zhu, X.; Liu, W.; Qi, K.; Xie, Z.; Zhang, S.; Wu, J.; Wang, P. Characterization of the REVEILLE family in Rosaceae and role of PbLHY in flowering time regulation. *BMC Genom.* **2023**, *24*, 49. [CrossRef]
80. Millar, A.J.; Kay, S.A. Integration of circadian and phototransduction pathways in the network controlling *CAB* gene transcription in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15491–15496. [CrossRef]
81. Yarkhunova, Y.; Guadagno, C.R.; Rubin, M.J.; Davis, S.J.; Ewers, B.E.; Weinig, C. Circadian rhythms are associated with variation in photosystem II function and photoprotective mechanisms. *Plant Cell Environ.* **2018**, *41*, 2518–2529. [CrossRef] [PubMed]
82. Jiang, Q.; Xu, Z.S.; Wang, F.; Li, M.Y.; Ma, J.; Xiong, A.S. Effects of abiotic stresses on the expression of *Lhcb1* gene and photosynthesis of *Oenanthe javanica* and *Apium graveolens*. *Biol. Plant.* **2014**, *58*, 256–264. [CrossRef]
83. Glossop, N.R.; Lyons, L.C.; Hardin, P.E. Interlocked feedback loops within the Drosophila circadian oscillator. *Science* **1999**, *286*, 766–768. [CrossRef] [PubMed]
84. Lee, K.; Loros, J.J.; Dunlap, J.C. Interconnected feedback loops in the Neurospora circadian system. *Science* **2000**, *289*, 107–110. [CrossRef] [PubMed]
85. Takata, N.; Saito, S.; Saito, C.T.; Nanjo, T.; Shinohara, K.; Uemura, M. Molecular phylogeny and expression of poplar circadian clock genes, *LHY1* and *LHY2*. *New Phytol.* **2009**, *181*, 808–819. [CrossRef] [PubMed]
86. Li, C.; Liu, X.J.; Yan, Y.; Alam, M.S.; Liu, Z.; Yang, Z.K.; Tao, R.F.; Yue, E.K.; Duan, M.H.; Xu, J.H. *OsLHY* is involved in regulating fowering through the Hd1- and Ehd1- mediated pathways in rice (*Oryza sativa* L.). *Plant Sci.* **2022**, *315*, 111145. [CrossRef]
87. Zhang, Y.; Xu, P.; Xue, W.; Zhu, W.; Yu, X. Diurnal gene oscillations modulated by RNA metabolism in tomato. *Plant J.* **2023**, *116*, 728–743. [CrossRef]
88. Davis, E.M.; Sun, Y.; Liu, Y.; Kolekar, P.; Shao, Y.; Szlachta, K.; Mulder, H.L.; Ren, D.; Rice, S.V.; Wang, Z.; et al. SequencErr: Measuring and suppressing sequencer errors in next-generation sequencing data. *Genome Biol.* **2021**, *22*, 37. [CrossRef]
89. Brown, J.; Pirrung, M.; Mccue, L.A. FQC Dashboard: Integrates FastQC results into a web-based, interactive, and extensible FASTQ qualaity control tool. *Bioinformatics* **2017**, *33*, 3137–3139. [CrossRef]
90. Si, Y.M.; Xing, Y.Q.; Cai, L. Differential splicing event analysis of liver tumor-educated blood platelets RNA-seq data with Hisat2 and MISO. *J. Inn. Mong. Univ. Sci. Technol.* **2016**, *34*, 274–279.
91. Liao, Y.; Smyth, G.K.; Shi, W. The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **2013**, *41*, e108. [CrossRef] [PubMed]
92. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [CrossRef] [PubMed]
93. 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]
94. Barsalobres-Cavallari, C.F.; Severino, F.E.; Maluf, M.P.; Maia, I.G. Identification of suitable internal control genes for expression studies in Coffea arabica under different experimental conditions. *BMC Mol. Biol.* **2009**, *10*, 1. [CrossRef] [PubMed]
95. Dragan, A.I.; Pavlovic, R.; McGivney, J.B.; Casas-Finet, J.R.; Bishop, E.S.; Strouse, R.J.; Schenerman, M.A.; Geddes, C.D. SYBR Green I: Fluorescence properties and interaction with DNA. *J. Fluoresc.* **2012**, *22*, 1189–1199. [CrossRef]
96. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, e45. [CrossRef]
97. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

Metabonomics and Transcriptomic Analysis of Free Fatty Acid Synthesis in Seedless and Tenera Oil Palm

Lu Wei [1,2,3,†], Cheng Yang [1,2,†], Jerome Jeyakumar John Martin [1,2], Rui Li [1,2], Lixia Zhou [1,2], Shuanghong Cheng [4], Hongxing Cao [1,2,*] and Xiaoyu Liu [1,2,*]

[1] Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang 571339, China; wl79912021@163.com (L.W.); 15730095249@163.com (C.Y.); jeromejeyakumarj@gmail.com (J.J.J.M.); lirui@catas.cn (R.L.); glzz_2009@163.com (L.Z.)
[2] National Key Laboratory for Tropical Crop Breeding, Haikou 571101, China
[3] School of Horticulture, Hainan University, Haikou 570228, China
[4] College of Tropical Crops, Yunnan Agricultural University, Pu'er 665000, China; 1999033@ynau.edu.cn
* Correspondence: caohx@catas.cn (H.C.); liuxy86@catas.cn (X.L.)
† These authors contributed equally to this work.

Abstract: Oil palm, a tropical woody oil crop, is widely used in food, cosmetics, and pharmaceuticals due to its high production efficiency and economic value. Palm oil is rich in free fatty acids, polyphenols, vitamin E, and other nutrients, which are beneficial for human health when consumed appropriately. Therefore, investigating the dynamic changes in free fatty acid content at different stages of development and hypothesizing the influence of regulatory genes on free fatty acid metabolism is crucial for improving palm oil quality and accelerating industry growth. LC-MS/MS is used to analyze the composition and content of free fatty acids in the flesh after 95 days (MS1 and MT1), 125 days (MS2 and MT2), and 185 days (MS3 and MT3) of Seedless (MS) and Tenera (MT) oil palm species fruit pollination. RNA-Seq was used to analyze the expression of genes regulating free fatty acid synthesis and accumulation, with differences in genes and metabolites mapped to the KEGG pathway map using the KEGG (Kyoto encyclopedia of genes and genomes) enrichment analysis method. A metabolomics study identified 17 types of saturated and 13 types of unsaturated free fatty acids during the development of MS and MT. Transcriptomic research revealed that 10,804 significantly different expression genes were acquired in the set differential gene threshold between MS and MT. The results showed that FabB was positively correlated with the contents of three main free fatty acids (stearic acid, myristate acid, and palmitic acid) and negatively correlated with the contents of free palmitic acid in the flesh of MS and MT. ACSL and FATB were positively correlated with the contents of three main free fatty acids and negatively correlated with free myristate acid. The study reveals that the expression of key enzyme genes, FabB and FabF, may improve the synthesis of free myristate in oil palm flesh, while FabF, ACSL, and FATB genes may facilitate the production of free palmitoleic acid. These genes may also promote the synthesis of free stearic acid and palmitoleic acid in oil palm flesh. However, the FabB gene may inhibit stearic acid synthesis, while ACSL and FATB genes may hinder myristate acid production. This study provides a theoretical basis for improving palm oil quality.

Keywords: oil palm; metabolomics; transcriptomics; lipid synthesis; free fatty acid

1. Introduction

Oil palm, a highly versatile crop, can be categorized into two primary resource types: African oil palm (*Elaeis guineensis* Jacq.) and American oil palm (*E. oleifera*, previously *E. melanococca*). African oil palm is the most widely cultivated and commercially significant variety, known for its high oil yield and adaptability to different climatic conditions. On the other hand, American oil palm is less commonly grown and is mainly utilized for breeding purposes to develop new hybrid varieties with improved traits [1]. Oil palm

is widely cultivated in Malaysia, Indonesia, Nigeria, and other countries has brought substantial commercial value. The large-scale cultivation of oil palm has proven to be highly profitable due to its high oil yield and versatility in various industries. However, at present, some countries in Southeast Asia expand the planting area of oil palm by reducing the area of forest, which directly poses a certain threat to biodiversity, reduces the habitat of organisms, and affects the formation of the greenhouse effect. Therefore, it is very important to maintain a good balance between the planting area of oil palm and the forest cover area, which is related to whether to maximize the economic value of oil palm under the premise of protecting the environment and biodiversity [2–4]. The classification of oil palm is based on the thickness of the fruit shell, which can be divided into three types: Dura, Pisifera, and Tenera. This categorization helps in understanding the different characteristics and uses of each type. In addition to the mentioned matrilineal varieties, there are also paternal varieties of certain crops. For instance, in the case of certain crops like Deli dura and Bamenda, they are primarily matrilineal varieties. On the other hand, Ekona and Nigeria are examples of paternal varieties. These different varieties have their own distinct characteristics, including traits such as yield, resistance to diseases, and adaptability to specific environmental conditions.

China's decision to import oil palm in the 1920s was a significant move that involved importing a substantial amount of germplasm from countries like Indonesia. This decision had far-reaching implications for China's agricultural landscape and economy. By importing oil palm, China was able to diversify its crop production and tap into the economic benefits associated with this versatile plant. Currently, the oil palm has been introduced and cultivated in various provinces in China, including Hainan, Yunnan, Guangdong, Guangxi, and Fujian [5]. The seedless oil palm variety mentioned in this study is a result of hybridization between O×G Amazon species and contains higher unsaturated fatty acids. The Tenera oil palm is a variety produced by Deli×Ghana hybridization and has a higher oil content in its fruits. Therefore, both varieties are good experimental materials and can be used for correlation analysis of their fatty acids.

Oil palm is one of the most productive tropical crops globally, with the highest oil production [6]. The oil is extracted from the fruit, known as Crude Palm Oil (CPO) [7]. CPO, or Crude Palm CPO, is not only edible, but also plays an important role in various industries. It is utilized in the production of cosmetics, detergents, chemicals, and pharmaceuticals [8,9]. CPO is mainly composed of saturated fatty acids with a significant portion of palmitic acid and oleic acid and unsaturated fatty acids, with oleic acid and linoleic acid being the most prominent ones. In addition to fatty acids, it also contains vitamins, fatty alcohols, carotenoids, and other nutrients [10]. The content of fatty acids in flesh and the content ratio of different components directly affect the yield and quality of CPO. The current focus of oil palm biotechnology research is to enhance the levels of unsaturated fatty acids such as oleic acid and linoleic acid in flesh [11], so as to improve the oil yield of oil palm fruit and the quality of CPO. In addition, the content of free fatty acids in palm oil serves as a crucial quality indicator for various aspects of its storage, marketing, and production time, as well as the price of palm oil products [12], and the high content of free fatty acids leads to poor fat quality [13]. Therefore, the quality of palm oil is closely related to the content of free fatty acids.

The synthesis of oil palm fatty acids is the most important phase in oil formation and serves as the basis for yield formation. Improving palm oil yield can be achieved by mining genes and transcription factors related to oil palm fatty acid synthesis and analyzing their regulatory effects [14–16]. At present, the genes related to fatty acid synthesis of oil palm have been studied. For instance, Cytochrome P450 is a superfamily of genes involved in various biochemical reactions [17,18], and plays an important role in plant fatty acid metabolism. According to Liang et al. [19], the concentration of fatty acids in the peel of oil palm increases consistently throughout its development, reaching its highest point at maturity. This similarity in expression pattern P450 genes in the same tissue, indicating that cytochrome P450 expression impacts the metabolism of oil palm fatty acids (oxidation,

epoxidation, hydrocarbylation, etc.), thereby influences the composition and content of oil palm fatty acids. During the rapid accumulation of fatty acids in African oil palms, a recent study found a significant increase in the expression levels of EgSAD1 and EgSAD2. Meanwhile, the hybrid oil palms exhibited a notably higher content of unsaturated fatty acids than the African oil palms. The differential expression of OeSADs in hybrid oil palms may contribute to the increased unsaturated fatty acids in these palms. This suggests that the high expression of SAD genes could potentially influence the fatty acid composition of hybrid oil palms [20]. The overexpression of EgFatB1, EgFatB2, and EgFatB3 in oil palm was successfully transferred to wild Arabidopsis Thaliana. This transfer revealed that the overexpression of these genes resulted in an increased palmitic acid content in transgenic seeds. EgFatB3 is an important enzyme that positively regulates medium-chain fatty acid synthesis. This enzyme is responsible for the production of new fatty acids, specifically lauric acid and myristate acid. Interestingly, out of the two, myristate acid tends to accumulate the most. It is speculated that EgFatB1, EgFatB2, and EgFatB3 in oil palm regulate the synthesis of palmitic acid, and EgFatB3 promotes the synthesis of medium-chain fatty acids [21].

The study aims to analyze the changes in free fatty acid content and key enzyme gene expression in Seedless (O×G Amazon) and Tenera (Deli×Ghana) oil palm fruits after pollination. The changes in free fatty acid content of two fruit varieties were examined at different time points after postharvest. The study reported on the levels of free fatty acids at 0 h, 24 h, and 36 h., Additionally, candidate genes that may play a role in regulating these changes in free fatty acid content were identified [22]. By combining metabolomics and transcriptomics analysis of changes in free fatty acids during the development stage of oil palm fruits is crucial in understanding the metabolic processes that occur during this period., By identifying the key genes involved in the production of free fatty acids, valuable insights into the factors that influence oil accumulation in oil palm fruits can be gained. This knowledge can then be utilized to improve oil palm breeding programs and enhance oil production in these two varieties. Therefore, understanding the dynamic changes of free fatty acids during the fruit development stage is an important step towards maximizing the potential of oil palm cultivation and meeting the growing demand for palm oil.

2. Results

2.1. Analysis of Dynamic Changes in Metabolites of Free Fatty Acids in the Flesh of Seedless and Tenera Oil Palm Species at Different Developmental Stages

After pollination, the mesocarp of oil palm fruits at different stages, namely 95d (MS1 and MT1), 125d (MS2 and MT2), and 185d (MS3 and MT3), were selected for metabolomic analysis. Three biological replicates were performed at each stage to ensure the reliability and accuracy of the results. Qualitative mass spectrometric analysis of sample metabolites was performed based on information obtained from a local lipid database, after which mass spectrometric peaks observed in different samples for each substance were corrected to ensure accurate quantification. The quantitative analysis of a randomly selected substance in different samples was integrated with the correction results (Figure S1). By performing overlapping display analysis on the total ion flow chromatogram (TIC) of mass spectrometry detection and analysis of the same quality control sample (Figure S2), the total ion flow chromatogram (TIC) of mass spectrometry detection and analysis of QC samples showed significant overlap. The CV value distribution map of the QC samples reveals that more than 75% of substances have CV values less than 0.3 (Figure S3), exhibiting that the metabolomics data are relatively stable and reliable. Principal component analysis (PCA) was used to segregate the six sets of data, namely MS1, MS2, MS3, MT1, MT2, and MT3, including quality control samples (Figure S4). The first principle component (PC1) and second principal component (PC2) were reported to account for 57.35% and 13.1% of the total variables, respectively.

The research on the free fatty acid metabolites in the flesh of Seedless and Tenera varieties at different development stages revealed that the main constituents were saturated

free fatty acids. In fact, the study identified a total of 17 different types of saturated free fatty acids, such as palmitic acid (C16:0), stearic acid (C18:0), decanoic acid (C10:0), and unsaturated free fatty acids, such as oleic acid (C18:1), linoleic acid (C18:2), linolenic acid (C18:3), and 13 other components (Table 1). As shown in Table 1, the total free fatty acid content of both the Seedless and Tenera varieties experienced a gradual increase as the flesh grew and developed. However, it is worth noting that there was a significant spike in the early to middle stages of growth. The oleic acid content in the flesh undergoes significant changes as it matures, with the highest level of free fatty acid content in ripe flesh. The data show a significant increase in the concentration of the substance from MS1 to MS3, with a rise from 277.30 nmol/g to 32,992.80 nmol/g. Similarly, in the MT group, the concentration increased from 4549.25 nmol/g in MT1 to 30,396.86 nmol/g in MT3. These findings suggest a substantial increase in the amount of the substance over time in both the MS and MT groups. According to the results, the linoleic acid content in Seedless oil palm pulp was found to be the least in MS1, the highest in MS3, and the next-highest in terms of free fatty acid content. However, Tenera oil palm pulp had the least amount of palmitic acid in MT1, the most in MT3, and the second highest amount in term of its free fatty acid content. As the Seedless oil palm fruit ripens from MS1 to MS3 and MT1 to MT2, the levels of palmitic acid in the flesh increase significantly. In fact, the palmitic acid content reaches its peak during this stage, making it one of the highest levels of free fatty acids found in the fruit. During the ripening stage, the Teneral oil palm's flesh experiences a significant increase in the amount of linoleic acid.

Table 1. Dynamic changes in the accumulation of respective free lipid metabolites in the mesocarp of Seedless and Tenera oil palm during different developmental periods (nmol/g).

Free Lipid Metabolites	Different Developmental Periods					
	Seedless			Tenera		
	MS1	MS2	MS3	MT1	MT2	MT3
C10:0 Decanoic acid	375.30	266.61	206.91	428.91	253.11	239.25
C12:0 Lauric acid	6.52	9.45	8.91	15.32	3.47	7.54
C14:0 Myristic acid	20.10	112.27	101.73	44.33	225.33	371.80
C15:0 Pentadecanoic acid	6.49	26.89	36.51	17.39	21.18	40.47
C16:0 Palmitic acid	401.84	9126.63	8060.21	1401.43	9585.10	14,345.15
C17:0 Heptadecanoic acid	14.20	96.47	103.22	48.70	68.62	137.42
C18:0 Stearic acid	320.25	1684.93	1756.83	451.25	1922.86	2988.75
C20:0 Arachidic acid	\	133.86	75.33	\	65.60	178.70
C22:0 Docosanoic acid	35.44	47.38	42.29	39.38	54.47	104.06
C24:0 Lignoceric acid	53.07	69.62	63.93	110.51	49.04	128.63
C28:0 Octacosanoic acid	42.90	65.53	78.27	26.15	121.22	102.38
C30:0Triacontanoic acid	51.42	57.87	11.78	13.10	9.59	3.52
C32:0 Lacceroic acid	12.25	37.60	7.80	9.84	5.54	2.13
C33:0 Psyllic acid	2.25	12.34	9.23	4.32	6.86	6.87
C34:0 Gheddic acid	\	7.89	7.88	\	5.13	7.68
C35:0 Tripentacontanoic acid	\	37.46	33.83	\	32.80	34.80
C36:0 Trihexadecanoic acid	\	30.67	29.48	\	29.70	30.63
Total of saturated free fatty acid	1342.05	11,823.46	10,634.14	2610.62	12,460.62	18,729.78
C16:1 Palmitoleic acid	3.22	182.91	247.14	25.22	80.37	118.02
C17:1 Heptadecenoic acid	\	54.57	125.01	\	30.06	41.74
C18:1 Oleic acid	277.30	31,536.14	32,992.80	4549.25	26,021.06	30,396.86
C18:2 Linoleic acid	88.94	9508.93	13,475.84	1586.44	6680.46	10,188.00
C18:3 Linolenic acid	50.36	107.42	237.34	452.89	121.90	189.08
C19:1 Nonadecenoic acid	\	16.40	29.67	\	8.45	12.10
C20:1 Eicosenoic acid	\	301.53	395.73	47.56	260.78	392.81
C20:2 Eicosadienoic acid	\	2.07	3.67	\	1.38	3.96
C22:3 Docosatrienoic acid	\	19.36	19.44	\	16.65	22.56
C22:4 Docosatetraenoic acid	\	6.05	5.60	\	5.02	6.14

Table 1. *Cont.*

Free Lipid Metabolites	Different Developmental Periods					
	Seedless			Tenera		
	MS1	MS2	MS3	MT1	MT2	MT3
C22:6 Docosahexaenoic acid	\	10.09	9.55	\	4.56	7.15
C24:5 Tactosapentaenoic acid	\	9.49	9.76	\	8.60	10.66
C24:6 Tacosahexaenoic acid	11.23	47.89	45.71	41.35	47.22	55.90
Total of unsaturated free fatty acid	431.06	41,802.85	47,597.27	6702.71	33,286.52	41,444.97
Total of free fatty acid	1773.11	53,626.311	58,231.41	9313.33	45,746.14	60,174.75

Note: "\" indicate there are no or very little free fatty acid content in this period here; MS1 and MT1: early fruit development; MS2 and MT2: middle fruit development; MS3 and MT3: late stage of fruit development. Cluster analysis of all the free fatty acids in the flesh of Seedless and Tenera oil palm seeds at three different development stages. It showed that the overall clustering of 30 free fatty acids treated by Z-score normalization fell into 2 categories (Figure 1). In Group 1, there were 24 different kinds of free fatty acids. In the beginning, they were minimal and then surged rapidly, reaching their peak at the end of the development phase. The contents of heptadecenoic acid, nonadecenoic acid and palmitoleic acid in Seedless oil palm changed most obviously in MS1-MS3, but the contents of docosahexaenoic acid, tridecanoic acid, and arachidonic acid in MS1-MS3 increased first and then decreased in MS1-MS3, with the highest value in the middle stage of development (MS2). The contents of docosanic acid, myristic acid, and arachidic acid in Tenera oil palm changed particularly in MT1-MT3, but the contents of octacosanoic acid in MT1-MT3 increased first and then decreased, peaking in the middle stage of development (MT2), and the contents of tridecanoic acid in MT2-MT3 showed a stable trend. Its content reaches its highest value in MT3. In Group 2, the six free fatty acids changed in diverse ways. The content of decanoic acid in Seedless oil palm decreased continuously, while the quantity of linolenic acid in MS1-MS3 increased. In MS1-MS3, the levels of triaconic acid, lauric acid, tetracosanoic acid, and triededecanoic acid exhibited a decreasing trend followed by an increase, reaching their highest levels in MT3.

