



# Article Comprehensive Analysis of bHLH Transcription Factors in *Ipomoea aquatica* and Its Response to Anthocyanin Biosynthesis

Zheng Liu <sup>1</sup>, Xiaoai Fu<sup>1</sup>, Hao Xu<sup>1</sup>, Yuxin Zhang<sup>1</sup>, Zhidi Shi<sup>1</sup>, Guangzhen Zhou<sup>2</sup> and Wenlong Bao<sup>1,3,\*</sup>

- Key Laboratory for Quality Regulation of Tropical Horticultural Crops of Hainan Province, School of Horticulture, Hainan University, Haikou 570228, China
- <sup>2</sup> College of Tropical Crops, Hainan University, Haikou 570228, China
- <sup>3</sup> Hainan Yazhou Bay Seed Laboratory, Sanya Nanfan Research Institute of Hainan University, Sanya 572025, China
- \* Correspondence: wlbao@hainanu.edu.cn

Abstract: The basic helix-loop-helix (bHLH) proteins compose one of the largest transcription factor (TF) families in plants, which play a vital role in regulating plant biological processes including growth and development, stress response, and secondary metabolite biosynthesis. Ipomoea aquatica is one of the most important nutrient-rich vegetables. Compared to the common green-stemmed I. aquatica, purple-stemmed I. aquatica has extremely high contents of anthocyanins. However, the information on bHLH genes in I. aquatica and their role in regulating anthocyanin accumulation is still unclear. In this study, we confirmed a total of 157 bHLH genes in the I. aquatica genome, which were classified into 23 subgroups according to their phylogenetic relationship with the bHLH of Arabidopsis thaliana (AtbHLH). Of these, 129 labHLH genes were unevenly distributed across 15 chromosomes, while 28 IabHLH genes were spread on the scaffolds. Subcellular localization prediction revealed that most IabHLH proteins were localized in the nucleus, while some were in the chloroplast, extracellular space, and endomembrane system. Sequence analysis revealed conserved motif distribution and similar patterns of gene structure within *labHLH* genes of the same subfamily. Analysis of gene duplication events indicated that DSD and WGD played a vital role in the *labHLH* gene family expansion. Transcriptome analysis showed that the expression levels of 13 IabHLH genes were significantly different between the two varieties. Of these, the *labHLH027* had the highest expression fold change, and its expression level was dramatically higher in purple-stemmed I. aquatica than that in green-stemmed I. aquatica. All upregulated DEGs in purple-stemmed I. aquatica exhibited the same expression trends in both qRT-PCR and RNA-seq. Three downregulated genes including *labHLH142*, IabHLH057, and IabHLH043 determined by RNA-seq had opposite expression trends of those detected by qRT-PCR. Analysis of the cis-acting elements in the promoter region of 13 differentially expressed genes indicated that light-responsive elements were the most, followed by phytohormone-responsive elements and stress-responsive elements, while plant growth and development-responsive elements were the least. Taken together, this work provides valuable clues for further exploring *IabHLH* function and facilitating the breeding of anthocyanin-rich functional varieties of I. aquatica.

**Keywords:** gene family; cis-acting element; gene duplication events; evolutionary relationship; gene expression

# 1. Introduction

The bHLH proteins comprise one of the largest transcription factor (TF) families in eukaryotes including plants, animals, and fungi. The bHLH TF contains two functional conserved domains with a total of approximately 60 amino acid residues, namely the N-terminal basic region composed of 13–17 amino acids and the C-terminal HLH region with 40–50 amino acids, respectively. Typically, the basic region of bHLH TF in plants possesses a highly conserved HER motif that can specifically bind to the cis-acting element E-box



Citation: Liu, Z.; Fu, X.; Xu, H.; Zhang, Y.; Shi, Z.; Zhou, G.; Bao, W. Comprehensive Analysis of bHLH Transcription Factors in *Ipomoea aquatica* and Its Response to Anthocyanin Biosynthesis. Int. J. Mol. Sci. 2023, 24, 5652. https://doi.org/ 10.3390/ijms24065652

Academic Editors: Jian Zhang and Zhiyong Li

Received: 7 February 2023 Revised: 11 March 2023 Accepted: 14 March 2023 Published: 15 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). (CANNTG) or G-box (CACGTG) within the promoter region of the target gene. The HLH region consists of two relatively conserved amphipathic  $\alpha$ -helices that are separated by an intervening loop of variable length, which is required for the formation of bHLH dimers through protein-protein interaction to modulate the expression of downstream genes that correlated with diverse signaling pathways [1–4]. With the increasing publications of plant genome data, bHLH TFs in numerous plant species have been identified at the whole genome level. For instance, there are 602 bHLH genes in the Brassica napus genome [5], 208 in the Zea mays genome [6], 162 in the Arabidopsis thaliana genome [7], 115 in the Vitis davidii genome [8], and 141 in the Hordeum vulgare genome [9]. The bHLHs exert a broad range of functions in various biological processes of diverse plant species [3,10,11]. For instance, Arabidopsis flowering bHLHs are involved in the regulation of flowering time by activating the gene encoding CONSTANS (CO) protein [12]. Rice bHLH gene INCREASED LAMINA INCLINATION1 (ILI1) participates in brassinosteroid-mediated plant development by serving as a downstream target of BRASSINAZOLE-RESISTANT1 (BZR1) TF [13]. In apple, MdbHLH33 facilitates plant cold stress tolerance and anthocyanin accumulation by positively regulating the expressions of the C-repeat binding factor (CBF) TF gene and dihydroflavonol 4-reductase (DFR) gene [14].

Anthocyanins are ubiquitous secondary metabolites generated from the branch of the phenylpropanoid biosynthesis pathway. As water-soluble pigments, anthocyanins are widely distributed in diverse plant organs and tissues and give rise to colors varying with pH value, playing a crucial role in plant pollination and seed dispersal. As natural antioxidants, anthocyanins can help plants defend against various stresses including drought, low temperature, and ultraviolet radiation [15,16]. Consuming a diet high in anthocyanins can enhance immunity and delay the aging process in humans [17]. The key structural genes and regulatory genes involved in the hierarchical regulatory networks governing the anthocyanin biosynthetic pathway are well-studied in many plant species [18,19]. The structural genes involved in the anthocyanin biosynthetic pathway are classified into two categories, namely the early biosynthetic genes (EBGs) and late biosynthetic genes (LBGs). Enzymes encoded by EGBs are responsible for producing large amounts of anthocyanin precursors while those encoded by LBGs are involved in the generation of various anthocyanins [20,21]. Studies showed that the expression of the EBGs is mainly activated or repressed by the MYB TFs, while the LBGs are coordinately modulated by the members of the MYB, bHLH, and WDR TF families [22–24]. Of these TFs, the bHLHs act as an essential cofactor of the MYBs and WDRs and play a central role in the late stage of anthocyanin biosynthesis.

