horticulturae

MDPI

Article

Genomic Survey and Expression Analysis of *GLKs* in Watermelon (*Citrullus lanatus*)

Qiangqiang Ding [1,2,†], Li Jia [1,2,†], Xiangting Jiang [1,2], Mingxia Wang [1,2], Yan Wang [1,2], Haikun Jiang [1,2], Feifei Yu [1,2] and Congsheng Yan [1,2,*]

[1] Institute of Horticulture, Anhui Academy of Agricultural Sciences, Hefei 230001, China
[2] Key Laboratory of Genetic Improvement and Ecophysiology of Horticultural Crop, Hefei 230001, China
* Correspondence: congshengyan@126.com
† These authors contributed equally to this work.

Abstract: *Golden2-like* (*GLK*) genes positively regulate chloroplast development, increase crop yields, and improve fruit quality. However, there has been no comprehensive identification and characterization of *GLKs* in watermelon. In this study, a total of 48 *ClGLKs* were identified in the watermelon genome. Based on phylogenetic analysis, they were divided into five groups. *ClGLKs* within the same group showed a similar gene structure and conserved motif compositions. Promoter analysis indicated that cis-elements responsive to light were the most abundant, though cis-elements associated with hormones, stress, and developmental regulation were also identified in *ClGLKs* promoters. Expression analysis indicated significant responses of some *ClGLKs* to drought and CGMMV stress, suggesting that these genes may participate in responses to biotic and abiotic stresses. Phenotypic analyses revealed enhanced chloroplast development and increased thylakoid density and chlorophyll content in the pericarp of a "dark green" watermelon cultivar. *ClGLK8* was identified as the homolog of *GLK1-2*, the genes that promote chloroplast development and chlorophyll biosynthesis in fruits, and showed significantly increased expression in accordance with chloroplast development and chlorophyll accumulation. Our results provide detailed knowledge of the *ClGLKs*, which will enhance efforts to further improve the fruit quality of watermelon.

Keywords: watermelon; GLKs; bioinformatics analysis; expression profiles

Citation: Ding, Q.; Jia, L.; Jiang, X.; Wang, M.; Wang, Y.; Jiang, H.; Yu, F.; Yan, C. Genomic Survey and Expression Analysis of *GLKs* in Watermelon (*Citrullus lanatus*). *Horticulturae* **2023**, *9*, 229. https://doi.org/10.3390/horticulturae9020229

Academic Editor: Jia-Long Yao

Received: 17 January 2023
Revised: 31 January 2023
Accepted: 7 February 2023
Published: 8 February 2023

1. Introduction

The chloroplast is an important organelle that is the basic functional unit of plant photosynthesis and the biological factory of organic synthesis in the plant cell. Chlorophyll housed in the chloroplast can use light energy to convert carbon dioxide and water into organic substances. The synthesized organic substances can be used for plant growth and development, and the photosynthates transported into the fruit can improve fruit quality, by the accumulation of sugar and the influence of crop yields. Aside from chlorophyll (Chl), carotenoids are the other main pigments in the chloroplast. The ratios of these pigments in the chloroplasts of different plant organs such as leaves or fruit result in the various observed colors. Thus, enhancing the chloroplast function and manipulating the transformation of its pigments may have great significance in increasing crop yields [1], improving fruit quality [2], and enhancing plant aesthetics [3].

The chloroplast is a semi-autonomous organelle whose development requires coordination between its own genome and the nuclear genome. In fact, the genome of the chloroplast encodes few genes. Most proteins that function in the chloroplast are encoded and transported from the nuclear genome and play important roles in the development, maintenance, and degradation of the chloroplast [4]. Among these, several transcription factors (TFs) have unique contributions to chloroplast development and pigment metabolism. To date, several TF families, such as Golden2-like (GLK) [5], NAM, ATAF, and CUC (NAC) [6], KNOTTED-like homeodomain (KNOX) [7], and BEL1-like homeodomain

(BELL) [8], have been found to regulate chloroplast biogenesis and chlorophyll metabolism. For example, members of the GLK and KNOX families have been shown to play positive roles in increasing the number and volume of chloroplasts and promoting Chl biosynthesis [5,7], however, the overexpression of *AtNAC046* in *Arabidopsis* led to an early-senescence phenotype and lower chlorophyll content [6].

GLKs belong to the GARP superfamily of MYB TFs, which encode proteins that usually contain a Myb-DNA binding domain (DBD) and a C-terminal GCT box [9,10]. Since *Golden2* was originally identified in maize, homologous genes have subsequently been cloned from many other plant species. To date, *GLKs* have been identified in *Arabidopsis* [5], rice [11], tomato [12], and many others. Most genetic and biochemical studies of *GLKs* have focused on their roles in regulating chloroplast development and functional maintenance. *AtGLK1*, *AtGLK2*, *SlGLK1*, and *SlGLK2* have been shown to play positive roles in increasing the number and volume of chloroplasts, increasing the density of thylakoids, and promoting the biosynthesis of chlorophyll [13,14]. In addition, recent research has found that the photosynthetic capacity of C3 crops can be modified by using the *GLKs* of the C4 plant, significantly increasing the photosynthetic efficiency of rice and increasing the rate of yield up to 70% [15,16]. Overexpression of *SlGLK1* and *SlGLK2* in tomato resulted in increases in the chloroplast number and chlorophyll content in immature fruits and significantly increased the soluble solid, sugar, and carotenoid content in mature fruits [12,13]. These results demonstrate the significant benefits of exploiting GLKs to breed high-yield crops and improve fruit quality.