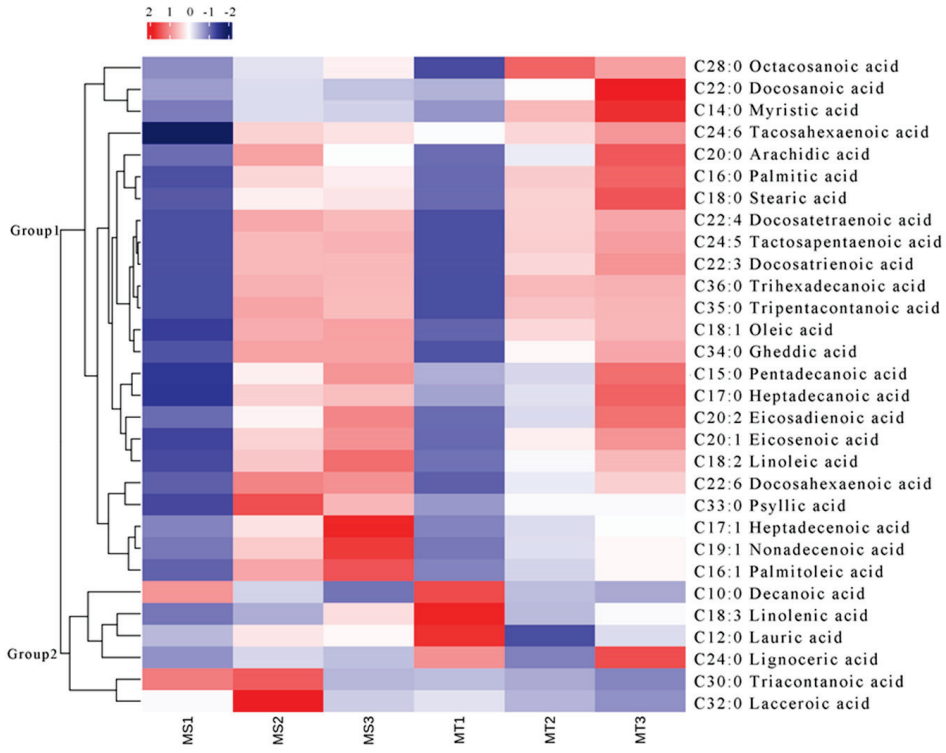


Figure 1. Heat map of metabolites clustering in the mesocarp of Seedless and Tenera oil palm during different developmental periods. The 6 columns in each heatmap represent the fruits development stages (MS1, MS2, MS3, MT1, MT2, and MT3).

The research findings indicate that Seedless oil palm has a distinct advantage over MS2 and MS3 in terms of its lower total free fatty acid content. Additionally, Seedless oil palm showcases a significantly higher percentage of saturated free fatty acid, reaching an impressive 75.69%. These results suggest that Seedless oil palm may be a more desirable option due to its reduced levels of free fatty acid and its higher proportion of saturated

free fatty acid when compared to MS2 and MS3. The unsaturated to saturated free fatty acid (UFFA/SFFA) ratio of 0.33 was found to be lower than MS2 and MS3. MT1 exhibited a lower total free fatty acid level compared to MT2 and MT3. Furthermore, the ratio of unsaturated to saturated free fatty acid content (UFFA/SFFA) in MT1 was 2.57, which was lower than that of MT2 but higher than MT3. The data presented in Figure 2 show a gradual accumulation of palm oil in both the Seedless and Tenera palm species. The concentration of free fatty acids in the Seedless oil palm has experienced a significant increase. It was 30.24 times the total content of all kinds of free fatty acids in the MS1 stage. Moreover, the percentage of unsaturated free fatty acids in the total content reached 77.96%. Additionally, the ratio of unsaturated to saturated free fatty acid (UFFA/SFFA) was recorded at 3.54, further highlighting the dominance of unsaturated fatty acids in the MS1 stage. The total content of free fatty acids in Tenera oil palm significantly increased to 45746.14 nmol/g. This was found to be 4.91 times higher than the total content of all kinds of free fatty acids observed in the MT1 stage. The percentage of unsaturated free fatty acids reached its peak at 72.70%. Additionally the ratio of unsaturated and saturated free fatty acids (UFFA/SFFA) reached the maximum at 2.67. The total content of free fatty acids in Seedless oil palm showed a significant increase, reaching 58,231.41 nmol/g, which was 32.84 times higher than the total content of all kinds of free fatty acids in MS1 and 1.09 times higher than in MS2. The ratio of unsaturated to saturated free fatty acids (UFFA/SFFA) reached its maximum (4.48). The total content of free fatty acids in Tenera oil palm increased to 60,174.75 nmol/g, which was 6.46 times higher than the total content of various free fatty acids in MT1 and 1.32 times higher than in MT2. The percentage of saturated free fatty acids reached its peak at 31.23%, surpassing the levels found in MT1 and MT2. The unsaturated free fatty acid content of MT3 is 2.21 times higher than the saturated free fatty acid content. The results indicated that unsaturated free fatty acids were the main sources of oil synthesis in the flesh of the soon-to-mature Seedless and Tenera oil palms.

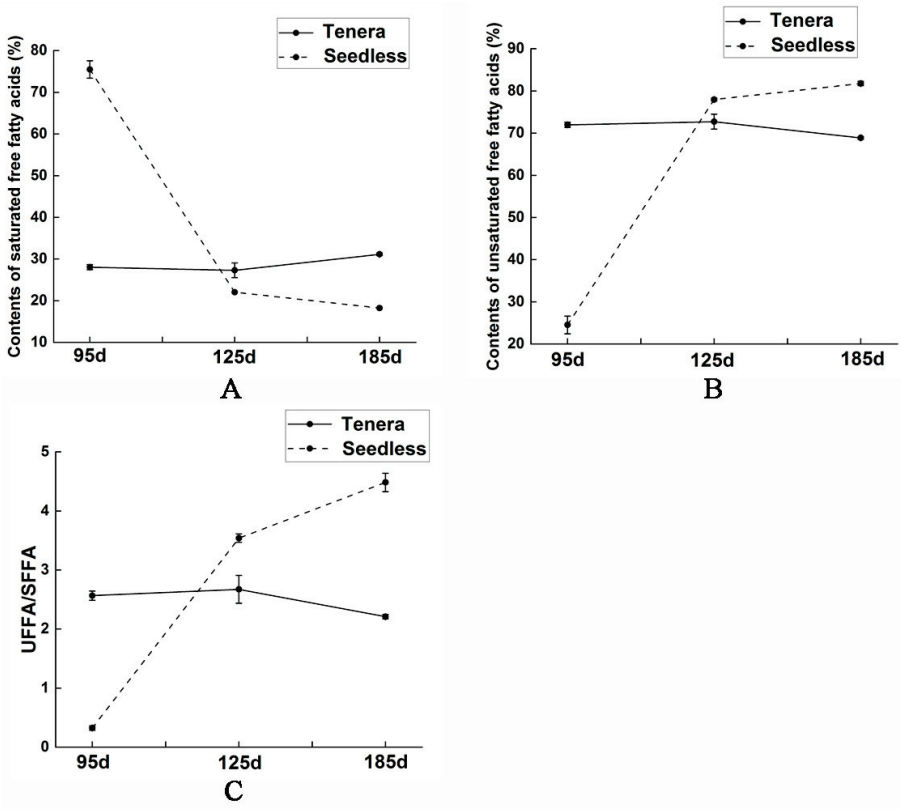


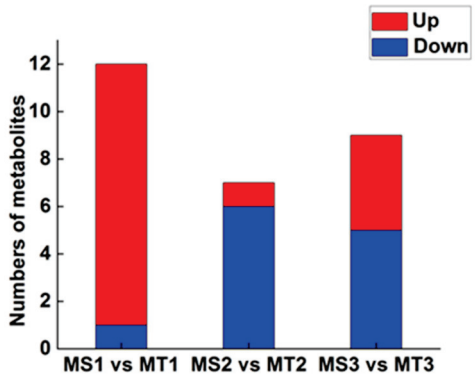
Figure 2. Dynamic changes of free fatty acid content in the mesocarp of Seedless and Tenera oil palm during different development periods. (**A**) changes in content of saturated fatty acids; (**B**) changes in

content of unsaturated fatty acids; (**C**) changes in the content ratio of unsaturated fatty acids to saturated fatty acids. Seedless and Tenera oil palm during different developmental periods In summary, the study focused on analyzing the dynamic changes in fatty acid content in the flesh of two varieties of oil palm seeds. The result revealed substantial changes in fatty acid levels as the seeds developed from the early growth stage to maturity (MS1 to MS3 and MT1 to MT3). The proportion of total unsaturated free fatty acids in the flesh of Seedless and Tenera oil palm was found to be 81.74% and 68.87%, respectively, indicating that the total unsaturated free fatty acids in Seedless oil palm were found to be relatively high, with a higher degree of unsaturated palm oil compared to Tenera oil palms.

The contents of decanoic acid, linolenic acid, lauric acid, triaconic acid, and tridecanoic acid in MT1-MT3 showed a continuous decrease and reached the highest value at MT1, while the contents of tetracosanoic acid in MT1-MT3 initially showed a decrease followed by an increase. Analyzing the dynamic changes in the metabolites of free fatty acids in three different developmental periods, namely, slow accumulation (MS1, MT1), rapid accumulation (MS2, MT2), and stable accumulation (MS3, MT3), provides valuable insights into the metabolic processes occurring at each stage. By studying the fluctuations of these metabolites, researchers can gain a better understanding into how free fatty acids are synthesized and utilized during different developmental periods.

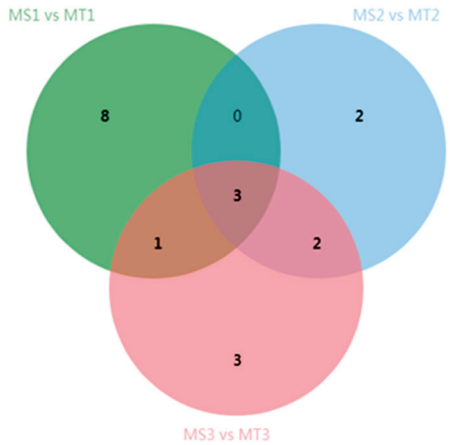
2.2. Analysis of Differential Metabolites of Free Fatty Acids in the Flesh of Oil Palm from Seedless Species and Tenera Species

The quantity of different metabolites in pairwise groups was studied by comparing the early, middle, and late stages of oil palm development. According to the variable importance in projection (VIP) ≥ 1, Fold Change ≥ 2 or Fold Change ≤ 0.5. A total of 19 metabolites of free fatty acids with significant differences were obtained. According to the comparison of differential metabolites in Seedless and Tenera oil palm species at different stages of pulp development (Figure 3A), there were 11 up-regulated metabolites and 1 down-regulated metabolite in the early stage (MS1 vs. MT1), 1 up-regulated metabolite and 6 down-regulated metabolites in the middle stage (MS2 vs. MT2), and 4 up-regulated metabolites and 5 down-regulated metabolites in the late stage (MS3 vs. MT3). The results show that there are notable differences in the levels of metabolites between MS1 vs. MT1 as well as MS3 vs. MT3. On the other hand, the comparison between MS2 and MT2 reveals a relatively low number of significantly different metabolites. Through Venn diagram analysis, it was observed that MS2 vs. MT2 and MS3 vs. MT3 exhibited the highest frequency of distinct metabolites, as depicted in Figure 3B. The results showed that the proportion of up-regulated metabolites was significantly higher than that of down-regulated metabolites, indicating that the content of most of the differential metabolites increased with the development days of flesh. In fact, it reached its maximum value in the late stage of flesh development.



(**A**)

Figure 3. *Cont.*



(**B**)

Figure 3. Oil palm differentially expressed metabolites statistics of MS and MT. (**A**) Venn diagram of differentially expressed metabolites; (**B**) histogram of significantly differentially expressed metabolites.

2.3. Analysis of Differential Genes in the Flesh of Oil Palm from Seedless Species and Tenera Species

We compared samples of Seedless and Tenera oil palm pulp at the same stage of growth (MS1 vs. MT1, MS2 vs. MT2, MS3 vs. MT3). The study focused on analyzing the relationship between specific genes and their functions in the synthesis and utilization of fatty acids to uncover the essential genetic factors involved. In this study, a total of 10804 significantly differentially expressed genes were identified in oil palm pulp. These genes were obtained based on the set differential gene threshold. The differential genes in Seedless and Tenera oil palm species were analyzed at different stages of pulp development (Figure 4A). The comparison of gene expression in the early stage (MS1 vs. MT1) revealed 3159 up-regulated genes and 3981 down-regulated genes. Similarly, during pulp development (MS2 vs. MT2), there were 2395 up-regulated genes and 2486 down-regulated genes. During the later stage of the experiment, it was observed that there were 2567 up-regulated genes and 2119 down-regulated genes when comparing MS3 to MT3. Therefore, MS1 vs. MT1 had the most significantly different genes, followed by MS2 vs. MT2, whereas MS3 vs. MT3 had the least number of significantly different genes. The number of differential genes in the three comparison groups gradually decreased with the increase in flesh development days, indicating that the expression of genes regulating the growth and development of oil palm flesh in Seedless and Tenera species tended to be stable. According to Venn diagram analysis (Figure 4B), there were 1401 common differential genes among the three comparison groups.

2.4. Enrichment Analysis of Significantly Differentially Expressed Genes

According to the KEGG analysis, the MS1 vs. MT1 group showed enrichment in 136 pathways. Interestingly, a total of 134 pathways were found to be enriched in both the MS2 vs. MT2 and MS3 vs. MT3 groups. The findings indicated that the distinct genes in MS and MT during various growth phases primarily occur in metabolic pathways, the biosynthensis of secondary metabolites, and plant-pathogen interactions. "Metabolic pathways" and "biosynthensis of secondary metabolites" are the main metabolic pathways in plant synthesis and metabolism, and are widely involved in the biosynthesis of plant metabolites.

The enrichment degree of fatty acid synthesis pathways in the three groups of comparison follows the pattern of MS3 vs. MT3 > MS1 vs. MT1 > MS2 vs. MT2, indicating that a large number of genes are expressed in the early to middle stages of MS and MT fruit development, as well as in the mature fruit to stimulate fatty acid synthesis. Therefore, exploring the expression patterns of genes involved in the fatty acid synthesis pathway is beneficial for the further exploration of oleic acid, stearic acid, palmitic acid. The synthesis and accumulation of fatty acids such as palmitoleic acid (Figure 5).

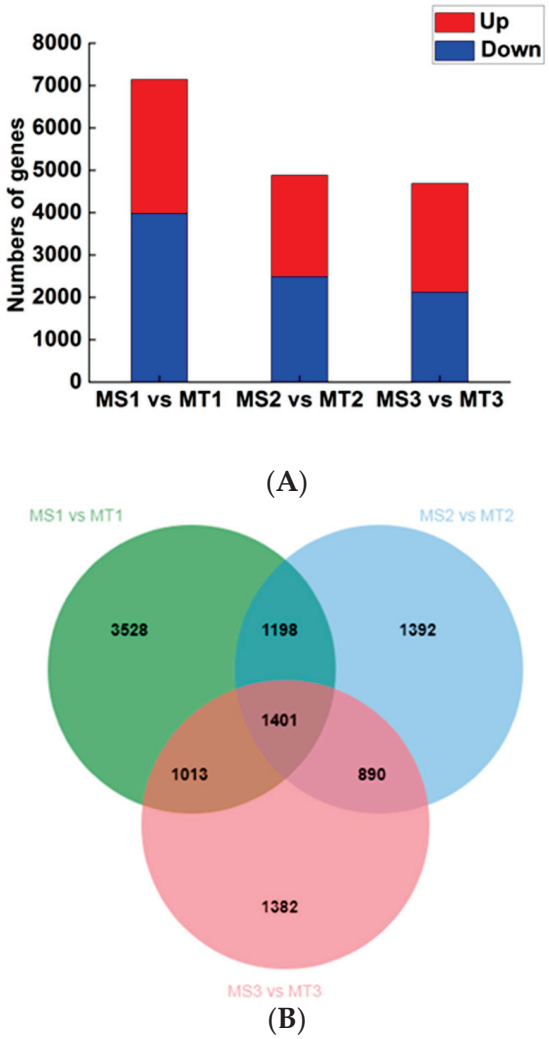


Figure 4. Oil palm differentially expressed genes statistics of MS and MT. (**A**) Venn diagram of differentially expressed genes; (**B**) histogram of significantly differentially expressed genes.

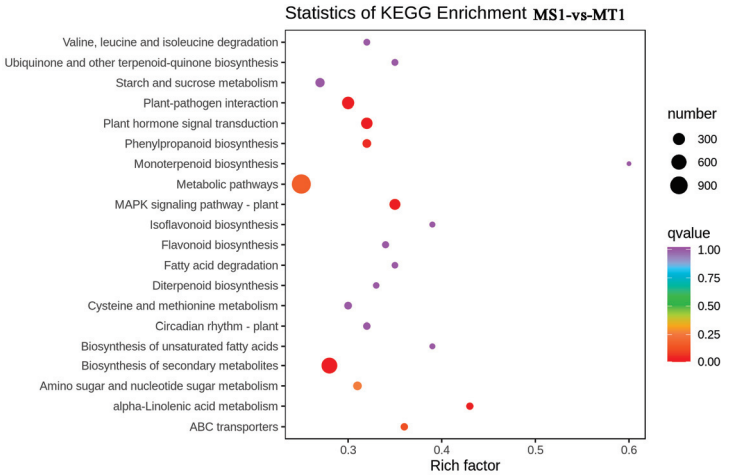


Figure 5. *Cont*.

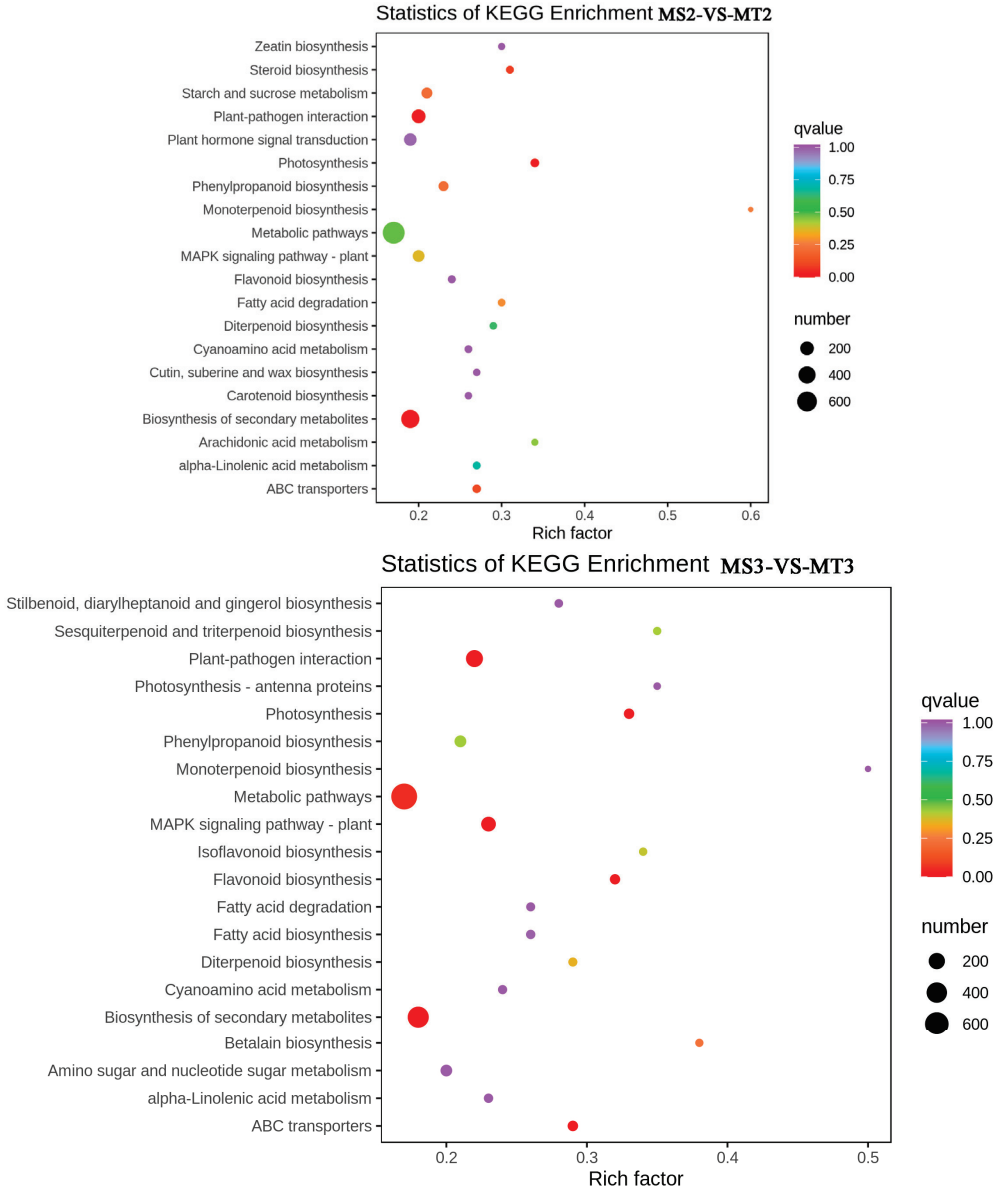


Figure 5. Top 20 of KEGG enrichment classification scatter chart for the pairwise comparisons of MS1 vs. MT1, MS2 vs. MT2, MS3 vs. MT3.