Since the first bHLH protein associated with anthocyanin biosynthesis was discovered in maize, a growing understanding of bHLH functions regulating anthocyanin production in plants was carried out. In Arabidopsis thaliana, a bHLH TF TRANSPARENT TESTA8 (TT8) regulates the expression of the LBG DFR and BANYULS (BAN) genes which in turn modulate the seed coat pigmentation [25]. In Chrysanthemum morifolium, CmbHLH2 serves as the essential partner of CmMYB6 to facilitate the expression of *CmDFR* and thereby enhance anthocyanin accumulation [26]. In Malus domestica, the MdbHLH3 gene was induced by low-temperature stress and promoted fruit coloration by activating the anthocyanin biosynthetic genes (ABGs) including MdUFGT and MdDFR [27]. In contrast, the StbHLH1 gene of *Solanum tuberosum* was down-regulated under high-temperature stress and led to the reduction of anthocyanin contents in flesh by affecting the expression of the ABGs [28]. In Populus, the PdbHLH (a homolog of TT8 in Arabidopsis) acts as an efficient enhancer of PdMYB118 to promote the expression of the ABGs and thereby promotes the anthocyanin production under wound [29]. In Vitis vinifera, the VvMYC1 that has been characterized as bHLH TF cooperates with several MYB TFs to participate in the transcriptional cascade involved in the regulation of anthocyanin accumulation in berries [30]. In mulberry fruits, the *bHLH3* was implicated in maintaining flavonoid homeostasis, which regulated the fruit coloration by affecting the anthocyanin compositions [31]. These studies demonstrated that bHLH TFs act as vital regulators of anthocyanin biosynthetic pathways in diverse plant species.

*I. aquatica* has a wide distribution in tropical and subtropical regions. As a healthpromoting leafy vegetable, *I. aquatica* is rich in essential amino acids, flavonoids, and diverse mineral elements such as calcium, potassium, and phosphorus [32,33]. Moreover, *I. aquatica* was used in traditional medicine as a remedy for different diseases including liver disorders, diabetes, and in the treatment of heavy metal intoxication [34–36]. Our previous study revealed that the purple-stemmed *I. aquatica* has extremely high anthocyanin contents relative to the common green-stemmed *I. aquatica* [37]. To understand the possible roles of bHLH TF related to anthocyanin biosynthesis in *I. aquatica*, we comprehensively analyzed the members of the *bHLH* gene family and their expression patterns by combining bioinformatics and experimental methods. Our study laid a foundation for further exploring the roles of *IabHLHs* in regulating anthocyanin biosynthesis in *I. aquatica*.

## 2. Results

## 2.1. Identification of IabHLH Genes in the I. aquatica Genome

To identify the members of the *labHLH* gene family, the HMM profile of bHLH and the AtbHLH protein sequences were used as queries to search against the *I. aquatica* genome. After combining the above results, the redundant sequences and the sequences without the bHLH domain were eliminated based on the running consequences of Pfam, CDD, and SMART programs (Figure 1A). A total of 157 *IabHLH* genes were identified, which were named IabHLH001–IabHLH157 according to their positions on the chromosomes and scaffolds (Table S1). The physicochemical properties of proteins play a crucial role in the bio-function and conformation of proteins. As shown in Figure 1B and Table S2, the length of IabHLHs ranged from 140 aa (IabHLH099) to 701 aa (IabHLH022), and their molecular weight (MW) varied from 15.8 kDa (IabHLH099) to 76.5 kDa (IabHLH014). The IabHLH023 had the highest theoretical isoelectric point (pI), with a value of 9.59, while IabHLH136 had the lowest, with a value of 4.8. Except for IabHLH027, IabHLH071, and IabHLH141, the remaining IabHLHs had instability index (Ii) values greater than 40. In addition, 66 of the 157 IabHLHs had aliphatic index (Ai) values lower than 71. Analysis of the grand average of hydropathicity index (GRAVY) showed that all IabHLHs had negative GRAVY values. Subcellular localization prediction of IabHLHs showed that 139, 13, 4, and one bHLH were localized in the nucleus, chloroplast, extracellular space, and endomembrane system, respectively (Table S2).



Figure 1. Cont.



**Figure 1.** Identification of *labHLH* genes. (**A**) Schematic diagram of the procedure for retrieving *labHLH* genes from the *l. aquatica* genome. (**B**) The physicochemical properties of 157 labHLH proteins.

#### 2.2. Synteny Analysis of IabHLHs

To understand the *labHLH* gene distribution in the *I. aquatica* genome, their genomic location was analyzed. A total of 129 of the 157 *labHLH* genes were unevenly distributed on 15 chromosomes, while the remaining genes were spread on the scaffolds (Figure 2A and Table S1).

To reveal the *labHLH* collinear gene pairs in the *I. aquatica* genome, the analysis of intraspecific collinearity was performed. A total of 61 *labHLH* colinear gene pairs were identified in the *I. aquatica* genome, of which 44 gene pairs were mapped on the chromosomes (Figure 2A). To explore the potential evolutionary process of *labHLH* genes, we conducted the analysis of interspecific syntenic relationships among *labHLH* genes and the other plant species' genomes, including plants of the *Convolvulaceae* family (*Ipomoea batatas*), *Brassicaceae* family (*A. thaliana*), and *Poaceae* family (*Oryza sativa*). As shown in Figure 2B, the highest number of homologous gene pairs were identified between *labHLH* genes and *I. batatas*, with 239, indicating the closest evolutionary distances between the *labHLH* gene family and *I. batatas*. As shown in Figure 2C,D, a total of 127 and 43 homologous gene pairs were identified between *I. aquatica* and *O. sativa*, respectively, suggesting a closer evolutionary relationship of the *labHLH* gene family with *A. thaliana* than with *O. sativa*. Moreover, 26 *labHLH* genes were shared among all homologous gene pairs (Table S3), inferring their relative conservation in the plant evolutionary process.

#### 2.3. Phylogenetic Relationships of IabHLHs and AtbHLHs

To gain further insights into the potential biological functions of IabHLHs, the phylogenetic relationships of IabHLHs and well-characterized AtbHLHs were analyzed. The topology of the phylogenetic tree indicated that the *IabHLH* genes could be clustered into 23 subfamilies, including Ia, Ib(1), Ib(2), II, III(a+c), III(d+e), IIIb, IIIf, IVa, IVb, IVc, IVd, Va, Vb, VII(a+b), VIIIa, VIIIb, VIIIc(1), VIIIc(2), IX, XI, XII, and orphans. The IabHLHs were not clustered together with AtbHLHs in the subfamilies X, XIII, XIV, and XV. In addition, 3 IabHLH orphans were not gathered in the same node (Figure 3A). The subfamily Ib(2) was the largest, with 27 *IabHLH* genes, followed by subfamily XII, with 22 *IabHLH* genes, while the subfamily II and VIIIc(1) were the smallest, with only one *IabHLH* gene (Figure 3A).



**Figure 2.** Analysis of *labHLH* genes at the chromosome level. (**A**) Chromosomal location of the *labHLH* genes. The gene pairs linked by red lines were the 44 *labHLH* collinear gene pairs located on chromosomes; the outer circle represents 15 chromosomes of *I. aquatica*, while the inner circle indicates the gene density of each chromosome. (**B**–**D**) Syntenic relationships between *labHLH* genes and the other plant species' genome. The gray lines among chromosomes indicate all homologous gene pairs at the whole-genome level, while the red lines among chromosomes show homologous gene pairs related to the *labHLH* genes. More information concerning the homologous gene pairs was shown in Table S3.

#### 2.4. The Sequence Features of the IabHLHs

To investigate the sequence characteristics of the IabHLHs, the multiple sequence alignments of 157 IabHLH protein sequences were carried out. The results showed that the IabHLH proteins contained the typical conserved structure of basic-helix-loop-helix (Figure 4A,B). A total of 16 amino acids had values of frequency greater than 50%. Of these, the highly conserved amino acids with a frequency greater than 85% were identified within the basic region (including Glu-5, Arg-8, and Arg-9), helix 1 region (including Leu-19 and Pro-24), and helix 2 region (Leu-51).