2.5. Metabolomics and Transcriptomic Association Analysis

KEGG is a comprehensive database that enables the systematic analysis of metabolism and gene function [23]. The lipid metabolome and the transcriptome data were examined together using the metabolic pathway information provided by KEGG. The study revealed that 30 types of free fatty acids, including palmitoleic acid, myristate acid, arachidonic acid, tridedecanoic acid, triacsanic acid, tetracosanoic acid, docosanic acid, eicosadienoic acid, stearic acid, octadecanoic acid, palmitoic acid, and docosahexaenoic acid were enriched in nine metabolic pathways related to lipid metabolism in oil palm (Figure 6). It was shown that seven sorts of saturated free fatty acids were significantly different in the unsaturated fatty acid biosynthesis pathway between Seedless species and Tenera species, including triacontane acid, tetracosanoic acid, arachidonic acid, tridedecanoic acid, and triacontane acid, and seven corresponding differentially expressed genes. Two free fatty acids with significant differences in the fatty acid biosynthesis pathway are myristic acid and palmitoleic acid. There are two non-significant differences in saturated free fatty acids,

stearic acid, and palmitoleic acid, and nineteen significantly different expression genes corresponding to these acids (Table 2).

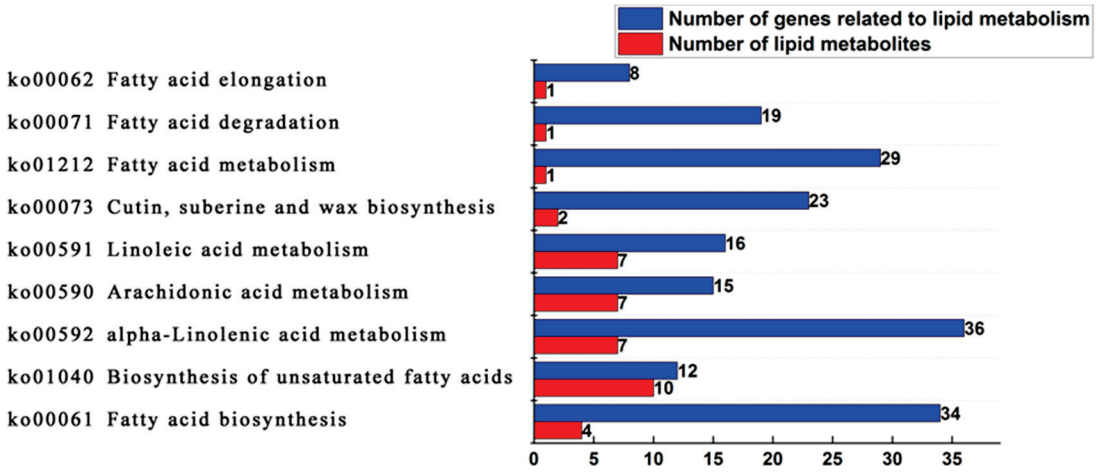


Figure 6. KEGG enrichment pathway statistics of lipid metabolites and genes related to lipid metabolism in the mesocarp of Seedless and Tenera oil palm during different developmental periods.

Table 2. Lipid metabolites and genes related to lipid metabolism statistics of the fatty acid biosynthesis (ko00061) pathway in the mesocarp of Seedless and Tenera oil palm during different developmental periods.

Comparable Group	Lipid Metabolites ID	Lipid Metabolites	Genes ID
MS1 vs. MT1	Lipid-B-N-0026	(FFA16:1) Palmitoleic acid	LOC105034507;LOC105034511;LOC105035520; LOC105039456;LOC105042279;LOC105061231; LOC105035641;LOC105040698;LOC105040700; LOC105044978;novel.3225;LOC105049274; ElguCp049;LOC105058068;LOC105037839; LOC105042027;LOC105056688;LOC105047747; LOC109506086;PTE;LOC105050310; LOC105048939
	Lipid-B-N-0005	(FFA14:0) Myristic acid	
MS2 vs. MT2	Lipid-B-N-0007	(FFA16:1) Palmitic acid	LOC105034507;LOC105034511;LOC105035520; LOC105039456;LOC105042279;LOC105033685; LOC105061231;LOC105056570;LOC105035641; LOC105040698;novel.3225;LOC105050555; ElguCp049;LOC105042027;LOC105056688; LOC105040257;LOC105047747;LOC109506086;PTE;LOC105048939
	Lipid-B-N-0009	(FFA18:0) Stearic acid	
	Lipid-B-N-0005	(FFA14:0) Myristic acid	
MS3 vs. MT3	Lipid-B-N-0026	(FFA16:1) Palmitoleic acid	LOC105034511;LOC105035520;LOC105058779;LOC105042279; LOC105035641;LOC105036494;LOC105039221;LOC105040698; LOC105032905;novel.3225;LOC105050555;ElguCp049; LOC105058068;LOC105051936;LOC105040922;LOC105042027; LOC105056688;LOC105040257;LOC105047747;LOC105050310; LOC105048939;LOC105049664;LOC105041644
	Lipid-B-N-0005	(FFA14:0) Myristic acid	

The Nr annotation results of 19 significantly differentially expressed genes in Seedless and Tenera oil palm, revealed that the enzymes during the flesh development were 3-oxoacyl-[acyl-carrier-protein] -synthase I, FabB), 3-oxoacyl-(acyl-carrier-protein) synthaseII (FabF), long-chain acyl-CoA synthetase (long-chain acyl-Coa synthetase), ACSL), and palmitoyl-[acyl-carrier-protein] -Thioesterase (FATB).

The dynamic changes in gene expression levels of four key enzymes in Seedless and Tenera oil palm species, as shown in Figure 7, revealed inconsistent patterns for the gene

expression of FabB, FabF, ACSL, and FATB. The expression levels of the FabB and ACSL genes in seedless oil palm pulp showed an interesting pattern throughout the growth stages. Initially, at the beginning of flesh development, the expression of these genes was low. However, as the process progressed, the expression levels gradually increased and reached their highest point at the middle and late stage of flesh development. During the development of Tenera palm pulp, the expression level of the FabB gene consistently increased, indicating its active role in this process. On the other hand, the expression level of the ACSL gene showed a sustained decrease, suggesting a potential downregulation. Expression levels of the FabF gene and FATB gene in Seedless oil palm showed an initial decrease followed by an increase during pulp development. In contrast, the expression levels of the FabF gene and FATB gene in Tenera palm exhibited an opposite trend, with an initial increase followed by a decrease during pulp development.

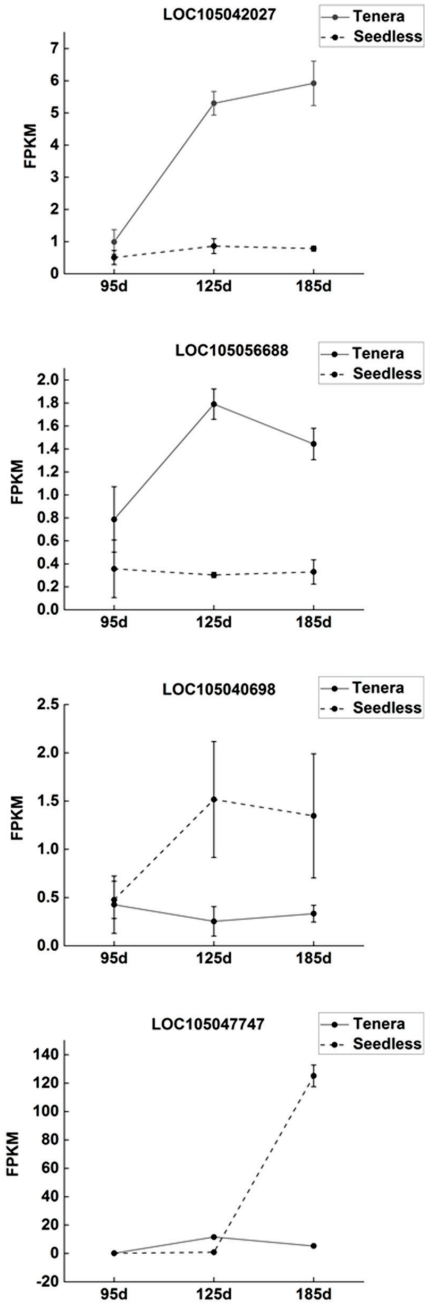


Figure 7. Dynamic changes in expression levels of key enzyme genes in the mesocarp of Seedless and Tenera oil palm during different developmental periods.

The study found that the contents of four major free fatty acids in oil palm pulp were significantly correlated with the expression levels of four key enzyme genes. The contents of FabB were found to have a positive correlation with the three main free fatty acids: stearic acid, myristate acid, and palmitic acid. FabF, an enzyme involved in fatty acid biosynthesis, has been found to have a positive correlation with the levels of four key free fatty acids: stearic acid, myristate acid, palmitoleic acid, and palmitic acid. Oil palm pulp contains four main free fatty acids, and correlation analysis showed that the content of these free fatty acids was significantly related to the expression level of four major enzyme genes (Figure 7; Table 3). The study found that ACSL and FATB showed a positive correlation with the levels of three primary free fatty acids: stearic acid, palmitoleic acid, and palmitic acid. On the other hand, there was a negative correlation between ACSL and FATB and the levels of free myristate acid. These results indicate that the expression of FabB and FabF enzyme genes may enhance the synthesis of free myristate in oil palm flesh, while FabF, ACSL, and FATB enzyme genes may promote the synthesis of free palmitoleic acid. However, the expression of the FabB gene may inhibit the synthesis of stearic acid, and the expression of the ACSL and FATB gene may inhibit the synthesis of myristate acid in oil palm flesh.

Table 3. Correlation analysis between the expression levels of key enzyme genes regulating free fatty acid synthesis and the contents of main free fatty acids in the mesocarp of Seedless and Tenera oil palm.

Gene ID	Enzyme Name	Stearic Acid	Myristic Acid	Palmitoleic ACID	Palmitic Acid
LOC105042027	FabB	0.75 **	0.917 **	−0.059	0.717 **
LOC105056688	FabF	0.517 *	0.722 **	0.252	0.490 *
LOC105040698	ACSL	0.060	−0.256	0.697 **	0.123
LOC105047747	FATB	0.167	−0.106	0.723 **	0.134

Note: "*" indicates significant correlation ($p < 0.05$), and "**"indicates extremely significant correlation ($p < 0.01$).

2.6. QRT-PCR of the Transcriptomic Data

To verify the accuracy of RNA-Seq expression profile sequencing, 10 differentially expressed genes were randomly selected from related metabolic pathways and analyzed by qRT-PCR (Figure 8). The results showed that the trend of qRT-PCR was consistent with that of RNA-Seq, and the coefficient of determination (R2) was greater than 0.8.

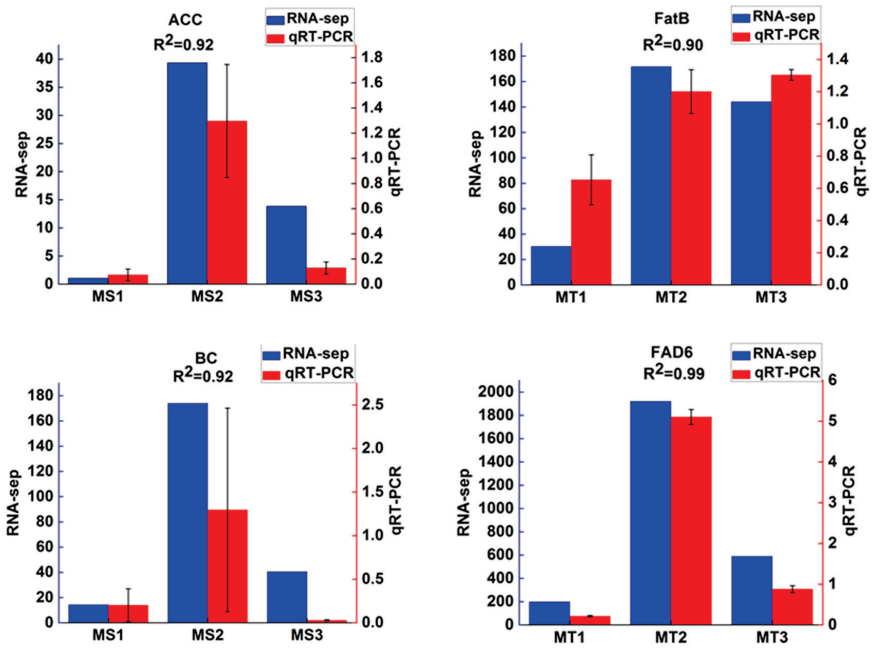


Figure 8. *Cont.*

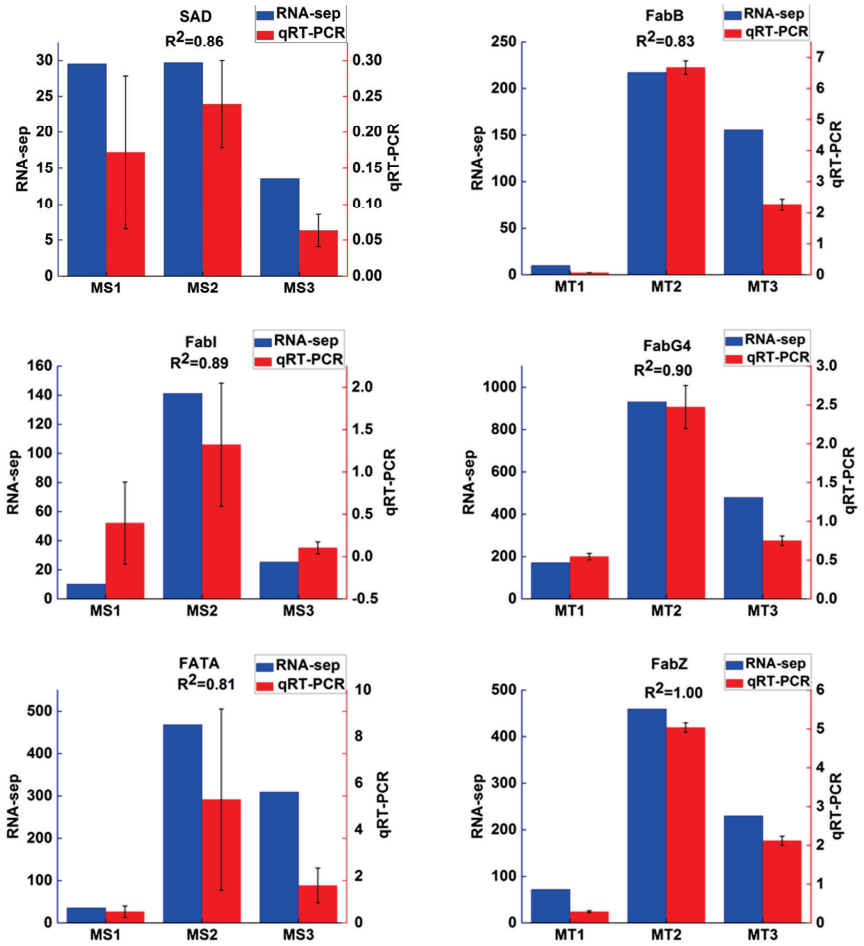


Figure 8. Relative expression levels of 10 selected genes during different developmental periods (MS1, MS2, MS3, MT1, MT2, and MT3). The $2^{-\Delta\Delta Ct}$ method was used to determine the relative expression levels of genes. The statistical differences were analyzed by ANOVA based on Duncan's multiple test ($p < 0.05$).

2.7. Analysis of Fatty Acid Synthesis Pathways in Oil Palm Flesh at Different Development Stages

The biosynthesis of fatty acids in oil palm fruit pulp is mainly enriched in the fatty acid biological pathway (ko00061) (Figure 9; Table 2), which is composed of Acetyl CoA as raw material and ultimately forms long-chain fatty acids through the action of some enzyme genes. After the combined analysis of fatty acid transcriptome and metabolome, we were able to identify the differential metabolites and genes related to the KEGG metabolic pathway in the fatty acid biological pathway. Based on the comprehensive analysis of gene FPKM values and correlation coefficients with stearic acid, myristic acid, palmitic acid, and palmitoleic acid (Table 3), it was finally found that FATB and ACSL are significantly positively correlated with palmitoleic acid; while FabB is significantly positively correlated with stearic acid, myristic acid, and palmitic acid; FabF is significantly positively correlated with myristic acid. By combining the specific positions of these enzyme genes on the pathway, it can be concluded that FATB may promote the synthesis of palmitic acid during the MS3 phase, ACSL may promote the synthesis of palmitic acid during the MS2 and MS3 phases, and FabB and FabF may promote the formation of stearic acid, myristic acid, and palmitic acid upstream of the fatty acid synthesis pathway during the MT2 and MT3 phases. In summary, the variation in gene expression patterns between the two varieties can explain the disparities in total fatty acid content and composition. During the development process of oil palm fruit, through the combined action of the aforementioned enzymes FATB, ACSL, FabB, FabF, etc., long-chain fatty acids rich in oil palm such as stearic acid, myristic acid,

palmitic acid, and palmitoleic acid are formed, thereby promoting the biosynthesis of oil palm fatty acids.

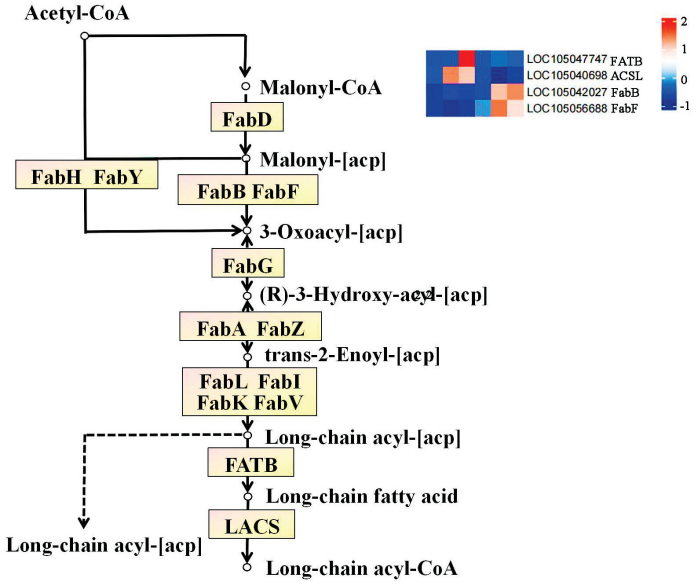


Figure 9. The main pathways and key enzymes of fatty acid biosynthesis. The yellow rectangle represents the enzyme. The rounded rectangle represents the connected channels. The white circle represents the compound.

3. Discussions

In this study, the lipid metabolites of Seedless and Tenera oil palm pulp at three developmental stages were analyzed by liquid chromatography-tandem mass spectrometry. The results of the study reveal a significant change in the total free fatty acid content of oil palm before and after the development of the flesh, and the total free fatty acid content in the flesh gradually increases as the growth and development of the flesh progresses. In particular, there is a sharp increase in the total free fatty acid content during the early to middle stages of development. This is in line with the evolution of fatty acids in camellia fruit as the oil, oleic, linoleic, and linolenic content increase as the fruit develops [24]. The total free fatty acid content of seedless oil palm remains stable throughout its mid to late development stage. This stability is crucial as it ensures that the oil extracted from the palm retains its quality and does not degrade over time. However, the most significant changes were observed in the contents of linolenic acid, linoleic acid, palmitic acid, and eicosenoic acid in different development stages. The contents of these four free fatty acids increased by 186.95, 13,386.90, 243.92, and 395.73 nmol/g. Significant changes were observed in the flesh of Tenera seed oil palm, with an increase in the content of myristic acid, stearic acid, palmitoleic acid, and linoleic acid by 327.47, 2537.50, 92.80, and 8601.56 nmol/g, respectively. The total content of free fatty acids increased from 1733.11 nmol/g to 58,231.41 nmol/g and 9313.33 nmol/g to 60,174.75 nmol/g. However, the increase in the total content of free fatty acids in Seedless oil palm was even more substantial, with a remarkable elevation of 56,498.30 nmol/g.

FabB is a β-oxyacyl-ACP synthase that uses acyl-Coenzyme A as an intermediate to catalyze the Claisen condensation reaction between acyl-ACP and malonyl-ACP [25,26]. It is found in bacteria as part of the fatty acid synthesis II pathway and is a potential target for antibiotics [27]. Enzymes encoded by the FabA and FabB genes have been found to catalyze the introduction of double bonds in decane precursors that are elongated into the 16:1Δ9 and 18:1Δ11 unsaturated fatty acyl chains required for functional membrane phospholipids [28]. In this study the FabB enzyme gene (ID:LOC105050891) was expressed in the flesh of Seedless and Tenera oil palm species, and positively correlated with the contents of free stearic acid, myristate acid, and palmitic acid, while negatively correlated

with the contents of free palmitic acid. The results indicated that the expression of the FabB gene may promote the biosynthesis of free stearic acid, myristate acid, and palmitic acid in oil palm flesh and inhibit the accumulation of free palmitic acid in oil palm flesh.

FabF is an essential enzyme for fatty acid synthesis and elongation. It plays a crucial role in the production of phospholipid membranes, lipoproteins, and lipopolysaccharides [29]. It plays a role in the FASII extension stage, where it converts malonyl ACP with growing ACP to form carbon dioxide and beta-ketoacP. Labeling it with [13C6]-glucose showed a significant decrease in the 13C flow to myristate and palmitic acid in the ΔFabB/F strain [30]. The study found that the expression of the FabF gene positively correlated with the content of free stearic acid, myristate acid, palmitic acid, and palmitic acid, suggesting that their synthesis and accumulation were promoted.