## 2.5. Conserved Motifs, and Exon-Intron Structures of IabHLH Genes

To further understand the functional and structural features of the IabHLH proteins, the MEME program was performed to analyze the conserved motifs within the IabHLHs. Ten conserved motifs were identified, of which motif 1 and motif 2 were present in all IabHLHs. The IabHLHs within the same group had similar motif distributions, such as the motifs of the IabHLHs in the subfamily Va, IVc, and Ib(1), indicating their similar functions. Motif 5 was only found in the IabHLHs of the subfamily Ib(2), which may impart a specific function to these IabHLHs (Figure 5A,B). The analysis of exon-intron structures within 157 *IabHLH* genes showed that the number of introns ranged from 0 to 11 (Figures 5C and S1, Table S1). The *IabHLH144* in the subfamilies III(d+e) and VIIIb were not disrupted by an intron. Most *IabHLH* genes from the same subfamily had similar exon-intron structures.



**Figure 3.** Phylogenetic relationships of IabHLHs and AtbHLHs (**A**) The phylogenetic tree contains 157 IabHLH proteins and 148 AtbHLH proteins. The hollow green circles represent IabHLHs, while the solid red pentagons represent AtbHLHs. The gray, yellow, and green squares represent the bootstrap values varied from 0 to 49, 50 to 79, and 80 to 100, respectively. (**B**) The number of IabHLHs and AtbHLHs in each subfamily.



**Figure 4.** The sequence characteristics of IabHLHs. **(A)** Three-dimensional structure of bHLH domain in the IabHLH003. **(B)** Sequence logos of the IabHLH domain. **(C)** The frequency of the most conserved amino acids at the respective position.



**Figure 5.** Phylogenetic tree, conserved motif, and exon-intron structure of IabHLHs. (**A**) The phylogenetic tree containing 157 IabHLH protein sequences; (**B**) distribution of conserved motifs within IabHLH proteins. The rectangles with different colors represent different motifs; (**C**) exon-intron structure of the 157 IabHLH genes. The green rectangles indicate exons, while the orange rectangles indicate untranslated regions (UTRs). The black solid lines among the rectangles indicate introns.

## 2.6. Analysis of Gene Duplication Events of IabHLH Genes

To explore the potential driven force of the *labHLH* gene family expansion, the Dup-Gen\_finder was performed to analyze gene duplication events, including WGD (whole-genome duplication), TD (tandem duplication), PD (proximal duplication), TRD (transposed duplication), and DSD (dispersed duplication). A total of 227 duplicated gene pairs derived from five duplication events were identified, with the maximum number of duplicated gene pairs derived from DSD duplication events (126 duplicated gene pairs), followed by WGD-derived duplicated gene pairs with 61. In contrast, TD and PD duplication events had the fewest number of duplicated gene pairs with only five. Moreover, we identified 30 TRD-derived duplicated gene pairs (Figure 6A and Table S4). These results suggested that the DSD and WGD duplication events played a pivotal role in the *labHLH* gene family expansion.



**Figure 6.** Gene duplication events of *IabHLH* genes. (**A**) The number distribution of *IabHLH* duplicated gene pairs in five duplication events. (**B**–**F**) The Ka, Ks, and Ka/Ks value distribution of the duplicated gene pairs in five gene duplication events. The DSD, WGD, TRD, PD, and TD events were from (**B**–**F**), respectively. More information concerning the duplicated genes in five gene duplication events was shown in Table S4.

To investigate the selection pressure on the *labHLH* genes, the values of the nonsynonymous substitution rate (Ka), the synonymous substitution rate (Ks), and Ka/Ks of the duplicated gene pairs derived from five replication events were calculated. As shown in Figure 6B–F, the values of Ka/Ks of all duplicated gene pairs were lower than 1, indicating these genes were subjected to the purifying selection.

## 2.7. Expression Levels of IabHLHs in Two Varieties

The purple-stemmed *I. aquatica*, which is abundant in anthocyanins, exhibits distinct phenotypic differences relative to the common green-stemmed *I. aquatica* (Figure 7A). To explore the role of the *IabHLH* genes in anthocyanin accumulation in the stems of *I. aquatica*, we analyzed the RNA-seq datasets from the purple-stemmed and green-stemmed varieties. Herein, the expression levels of 157 *IabHLH* genes in two varieties were quantified by

calculating gene FPKM values according to the RNA-seq datasets. At this end, the IabHLHs with expression levels that met the criteria ( $|\log_2 FC| > 1$ , FDR < 0.05, and *p*-value < 0.05) were considered differentially expressed genes (DEGs). A total of 13 DEGs were identified between the two varieties, including six upregulated DEGs (priority of the fold change: *IabHLH027* > *IabHLH130* > *IabHLH084* > *IabHLH022* > *IabHLH112* > *IabHLH085*) and seven downregulated DEGs (priority of the fold change: *IabHLH037 > IabHLH114 > IabHLH142* > *IabHLH057* > *IabHLH030* > *IabHLH043* > *IabHLH116*) in the purple-stemmed *I. aquatica* compared to the green-stemmed *I. aquatica* (Table S5). Of these, the log-transformed fold change of *IabHLH027* belonging to subfamily XII between the two varieties was the highest, with 11.3, suggesting its potential positive role in regulating anthocyanin biosynthesis in I. aquatica. Contrarily, the *labHLH037* belonging to subfamily Ia was significantly downregulated in purple-stemmed *I. aquatica*. The expression trends of 13 DEGs determined by RNA-seq were validated using qRT-PCR. As shown in Figure 7B–G, the expression trends of all upregulated DEGs were highly consistent with those detected by qRT-PCR. However, three downregulated DEGs including *labHLH142* (Figure 7J), *labHLH057* (Figure 7K), and IabHLH043 (Figure 7M) determined by RNA-seq had opposite expression trends of those detected by qRT-PCR.



**Figure 7.** Expression analysis of the 13 DEGs. (**A**) The phenotype of two varieties. The G\_S represents green-stemmed *I. aquatica*, while the P\_S represents purple-stemmed *I. aquatica*. Scale bars are 2 cm. (**B–N**) The qRT-PCR validation of the expression trends of 13 DEGs was determined by RNA-seq. The left Y-axis and histogram indicated the relative expression levels of the *IabHLH* gene quantified by qRT-PCR. The right Y-axis and histogram indicate the FPKM values of the *IabHLH* gene obtained by RNA-seq datasets. Statistical analysis was performed using Student's *t*-test. Data are shown as mean  $\pm$  SD of three biological replicates. The \*, \*\*, and \*\*\* indicate *p* < 0.05, *p* < 0.01, and *p* < 0.001, respectively.

## 2.8. Analysis of Cis-Acting Elements of DEGs

To gain more information concerning the potential regulatory functions of 13 DEGs, the cis-acting elements within their promoter regions were analyzed by using the PlantCARE program. Except for the common cis-acting elements (TATA-box and CAAT-box) and ones of unknown function, a total of 338 cis-acting elements were detected in 13 DEGs, which were categorized as four classes, including stress-responsive elements, phytohormoneresponsive elements, and plant growth and development-responsive elements (Table S6). Of these, the light-responsive elements were the largest, with 20 types including 172 elements, followed by the phytohormone-responsive elements, with 10 types including 104 elements, while the plant growth and development-responsive elements were the least, with only two types including 12 elements. The stress responsive-elements were 50, which were classified into five types. The number distribution of each cis-acting element differs in different DEGs. The light-responsive elements and the phytohormone-responsive elements were detected in all DEGs. Except for the *IabHLH114*, all DEGs had stress-responsive elements (Figure 8A), indicating their importance in plant stress response. The *labHLH027* with the highest expression fold change between the two varieties was rich in light-responsive elements and phytohormone-responsive elements, suggesting a critical role in regulating light and phytohormone signaling pathways in *I. aquatica*.