ACS can be divided into very long-chain acyl-CoA synthases (ACSVL) and long-chain acyl-Co synthases (ACSL), medium-chain acyl-Co synthases (AC-SM), and short-chain acyl-Co synthases (ACSS). ACSLs are a family of enzymes that can convert saturated and unsaturated FAs with a chain length of 8~22 into fatty acyl-CoA esters [31]. In this study, the ACSL enzyme gene showed a positive correlation with the contents of free stearic acid and palmitic acid, and a negative correlation with the contents of free myristate acid. The expression of the ACSL enzyme gene was found to have a significant impact on the fatty acid composition in the flesh of Seedless and Tenera varieties. It was speculated that the expression of this gene could promote the accumulation of free stearic acid, palmitic acid, and palmitic acid in the flesh of Seedless and Tenera varieties, and could inhibit the synthesis of free myristate acid. These findings highlight the role of the ACSL enzyme gene in influencing the fatty acid profile of these varieties and suggest its potential as a target for manipulating the fatty acid composition in agricultural crops.

Fatty acyl-ACP thioesterases in higher plants can be categorized into FATA and FATB based on amino acid sequence comparison and substrate specificity [32]. The C16:0-ACP thioesterase is a common FATB. The highest thioesterase activity was found in crude extracts of oil palm pulp [33]. Wang [34] cloned FATB from the seed kernel of LpFATB by the homologous cloning method, and found that the contents of eicosenoic acid (20:1) and oleic acid (18:1) in the seeds of LPFATB transformed strains with Col-0 as the backgrounds were decreased to a certain extent, while the contents of palmitic acid (16:0) were significantly increased. In this study, the FATB enzyme gene was highly expressed in the flesh of Seedless and Tenera oil palm species and was positively correlated with the contents of free stearic acid and palmitic acid, while negatively correlated with the contents of free myristate acid. Dussert et al. [35] found that FATB in oil palm pulp can promote the synthesis of fatty acids, and EgWRI1-1 and EgWRI1-2 can combine FATB1, FATB1, and FATB3 to further promote the increase in fatty acid content in the TAG pathway. This result once again confirmed that the FATB gene is the key gene for fatty acid synthesis in oil palm pulp.

4. Materials and Methods

4.1. Plant Materials

In this experiment, the fruits of Seedless and Tenera at different developmental stages were selected from the experimental base of oil palm (19°33′ N, 110°47′ E) from the Institute of Coconut Research, Chinese Academy of Tropical Agricultural Sciences, Wenchang City, China. The pulp of each group at 95 days (early development stage: MS1 and MT1), 125 days (middle development stage: MS2 and MT2), and 185 days (late development stage: MS3 and MT3) after pollination were selected and frozen with liquid nitrogen and stored at −80 °C for the detection of metabolome and transcriptome. For each biological replicate, three independent technical replicates were analyzed for each different developmental stages.

4.2. Methods

4.2.1. Metabolite Extraction

To determine the metabolite, 20 mg of sample was thawed into a centrifuge tube, containing 1 mL of lipid extract (methyl tert-butyl ether: methanol = 3:1, v/v, including internal standard mixture), and shaken for 30 min. Then, 300 µL of ultrapure water was added, shaken for 1 min, and placed at 4 °C for 10 min. Then, the tube was centrifuged for 3 min (12,000 rpm, 4 °C) and 400 µL of supernatant was transferred into a new centrifuge tube and concentrated at 20° C for 2 h, or until completely dry. Then, 200 µL of lipid complex solution (acetonitrile: isopropanol = 1:1, v/v) was added to the solution, and the solution was vortexed for 3 min and centrifuged for 10 min (12,000 rpm, 4 °C). Subsequently, 120 µl of the supernatant was transferred into the glass-lined tube until used.

4.2.2. Metabolomics Analysis and Data Processing

The instrument platform for LC-MS/MS analysis was composed of ExionLC ultra-performance liquid chromatography in tandem with SCIEX QTRAP 6500+ mass spectrometer, and the chromatographic column was Thermo AccucoreC30 column (2.6 µm, 2.1 mm × 100 mm i.d.). The software Analyst 1.6.3 was used to process the mass spectral data and combined with the information of the local lipid database, mass spectral qualitative analysis was performed on the metabolites of the samples. The mass spectral peaks detected for the same metabolite in different samples were corrected, and the integrated peak areas of all samples detected were calculated to obtain the absolute content of metabolites. The MRM model results are shown in Supplementary Figure S2. Differential metabolites between groups were screened using a combination of orthogonal partial least square discriminant analysis (OPLS-DA) and univariate analysis of p-value or Fold Change. The screening criteria were variable importance in projection (VIP) ≥ 1, Fold Change ≥ 2, or Fold Change ≤ 0.5 to obtain differential metabolites. The variation trend of different differential metabolite content can be obtained by analyzing the clustering heat map of differential metabolite content under Z-Score normalization treatment.

4.2.3. Total RNA Extraction and High-Throughput Sequencing

RNA from oil palm pulp was extracted using the Plant Total RNA Extraction Kit for Tiangen Biota, and sequencing was performed using the Illumina HiSeq platform. Using HISAT2 [36], Clean Reads were sequentially compared with the reference genome (Genome assembly EG5 http://www.ncbi.nlm.nih.gov/assembly/GCA-000442705, (accessed on 27 January 2024)) to obtain location information on the reference genome or gene, as well as specific sequence characteristics of the sequenced samples. Gene expression was calculated using the FPKM (fragments per kilobase of transcript per million fragments mapped) method [37]. The quality of RNA-seq is shown in Supplementary Table S1.

4.2.4. Transcriptome Data Processing and Differentially Expressed Genes Analysis

DESeq2/edgeR [38] was used to analyze the differentially expressed genes between the groups at different developmental stages, including MS1 vs. MS2, MS1 vs. MS3, MS2 vs. MS3, MT1 vs. MT2, MT1 vs. MT3, and MT2 vs. MT3; the screening threshold of the FDR (False Discovery Rate) is less than 0.05 with log2-Fold Change$|\geq$1. After that, the selected differentially expressed genes were analyzed by Nr functional annotation, enrichment analysis, and other bioinformatics.

4.2.5. Combined Metabolome and Transcriptome Analysis

Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis: This experiment compares metabolite analysis results and genetic analysis results for the same gene, comparing the differences between the groups of metabolites and mapped to the KEGG pathway chart (https://www.genome.jp/kegg), can be found in the two groups to learn common enrichment pathway. Correlation analysis: The Pearson correlation coefficient was used to calculate the quantitative correlation between the absolute content of free fatty

acid metabolites and the expression of related genes. Correlation results with a correlation coefficient > 0.80 and p-value < 0.05 were selected.

4.2.6. Real-Time Fluorescence Quantitative PCR Verification

Ten genes with potential roles during anthocyanoside synthesis in oil palm pericarp were selected for real-time fluorescence quantitative PCR reactions. qRT-PCR primers were designed using NCBI (Supplementary Table S2) and Actin was used as an internal reference gene [39], and the QuantStudioTM7 Real-Time PCR Instrument was used along with the Real-Time Fluorescence Quantitative PCR Instrument to determine the relative expression of selected genes in 96-well microtiter plates. The relative expression of the selected genes was calculated using the $2^{-\Delta\Delta Ct}$ method. The qRT-PCR reaction system used in this study was 20 µL, consisting of SYBR® Select Master Mix (2X): 10 µL, cDNA template: 1 uL, positive primer: 1 µL, water (without RNAase): 7 µL, and negative primer: 1uL. The reaction conditions were as follows: phase I UDG activation at 50 °C for 2 min; phase II UP activation at 95 °C for 2 min; phase III denaturation at 95 °C for 15 s for a total of 40 cycles; and phase IV annealing/extension at 60 °C for 1 min.

5. Conclusions

Oil palm fruit has the highest oil content at 185 days after pollination, which is the optimum harvesting time. Total free fatty acid levels in oil palm fruit are divided into three distinct phases: slow accumulation, rapid accumulation, and stable accumulation. These phases indicate different stages of development. During the slow accumulation phase, the levels of free fatty acids in the fruit gradually increase at a relatively slow pace. This is followed by the rapid accumulation phase, where the levels of free fatty acids increase rapidly, indicating a critical stage of development. Finally, the fruit enters the stable accumulation phase, where the levels of free fatty acids stabilize, indicating the fruit's maturity. The expression patterns of key enzyme genes have a significant correlation with the synthesis and accumulation of free fatty acids. FabB was positively correlated with the contents of stearic acid, myristic acid, and palmitic acid and negatively correlated with the contents of free palmitic acid. FabF, an enzyme involved in fatty acid biosynthesis, has been found to exhibit a positive correlation with the levels of four main free fatty acids: stearic acid, myristate acid, palmitoleic acid, and palmitic acid. ACSL and FATB were positively correlated with the contents of three main free fatty acids (stearic acid, palmitoleic acid, and palmitic acid), and negatively correlated with the contents of free myristate acid. These results indicate that the expression of FabB and FabF enzyme genes may promote the synthesis of free myristate in oil palm flesh. The expression of FabF, ACSL, and FATB enzyme genes may promote the synthesis of free palmitoleic acid in oil palm flesh, and the expression of these four key enzyme genes may promote the synthesis of free stearic acid and palmitoleic acid in oil palm flesh. The expression of the FabB gene may inhibit the synthesis of stearic acid in oil palm flesh and the expression of the ACSL and FATB gene may inhibit the synthesis of myristate acid in oil palm flesh. The effects of these genes on the synthesis pathway of free fatty acids in oil palm were thoroughly examined. This research shed light on how these genes contribute to the production of free fatty acids, which are crucial components of oil palm fruit. Additionally, the study also investigated the dynamic evolution of free fatty acid contents in different stages of oil palm fruit development. By analyzing the changes in free fatty acid levels, researchers were able to gain a deeper understanding of the metabolic processes occurring during different developmental stages of oil palm fruit. This knowledge can be valuable for optimizing oil palm cultivation and improving the quality of oil palm products.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25031686/s1.

Author Contributions: L.W., C.Y., J.J.J.M., R.L., L.Z. and S.C. conducted the experiments and performed the data analysis; H.C. organized and supervised the overall project; X.L. guided the analysis of the experimental data; L.W. and C.Y. wrote the manuscript; L.W. and J.J.J.M. edited the manuscript. All authors contributed to the article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was financially supported by the National Key R&D program of China (2023YFD2200700), Hainan Provincial Natural Sciences Foundation of China (323MS073), China Agriculture Research System (CARS-14-2-3), Central Public Interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (No. 1630152022001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

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. Perera, S.A.C.N. Oil Palm and Coconut. In *Alien Gene Transfer in Crop Plants*, 2nd ed.; Pratap, A., Kumar, J., Eds.; Springer: New York, NY, USA, 2014; Volume 2, pp. 231–252. [CrossRef]
2. Fitzherbert, E.B.; Struebig, M.J.; Morel, A.; Danielsen, F.; Brühl, C.A.; Donald, P.F.; Phalan, B. How will oil palm expansion affect biodiversity? *Trends Ecol. Evol.* **2008**, *23*, 538–545. [CrossRef] [PubMed]
3. Koh, L.P.; Wilcove, D.S. Is oil palm agriculture really destroying tropical biodiversity? *Conserv. Lett.* **2008**, *1*, 60–64. [CrossRef]
4. Wilcove, D.S.; Koh, L.P. Addressing the threats to biodiversity from oil-palm agriculture. *Biodivers. Conserv.* **2010**, *19*, 999–1007. [CrossRef]
5. Feng, M.L.; Zhang, H.Q.; Cao, H.X. Evaluation of agronomic traits of six imported oil palm varieties. *Chin. J. Trop. Crops* **2022**, *43*, 940–947. [CrossRef]
6. Mahlia, T.M.I.; Ismail, N.; Hossain, N.; Silitonga, A.S.; Shamsuddin, A.H. Palm oil and its wastes as bioenergy sources: A comprehensive review. *Environ. Sci. Pollut. Res. Int.* **2019**, *26*, 14849–14866. [CrossRef]
7. Saulia, L.; Subrata, I.D.M.; Yusuf, M.M. Postural analysis in design evaluation of oil palm loose fruits collector machine. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *1038*, 012060. [CrossRef]
8. Paterson, R.R.M.; Sariah, M.; Lima, N. How will climate change affect oil palm fungal diseases ? *Crop Prot.* **2013**, *46*, 113–120. [CrossRef]
9. Paterson, R.R.M.; Lima, N. Climate change affecting oil palm agronomy, and oil palm cultivation increasing climate change, require amelioration. *Ecol. Evol.* **2018**, *8*, 452–461. [CrossRef]
10. Nainggolan, M.; Sinaga, A.G.S. Characteristics of fatty acid composition and minor constituents of red palm olein and palm kernel oil combination. *J. Adv. Pharm. Technol. Res.* **2021**, *12*, 22–26. [CrossRef]
11. Parveez, G.K.; Rasid, O.A.; Masani, M.Y.; Sambanthamurthi, R. Biotechnology of oil palm: Strategies towards manipulation of lipid content and composition. *Plant Cell Rep.* **2015**, *34*, 533–543. [CrossRef]
12. Ahmad, N.A.F.; Muthiyah, K.; Mahmud, M.S. Free fatty acid formation in oil palm fruits during storage. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *991*, 012009. [CrossRef]
13. Nurulain, S.; Aziz, N.A.; Najib, M.S.; Salim, M.R.; Manap, H.A. Review of free fatty acid determination methods for palm cooking oil. *J. Phys. Conf. Ser.* **2021**, *1921*, 012055. [CrossRef]
14. Li, S.Y.; Zhang, Q.; Jin, Y.H.; Zou, J.X.; Zheng, Y.S.; Li, D.D. A MADS-box gene, EgMADS21, negatively regulates EgDGAT2 expression and decreases polyunsaturated fatty acid accumulation in oil palm (*Elaeis guineensis* Jacq.). *Plant Cell Rep.* **2020**, *39*, 1505–1516. [CrossRef]
15. Ting, N.C.; Sherbina, K.; Khoo, J.S.; Kamaruddin, K.; Chan, P.L.; Chan, K.L.; Halim, M.A.A.; Sritharan, K.; Yaakub, Z.; Mayes, S.; et al. Expression of fatty acid and triacylglycerol synthesis genes in interspecific hybrids of oil palm. *Sci. Rep.* **2020**, *10*, 16296. [CrossRef] [PubMed]
16. Xu, X.; Li, M.; Zou, J.X.; Zheng, Y.S.; Li, D.D. EgMYB108 regulates very long-chain fatty acid (VLCFA) anabolism in the mesocarp of oil palm. *Plant Cell Rep.* **2022**, *41*, 1449–1460. [CrossRef] [PubMed]
17. Whitbred, J.M.; Schuler, M.A. Molecular characterization of CYP73A9 and CYP82A1 P450 genes involved in plant defense in pea. *Plant Physiol.* **2000**, *124*, 47–58. [CrossRef]
18. Pinot, F.; Beisson, F. Cytochrome P450 metabolizing fatty acids in plants: Characterization and physiological roles. *FEBS J.* **2010**, *278*, 195–205. [CrossRef] [PubMed]
19. Liang, Y.; Yuan, Y.; Bao, Y.; Li, D. Dynamic expression of one cytochrome P45O-like gene in mesocarp of oil palm nut. *J. Huazhong Agric. Univ.* **2012**, *31*, 419–422. [CrossRef]

20. Wang, M.; Li, R.; Cao, H.; Jin, L.; Li, X. Identification and expression analysis of Δ9 stearoyl-ACP desaturase gene in oil palm. *Chin. J. Trop. Crops* **2022**, *43*, 235–243. [CrossRef]
21. Luo, T. Analysis of Fatty Acid, Vitamin E Components Content and Related Genes Function in Oil Palm Mesocarp. Master's Thesis, Huazhong Agricultural University, Wuhan, China, 2017.
22. Zhang, S.; Zhang, W.; Martin, J.J.J.; Qadri, R.; Fu, X.; Feng, M.; Wei, L.; Zhang, A.; Yang, C.; Cao, H. Differential analysis of transcriptomic and metabolomic of free fatty acid rancidity process in oil palm (*Elaeis guineensis*) fruits of different husk types. *Front. Plant Sci.* **2023**, *14*, 1132024. [CrossRef]
23. 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. *Acids Res.* **2007**, *36*, 480–484. [CrossRef] [PubMed]
24. Peng, S.; Lu, J.; Zhang, Z.; Ma, L.; Liu, C.; Chen, Y. Global Transcriptome and Correlation Analysis Reveal Cultivar-Specific Molecular Signatures Associated with Fruit Development and Fatty Acid Determination in *Camellia oleifera* Abel. *Int. J. Genom.* **2020**, 6162802. [CrossRef] [PubMed]
25. Chen, X. Screening of In Vivo Induced Antigen of *Haemophilus parasuis* and Construction of Eukaryotic Expression Library. Master's Thesis, Huazhong Agricultural University, Wuhan, China, 2008.
26. Lee, S.; Lee, S.; Yoon, Y.J.; Lee, J. Enhancement of long-chain fatty acid production in *Escherichia coli* by coexpressing genes, including fabF, involved in the elongation cycle of fatty acid biosynthesis. *Appl. Biochem. Biotechnol.* **2013**, *169*, 462–476. [CrossRef] [PubMed]
27. Yadrykhins'ky, V.; Georgiou, C.; Brenk, R. Crystal structure of *Pseudomonas aeruginosa* FabB C161A, a template for structure-based design for new antibiotics. *F1000Research* **2021**, *10*, 1–22. [CrossRef]
28. Dong, H.; Wang, H.; Cronan, J.E. Divergent unsaturated fatty acid synthesis in two highly related model pseudomonads. *Mol. Microbiol.* **2023**, *119*, 252–261. [CrossRef] [PubMed]
29. Soares da Costa, T.P.; Nanson, J.D.; Forwood, J.K. Structural characterisation of the fatty acid biosynthesis enzyme FabF from the pathogen *Listeria monocytogenes*. *Sci. Rep.* **2017**, *7*, 39277. [CrossRef] [PubMed]
30. Xu, X.P.; Elsheikha, H.M.; Liu, W.G.; Zhang, Z.W.; Sun, L.X.; Liang, Q.L.; Zhu, X.Q. The Role of Type II Fatty Acid Synthesis Enzymes FabZ, ODSCI, and ODSCII in the Pathogenesis of *Toxoplasma gondii* Infection. *Front. Microbiol.* **2021**, *12*, 703059. [CrossRef]
31. Tang, Y.; Zhou, J.; Hooi, S.C.; Jiang, Y.M.; Lu, G.D. Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-Co A synthetases. *Oncol. Lett.* **2018**, *16*, 1390–1396. [CrossRef]
32. Wang, K. Cloning and Functional Analysis of FATB Gene of *Litsea pungens*. Master's Thesis, Northwest A&F University, Xianyang, China, 2019.
33. Salas, J.J.; Ohlrogge, J.B. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Arch. Biochem. Biophys.* **2002**, *403*, 25–34. [CrossRef]
34. Wang, Q.; Huang, W.D.; Jiang, Q.; Qian, W.; Wenda, H.; Qi, J.; Jinpan, L.; Jianli, S.; Hong, X.; Hui, X.Z.; et al. Lower levels of expression of FATA2 gene promote longer siliques with modified seed oil content in *arabidopsis thaliana*. *Plant Mol. Biol. Report.* **2013**, *31*, 1368–1375. [CrossRef]
35. Dussert, S.; Guerin, C.; Andersson, M.; Joët, T.; Tranbarger, T.J.; Pizot, M.; Sarah, G.; Omore, A.; Durand-Gasselin, T.; Morcillo, F. Comparative transcriptome analysis of three oil palm fruit and seed tissues that differ in oil content and fatty acid composition. *Plant Physiol.* **2013**, *162*, 1337–1358. [CrossRef] [PubMed]
36. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef]
37. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.-C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from rna-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]
38. 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]
39. Xia, W.; Mason, A.S.; Xiao, Y.; Liu, Z.; Yang, Y.D.; Lei, X.T.; Wu, X.M.; Ma, Z.L.; Peng, M. Analysis of multiple transcriptomes of the African oil palm (*Elaeis guineensis*) to identify reference genes for RT-qPCR. *J. Biotechnol.* **2012**, *184*, 63–73. [CrossRef] [PubMed]

International Journal of
Molecular Sciences

MDPI

Article

Morphological Analyses and QTL Mapping of Mottled Leaf in Zucchini (*Cucurbita pepo* L.)

Kexin Wang [1,2], Xinyu Wang [1,2], Lijing Zhang [1,2], Yichen Chi [1,2], Yusong Luo [1,2], Wenlong Xu [1,2], Yunli Wang [1,2,*] and Shuping Qu [1,2,*]

[1] Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (Northeast Region), Ministry of Agriculture and Rural Affairs, Northeast Agricultural University, Harbin 150030, China; wkx15104594791@126.com (K.W.); 15765788116@163.com (X.W.); 18804658660@163.com (L.Z.); 18845100319@163.com (Y.C.); yelingyuese@126.com (Y.L.); xwl@neau.edu.cn (W.X.)
[2] College of Horticulture and Landscape Architecture, Northeast Agricultural University, Harbin 150030, China
* Correspondence: wangyunli@neau.edu.cn (Y.W.); spqu@neau.edu.cn (S.Q.); Tel.: +86-451-55190204 (Y.W.); +86-451-55190204 (S.Q.)

Abstract: The mottled leaf is one of the agronomic traits of zucchini and can be applied as a marker trait in aggregation breeding. However, the genetic mechanism responsible for mottled leaf has yet to be elucidated. In the present study, we used two inbred lines (line '19': silver mottled leaf; line '113': normal leaf) as parents for the physiological and genetic analysis of mottled leaf. The synthesis and net photosynthetic rate of chlorophyll were not significantly affected in the mottled areas of leaves. However, we detected a large space between the palisade parenchyma in the leaf mottle area of line '19', which may have caused the mottled leaf phenotype. Light also plays an important role in the formation of mottled leaf, and receiving light during the early stages of leaf development is a necessary factor. Genetic analysis has previously demonstrated that mottled leaf is a quantitative trait that is controlled by multiple genes. Based on the strategy of quantitative trait locus sequencing (QTL-seq), two QTLs were identified on chromosomes 1 and 17, named *CpML1.1* and *CpML17.1*, respectively. Two major loci were identified using R/qtl software version 1.66 under greenhouse conditions in April 2019 (2019A) and April 2020 (2020A) and under open cultivation conditions in May 2020 (2020M). The major QTL, *CpML1.1*, was located in a 925.2-kb interval on chromosome 1 and explained 10.51%-24.15% of the phenotypic variation. The *CpML17.1* was located in a 719.7-kb interval on chromosome 17 and explained 16.25%-38.68% of the phenotypic variation. Based on gene annotation, gene sequence alignment, and qRT–PCR analysis, the *Cp4.1LG01g23790* at the *CpML1.1* locus encoding a protein of the TPX2 family (target protein of Xklp2) may be a candidate gene for mottled leaf in zucchini. Our findings may provide a theoretical basis for the formation of mottled leaf and provide a foundation for the fine mapping of genes associated with mottled leaf. Molecular markers closely linked to mottled leaf can be used in molecular-assisted selection for the zucchini mottled leaf breeding.

Keywords: *Cucurbita pepo* L.; QTL analysis; mottled leaf

1. Introduction

Leaf color arises from the dynamic balance of various photosynthetic pigments, including chlorophyll and carotenoids, of which chlorophyll accounts for the largest proportion. Therefore, the color of the leaves is mostly green. When the proportion of photosynthetic pigments in plants changes, healthy leaves will appear in different colors. According to the leaf color phenotype, the color of leaves can be divided into pale green, stripe, spotted leaf, yellow-green, white-green, and other types [1]. According to their anatomical structure, they can be divided into chlorophyll type, air space type, epidermis type, pigment type, and appendage type [2]. Overall, the diversity of leaf colors is beneficial for enhancing the

ornamental value of plants. Therefore, some plants intentionally cultivate different leaf color varieties, but usually, only a portion of seeds can transmit traits.

At present, many leaf color mutants have been reported in *Arabidopsis*. The *im* mutant exhibits white-green variation and possesses a small amount of carotenoids, abnormal plastids, and palisade parenchyma in the white area. The *IM* gene encodes a plastid terminal oxidase (PTOX), which participates in the synthesis of carotenoids. The formation of white areas may be due to insufficient carotenoid synthesis in *im* mutants; this can lead to chlorophyll photooxidation and influence the expression of chloroplast genes, ultimately resulting in the formation of leaf variation [3–5]. Similar to the *im* mutant, the *var3*, *msl-1*, and *msl-2* mutants also exhibit abnormal palisade tissue [6,7]. Both *var1* and *var2* exhibit a yellow-green-white mottled phenotype. *VAR1* and *VAR2* encode a similar chloroplast FtSH protease, which plays an important role in the degradation of photodamage subunits in photosystem II. *VAR1* and *VAR2* are known to exert synergistic effects. Furthermore, the absence of *VAR1* and *VAR2* may damage the photoprotection mechanism and the development of thylakoids, resulting in variegated leaves [8–11]. White-green leaf mutants have also been reported in rice. Both *zl16* and *zebra524* exhibited a white-green leaf phenotype, encoding β-hydroxyl ACP dehydratase (HAD) and lycopene β-cyclizing enzymes, respectively, which play an important role in the synthesis of photosynthetic pigments [12,13]. The plastid ribosomal proteins (PRPs) and phosphate ribosamine glycine ligase (PurD) are involved in regulating chloroplast development and chlorophyll metabolism during leaf development to regulate rice leaf color [14,15]. RLK proteins also participate in the whitening of rice leaves through other pathways [16]. Similar phenotypes have also been reported in maize [17].

Among Cucurbitaceae crops, there are the most reports on the leaf color of cucumbers and pumpkins. Yellow-green leaf mutants are commonly found in cucumbers. Both mutants *C777* and *Ygl1* exhibited a yellow-green leaf phenotype. Their candidate genes are involved in chloroplast development and chlorophyll metabolism pathways to regulate cucumber leaf color [18,19]. *Cucurbita pepo* L. is an important commercial crop with good nutritional and medicinal value. *C. pepo* is rich in resources; most germplasm resources possess green leaves, although some plants have yellowing, albinism, or mottled leaves. There are two types of silver mottle on leaves in zucchini; one is a hereditary form of leaf mottle, and the other is caused by whitefly. Hereditary leaf mottle is regulated by a dominant gene which shows an irregular shape and grows along the direction of the leaf vein [20–22]. The lower epidermis of the mottled leaves remains green. The expression of the leaf mottle is regulated by the *M* gene, the expression of which can be affected in many ways. Firstly, the *M* gene is regulated by the modified gene, which can affect the expression of *M* in terms of both time and degree. In addition, environmental factors and expression positions play an important role in the expression of the *M* gene. The time at which *M* is first expressed during plant development can also impact the expression levels of the *M* gene [23–25]. Furthermore, the *M* gene can cause air space between the palisade parenchyma and the upper epidermal cells, which causes changes in light refraction, making the leaves appear silver. Light reflection from the leaf mottle area has been shown to range from 400 to 700 nm, which is higher than that of the green area of leaves [26]. In addition to hereditary leaf mottle, leaf mottle can also be caused by whitefly, which can be affected by light, temperature, humidity, and insect density. Chloroplasts in palisade cells and the plasma membrane around some vascular cells exhibited slight ultrastructural damage, thereby affecting photosynthesis and the agronomic traits of plants [27,28]. The structures of hereditary leaf mottle and the leaf mottle caused by whitefly are similar, with spaces between the palisade cells and the upper epidermis. However, few studies have investigated the location of mottled leaf in *C. pepo*. A genetic linkage map of *C. pepo* was constructed using BC_1 and showed that the dominant gene *M* was located in the linkage group 6 [29]. Through a high-density genetic map of *C. pepo* using an RIL population, a major QTL related to leaf mottle was located in a 0.29 Mb interval on chromosome 17 that explained 23.3% of the phenotypic variation. Two minor QTLs were also found to be

located on chromosome 1 (with a 0.18 Mb interval) and chromosome 13 (with a 1.67 Mb interval); these explained 3.83% and 8.01% of the phenotypic variation, respectively [30].

In this study, we used lines '19' and '113' as materials to observe mottled leaf phenotype, constructed F_2 populations for QTL analysis and stability analysis, and analyzed the interaction of mottled leaf QTLs. Our findings indicate that molecular markers closely linked to mottled leaf can be used for molecular marker-assisted selection breeding in zucchini and provide a foundation for the fine mapping of genes and the molecular mechanistic research of mottled leaf.

2. Results

2.1. Observation of the Phenotypes Associated with the Mottled Leaf Trait

We observed the mottled leaf traits of lines '19' and '113' during growth and observed that leaf mottle appeared on the leaves of line '19' during the 8-leaf stage. There was no significant leaf mottle at this stage, although chlorosis began on the first and second leaves under the growing tip. Subsequently, the area of chlorosis gradually turned silver, and leaf mottle was formed until the fifth leaf stage under the growing tip. As the plant grew, the leaf mottle became more prominent. Subsequently, all leaves were covered with leaf mottle. No chlorosis and leaf mottle appeared in line '113' during the growth period. The appearance period of leaf mottle in the F_1 generation was similar to that of line '19', although the grade of leaf mottle was lower compared to line '19' (Figure 1). The mottled leaf phenotype was most significant at the 20-leaf stage. Lines '19' and '113', and the F_1 generation, were classified as grade 3, grade 0, and grade 2, respectively.

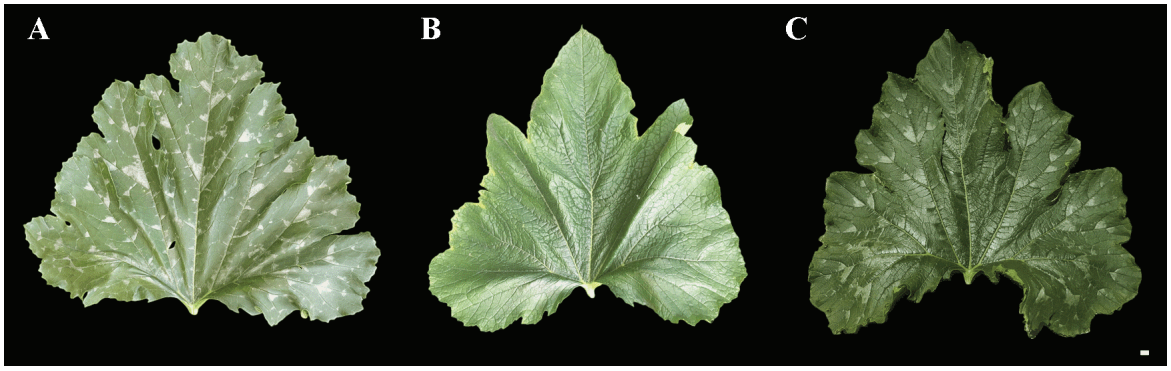


Figure 1. The phenotypes of '19', '113' and F_1. (**A**) 19. (**B**) 113. (**C**) F_1. Scale = 1 cm.

2.2. Determination of the Content of Photosynthetic Pigment in Leaves

In order to investigate the changes in pigment content, we selected the mottled area and green area of parent '19' and the corresponding position of '113' at the 20-leaf-stage to determine pigment content.

The first, second, third, and fifth leaves under the growing tip at the 20-leaf-stage were selected for the determination of pigment content. The color of the first and second leaves of parent '19' was an inconspicuous chlorotic mottle. The pigment content of all the leaves was measured collectively. Leaf mottle was evident on the third and fifth leaves, and the pigment content was measured separately for the green and mottle areas. The content of chlorophyll a in the two parents was always higher than that of chlorophyll b, with a ratio between 4.1 and 5.7. The ratio of carotenoid to total chlorophyll was stable between 0.22 and 0.27. Due to the higher chlorophyll content of the two parents, the leaves were predominantly green. Except for the significantly higher carotenoid content in the green area of line '19' of the third and fifth leaves when compared to the corresponding area of line '113', there was no significant difference in chlorophyll content and carotenoid content between the two lines of plants (Figure 2). There was no significant difference in the levels of chlorophyll a, chlorophyll b, carotenoid, and total chlorophyll in the first, second, third,

and fifth leaves of the mottled area in line '19' and the green area in line '19'. Moreover, there was no significant difference in the chlorophyll content when comparing the mottle area and the green area in line '19', thereby indicating that the formation of leaf mottle did not affect the synthesis of carotenoid and chlorophyll.

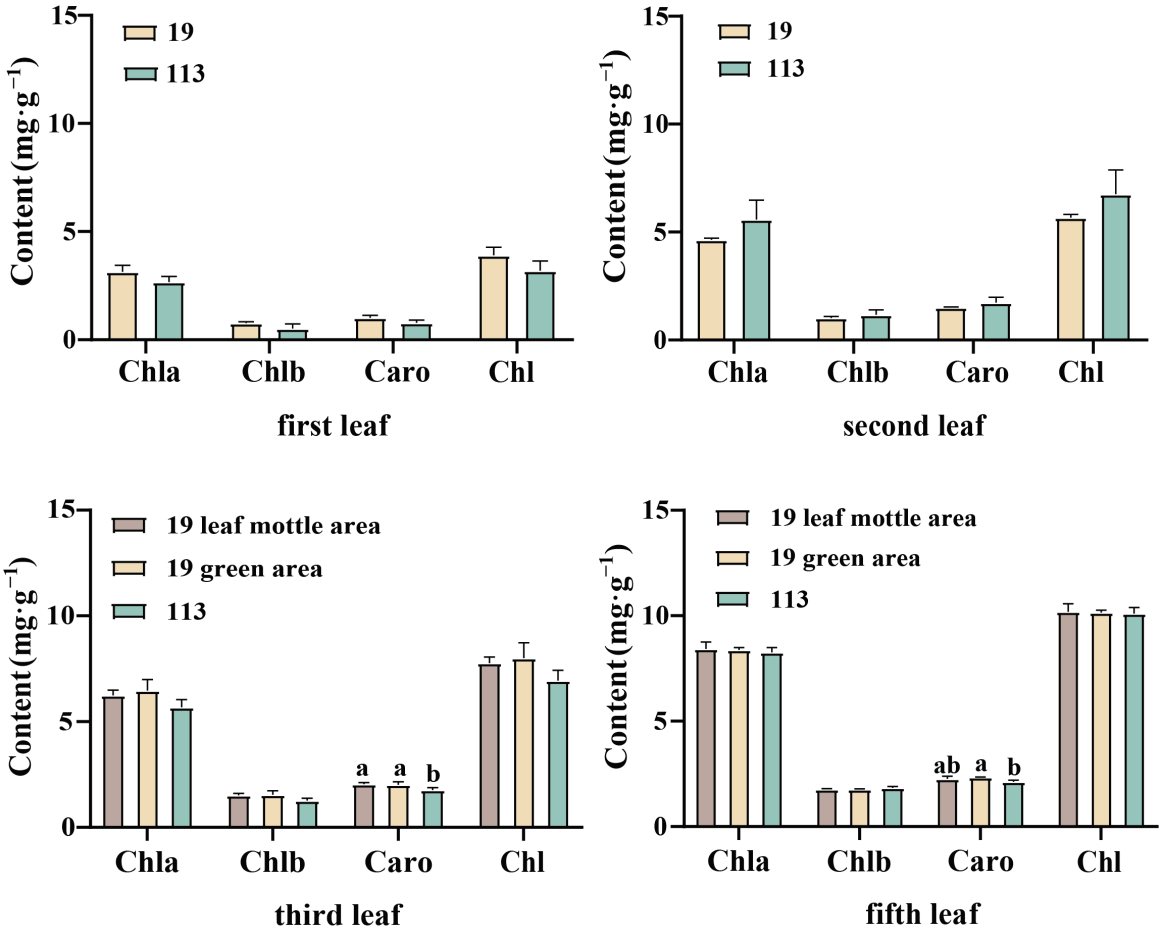


Figure 2. Photosynthetic pigment content of parent in different stages. Chl a: chlorophyll a, Chl b: chlorophyll b, Caro: carotenoid, Chl: chlorophyll. Different letters indicate significant differences at 0.05 level. Columns are the mean values ± SD. Each column has been tested in three replicates.

2.3. Analysis of Photosynthetic Parameters in Leaves

Next, we determined the net photosynthetic rate (Pn), transpiration rate (Ti), stomatal conductance (Cs), and intercellular CO_2 concentration (Ci) of lines '19' and '113' at the 20-leaf-stage. Analysis showed that there was no significant difference in net photosynthetic rate between the two parents (Figure 3A). Compared with the green area of lines '19' and '113', the Ti, Cs, and Ci of the mottle area in line '19' were all significantly reduced (Figure 3B–D). The formation of leaf mottle, therefore, affects Ti, Cs, and Ci but has no significant impact on Pn.

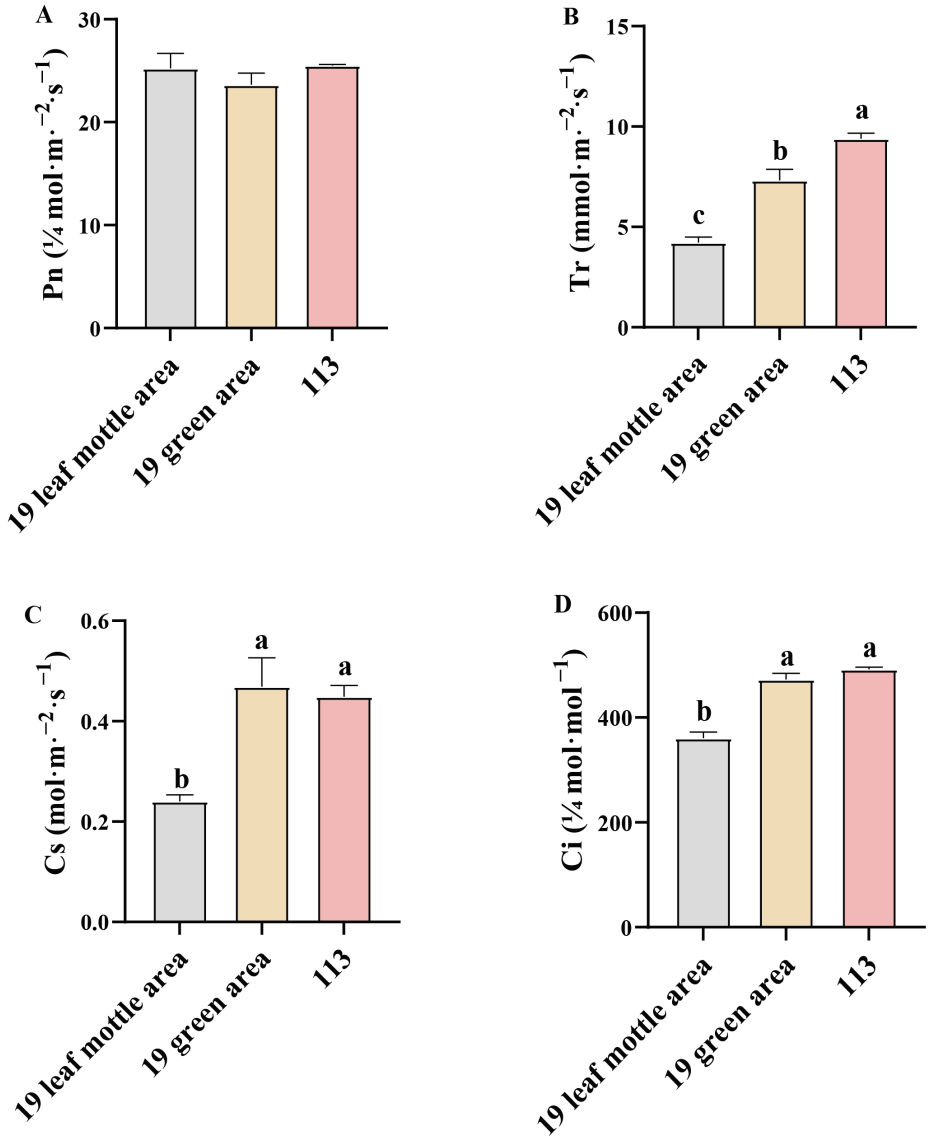


Figure 3. Photosynthetic parameters in different leaf areas. (**A**) net photosynthetic rate of parents. (**B**) Transpiration rate of parents. (**C**) Stomatal conductance rate of parents. (**D**) Intercellular CO_2 concentration rate of parents. Different letters indicate significant differences at 0.05 level. Columns are the mean values ± SD. Each column has been tested in three replicates.

2.4. Shading Analysis of Leaf

In order to investigate whether light exerts an impact on the formation of leaf mottle, we applied shading treatment to the leaves of line '19'. We selected plants with the same extent of growth, covered the first true leaf under the growing tip, and plants with all leaves exposed to natural light acted as controls. Analysis showed that no leaf mottle appeared on the leaves under the 5-days shading treatment. Subsequently, the cover was removed, and all leaves were exposed to natural light. After 10-days, no mottle had appeared on the leaves (Figure 4A–C). The control group maintains a mottled leaf phenotype during these stages (Figure 4D–F). Our findings indicate that light is involved in the formation of leaf mottle, and that receiving light during the early stages of leaf development is a necessary factor for the formation of leaf mottle.

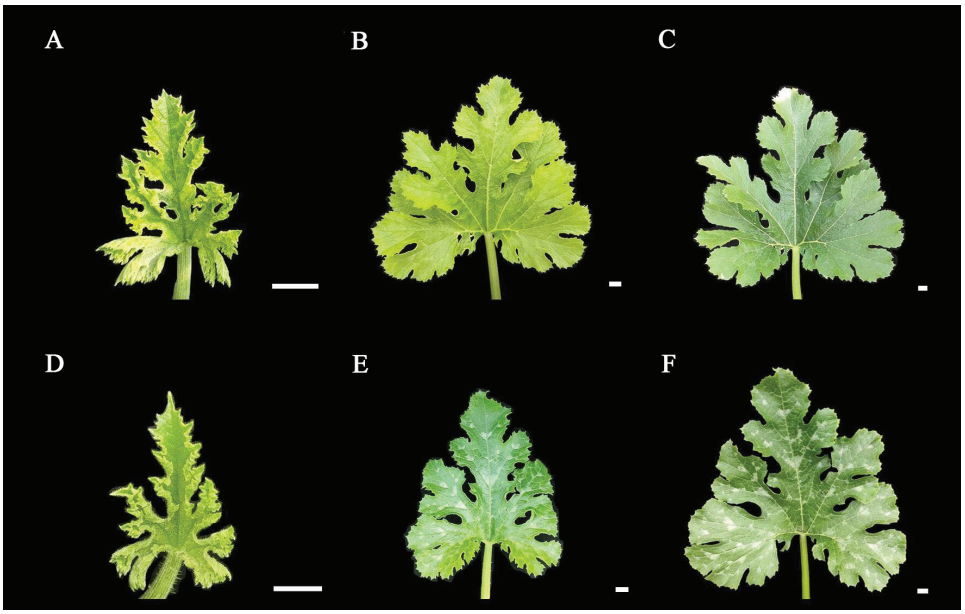


Figure 4. Results of leaves under shading treatments. (**A**) The leaf without shading treatment. (**B**) The leaf under 5-days of shading. (**C**) The leaf under 5-days of shading, then 10-days of light restoration. (**D**–**F**) Leaves growing under light in the same periods as (**A**–**C**). Scale = 1 cm.

2.5. Analysis of Differences in the Microstructure of Leaves

In order to determine whether there was a structural difference between the mottled area and the green area of leaves, we selected the fifth leaf (a mature and functional leaf) under the growing tip at the 20-leaf-stage for analysis. At this time, leaf mottle had developed obviously on the leaves. The green area and mottle area of line '19' and the corresponding area of line '113' were selected to prepare paraffin-embedded sections. The zucchini leaves were composed of an upper epidermal layer, a palisade parenchyma, a sponge parenchyma, and a lower epidermal layer (Figure 5). In the green area of line '19' and line '113', the palisade parenchyma was closely arranged (Figure 5B,C), while the palisade parenchyma in the mottled area of leaves from line '19' was loosely arranged and contained large air space between cells. The sponge parenchyma in the green area of the leaves from line '19' and line '113' was tighter than in the mottled area on leaves from line '19' (Figure 5). Next, we measured the leaf thickness and the cell thickness of each layer of the two parents. Analysis showed that the thickness of cells in the upper epidermis cells and the sponge parenchyma in line '19' exhibited significant differences when compared between the mottle area and green area (Table 1).