**Figure 8.** Cis-acting elements within the promoter region of the 13 DEGs. (**A**) Heatmap of the cisacting elements. The stacked bars on the right side indicate the number distribution of different cisacting elements in the 13 DEGs; (**B**) percentage of different cis-acting elements in the respective class.

In addition, the number of elements that belonged to the respective class varied greatly. The top five elements of the light-responsive elements were G-Box, Box 4, GT1-motif, TCT-motif, and AE-box. The AREs were the most abundant elements of stress-responsive elements, accounting for 42%. Among the phytohormone-responsive elements, the ABREs

were the most numerous elements, accounting for 46%. Only one circadian-related element was detected in the plant growth and development-responsive elements (Figure 8B).

#### 3. Discussion

Basic-helix-loop-helix (bHLH) proteins are widely found in eukaryotic organisms, which compose one of the largest transcription factor families in plants [38]. bHLH proteins play a vital role in regulating plant biological processes, including plant growth and development, metabolic activities, and various stress responses [39,40]. The increasing release of numerous plant genomic databases facilitates genome-wide identification of transcription factors. To date, large amounts of bHLH transcription factors have been identified in diverse plant species at the whole genome level, including model plants such as Arabidopsis thaliana [7] and Nicotiana tabacum [41], grain crops such as Oryza sativa [42] and Zea mays [6]; vegetable crops such as Brassica rapa [43] and Solanum lycopersicum [44], and fruit trees such as *Vitis vinifera* [45] and *Malus×domestica* [46]. Intriguingly, the number of *bHLH* genes varied among different plants, while it was not proportional to their genome sizes. Moreover, studies showed that the number of *bHLH* genes in higher plants is higher than that in lower plants [5,47], suggesting a pivotal role of *bHLHs* in higher plants. In this study, a total of 157 bHLH genes were detected in the *I. aquatica* genome based on its high-quality genome data [33]. Most IabHLH proteins were acidic and thermostable. All IabHLH proteins were hydrophilic, a principal property for transcription factors to perform their biological functions. Subcellular localization prediction showed that most IabHLH proteins were localized in the nucleus, while some were localized in the chloroplast, endomembrane system, and extracellular space, inferring that IabHLH proteins might play a distinct role in different plant organelles. These results laid a foundation for further exploring the biological function of *IabHLH* genes.

Studies have shown that the *bHLH* gene family was subdivided into a distinct number of subgroups in plants, probably due to the lack of definite classification criteria for lowconserved sequences outside the bHLH domain. Yang et al. identified 175 bHLH genes in *Malus*  $\times$  *domestica* and classified them into 23 subgroups [46]. Sun et al. subdivided 159 bHLH genes of Solanum lycopersicum into 21 clades [44]. A total of 115 bHLH genes in Vitis davidii were gathered into 25 branches [8], while 94 bHLH genes in Vitis vinifera were clustered into 15 subclasses [45]. In the present study, phylogenetic analysis of bHLH sequences from I. aquatica and A. thaliana indicated that the 157 IabHLH genes could be subdivided into 23 subgroups, with three orphan genes. Compared to the AtbHLH gene family, although the members of the *labHLH* gene family were lost in X, XIII, XIV, and XV subfamilies, the *IabHLH* genes in 11 subgroups (including Ia, Ib (1), Ib (2), III (a+c), IIIb, IVa, IVc, IVd, Va, Vb, IX, XII) have expanded, suggesting their key roles in the biological activities of I. aquatica. Moreover, three IabHLH proteins (including IabHLH014, IabHLH089, and IabHLH138) were found in the IIIf subfamily. Some reports showed that the *bHLHs* in O. sativa belonging to the IIIf subfamily were associated with the regulation of anthocyanin biosynthesis [48,49]. However, the expression levels of these three *labHLH* genes have no significant difference between purple-stemmed I. aquatica and green-stemmed I. aquatica, probably other *labHLH* genes participate in the regulation of purple color formation in the I. aquatica stems.

Similar to the previous reports, two adjacent highly conserved motifs and highly conserved gene structural patterns within each subclass were identified [5,50]. Studies have shown that Glu-5 and Arg-8/Arg-9 in the basic region of the bHLH structural domain play an important role in DNA binding, while Leu-19 and Leu-51 in the HLH region may be required for dimerization [1,51]. In the present study, the highly conserved amino acids were identified within the basic region (including Glu-5, Arg-8, and Arg-9), helix 1 region (Leu-19), and helix 2 region (Leu-51) (Figure 4).

Gene duplication events play an indispensable driving role in gene family expansion, which includes WGD, DSD, PD, TRD, and TD [52]. The expansions of the F-box gene family in *Gossypium hirsutum* were mainly impacted by the WGD and TD [53], while

the WGD, TRD, and DSD play a dominant role in the expansions of the *P-type ATPase* gene family in *Pyrus bretschneideri* [54]. Similar to previous studies on *BAHD* and *MYB* gene families [52,55], the DSD and WGD made significant contributions to the expansions of the *IabHLH* gene family. The selection pressure could be estimated by calculating the Ka/Ks value. The Ka/Ks value of one, greater than one, and less than one indicated neutral evolution, positive selection (Darwinian selection), and negative selection (purifying selection), respectively [56,57]. The Ka/Ks values of all duplicated gene pairs in five gene duplication events were less than 1, indicating that *IabHLH* genes were subjected to the purifying selection. In addition, the syntenic relationship analysis between *I. aquatica* and other plant species verified closer evolutionary distance between similar plant taxa, suggesting the reliability of the analysis approach.

As transcription factors, the bHLH proteins directly or indirectly regulate the expression level of structural genes and thereby modulating anthocyanin biosynthesis. In grape berries, the expression level of *VdbHLH037* was highly correlated with anthocyanin accumulation in different varieties. Transgenic expression of VdbHLH037 into Arabidopsis resulted in significant upregulation of structural genes involved in anthocyanin biosynthesis and high accumulation of anthocyanins. In strawberry fruits, FvbHLH9 positively regulates light-dependent anthocyanin biosynthesis by forming a heterodimer with HY [58], while in rapeseed, BnbHLH92a impinges on the biosynthesis of anthocyanins by interacting with the BnTTG1 [59]. In the present study, we found that the expression level of 13 *labHLH* genes exhibited significant differences between purple-stemmed *I. aquatica* and green-stemmed *I.* aquatica. Of these, six DEGs were dramatically upregulated in purple-stemmed I. aquatica compared to green-stemmed *I. aquatica*. In particular, the *IabHLH027* belonging to the subgroup XII exhibited approximately 8-fold higher in purple-stemmed I. aquatica than that in green-stemmed *I. aquatica* (Figure 7B). The bHLHs of subgroup XII were correlated with positive regulation of the brassinosteroid (BR) signaling pathway [60]. Previous reports showed that BR together with other phytohormone signaling led to the enhancement of anthocyanin accumulation in plants and this process was potentially regulated by ternary MYB-bHLH-WD transcriptional complexes [61,62]. RNA-seq is widely recognized as a highly reliable and accurate method for quantifying gene expression levels, and therefore, additional qRT-PCR validation on RNA-seq results may not always be necessary when using the same samples as those used in RNA-seq. However, given the limited availability of RNA-seq samples and the potential for depletion, qRT-PCR validation of the gene of interest would be valuable when analyzing additional samples grown under the same conditions [63]. In this study, we conducted qRT-PCR analysis on additional samples and observed discrepancies in the expression trends of three downregulated differentially expressed genes (DEGs) between RNA-seq and qRT-PCR. These inconsistencies may be attributed to methodological biases that can affect the accuracy of the results obtained from each method [64–66]. The expression profiles of genes are closely related to their cis-regulatory elements. For instance, some bHLH genes in Hordeum vulgare containing several drought induction-associated cis-elements within the promoter regions significantly responded to drought stress stimulation [9]. The Brassica napus bHLH genes that possess abundant phytohormone-responsive elements were positively regulated upon exposure to more than one phytohormone, such as auxin (IAA), gibberellin (GA3), cytokinin (6-BA), abscisic acid (ABA), and ethylene (ACC) [5]. In the present study, we discovered that all DEGs had plentiful light-responsive elements and phytohormone-responsive elements, suggesting their essential function in *I. aquatica*'s response to light and phytohormone signaling. Taken together, these findings provide us valuable clues for further exploring the molecular function of *IabHLH* in regulating purple color formation in *I. aquatica* stems.

#### 4. Materials and Methods

## 4.1. Identification of bHLH Genes in the I. aquatica Genome

The whole genome database of *I. aquatica* (BioProject: PRJCA002216) was used as the reference genome. The full-length sequences of AtbHLH proteins were retrieved from

the TAIR (https://www.arabidopsis.org/. accessed on 1 August 2022). The BLASTP was performed to search against the *I. aquatica* proteome using AtbHLHs as queries. The HMM (Hidden Markov Model) profile of the bHLH domain (PF00010) was downloaded from the Pfam database (http://pfam-legacy.xfam.org/family/PF00010. accessed on 1 August 2022), which was used to retrieve putative *bHLH* genes from the *I. aquatica* genome. The results of the two searches were integrated to obtain candidate *IabHLH* genes. The online program CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. accessed on 8 August 2022), SMART (http://smart.embl-heidelberg.de/. accessed on 7 August 2022), and Pfam (http://pfam-legacy.xfam.org/search#tabview=tab1. accessed on 7 August 2022) were used to verify the structure of candidate *IabHLH* genes, while the redundant sequences and the sequences without bHLH structural domains were manually deleted.

## 4.2. Analysis of Physicochemical Parameters of IabHLH Proteins

The physicochemical characteristics of the IabHLH proteins were analyzed using the ExPASy program (http://web.expasy.org/protparam/. accessed on 22 November 2022), including the number of amino acids, molecular weight (kDa), theoretical isoelectric point (pI), instability index (Ii), aliphatic index (Ai), and grand average of hydropathicity index (GRAVY). Subcellular localization prediction of IabHLHs was carried out using BUSCA (http://busca.biocomp.unibo.it/. accessed on 13 January 2023).

#### 4.3. Analysis of Phylogenetic Relationships of IabHLHs and AtbHLHs

The AtbHLH protein sequences without conserved bHLH domain were excluded [43,60,67,68]. The ClustalW2.0 program with default parameters was used to perform multiple sequence alignment of IabHLHs and AtbHLHs, and then the phylogenetic tree was constructed using the neighbor-joining (NJ) method within MEGA-X (bootstrap value: 1000) [69]. The phylogenetic tree was visualized using the online tool Evolview-v2 [70] (https://evolgenius.info/. accessed on 13 January 2023). The TBtools (version 1.108) software was used to visualize the sequence logo of the IabHLH domain.

The SWISS-MODEL (https://www.swissmodel.expasy.org/. accessed on 14 January 2023) homology modeling was performed to predict protein tertiary structures. The PyMOL (version 4.6) was used to visualize the protein tertiary structural models [71].

#### 4.4. Analysis of Conserved Motif, Gene Structure, and Cis-Acting Elements

The conserved motif within IabHLH proteins was analyzed using the online program MEME (https://meme-suite.org/meme/tools/meme. accessed on 10 December 2022). The GFF annotation file of *I. aquatica* was used to obtain the intron-exon distributions of the *IabHLHs*.

The promoter sequences of the *IabHLH* genes (the upstream 2000 bp sequences from CDS of *IabHLH* genes) were extracted using TBtools based on the full-length DNA sequences of the *I. aquatica* genomes. The PlantCare online database (http://bioinformatics.psb.ugent.be/webtools/PlantCare/html/, accessed on 1 January 2023) was used to predict the cis-acting elements within promoter regions of *IabHLHs*. The results were visualized using TBtools.

#### 4.5. Analysis of Chromosomal Distribution, Collinearity, and Gene Duplication Events

The chromosomal distribution of *IabHLHs* and the gene density within the chromosome of *I. aquatica* were visualized using TBtools. The collinearity relationships of *I. aquatica*, *A. thaliana*, *Ipomoea batatas*, and *Oryza sativa* were analyzed using Multiple Collinearity Scan Toolkit (MCScanX), The results were visualized using the "Advanced Circos" function within TBtools [72].

Gene duplication events of *labHLHs* were analyzed using the dupGen\_finder pipeline, which includes WGD (whole-genome duplication), TD (tandem duplication), PD (proximal duplication), TRD (transposed duplication), and DSD (dispersed duplication). The MYN method within KaKs\_Calculator 2.0 (https://GitHub.com/qiao-xin/scripts\_for\_gb. accessed on 23 December 2022) was used to calculate the value of Ka, Ks, and Ka/Ks.

## 4.6. Plant Materials and Sampling

The purple-stemmed *I. aquatica* and green-stemmed *I. aquatica* were cultivated by cutting in Hoagland nutrient solution at 28 °C with a light cycle of 16 h (light)/8 h (dark). The uniformly colored stems of two varieties were harvested after four weeks of culture and immediately frozen in liquid nitrogen for a moment, which was then stored at -80 °C until use.

## 4.7. Transcriptome Analysis of IabHLH Genes

The transcript abundance of IabHLH genes was quantified by calculating the value of FPKM (fragments per kilobase per million mapped reads) obtained from the RNA-seq data (BioProject: PRJNA814206). The genes with expression levels that met the criteria ( $|log_2FC| > 1$ , FDR < 0.05, and *p*-value < 0.05) were considered differentially expressed genes (DEGs).

# 4.8. qRT-PCR Analysis

Total RNA was extracted from the stems of two varieties using the RNAprep Pure extraction kit (Tiangen, Beijing, China) according to the manufacturer's instructions. The RNA quality was assessed using the microspectrophotometer K5600C (KAIAO, Beijing, China) and agarose gel electrophoresis. The cDNA was synthesized using the FastKing cDNA kit (Tiangen, Beijing, China). The gene-specific primers were designed using Primer Premier 6.0 (Table S7). The reaction mixture of qRT-PCR includes 10  $\mu$ L 2xSYBR green mix (ChamQ Universal SYBR qPCR Master Mix), 1  $\mu$ L cDNA, 1  $\mu$ L primers (0.5  $\mu$ L forward primer and 0.5  $\mu$ L reverse primer), and 8  $\mu$ L PCR-grade water. The qRT-PCR was conducted on the QuantStudio Real-Time Fluorescence PCR System (ThermoFisher, MA, USA). The conditions of two-step qRT-PCR are as follows: 95 °C 3 min followed by 40 cycles of 95 °C 10 s and 55 °C 30 s. The *GAPDH* gene of *I. aquatica* was used as the internal reference gene. The relative expression of the target gene was calculated using the 2<sup>- $\Delta\Delta$ Ct</sup> method. Data were plotted using GraphPad Prism 8.