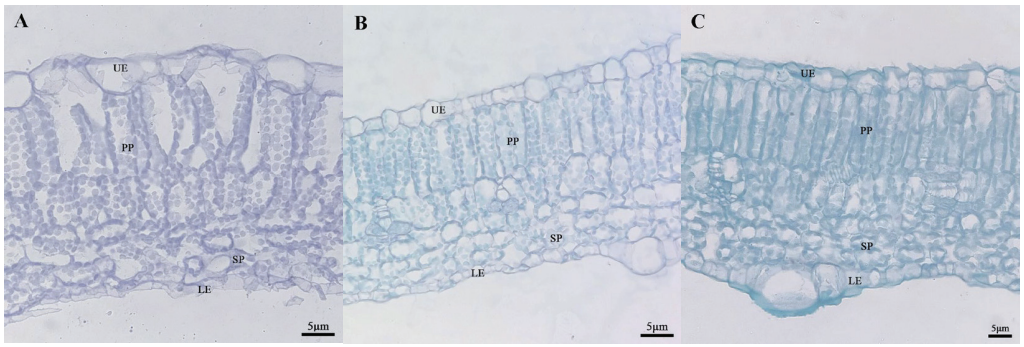


Figure 5. The cross-sectional structure of parental. (**A**) Silver area of '19' (40×). (**B**) Green area of '19' (40×). (**C**) '113' (40×). UE: upper epidermis; PP: palisade parenchyma; SP: sponge parenchyma; LE: Lower epidermis. Scale = 5 μm.

Table 1. Microstructure analysis between '19' and '113'.

Material	Leaf Thickness	Thickness of Upper Epidermis	Thickness of Palisade Parenchyma	Thickness of Sponge Parenchyma	Thickness of Lower Epidermis	CTR (%)	SR (%)
'19' mottle area	45.12 ± 0.09 a	4.64 ± 0.16 a	26.24 ± 0.10 a	8.69 ± 0.05 c	1.99 ± 0.10 a	58.15 ± 0.00 a	19.26 ± 0.00 c
'19' green area	45.09 ± 0.16 a	3.82 ± 0.08 b	26.22 ± 0.35 a	11.20 ± 0.06 a	2.32 ± 0.04 a	58.18 ± 0.01 a	24.83 ± 0.00 a
'113'	43.66 ± 0.14 b	3.80 ± 0.04 bB	22.26 ± 0.20 b	10.24 ± 0.24 b	2.32 ± 0.41 a	50.98 ± 0.01 b	23.45 ± 0.01 b

Note: Values are the average ± SD on three replicates. Different letters indicate significant differences at the 0.05 level.

2.6. Genetic Analysis of Mottled Leaf Trait

F_1 populations derived by crossing lines '19' and '113' showed significant leaf mottle on their leaves (grade 2) (Figure 1). An F_2 generation was constructed by F_1 self-crossing. The leaf mottle grades of the F_2 population planted in different environments were divided into three grades (grade 0, grade 2, and grade 3) for 2019A and four grades (grade 0, grade 1, grade 2, and grade 3) for 2020A and 2020M. It was found that the absolute values of skewness and kurtosis of mottled leaf traits in the F_2 population under different environments were <1 (Table S1), and the population phenotypic distribution conformed to a typical normal distribution (Figure 6). Therefore, the mottled leaf trait was a quantitative trait.

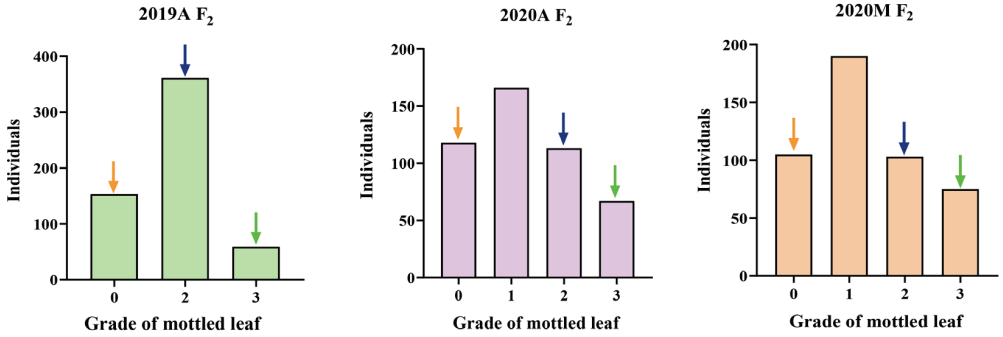


Figure 6. Frequency histogram of leaf mottle grades in different F_2 populations. The orange arrows indicate mottled leaf phenotype in '19', the green arrows indicate mottled leaf phenotype in '113', and the blue arrows indicate mottled leaf phenotype in F_1.

2.7. QTL Analysis of the Mottled Leaf Trait

In total, 49,104,069 clean reads and 49,745,164 clean reads were obtained from the M-pool (28× read depth with 97.21%) and the N-pool (28× read depth with 97.22%) using QTL-seq. Q30 values reached 88.18% and 87.99%, respectively. These results indicated that the sequencing results were reliable for gene mapping. A total of 979,650 SNPs and 324,752 InDels were identified on all 20 chromosomes. SNPs/InDels with multiple genotypes, a read depth <4 in both pools, or consistent genotypes in both pools were removed, thus leaving 617,314 high-quality SNPs, and then 191,073 high-quality InDels were obtained. Next, the ED algorithm was employed to identify significantly different InDels between the M-pool and N-pool based on sequencing data to predict candidate regions of the mottled leaf. A 1.26-Mb region and a 4.30-Mb region were identified on chromosomes 1 and 17, respectively (Figure 7A). The ED algorithm was also employed to identify significantly differential SNPs between the M-pool and N-pool based on sequencing data to predict candidate regions of the mottled leaf. A 1.35-Mb region and a 4.35-Mb region were identified on chromosomes 1 and 17, respectively (Figure 7B). The Δ(InDel-index) and Δ(SNP-index) were calculated and plotted by comparing the results of the InDel-index and SNP-index of the M-pool and N-pool in the genomic positions. The threshold of the Δ(InDel-index) value was 0.45; only one region was found on chromosome 17 (total length 2.10 Mb) (Figure 7C). The threshold of the Δ(SNP-index) value was 0.44, and two regions were found on chromosome 1 (total length 0.06 Mb) and chromosome 17 (total length:

2.03 Mb), respectively (Figure 7D). Combining the results of these analyses, the candidate regions associated with the mottled leaf trait were positioned within the 19.71–20.97 Mb intervals on chromosome 1 and between the 3.83–6.14 Mb intervals on chromosome 17. Therefore, we named these loci as *CpML1.1* and *CpML17.1*, respectively.

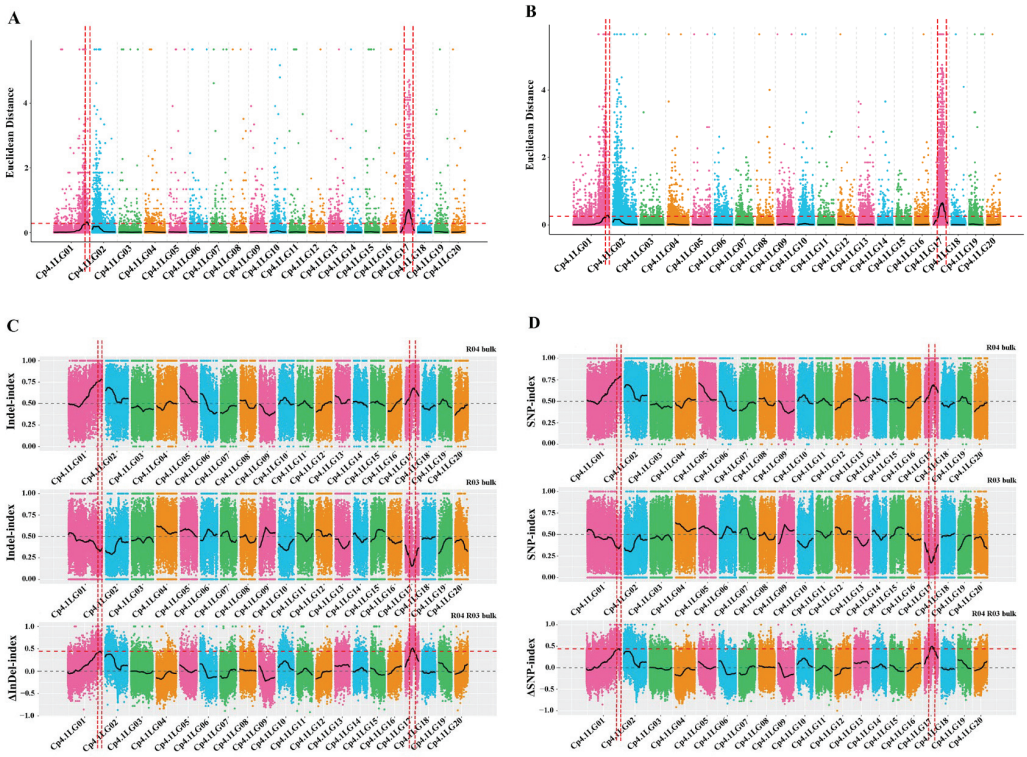


Figure 7. The results of QTL-seq. (**A**) Distribution of the ED-based linkage value with InDels on chromosomes. (**B**) Distribution of the ED-based linkage value with SNPs on chromosomes. (**C**) Distribution of InDel-index correlation values on chromosomes. (**D**) Distribution of SNP-index correlation values on chromosomes. In (**A**,**B**), the medians +3 SD of the fitted value of all sites were defined as the linkage threshold and shown by red lines. In (**C**,**D**), the top panel is a distribution map of the SNP-index (InDel-index) values of the M-pool, the middle panel is a distribution map of the SNP-index (InDel-index) values of the N-pool, and the bottom panel is a distribution map of the Δ(SNP-index)/Δ(InDel-index) values. The red line represents the confidence threshold line for the 99 percentile.

Based on the InDels in the candidate region between parental genomes, 25 polymorphic markers were developed on chromosomes 1 and 17 (Table S2). To test the stability of the QTLs *CpML1.1* and *CpML17.1*, we analyzed 580 F_2 plants in 2019A, 580 F_2 plants in 2020A, and 480 F_2 plants in 2020M. R/qtl analysis showed that a total of two QTLs associated with the mottled leaf trait were detected on chromosomes 1 and 17 (Figure 8). In 2019A, 2020A, and 2020M, *CpML1.1* was located between markers Chr01_18713051 and Chr01_19638223, Chr01_18590758 and Chr01_19419173, Chr01_18713051 and Chr01_19638223, respectively, with LOD scores of 43.79, 31.18, and 14.02. These scores explained 24.15%, 23.82%, and 10.51% of the phenotypic variation. With regards to the intersection of associated regions from QTL mapping, we used the CIM method in the three environments and QTL-seq analysis and determined that *CpML1.1* was located in a 925.2-kb region between markers Chr01_18713051 and Chr01_19638223. *CpML17.1* was located between markers Chr17_5225091 and Chr17_5944838, Chr17_5225091 and Chr17_5944838, Chr17_5225091 and Chr17_5944838, respectively, with LOD scores of 38.70, 20.65 and 52.99, thereby explaining 20.80%, 16.25%, and 38.68% of phenotypic variation. Combining the mapping

results of different populations, *CpML17.1* was located in the 719.7-kb region between markers Chr17_5225091 and Chr17_5944838 (Table 2).

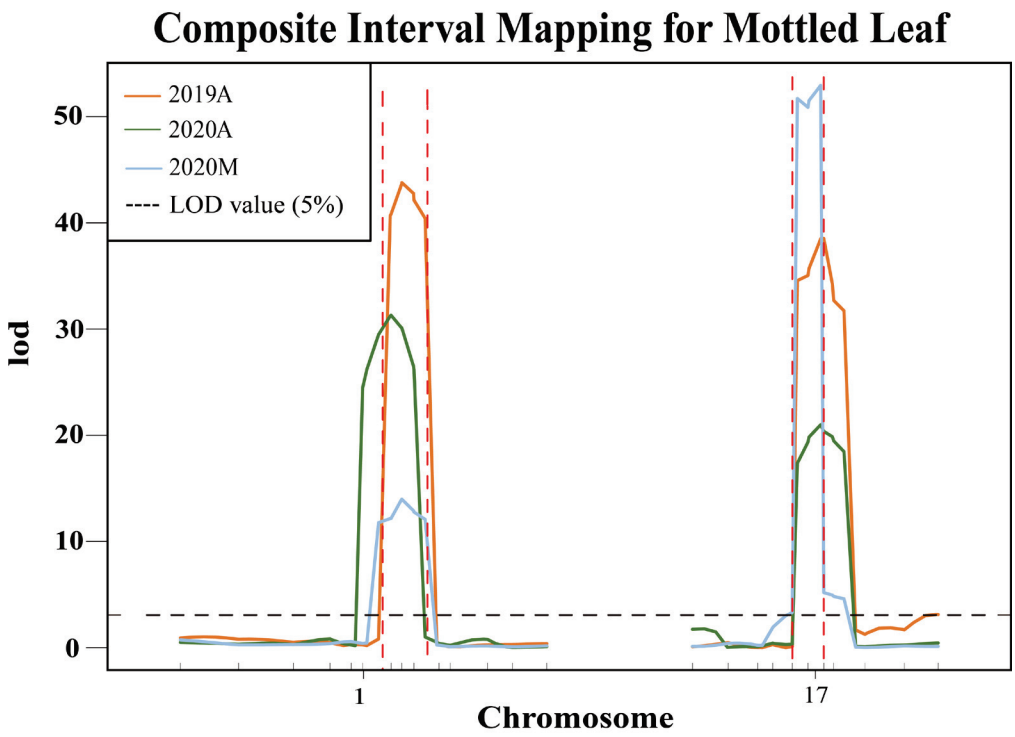


Figure 8. Verification of QTL stability. The area between the red dashed lines represents the QTL location.

Table 2. QTL mapping of the mottled leaf trait in the F_2 population.

QTL	Environment	Position Interval	Physical Distance (Kb)	LOD	Add	Dom	PVE (%)
CpML1.1	2019A	CHr01_18713051-CHr01_19638223	925.2	43.79	0.64	0.36	24.15
	2020A	CHr01_18590758-CHr01_19419173	828.4	31.18	0.56	0.04	23.82
	2020M	CHr01_18713051-CHr01_19638223	925.2	14.02	0.43	0.15	10.51
CpML17.1	2019A	CHr17_5225091-CHr17_5944838	719.7	38.70	0.63	0.03	20.80
	2020A	CHr17_5225091-CHr17_5944838	719.7	20.65	0.47	-0.01	16.25
	2020M	CHr17_5225091-CHr17_5944838	719.7	52.99	0.85	-0.19	38.68

2.8. QTL Interaction Analysis of Mottled Leaf Trait

Both *CpML1.1* and *CpML17.1* exhibited a positive additive effect on mottled leaf traits. In order to study the interaction effect of *CpML17.1* and *CpML1.1*, we analyzed the genotype and phenotype of the two loci (Figure 9). When both loci carry homozygous 'A' (from mottled leaf material from line '19'), the grade of mottled leaf was the highest, and when both loci carried the homozygous 'B' (from no-mottled-leaf material from line '113'), the grade of mottled leaf was the lowest. Under different environmental conditions, in *CpML1.1 BB* plants (the *CpML1.1* locus genotype was *BB*), the mottled leaf grade was *CpML1.1 AA* > *CpML1.1 AB* > *CpML1.1 BB*. In *CpML17.1 BB* plants (the *CpML17.1* locus genotype was *BB*), the mottled leaf grade was *CpML17.1 AA* > *CpML17.1 AB* > *CpML17.1 BB*.

These results showed that *CpML1.1* and *CpML17.1* exerted cumulative effects on mottled leaf traits.

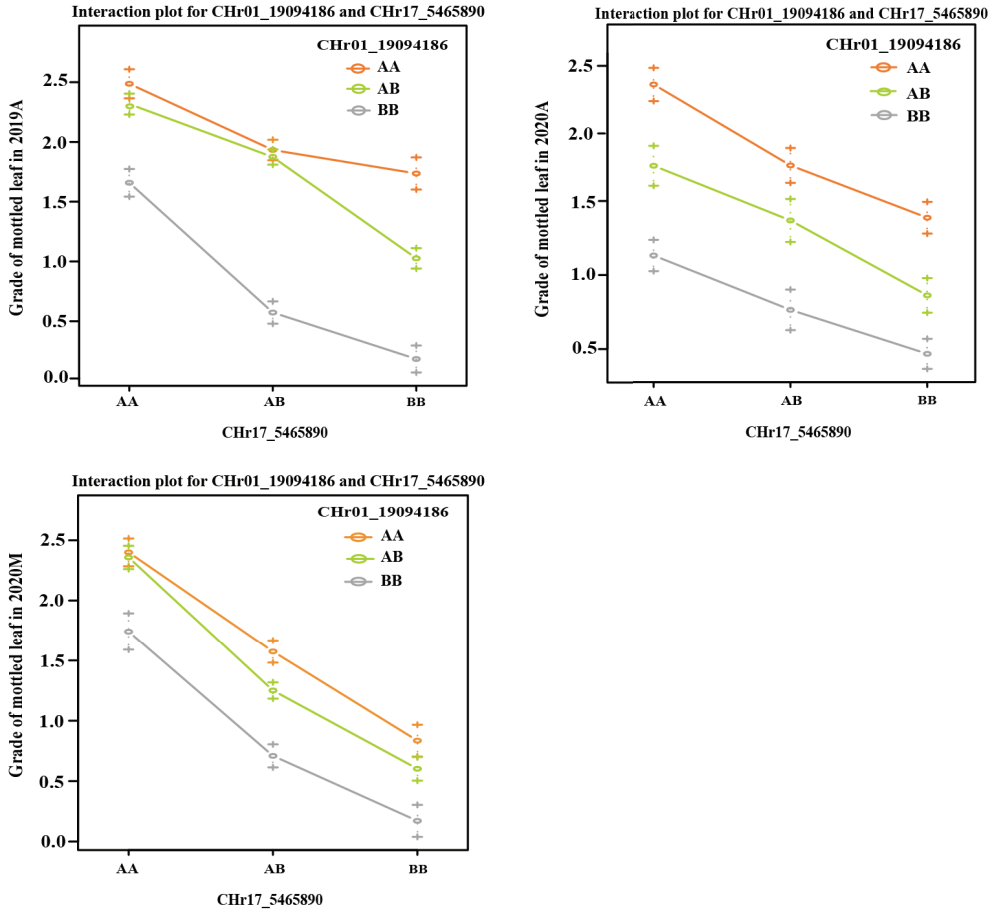


Figure 9. Interaction plots of the QTL pairs detected in the F_2 and RIL populations for mottled leaf.

2.9. Candidate Gene Analysis of Major QTLs

According to the Cucurbit Genomics Database and SoftBerry, a total of 105 putative genes were identified in the *CpML1.1* region (925.2 kb) and 41 putative genes in the *CpML17.1* region (719.7 kb). Of these, 12 genes in the *CpML1.1* region and 15 genes in the *CpML17.1* region had non-synonymous changes in the CDS region and InDels/SNPs in the promoter element when comparing lines '19' and '113' (Tables S3 and S4). In addition, we found that changes in the arrangement of palisade cells were the main reason underlying the mottled leaves. Therefore, candidate genes may be involved in the development of palisade cells. In addition, previous research reported that E3 ubiquitin ligase also participates in the formation of air-space type leaf color mutants [31]. Therefore, we first selected genes related to cell development and encoding E3 ubiquitin ligase for further validation. *Cp4.1LG17g08140* encodes an E3 ubiquitin ligase, while *Cp4.1LG17g08260* encodes a TBL protein that has been reported to participate in the growth and development of cell walls. *Cp4.1LG17g08300* encodes a beta-glucosidase, which represents one of the important components of cellulase that is mainly involved in cell wall degradation. *Cp4.1LG01g23790* encodes a TPX2 family protein (target protein of Xklp2), which is a microtubule-associated protein. The functions of these three genes are all related to cellular development processes. We compared the protein sequence differences of these four genes in the parents, and all four genes have non-synonymous mutations (Figures S1–S4).

We selected parental and F_2 leaves for qRT-PCR. The expression levels of *Cp4.1LG17g08260* in the first and second leaves of line '19' were significantly lower than those of line '113'. An opposite trend of expression was observed for *Cp4.1LG17g08260* in the fifth and seventh leaves. In the F_2 population, the expression levels of the *Cp4.1LG17g08260* only showed a significant difference in the leaf mottle area and green area during the first leaf period (Figure 10A,E). During the first to fifth leaf period, *Cp4.1LG17g08300* was only expressed at a low expression level in the leaf mottle area of line '19' and was significantly lower than in line '113'. However, during the seventh leaf period, the expression level of *Cp4.1LG17g08300* suddenly increased in the leaf mottle area of line '19', significantly higher than the level detected in line '113'. In the F_2 population, the expression level of *Cp4.1LG17g08300* in the leaf mottle area showed a gradually decreasing trend, and the expression level in the leaf mottle area was higher than that in the green area during the first and fifth leaf periods (Figure 10B,F). The expression level of *Cp4.1LG17g08140* was lower than line '113' in the second and third leaves of line '19' and significantly higher than line '113' in the other periods. The trend in expression levels of *Cp4.1LG17g08140* in the F_2 population was opposite to the parents in the third leaves (Figure 10C,G). The expression level of *Cp4.1LG01g23790* was significantly higher than line '113' in the first and second leaf periods of line '19' but significantly lower than line '113' in the third, fifth, and seventh leaf periods. In the F_2 population, *Cp4.1LG01g23790* was expressed at higher levels in the leaf mottle area during the first leaf period and at lower levels during the third and fifth leaf periods (Figure 10D,H).