## 5. Conclusions

In this study, a total of 157 *bHLH* genes were identified in *I. aquatica* at the whole genome level, which were subdivided into 23 subgroups. The *bHLHs* within the identical subgroup had similar motifs and gene structures. The DSD and WGD duplication events made significant contributions to the expansion of the *labHLH* gene family. All duplicated gene pairs were subjected to the purifying selection. A total of 13 DEGs were identified between the purple-stemmed and green-stemmed varieties. The expression profiling revealed that the *labHLH027* belonging to subgroup XII was dramatically upregulated in the purple-stemmed *I. aquatica* relative to the green-stemmed *I. aquatica*, suggesting a potential positive role in regulating anthocyanin accumulation in the *I. aquatica* stems. Our study provides important gene candidates for further exploring the regulatory network of anthocyanin biosynthesis of *I. aquatica*.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24065652/s1.

**Author Contributions:** Conceptualization, W.B.; methodology, W.B. and Z.L.; software, Z.L. and X.F.; validation, Z.L., X.F. and H.X.; formal analysis, Z.L., X.F., H.X., Y.Z., Z.S. and G.Z.; investigation, Z.L., X.F., H.X., Y.Z., Z.S. and G.Z.; writing—original draft preparation, Z.L. and X.F.; writing—review and editing, W.B.; funding acquisition, W.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Hainan Provincial Natural Science Foundation of China (2019RC059 and 320MS008) and the Initial Funds for the High-level Talents of Hainan University (KYQD(ZR)1935).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Materials.

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

### References

- Carretero-Paulet, L.; Galstyan, A.; Roig-Villanova, I.; Martinez-Garcia, J.F.; Bilbao-Castro, J.R.; Robertson, D.L. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. *Plant Physiol.* 2010, 153, 1398–1412. [CrossRef] [PubMed]
- Zhang, X.Y.; Qiu, J.Y.; Hui, Q.L.; Xu, Y.Y.; He, Y.Z.; Peng, L.Z.; Fu, X.Z. Systematic analysis of the basic/helix-loop-helix (bHLH) transcription factor family in pummelo (Citrus grandis) and identification of the key members involved in the response to iron deficiency. *BMC Genom.* 2020, 21, 233. [CrossRef]
- Hao, Y.; Zong, X.; Ren, P.; Qian, Y.; Fu, A. Basic Helix-Loop-Helix (bHLH) Transcription Factors Regulate a Wide Range of Functions in Arabidopsis. *Int. J. Mol. Sci.* 2021, 22, 7152. [CrossRef] [PubMed]
- An, F.; Xiao, X.; Chen, T.; Xue, J.; Luo, X.; Ou, W.; Li, K.; Cai, J.; Chen, S. Systematic Analysis of bHLH Transcription Factors in Cassava Uncovers Their Roles in Postharvest Physiological Deterioration and Cyanogenic Glycosides Biosynthesis. *Front. Plant Sci.* 2022, *13*, 901128. [CrossRef] [PubMed]
- Ke, Y.Z.; Wu, Y.W.; Zhou, H.J.; Chen, P.; Wang, M.M.; Liu, M.M.; Li, P.F.; Yang, J.; Li, J.N.; Du, H. Genome-wide survey of the bHLH super gene family in Brassica napus. *BMC Plant Biol.* 2020, 20, 115. [CrossRef] [PubMed]
- 6. Zhang, T.; Lv, W.; Zhang, H.; Ma, L.; Li, P.; Ge, L.; Li, G. Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize. *BMC Plant Biol.* **2018**, *18*, 235. [CrossRef]
- Bailey, P.C.; Martin, C.; Toledo-Ortiz, G.; Quail, P.H.; Huq, E.; Heim, M.A.; Jakoby, M.; Werber, M.; Weisshaar, B. Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana. *Plant Cell* 2003, *15*, 2497–2502. [CrossRef]
- Li, M.; Sun, L.; Gu, H.; Cheng, D.; Guo, X.; Chen, R.; Wu, Z.; Jiang, J.; Fan, X.; Chen, J. Genome-wide characterization and analysis of bHLH transcription factors related to anthocyanin biosynthesis in spine grapes (Vitis davidii). *Sci. Rep.* 2021, *11*, 6863. [CrossRef]
- 9. Ke, Q.; Tao, W.; Li, T.; Pan, W.; Chen, X.; Wu, X.; Nie, X.; Cui, L. Genome-wide Identification, Evolution and Expression Analysis of Basic Helix-loop-helix (bHLH) Gene Family in Barley (*Hordeum vulgare* L.). *Curr. Genom.* **2020**, *21*, 621–644. [CrossRef]
- Guo, J.; Sun, B.; He, H.; Zhang, Y.; Tian, H.; Wang, B. Current Understanding of bHLH Transcription Factors in Plant Abiotic Stress Tolerance. Int. J. Mol. Sci. 2021, 22, 4921. [CrossRef]
- Sun, X.; Wang, Y.; Sui, N. Transcriptional regulation of bHLH during plant response to stress. *Biochem. Biophys. Res. Commun.* 2018, 503, 397–401. [CrossRef] [PubMed]
- Ito, S.; Song, Y.H.; Josephson-Day, A.R.; Miller, R.J.; Breton, G.; Olmstead, R.G.; Imaizumi, T. FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 2012, 109, 3582–3587. [CrossRef] [PubMed]
- Zhang, L.Y.; Bai, M.Y.; Wu, J.; Zhu, J.Y.; Wang, H.; Zhang, Z.; Wang, W.; Sun, Y.; Zhao, J.; Sun, X.; et al. Antagonistic HLH/bHLH transcription factors mediate brassinosteroid regulation of cell elongation and plant development in rice and Arabidopsis. *Plant Cell* 2009, 21, 3767–3780. [CrossRef] [PubMed]
- 14. 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]
- Chen, W.; Miao, Y.; Ayyaz, A.; Hannan, F.; Huang, Q.; Ulhassan, Z.; Zhou, Y.; Islam, F.; Hong, Z.; Farooq, M.A.; et al. Purple stem Brassica napus exhibits higher photosynthetic efficiency, antioxidant potential and anthocyanin biosynthesis related genes expression against drought stress. *Front. Plant Sci.* 2022, *13*, 936696. [CrossRef] [PubMed]
- 16. Zhang, Q.; Zhai, J.; Shao, L.; Lin, W.; Peng, C. Accumulation of Anthocyanins: An Adaptation Strategy of Mikania micrantha to Low Temperature in Winter. *Front. Plant Sci.* **2019**, *10*, 1049. [CrossRef] [PubMed]
- 17. 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]
- Liu, Y.; Ma, K.; Qi, Y.; Lv, G.; Ren, X.; Liu, Z.; Ma, F. Transcriptional Regulation of Anthocyanin Synthesis by MYB-bHLH-WDR Complexes in Kiwifruit (*Actinidia chinensis*). J. Agric. Food Chem. 2021, 69, 3677–3691. [CrossRef]
- 19. Jaakola, L. New insights into the regulation of anthocyanin biosynthesis in fruits. Trends Plant Sci. 2013, 18, 477–483. [CrossRef]
- Wan, L.; Li, B.; Lei, Y.; Yan, L.; Huai, D.; Kang, Y.; Jiang, H.; Tan, J.; Liao, B. Transcriptomic profiling reveals pigment regulation during peanut testa development. *Plant Physiol. Biochem.* 2018, 125, 116–125. [CrossRef]

- Zhang, N.; Qi, Y.; Zhang, H.J.; Wang, X.; Li, H.; Shi, Y.; Guo, Y.D. Genistein: A Novel Anthocyanin Synthesis Promoter that Directly Regulates Biosynthetic Genes in Red Cabbage in a Light-Dependent Way. *Front. Plant Sci.* 2016, 7, 1804. [CrossRef] [PubMed]
- 22. Liu, Y.; Tikunov, Y.; Schouten, R.E.; Marcelis, L.F.M.; Visser, R.G.F.; Bovy, A. Anthocyanin Biosynthesis and Degradation Mechanisms in Solanaceous Vegetables: A Review. *Front. Chem.* **2018**, *6*, 52. [CrossRef] [PubMed]
- Gonzalez, A.; Zhao, M.; Leavitt, J.M.; Lloyd, A.M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 2008, 53, 814–827. [CrossRef] [PubMed]
- 24. 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]
- 25. Nesi, N.; Debeaujon, I.; Jond, C.; Pelletier, G.; Caboche, M.; Lepiniec, L. The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *Plant Cell* **2000**, *12*, 1863–1878. [CrossRef]
- 26. Xiang, L.L.; Liu, X.F.; Li, X.; Yin, X.R.; Grierson, D.; Li, F.; Chen, K.S. A Novel bHLH Transcription Factor Involved in Regulating Anthocyanin Biosynthesis in Chrysanthemums (*Chrysanthemum morifolium Ramat.*). *PLoS ONE* **2015**, *10*, e0143892. [CrossRef]
- Xie, X.B.; Li, S.; Zhang, R.F.; Zhao, J.; Chen, Y.C.; Zhao, Q.; Yao, Y.X.; You, C.X.; Zhang, X.S.; Hao, Y.J. The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ.* 2012, 35, 1884–1897. [CrossRef]
- Liu, Y.; Lin-Wang, K.; Espley, R.V.; Wang, L.; Li, Y.; Liu, Z.; Zhou, P.; Zeng, L.; Zhang, X.; Zhang, J.; et al. StMYB44 negatively regulates anthocyanin biosynthesis at high temperatures in tuber flesh of potato. *J. Exp. Bot.* 2019, *70*, 3809–3824. [CrossRef]
- Wang, H.; Wang, X.; Yu, C.; Wang, C.; Jin, Y.; Zhang, H. MYB transcription factor PdMYB118 directly interacts with bHLH transcription factor PdTT8 to regulate wound-induced anthocyanin biosynthesis in poplar. *BMC Plant Biol.* 2020, 20, 173. [CrossRef]
- Hichri, I.; Heppel, S.C.; Pillet, J.; Léon, C.; Czemmel, S.; Delrot, S.; Lauvergeat, V.; Bogs, J. The Basic Helix-Loop-Helix Transcription Factor MYC1 Is Involved in the Regulation of the Flavonoid Biosynthesis Pathway in Grapevine. *Mol. Plant* 2010, *3*, 509–523. [CrossRef]
- Li, H.; Yang, Z.; Zeng, Q.; Wang, S.; Luo, Y.; Huang, Y.; Xin, Y.; He, N. Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hortic. Res.* 2020, 7, 83. [CrossRef] [PubMed]
- Hefny Gad, M.; Tuenter, E.; El-Sawi, N.; Younes, S.; El-Ghadban, E.M.; Demeyer, K.; Pieters, L.; Vander Heyden, Y.; Mangelings, D. Identification of some Bioactive Metabolites in a Fractionated Methanol Extract from Ipomoea aquatica (Aerial Parts) through TLC, HPLC, UPLC-ESI-QTOF-MS and LC-SPE-NMR Fingerprints Analyses. *Phytochem. Anal.* 2018, 29, 5–15. [CrossRef] [PubMed]
- 33. Hao, Y.; Bao, W.; Li, G.; Gagoshidze, Z.; Shu, H.; Yang, Z.; Cheng, S.; Zhu, G.; Wang, Z. The chromosome-based genome provides insights into the evolution in water spinach. *Sci. Hortic.* **2021**, *289*, 110501. [CrossRef]
- Dua, T.K.; Dewanjee, S.; Khanra, R.; Bhattacharya, N.; Bhaskar, B.; Zia-Ul-Haq, M.; De Feo, V. The effects of two common edible herbs, Ipomoea aquatica and Enhydra fluctuans, on cadmium-induced pathophysiology: A focus on oxidative defence and anti-apoptotic mechanism. *J. Transl. Med.* 2015, 13, 245. [CrossRef] [PubMed]
- Lawal, U.; Maulidiani, M.; Ismail, I.S.; Khatib, A.; Abas, F. Discrimination of Ipomoea aquatica cultivars and bioactivity correlations using NMR-based metabolomics approach. *Plant Biosyst.—Int. J. Deal. All Asp. Plant Biol.* 2016, 151, 833–843. [CrossRef]
- 36. Alkiyumi, S.S.; Abdullah, M.A.; Alrashdi, A.S.; Salama, S.M.; Abdelwahab, S.I.; Hadi, A.H. Ipomoea aquatica extract shows protective action against thioacetamide-induced hepatotoxicity. *Molecules* **2012**, *17*, 6146–6155. [CrossRef]
- Yuxin, Z.; Zheng, L.; Tangquan, Z.; Yuanyuan, H.; Shanhan, C.; Zhiwei, W.; Jie, Z.; Wuqiang, M.; Wenlong, B. Genome-wide Identification and Expression Analysis of WRKY Gene Family in Water Spinach (Ipomoea aquatica). *Mol. Plant Breed.* 2023, 21, 55–66. (In Chinese) [CrossRef]
- Xiong, Y.; Liu, T.; Tian, C.; Sun, S.; Li, J.; Chen, M. Transcription factors in rice: A genome-wide comparative analysis between monocots and eudicots. *Plant Mol. Biol.* 2005, 59, 191–203. [CrossRef]
- Samira, R.; Li, B.; Kliebenstein, D.; Li, C.; Davis, E.; Gillikin, J.W.; Long, T.A. The bHLH transcription factor ILR3 modulates multiple stress responses in Arabidopsis. *Plant Mol. Biol.* 2018, 97, 297–309. [CrossRef]
- 40. 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]
- Bai, G.; Yang, D.H.; Chao, P.; Yao, H.; Fei, M.; Zhang, Y.; Chen, X.; Xiao, B.; Li, F.; Wang, Z.Y.; et al. Genome-wide identification and expression analysis of NtbHLH gene family in tobacco (Nicotiana tabacum) and the role of NtbHLH86 in drought adaptation. *Plant Divers.* 2021, 43, 510–522. [CrossRef] [PubMed]
- 42. Wei, K.; Chen, H. Comparative functional genomics analysis of bHLH gene family in rice, maize and wheat. *BMC Plant Biol.* **2018**, 18, 309. [CrossRef] [PubMed]
- Song, X.M.; Huang, Z.N.; Duan, W.K.; Ren, J.; Liu, T.K.; Li, Y.; Hou, X.L. Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Mol. Genet. Genom.* 2014, 289, 77–91. [CrossRef] [PubMed]
- Sun, H.; Fan, H.J.; Ling, H.Q. Genome-wide identification and characterization of the bHLH gene family in tomato. *BMC Genom.* 2015, 16, 9. [CrossRef] [PubMed]