Thus, our analysis revealed that the expression trends for *Cp4.1LG17g08260*, *Cp4.1LG17g08300*, and *Cp4.1LG17g08140* between the leaf mottle area and green area in the F_2 population were opposite to those of their parents. The expression trend for *Cp4.1LG01g23790* followed the same trend as their parents. Therefore, we speculated that *Cp4.1LG01g23790* is a possible candidate gene that plays a key role in the formation of mottled leaves.

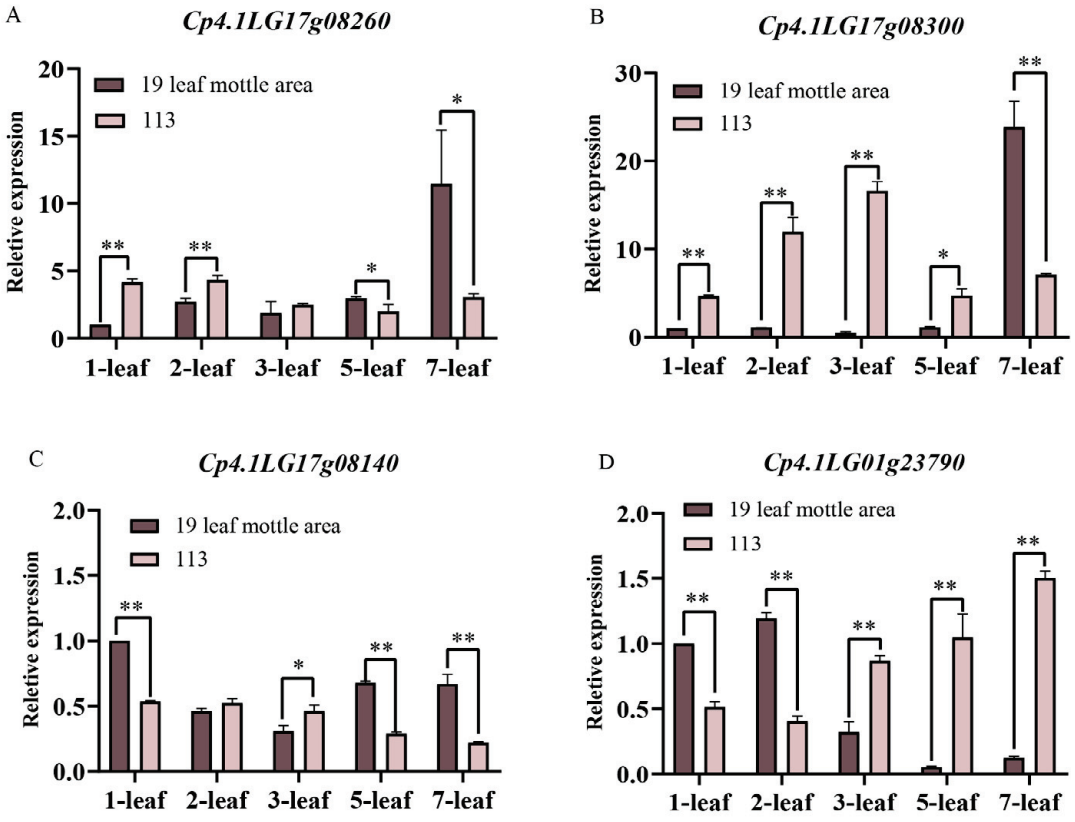


Figure 10. *Cont.*

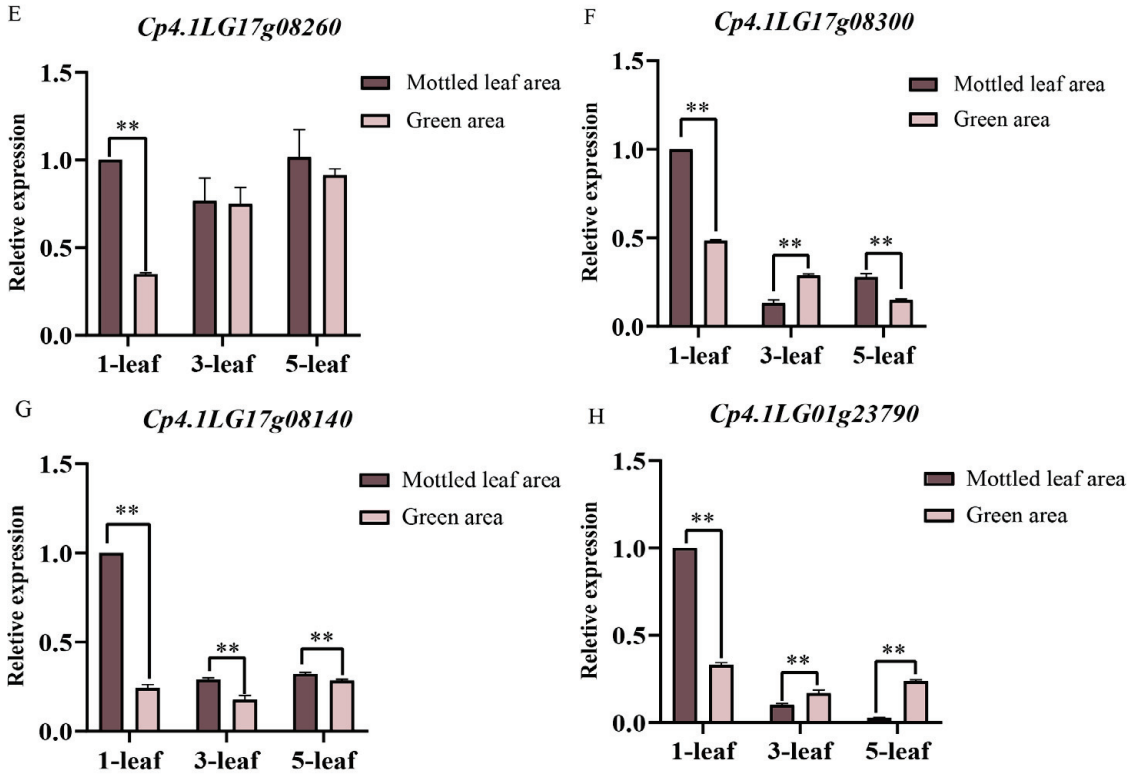


Figure 10. Relative expression levels of four candidate genes in the '19' and '113' and F_2 populations. (**A**) Relative expression levels of *Cp4.1LG17g08260*. (**B**) Relative expression levels of *Cp4.1LG17g08300*. (**C**) Relative expression levels of *Cp4.1LG17g08140*. (**D**) Relative expression levels of *Cp4.1LG01g23790*. (**A**–**D**) Relative expression levels of four candidate genes in '19' and '113'. (**E**) Relative expression levels of *Cp4.1LG17g08260*. (**F**) Relative expression levels of *Cp4.1LG17g08300*. (**G**) Relative expression levels of *Cp4.1LG17g08140*. (**H**) Relative expression levels of *Cp4.1LG01g23790*. (**E**–**H**) Relative expression levels of four candidate genes in the F_2 populations. The relative expression levels of four genes were quantified using the $2^{-\Delta\Delta CT}$ method. For the two parents, the relative expression level at the 1-leaf-stage in '19' was set to a value of 1 and used as a reference, respectively. For the F_2, the relative expression level at the 1-leaf-stage in mottled leaf area was set to a value of 1 and used as a reference, respectively. ** indicates an extremely significant difference, $p < 0.01$; * indicates a significant difference, $p < 0.05$.

3. Discussion

Various factors are responsible for leaf color variation and can be divided into the pigment type and non-pigmented type. Most leaf color mutations are mainly caused by changes in chlorophyll, carotenoid, and anthocyanin content. In the present study, we found that silver mottle appeared on zucchini leaves. To determine whether the appearance of the mottled leaf is related to chlorophyll and carotenoids, we measured the content of the photosynthetic pigment in the leaves of zucchini. The chlorophyll contents of the lines '19' and '113' were not significantly different. The differences in photosynthetic rate between these three areas were not significant. Therefore, the formation of leaf mottle in '19' does not affect chlorophyll content or chloroplast development. This indicates that mottled leaf may not affect yield. In addition, we constructed nearly isogenic lines (NILs) for the two QTL loci *CpML1.1* and *CpML17.1* identified in the article, using line '113' as the recurrent parent—both NIL-ML1.1 and NIL-ML17.1 exhibit mottled leaf phenotype. We simultaneously planted NIL-ML1.1, NIL-ML17.1, and recurrent parent '113' for unified field management. Compared with '113', the number of fruits and single fruit weight of

NIL-ML1.1 and NIL-ML17.1 did not significantly decrease. This result further confirms this speculation.

The morphology analysis of mottled leaves and green leaves showed that the palisade tissue was closely arranged in the green area of line '19' and '113', but was loosely arranged and featured air space between cells in the mottled area of leaves from line '19'. Therefore, the air space between the palisade tissue in the leaf mottle area of line '19' may be the main reason for the mottled leaf phenotype. Our findings are similar to those of Scarchuk, who previously reported the presence of air space between the palisade cells from the leaf mottle area and that these cells were not tightly associated with epidermal cells. These air spaces were found to be responsible for the silver leaf mottle phenotype [32]. A similar phenomenon has also been reported in ornamental plants. Air spaces between the cells were reported to cause leaf variegation, confirmed in *Erythronium dens-canis* L. and *Egonia rexPutz* [33,34]. In the present study, we found that the formation of mottled leaf requires the participation of light. However, the growth of leaf mottle varied between parental materials grown in different environments. Therefore, the formation of mottled leaf may be related to factors such as light quality, light intensity, and light duration. In Cucurbitaceae crops, the formation of mottled leaves in cucumber is also light-dependent and is most sensitive to ultraviolet light in sunlight [35]. In the future, we plan to investigate the effects of different light qualities on the formation of mottled leaf in zucchini to explore the most suitable light environment for the growth of mottled leaf in zucchini.

QTL-seq and the detection of molecular markers reveal that the silver leaf gene was located at an interval of 287.15 kb on chromosome 4 in *Cucurbita moschata*. The segregation ratio of green leaf to silver leaf was 3:1, showing that the silver leaf trait was inherited in a completely recessive manner. The gene *CmoCh04G023390* gene was located in the candidate interval of chromosome 4 in *C. moschata* and can be mapped to chromosome 1 in zucchini at the physical position of 16561706 bp to 16565740 bp. A major QTL for *Cucurbita pepo* L. silver leaf *Sl12_1* was detected on chromosome 17 at a physical position of 1.72 Mb-2.01 Mb. Two minor QTLs for silver leaf *Sl1_1* and *Sl16_1* were detected on chromosome 1 and chromosome 13 at the physical locations of 0.54 Mb-0.72 Mb and 0.1 Mb-1.77 Mb, respectively [30]. In our present research, we found that the mottled leaf trait was a quantitative trait. Two major QTLs were detected using composite interval mapping. *CpML1.1* was located on chromosome 1 at a physical position of 18.7 Mb-19.6 Mb. *CpML17.1* was located on chromosome 17 at a physical position of 5.22 Mb-5.94 Mb. Compared with previous results, *Sl1_1* and *CpML1.1* were both located on chromosome 1, while *Sl12_1* and *CpML17.1* were both located on chromosome 17. However, the physical positions of the two QTLs identified in this study were located far from *Sl1_1* and *Sl12_1*. Therefore, we hypothesize that *CpML1.1* and *CpML17.1* may represent two new QTLs that control mottled leaf.

To verify the accuracy of the QTL-seq results, we developed InDel molecular markers in the candidate regions to test the stability of the loci and narrow the location intervals. We found that the inheritance of *CpML1.1* and *CpML17.1* were stable in different environments. In different environments, when both loci carry a homozygous 'A', the mottled leaf is classified as grade 2–3. When both loci carry a homozygous 'B', the grade of the mottled leaf was 0–1 grade. When one of the two loci was homozygous for 'A', the phenotype of the mottled leaf was grade 1–3, which indicates that the two loci have an additive effect on the formation of leaf mottle. The more homozygote 'A' carried, the higher the grade of the mottled leaf. In 2019A, the *CpML1.1* and *CpML17.1* explained 24.15% and 20.80% of the phenotypic variation, respectively. When *CpML1.1* and *CpML17.1* carry different homozygosity, the mottled leaf grade was concentrated in grade 2, and the influence between the two loci is not significant. In 2020A, when *CpML1.1* and *CpML17.1* exhibited different homozygosity, the highest grade of mottled leaf was grade 2. However, when *CpML1.1* was homozygous for 'B', and *CpML17.1* was homozygous for 'A', there were more plants that were grade 0. The phenotypic variation of *CpML1.1* was also higher than that of *CpML17.1*. Therefore, the effect of *CpML1.1* was slightly stronger than that of *CpML17.1*.

In 2020M, when *CpML1.1* was homozygous for 'A' and *CpML17.1* was homozygous for 'B', the mottled leaf grade was mainly concentrated in grades 0–1. When *CpML1.1* was homozygous for 'B', and *CpML17.1* was homozygous for 'A', the mottled leaf grades were mainly concentrated in grades 1–2. In this environment, the LOD score for *CpML17.1* was as high as 52.99, explaining the phenotypic variation of 38.68%. *CpML1.1* explained 10.51% of the phenotypic variation. This phenomenon indicates that *CpML17.1* plays a stronger role than *CpML1.1* (Figure 11). Thus, the strength of *CpML1.1* and *CpML17.1* will be affected in different environments, but both loci will exert their effectiveness. Due to the limitations of F_2 population mapping, it is impossible to locate a single QTL precisely. Many quantitative traits are currently mapped by applying the NIL method. In order to investigate the relationship between *CpML1.1* and *CpML17.1* further, it was necessary to construct NILs for the two QTLs, respectively. These two NILs can be used for fine mapping and provide a foundation for gene cloning.

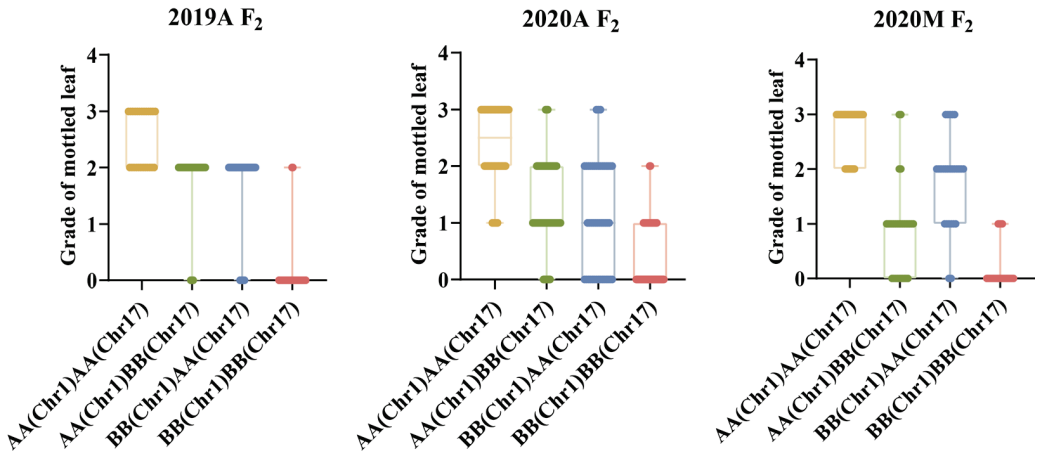


Figure 11. Results of the mottled leaf grade when *CpML1.1* and *CpML17.1* carry different homozygotes.

Cell division is crucial for plant growth and development, and microtubules are essential for eukaryotic cell division, expansion, and differentiation [36,37]. Plants have unique microtubule arrays that control the direction of cell division and expansion, mainly regulated by microtubule-associated proteins (MAPs). There are various conserved microtubule-related proteins in eukaryotes, such as augmin, TPX2, CLASP and EB1. By applying gene annotation, gene sequence alignment, and qRT-PCR analysis, we identified a *Cp4.1LG01g23790* gene, which encodes a TPX2 family protein (target protein of Xklp2). TPX2 (target protein Xklp2) is an evolutionarily conserved microtubule-associated protein and a key factor for mitotic spindle assembly factor. To date, all plant TPX2-family proteins have been shown to bind to microtubules and function in distinct processes such as cell division and the regulation of hypocotyl cell elongation by hormones and light signals [38,39]. Thus far, several TPX2 family proteins have been reported, including TPX2, WAVE DAMPENED 2 (WVD2), WAVE DAMPENED 2 LIKE (WDL) 1, 2, and 3 and MAP20. In *Arabidopsis*, WAVE-DAMPENED2-LIKE5 (WDL5) is a microtubule-stabilizing protein that plays a positive role in ethylene-regulated hypocotyl cell extension [40]. The overexpression of *EgMAP20* and *EgWDL3L* in *Arabidopsis* leads to changes in cell morphology and results in organ-twisting phenotypes [41]. However, there have been no reports on how TPX2 could influence the arrangement of leaf cells. It has been reported that the microtubule protein CLASP can influence the arrangement of leaf cells. In *Arabidopsis*, the *clasp-1* mutant exhibits defects in cell-directed amplification [42]. The lack of microtubule-related protein CLSAP leads to changes in cell division patterns, resulting in significant distortions in the topological relationships between cells and intercellular spaces and changes in their relative abundance [43]. In the present study, we detected a difference in the expression of TPX2 between the mottled area and green area; this may influence cell division and lead to mottled leaf.

4. Materials and Methods

4.1. Plant Material

The two *C. pepo* inbred lines '19' and '113' were originally developed by the Laboratory of Pumpkin Molecular Genetic Breeding, Northeast Agricultural University, Harbin, China. Inbred line '19' showed significant silver mottle on leaves; there was no leaf mottle on inbred line '113'. The F_1 and F_2 populations derived by crossing '19' and '113' were constructed to analyze the inheritance and QTL analysis for mottled leaves. 40 '19', 40 '113', 40 F_1, and 580 F_2 were planted in a greenhouse in April 2019 (2019A). 40 '19', 40 '113', 40 F_1, and 580 F_2 were planted in a greenhouse in April 2020 (2020A). 40 '19', 40 '113', 40 F_1, and 480 F_2 were planted under open cultivation in May 2020 (2020M). All plants were grown at the Xiangyang Base of Northeast Agricultural University, Harbin, China (N45°77′, E126°92′).

4.2. Determination of Photosynthetic Pigments

The first, second, third, and fifth leaves from the growing tip were selected as the material during the 20-leaf-stage. The mottle area and green area of '19' and the green leaves in the same area for '113' were collected. The pigment in the leaves was extracted using 80% acetone. The absorbance of the extract was measured at 663, 645, and 470 nm using a microplate reader. The control was 80% acetone. Three biological repeats were determined for each material, and technical repeats were performed three times in each biological repeat. The formula for calculating pigment content is as follows:

$$C_{chla}\ (mg/g) = 12.21 \times A663 - 2.81 \times A645$$

$$C_{chlb}\ (mg/g) = 20.13 \times A645 - 5.03 \times A663$$

$$C_{caro}\ (mg/g) = (1000 \times A470 - 3.27 \times Ca - 104 \times Cb)/229$$

$$\text{Total Chl} = C_{chla} + C_{chlb}$$

4.3. Determination of Photosynthetic Parameters

The fifth true leaf from the growing tip of the 20-leaf-stage was selected to measure the photosynthetic parameters. The mottle area and green area of '19' and the green leaves in the same area for '113' were collected. Net photosynthetic rates (Pn; μmol CO_2 $m^{-2}s^{-1}$), stomatal conductance (Gs; mol H_2O $m^{-2}s^{-1}$), intercellular carbon dioxide concentrations (Ci; μmol CO_2 mol^{-1}), and transpiration rates (Ts; mmol H_2O $m^{-2}s^{-1}$) were measured, using CI-340. Three biological repeats were determined for each material, and technical repeats were performed three times in each biological repeat.

4.4. Leaf Anatomy Assay

At the stage of twenty leaves, the fifth leaf under the growing tip is used for leaf structure analysis. The mottle area and green area of '19' and the green leaves in the same area for '113' were collected. Prepare FAA and different concentrations of alcohol to fix and dehydrate the leaves and dye them with safranin-fast green. After taking photos, use ImageJ to measure the thickness of the upper epidermis, palisade parenchyma, sponge parenchyma, and lower epidermis of leaves.

$$CTR = (\text{thickness of palisade parenchyma}/\text{leaf thickness}) \times 100\%$$

$$SR = (\text{thickness of sponge parenchyma}\ /\text{leaf thickness}) \times 100\%$$

4.5. Inheritance Analysis of the Mottled Leaf Trait

F_2 population of '19' and '113' was constructed for inheritance analysis of the mottled leaves. We conducted a survey on mottled leaf grade during the 20-leaf-stage. According to the percentage of mottled area to total leaf area, the grade of mottled leaf was divided into three grades in 2019 and four grades in 2020. The grade 0 plant showed no leaf mottle. The leaf mottle area of grades 1 to 3 accounts for 20%, 50%, and 80% of the total leaf area, respectively (Figure 12).