- Wang, P.; Su, L.; Gao, H.; Jiang, X.; Wu, X.; Li, Y.; Zhang, Q.; Wang, Y.; Ren, F. Genome-Wide Characterization of bHLH Genes in Grape and Analysis of their Potential Relevance to Abiotic Stress Tolerance and Secondary Metabolite Biosynthesis. *Front. Plant Sci.* 2018, *9*, 64. [CrossRef]
- 46. Yang, J.; Gao, M.; Huang, L.; Wang, Y.; van Nocker, S.; Wan, R.; Guo, C.; Wang, X.; Gao, H. Identification and expression analysis of the apple (Malus × domestica) basic helix-loop-helix transcription factor family. *Sci. Rep.* **2017**, *7*, 28. [CrossRef]
- Wang, X.J.; Peng, X.Q.; Shu, X.C.; Li, Y.H.; Wang, Z.; Zhuang, W.B. Genome-wide identification and characterization of PdbHLH transcription factors related to anthocyanin biosynthesis in colored-leaf poplar (*Populus deltoids*). BMC Genom. 2022, 23, 244. [CrossRef]
- Spelt, C.; Quattrocchio, F.; Mol, J.N.M.; Koes, R. anthocyanin1 of Petunia Encodes a Basic Helix-Loop-Helix Protein That Directly Activates Transcription of Structural Anthocyanin Genes. *Plant Cell* 2000, 12, 1619–1631. [CrossRef]
- 49. Hu, J.; Reddy, V.S.; Wessler, S.R. The rice R gene family: Two distinct subfamilies containing several miniature inverted-repeat transposable elements. *Plant Mol. Biol.* **2000**, *42*, 667–678. [CrossRef]
- Hong, Y.; Ahmad, N.; Tian, Y.; Liu, J.; Wang, L.; Wang, G.; Liu, X.; Dong, Y.; Wang, F.; Liu, W.; et al. Genome-Wide Identification, Expression Analysis, and Subcellular Localization of Carthamus tinctorius bHLH Transcription Factors. *Int. J. Mol. Sci.* 2019, 20, 3044. [CrossRef]
- 51. Atchley, W.R.; Fitch, W.M. A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci.* USA **1997**, *94*, 5172–5176. [CrossRef] [PubMed]
- Liu, Z.; Zhang, Y.; Altaf, M.A.; Hao, Y.; Zhou, G.; Li, X.; Zhu, J.; Ma, W.; Wang, Z.; Bao, W. Genome-wide identification of myeloblastosis gene family and its response to cadmium stress in Ipomoea aquatica. *Front. Plant Sci.* 2022, 13, 979988. [CrossRef] [PubMed]
- Zhang, S.; Tian, Z.; Li, H.; Guo, Y.; Zhang, Y.; Roberts, J.A.; Zhang, X.; Miao, Y. Genome-wide analysis and characterization of F-box gene family in *Gossypium hirsutum* L. *BMC Genom.* 2019, 20, 993. [CrossRef] [PubMed]
- 54. Zhang, Y.; Li, Q.; Xu, L.; Qiao, X.; Liu, C.; Zhang, S. Comparative analysis of the P-type ATPase gene family in seven Rosaceae species and an expression analysis in pear (*Pyrus bretschneideri* Rehd.). *Genomics* **2020**, *112*, 2550–2563. [CrossRef]
- Liu, C.; Qiao, X.; Li, Q.; Zeng, W.; Wei, S.; Wang, X.; Chen, Y.; Wu, X.; Wu, J.; Yin, H.; et al. Genome-wide comparative analysis of the BAHD superfamily in seven Rosaceae species and expression analysis in pear (*Pyrus bretschneideri*). *BMC Plant Biol.* 2020, 20, 14. [CrossRef]
- 56. Dong, H.; Chen, Q.; Dai, Y.; Hu, W.; Zhang, S.; Huang, X. Genome-wide identification of PbrbHLH family genes, and expression analysis in response to drought and cold stresses in pear (Pyrus bretschneideri). *BMC Plant Biol.* **2021**, *21*, 86. [CrossRef]
- 57. Starr, T.K.; Jameson, S.C.; Hogquist, K.A. Positive and negative selection of T cells. *Annu. Rev. Immunol.* 2003, 21, 139–176. [CrossRef]
- 58. Li, Y.; Xu, P.; Chen, G.; Wu, J.; Liu, Z.; Lian, H. FvbHLH9 Functions as a Positive Regulator of Anthocyanin Biosynthesis by Forming a HY5–bHLH9 Transcription Complex in Strawberry Fruits. *Plant Cell Physiol.* **2020**, *61*, 826–837. [CrossRef]
- 59. Hu, R.; Zhu, M.; Chen, S.; Li, C.; Zhang, Q.; Gao, L.; Liu, X.; Shen, S.; Fu, F.; Xu, X.; et al. BnbHLH92a negatively regulates anthocyanin and proanthocyanidin biosynthesis in *Brassica napus*. *Crop J*. **2022**. [CrossRef]
- 60. Pires, N.; Dolan, L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol. Biol. Evol.* **2010**, 27, 862–874. [CrossRef]
- 61. Peng, Z.; Han, C.; Yuan, L.; Zhang, K.; Huang, H.; Ren, C. Brassinosteroid enhances jasmonate-induced anthocyanin accumulation in Arabidopsis seedlings. *J. Integr. Plant Biol.* **2011**, *53*, 632–640. [CrossRef] [PubMed]
- 62. Yuan, L.B.; Peng, Z.H.; Zhi, T.T.; Zho, Z.; Liu, Y.; Zhu, Q.; Xiong, X.Y.; Ren, C.M. Brassinosteroid enhances cytokinin-induced anthocyanin biosynthesis in Arabidopsis seedlings. *Biol. Plant.* **2015**, *59*, 99–105. [CrossRef]
- 63. Coenye, T. Do results obtained with RNA-sequencing require independent verification? *Biofilm* **2021**, *3*, 100043. [CrossRef] [PubMed]
- 64. Everaert, C.; Luypaert, M.; Maag, J.L.V.; Cheng, Q.X.; Dinger, M.E.; Hellemans, J.; Mestdagh, P. Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci. Rep.* **2017**, *7*, 1559. [CrossRef]
- 65. Love, M.I.; Hogenesch, J.B.; Irizarry, R.A. Modeling of RNA-seq fragment sequence bias reduces systematic errors in transcript abundance estimation. *Nat. Biotechnol.* **2016**, *34*, 1287–1291. [CrossRef]
- 66. Minshall, N.; Git, A. Enzyme- and gene-specific biases in reverse transcription of RNA raise concerns for evaluating gene expression. *Sci. Rep.* **2020**, *10*, 8151. [CrossRef]
- 67. Toledo-Ortiz, G.; Huq, E.; Quail, P.H. The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell* **2003**, *15*, 1749–1770. [CrossRef]
- 68. Heim, M.A.; Jakoby, M.; Werber, M.; Martin, C.; Weisshaar, B.; Bailey, P.C. The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* **2003**, *20*, 735–747. [CrossRef]
- 69. 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]
- He, Z.; Zhang, H.; Gao, S.; Lercher, M.J.; Chen, W.H.; Hu, S. Evolview v2: An online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* 2016, 44, W236–W241. [CrossRef]

- Seeliger, D.; de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J. Comput. Aided Mol. Des. 2010, 24, 417–422. [CrossRef] [PubMed]
- 72. Chen, C.; Wu, Y.; Xia, R. A painless way to customize Circos plot: From data preparation to visualization using TBtools. *iMeta* **2022**, *1*, e35. [CrossRef]

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