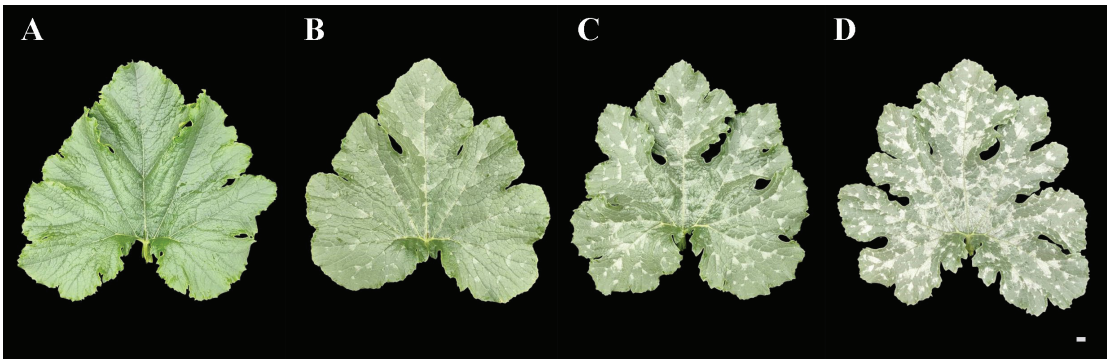


Figure 12. Grading standard of F_2 mottled leaf. (**A**) 0 grade. (**B**) 1 grade. (**C**) 2 grade. (**D**) 3 grade. Scale = 1 cm.

4.6. Pool Construction and QTL-seq Analysis

Genomic DNA was extracted from '19', '113', F_1, and F_2 using the modified cetyltrimethylammonium bromide (CTAB) method. The mottled leaf pool (M-Pool) and no-mottled-leaf pool (N-Pool) were constructed by mixing equal amounts of DNA from 30 extremely mottled leaf plants (grade 3) and 30 extremely without mottled leaf plants (grade 0) from the 2019A F_2 population. The parental DNA pools were constructed by mixing equal amounts of DNA from the 30 '19' and 30 '113' plants. DNA libraries were sequenced on the Illumina Hiseq 2000 platform in BioMaker (Peking, China). In order to ensure the quality of information analysis, the Raw reads were filtered to obtain clean reads. The clean reads of all samples were compared with the reference genome of zucchini (http://cucurbitgenomics.org/organism/14, accessed on 2 July 2019) using the BWA software (https://bio-bwa.sourceforge.net/bwa.shtml, accessed on 2 July 2019) for subsequent variance analysis [44,45]. The SNPs and InDels were detected and filtered using GATK. The SNP/InDel index of the Euclidean distance (ED) and Δ(SNP/InDel-index) were calculated for all positions to determine the region associated with the mottled leaf.

4.7. QTL Analysis with Molecular Markers

Polymorphic InDel markers were developed in the candidate regions of mottled leaf based on the parent resequencing. InDel markers were designed with Primer Premier 5.0 (Table S2). PCR was carried out using 10 μL samples containing ~40 ng of genomic DNA, each primer at 0.5 μM, 200 μM dNTPs, 1× reaction buffer, and 0.5 U of Taq DNA polymerase (Aidlab Biotechnologies, Beijing, China). PCR amplification was performed using the following program: 94 °C for 5 min; 35 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s; and 72 °C for 5 min [46,47]. The primers used in this study were synthesized using the BGI gene. Products were separated on an 8% polyacrylamide gel by electrophoresis. After electrophoresis at 220 V for 2 h, the gel was stained with 0.3% $AgNO_3$ solution, and the silver-stained DNA bands were revealed.

4.8. Data Analysis

A, B, and H are homozygous '19', homozygous '113', and heterozygous F_1 genotypes, respectively. We used JoinMap 4.0 to perform genetic mapping of F_2 individuals using

InDel markers on target chromosomes and obtained genetic distances. In R/qtl software (http://www.rqtl.org/, accessed on 2 July 2019), composite interval mapping (CIM) was used to test the significance of QTL [48,49]. The significance of each QTL was tested using LOD thresholds ($p < 0.05$), which were determined using 1000 permutations. For each detected QTL, a 2-LOD-support interval was calculated and defined by left and right markers. The QTLs were named according to chromosome locations and different environments.

4.9. Prediction of Candidate Genes and qRT-PCR

The gene ID, function, and structure data were obtained from the gourd genome database [44]. SoftBerry online software (http://linux1.softberry.com/, accessed on 22 November 2022) was used for gene structure prediction based on two family plants (*Arabidopsis*). The first, second, third, fifth, and seventh leaves under the growing tip at the 20-leaf stage of parents material and the first, third, and fifth leaves under the growing tip at the 20-leaf stage in the F_2 population were selected to determine the expression of candidate genes. Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Three pairs of primers for candidate genes were designed in total. Taq SYBR Green qPCR Premix (Yugong Biolabs, Inc., Jiangsu, China) was applied to perform qRT-PCR. The program used in this assay was as follows: 96 °C for 1 min; 30 cycles of 95 °C for 15 s, 56 °C for 15 s, and 72 °C for 45 s [50,51]. The Actin gene was used as the internal reference. Three technical replicates were set for each sample, and relative expression levels were quantified using the $2^{-\Delta\Delta CT}$ method.

5. Conclusions

In summary, our findings indicate that the formation of mottled leaves did not affect the pigment content or photosynthesis in zucchini leaves. The large air space between cells in the palisade parenchyma was identified as the main factor underlying the formation of mottled leaf. During the early stages of leaf development, light is an important condition for the appearance of the mottled leaf. Genetic analysis revealed that mottled leaf is an inheritable quantitative trait and is controlled by multiple genes. Based on QTL-seq and the development of InDel markers, two major QTLs, *CpML1.1* and *CpML17.1*, were detected in different environments. They were located in a 925.2-kb interval on chromosome 1 and a 719.7-kb interval on chromosome 17, respectively. By performing gene annotation, gene sequence alignment, and qRT-PCR analysis, we found that the *Cp4.1LG01g23790* gene may be a candidate gene for mottled leaf in zucchini. Our results provide a suitable foundation for the fine mapping and mechanistic research of mottled leaf traits.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25052491/s1.

Author Contributions: K.W. performed the data analysis and prepared the manuscript. X.W. and L.Z. contributed to collecting phenotypic characteristics and DNA extraction. Y.C. contributed to growing plants. Y.L. contributed to RNA extraction and qRT-PCR test. W.X., Y.W., and S.Q., the corresponding authors, oversaw all project implementation and manuscript development activities. 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 (32272723 and 32072590).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw datasets of QTL-seq generated during the current study are available in the NCBI repository, BioProject ID: PRJNA826826 ('19' and '113') and PRJNA1077909 (M-pool and N-pool).

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

References

1. Vairam, N.; Ibrahim, S.-M.; Vanniarajan, C. Frequency and spectrum of chlorophyll mutations in greengram [*Vigna radiata* (L.) Wilczek]. *Asian J. Biol. Sci.* **2014**, *9*, 204–207. [CrossRef]
2. Zhang, J.-H.; Zeng, J.-C.; Wang, X.-M.; Chen, S.-F.; Albach, D.C.; Li, H.-Q. A revised classification of leaf variegation types. *Flora* **2020**, *272*, 151703. [CrossRef]
3. Aluru, M.-R.; Bae, H.; Wu, D.; Rodermel, S.-R. The *Arabidopsis immutans* mutation affects plastid differentiation and the morphogenesis of white and green sectors in variegated plants. *Plant Physiol.* **2001**, *127*, 67–77. [CrossRef]
4. Wetzel, C.-M.; Jiang, C.-Z.; Meehan, L.-J.; Voytas, D.F.; Rodermel, S.-R. Nuclear-organelle interactions: The *immutans* variegation mutant of *Arabidopsis* is plastid autonomous and impaired in carotenoid biosynthesis. *Plant J.* **1994**, *6*, 161–175. [CrossRef]
5. Carol, P.; Stevenson, D.; Bisanz, C.; Breitenbach, J.; Sandmann, G.; Mache, R.; Coupland, G.; Kuntz, M. Mutations in the Arabidopsis gene *IMMUTANS* cause a variegated phenotype by inactivating a chloroplast terminal oxidase associated with phytoene desaturation. *Plant Cell* **1999**, *11*, 57–68. [CrossRef] [PubMed]
6. Næsted, H.; Holm, A.; Jenkins, T.; Nielsen, H.-B.; Harris, C.-A.; Beale, M.-H.; Andersen, M.; Mant, A.; Scheller, H.; Camara, B.; et al. *Arabidopsis VARIEGATED 3* encodes a chloroplast-targeted, zinc-finger protein required for chloroplast and palisade cell development. *J. Cell Sci.* **2004**, *117*, 4807–4818. [CrossRef]
7. Haswell, E.-S.; Meyerowitz, E.-M. MscS-like proteins control plastid size and shape in *Arabidopsis thaliana*. *Curr. Biol.* **2006**, *16*, 1–11. [CrossRef] [PubMed]
8. Sakamoto, W.; Tamura, T.; Hanba-Tomita, Y.; Murata, M.; Sodmergen. The *VAR1* locus of *Arabidopsis* encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. *Genes Cells* **2002**, *7*, 769–780. [CrossRef]
9. Chen, M.; Choi, Y.; Voytas, D.-F.; Rodermel, S. Mutations in the Arabidopsis VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. *Plant J.* **2000**, *22*, 303–313. [CrossRef]
10. Bailey, S.; Thompson, E.; Nixon, P.-J.; Horton, P.; Mullineaux, C.-W.; Robinson, C.; Mann, N. A critical role for the Var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle in vivo. *J. Biol. Chem.* **2002**, *277*, 2006–2011. [CrossRef] [PubMed]
11. Liu, X.; Yu, F.; Rodermel, S. Arabidopsis chloroplast FtsH, *var2* and suppressors of *var2* leaf variegation: A review. *J. Integr. Plant Biol.* **2010**, *52*, 750–761. [CrossRef] [PubMed]
12. Liu, Z.-W.; Wang, Z.-Y.; Gu, H.; You, J.; Hu, M.-M.; Zhang, Y.-J.; Zhu, Z.; Wang, Y.-H.; Liu, S.-J.; Chen, L.-M.; et al. Identification and Phenotypic Characterization of ZEBRA LEAF16 Encoding a β-Hydroxyacyl-ACP Dehydratase in Rice. *Front. Plant Sci.* **2018**, *9*, 782. [CrossRef] [PubMed]
13. Li, Y.-Q.; Zhong, P.; Gao, Z.-Y.; Zhu, B.-Y.; Chen, D.; Su, C.-H.; Wang, P.-R.; Deng, X.-J. Morphological characterization and candidate gene analysis of zebra leaf mutant zebra524 in rice. *Sci. Agric. Sin.* **2014**, *47*, 2907–2915.
14. Qiu, Z.-N.; Chen, D.-D.; He, L.; Zhang, S.; Yang, Z.-N.; Wang, Z.-W.; Ren, D.-Y.; Qian, Q.; Guo, L.-B.; Zhu, L. The rice white green leaf 2 gene causes defects in chloroplast development and affects the plastid ribosomal protein S9. *Rice* **2018**, *11*, 39. [CrossRef] [PubMed]
15. Zhang, T.; Feng, P.; Li, Y.-F.; Yu, G.-L.; Sang, X.-C.; Ling, Y.-H.; Zeng, X.-Q.; Li, Y.-D.; Huang, J.-Y. VIRESCENT-ALBINO LEAF 1 regulates leaf colour development and cell division in rice. *J. Exp. Bot.* **2018**, *69*, 4791–4804. [CrossRef]
16. Feng, P.; Shi, J.-Q.; Zhang, T.; Zhong, Y.-Q.; Zhang, L.-S.; Yu, G.-L.; Zhang, T.-Q.; Zhu, X.-Y.; Xing, Y.-D.; Yin, W.-Z.; et al. Zebra leaf 15, a receptor-like protein kinase involved in moderate low temperature signaling pathway in rice. *Rice* **2019**, *12*, 1–13. [CrossRef]
17. Li, Q.-G.; Du, J.-Y.; Qiao, Z.-H.; Pan, C.; He, W.-Q.; Zhang, L.; Li, X.-H.; Nie, Y.-X.; Li, X.-Z.; Pan, G.-T.; et al. White and green striate leaves 1, predicted to encode a 16S rRNA processing protein, plays a critical role in the processing of chloroplast ribosomes in maize (*Zea mays* L.). *Mol. Breed.* **2023**, *43*, 65. [CrossRef]
18. Hu, L.-L.; Zhang, H.-Q.; Xie, C.; Wang, J.; Zhang, J.-Y.; Wang, H.; Weng, Y.-Q.; Chen, P.; Li, Y.-H. A mutation in CsHD encoding a histidine and aspartic acid domain-containing protein leads to yellow young leaf-1 (yyl-1) in cucumber (*Cucumis sativus* L.). *Plant Sci.* **2020**, *293*, 110407. [CrossRef]
19. Ding, Y.; Yang, W.; Su, C.-G.; Ma, H.-H.; Pan, Y.; Zhang, X.-G.; Li, J.-H. Tandem 13-lipoxygenase genes in a cluster confers yellow-green leaf in cucumber. *Int. J. Mol. Sci.* **2019**, *20*, 3102. [CrossRef]
20. Scarchuk, J. Fruit and leaf characters in summer squash: The interrelationship of striped-fruit and mottled-leaf. *J. Hered.* **1954**, *45*, 295–297. [CrossRef]
21. Scott, D.-H.; Riner, M.-E. A Mottled-Leaf Character in Winter Squash: Inherited as a Dominant Mendelian Character. *J. Hered.* **1946**, *37*, 27–28. [CrossRef] [PubMed]
22. Lopez-Anido, F.; Cointry, E.; Firpo, I.; García, S.-M.; Gattuso, S. Inheritance of gray leaf color in a material derived from a *Cucurbita maxima* Duch. x *C. moschata* Duch. hybrid. *Rep. Cucurbit Genet. Coop.* **2002**, *25*, 46–48.
23. Ribeiro, A.; da Costa, C.-P. Inheritance of mottled leaf in *Cucurbita moschata*. *Cucurbit Genet. Coop.* **1989**, *29*, 70–71.
24. Shifriss, O. Further notes on the silvery-leaf trait in Cucurbita. *Cucurbit Genet Coop. Rep.* **1984**, *7*, 81–83.
25. Shifriss, O. Do Cucurbita plants with silvery leaves escape virus infection? Origin and characteristics of NJ260. *Cucurbit Genet. Coop.* **1981**, *4*, 42–43.
26. Shifriss, O. Reflected light spectra from silvery and nonsilvery leaves of *Cucurbita pepo*. *Cucurbit Genet. Coop.* **1983**, *44*, 89–90.
27. Mcauslane, H.-J.; Chen, J.; Carle, R.-B.; Schmalstig, J. Influence of *Bemisia argentifolii* (Homoptera: Aleyrodidae) infestation and squash silverleaf disorder on zucchini seedling growth. *J. Econ. Entomol.* **2004**, *97*, 1096–1105. [CrossRef]

28. Jiménez, D.-R.; Yokomi, R.-K.; Mayer, R.-T.; Shapiro, J.-P. Cytology and physiology of silver leaf whitefly-induced squash silver leaf. *Physiol. Mol. Plant Pathol.* **1995**, *46*, 227–242. [CrossRef]
29. Brown, R.-N.; Myers, J.-R. A genetic map of squash (*Cucurbita* sp.) with randomly amplified polymorphic DNA markers and morphological markers. *J. Am. Soc. Hortic. Sci.* **2002**, *127*, 568–575. [CrossRef]
30. Montero-Pau, J.; Blanca, J.; Esteras, C.; Martínez-Pérez, E.-M.; Gómez, P.; Monforte, A.-J.; Cañizares, J.; Picó, B. An SNP-based saturated genetic map and QTL analysis of fruit-related traits in Zucchini using Genotyping-by-sequencing. *BMC Genom.* **2017**, *18*, 94. [CrossRef]
31. Li, K.-T.; Zhang, J.; Kang, Y.-H.; Chen, M.-C.; Song, T.-T.; Geng, H.; Tian, J. McMYB10 modulates the expression of a Ubiquitin Ligase, McCOP1 during leaf coloration in crabapple. *Front. Plant Sci.* **2018**, *9*, 704. [CrossRef] [PubMed]
32. Scarchuk, J.; Lent, J.-M. The structure of mottled-leaf summer squash. *J. Hered.* **1965**, *56*, 167–168. [CrossRef]
33. Zhang, Y.; Hayashi, T.; Hosokawa, M.; Yazawa, S.; Li, Y. Metallic lustre and the optical mechanism generated from the leaf surface of *Begonia rex* Putz. *Sci. Hortic.* **2009**, *121*, 213–217. [CrossRef]
34. La Rocca, N.; Pupillo, P.; Puppi, G.; Rascio, N. *Erythronium dens-canis* L.(Liliaceae): An unusual case of change of leaf mottling. *Plant Physiol. Biochem.* **2014**, *74*, 108–117. [CrossRef]
35. Cao, W.; Du, Y.; Wang, C.; Xu, L.; Wu, T. Cscs encoding chorismate synthase is a candidate gene for leaf variegation mutation in cucumber. *Breed. Sci.* **2018**, *68*, 571–581. [CrossRef]
36. Hamada, T. Microtubule organization and microtubule-associated proteins in plant cells. *Int. Rev. Cell Mol. Biol.* **2014**, *312*, 1.
37. Hashimoto, T. Microtubules in plants. *Arab. Book Am. Soc. Plant Biol.* **2015**, *13*, e0179. [CrossRef] [PubMed]
38. Smertenko, A.; Clare, S.-J.; Effertz, K.; Parish, A.; Ross, A.; Schmidt, S. A guide to plant TPX2-like and WAVE-DAMPENED2-like proteins. *J. Exp. Bot.* **2021**, *72*, 1034–1045. [CrossRef]
39. Dvořák Tomaštíková, E.; Rutten, T.; Dvořák, P.; Tugai, A.; Ptosková, K.; Petrovská, B.; van Damme, D.; Houben, A.; Dolezel, J.; Demidov, D. Functional divergence of microtubule-associated TPX2 family members in Arabidopsis thaliana. *Int. J. Mol. Sci.* **2020**, *21*, 2183. [CrossRef]
40. Sun, J.; Ma, Q.; Mao, T. Ethylene regulates the *Arabidopsis* microtubule-associated protein WAVE-DAMPENED2-LIKE5 in etiolated hypocotyl elongation. *Plant Physiol.* **2015**, *169*, 325–337. [CrossRef]
41. Du, P.-Z.; Kumar, M.; Yao, Y.; Xie, Q.-L.; Wang, J.-Y.; Zhang, B.-L.; Gan, S.-M.; Wang, Y.-Q.; Wu, A.-M. Genome-wide analysis of the TPX2 family proteins in *Eucalyptus grandis*. *BMC Genom.* **2016**, *17*, 967. [CrossRef]
42. Ambrose, J.-C.; Shoji, T.; Kotzer, A.-M.; Pighin, J.-A.; Wasteneys, G.-O. The *Arabidopsis CLASP* Gene Encodes a Microtubule-Associated Protein Involved in Cell Expansion and Division. *Plant Cell* **2007**, *19*, 2763–2775. [CrossRef] [PubMed]
43. Zhang, L.; Ambrose, C. CLASP balances two competing cell division plane cues during leaf development. *Nat. Plants* **2022**, *8*, 682–693. [CrossRef]
44. Montero-Pau, J.; Blanca, J.; Bombarely, A.; Ziarsolo, P.; Esteras, C.; Marti-Gómez, C.; Ferriol, M.; Gómez, P.; Jamilena, M.; Mueller, L.; et al. De novo assembly of zucchini genome reveals a whole-genome duplication associated with the origin of the *Cucurbita* genus. *Plant Biotechnol. J.* **2018**, *16*, 1161–1171. [CrossRef] [PubMed]
45. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]
46. Qu, S.-P.; Dan, Y.; Yu, H.-Y.; Chen, F.-Y.; Wang, K.-X.; Ding, W.-Q.; Xu, W.-L.; Wang, Y.-L. QTL analysis of early flowering of female flowers in zucchini (*Cucurbita pepo* L.). *J. Integr. Agric.* **2023**, *22*, 3321–3330. [CrossRef]
47. Wang, Y.-L.; Wang, G.-C.; Lin, D.-J.; Luo, Q.-F.; Xu, W.-L.; Qu, S.-P. QTL mapping and stability analysis of trichome density in zucchini (*Cucurbita pepo* L.). *Front. Plant Sci.* **2023**, *14*, 1232154. [CrossRef]
48. Wang, G.-C.; Dai, D.-Y.; Wang, L.; Sheng, Y.-Y.; Wang, D.; Li, D.-D.; Tian, L.-M.; Luan, F.-S. QTL analysis of flowering-related traits by specific length amplified fragment sequencing in melon. *Crop Sci.* **2022**, *62*, 203–215. [CrossRef]
49. Weng, Y.-Q.; Colle, M.; Wang, Y.-H.; Yang, L.-M.; Rubinstein, M.; Sherman, A.; Ophir, R.; Grumet, R. QTL mapping in multiple populations and development stages reveals dynamic QTL for fruit size in cucumbers of different market classes. *Theor. Appl. Genet.* **2015**, *128*, 1747–1763. [CrossRef]
50. Ding, W.-Q.; Wang, Y.-L.; Qi, C.; Luo, Y.-S.; Wang, C.-J.; Xu, W.-L.; Qu, S.-P. Fine mapping identified the gibberellin 2-oxidase gene CpDw leading to a dwarf phenotype in squash (*Cucurbita pepo* L.). *Plant Sci.* **2021**, *306*, 110857. [CrossRef]
51. Luo, Y.-S.; Wang, C.-J.; Wang, M.-M.; Wang, Y.-L.; Xu, W.-L.; Han, H.-Y.; Wang, Z.-C.; Zhong, Y.-J.; Huang, H.-X.; Qu, S.-P. Accumulation of carotenoids and expression of carotenoid biosynthesis genes in fruit flesh during fruit development in two Cucurbita maxima inbred lines. *Hortic. Plant J.* **2021**, *7*, 529–538. [CrossRef]

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

International Journal of Molecular Sciences Editorial Office
E-mail: ijms@mdpi.com
www.mdpi.com/journal/ijms

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-5856-9