Watermelon is an important horticultural crop due to its sweet taste and rich source of nutrients. The sweetness is a vital characteristic for evaluating the quality of watermelon fruit and is determined by the total sugar content and ratios of the main sugar forms of glucose, sucrose, and fructose [17]. Photosynthesis is the critical mechanism for sugar accumulation in watermelon. Enhanced chloroplast development and chlorophyll content could increase the photosynthetic capacity, leading to a greater accumulation of sugar in fruit [12,18]. Therefore, *GLKs* could be used in breeding programs to improve the sugar content in watermelon fruit. In this study, *GLKs* were identified and characterized in watermelon. Then, the expression patterns of *GLKs* transcripts were analyzed in the fruit to better understand their roles in chloroplast development and Chl biosynthesis. The results of this study could provide candidate genes to improve the quality of watermelon fruit.

2. Materials and Methods

2.1. Identification and Characterization of GLKs in Watermelon

The watermelon (97103) v2 protein file was downloaded from the Cucurbit Genomics Database (http://cucurbitgenomics.org/, accessed on 7 December 2022). To identify GLKs in watermelon, the GLKs of *Arabidopsis* and tomato were used as query sequences to BLAST against the watermelon protein database. Alignments with E values less than E-5 were considered as candidate GLKs. The candidate proteins were then confirmed to be ClGLKs using SMART (https://smart.embl.de/, accessed on 20 November 2022) and Pfam (http://pfam.xfam.org/, accessed on 20 November 2022) to identify the conserved GLK domains. The detailed characteristics of putative ClGLKs were predicted as follows: the lengths of the amino acids were obtained from the protein data. The molecular weights (MW) and isoelectric points (pI) of the proteins were calculated using the Expasy site (http://web.expasy.org/protparam/, accessed on 21 November 2022). The signal peptides of proteins were predicted using the SignalP 5.0 Server (https://services.healthtech.dtu.dk/service.php?SignalP-5.0, accessed on 21 November 2022). Finally, all proteins were submitted to WoLF PSORT (http://wolfpsort.org/, accessed on 22 November 2022) to predict the subcellular localizations. All of the detailed information about *ClGLKs* is listed in Table S1.

2.2. Phylogenetic Analysis of ClGLKs

A phylogenetic tree of GLKs in watermelon, *Arabidopsis*, and tomato was constructed using MEGA-X with the maximum likelihood method. The phylogenetic tree was visualized using the online software tool EvolView (http://www.evolgenius.info/, accessed on 1 December 2022).

2.3. Gene Structure and Motif analysis

The information about the exon-intron organization of *ClGLKs* was obtained from the genomic database. Protein sequences were submitted to the MEME website (http://meme-suite.org/tools/meme, accessed on 5 December 2022) to identify conserved motifs. Finally, the gene structure and motif composition of *ClGLKs* were visualized using TBtools software [19].

2.4. Promoter Analysis of ClGLKs

For promoter analysis, sequences 2 kb upstream of the translation start site of *ClGLKs* were extracted from the genome and analyzed using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 6 December 2022) to identify putative cis-elements. TBtools was then used to visualize the results of the cis-elements analysis.

2.5. Expression Analysis of Response of ClGLKs to Biotic and Abiotic Stresses

The transcriptome data of uninoculated watermelon fruits and watermelons inoculated with THE cucumber green mottle mosaic virus (CGMMV) obtained from the Cucurbit Genomics Database (http://cucurbitgenomics.org/, accessed on 7 December 2022) (BioProject: PRJNA389184) were used to analyze the expression of *ClGLKs* in response to biotic stress. The transcriptome data of the watermelons treated with PEG6000 applied to roots for six hours (BioProject: PRJNA326331) were used to analyze the expression of *ClGLKs* in response to abiotic stress.

2.6. Plant Materials and Sample Collection

Two watermelon cultivars of "W-22-72" and "W-22-301" were used as the plant materials in this study. Thirty plants per cultivar were grown in a greenhouse located in Gangji Country (Hefei City) with about a 30 °C/18 °C day/night cycle. The skin color of the cultivar "W-21-72" changed from green to dark green as it started to mature, while cultivar "W-21-301" kept a light green pericarp during the entire growth process. The pericarp of the fruit was collected at three stages (before, during, and after the change in pericarp color, which corresponded to about 7, 9, and 16 days after pollination of the fruit, respectively) in "W-22-72", while the "W-22-301" samples were also collected at 7, 9, and 16 days after pollination. Each sample was collected in three biological replicates, and all samples were frozen in liquid nitrogen and immediately stored at −80 °C for further analysis.

2.7. Chl Content and Chloroplast Analysis

The Chl content samples was sent to KeMing Bio Co., Ltd. (Suzhou, China) to measure the Chl content. The fruit pericarp at different stages was soaked in FAA fixative solution, and the fixative solution was vacuum infiltrated into the tissue for 30 min. The treated samples were sent to Service Bio Co. Ltd (Wuhan, China) to analyze the chloroplast structure. The sample treatment included dehydration, resin penetration and embedding, polymerization, ultrathin section with 60–80 nm, and staining. Finally, the chloroplast ultrastructure was observed under transmission electron microscopy (TEM) and images taken with a 7000 times lens.

2.8. Expression Patterns Analysis of ClGLKs in the Pericarp of the Fruit

RNA-Seq was performed on the pericarp samples collected from two watermelon cultivars at three stages to analyze the expression profiles of *ClGLKs*. Genes with fold-changes $\geq$2 and p-values $\leq$ 0.05 were identified as differentially expressed genes (DEGs).

The log_2-transformed RPKM expression values of the genes were used to generate a heatmap using TBtools.

2.9. RNA Extraction, cDNA Synthesis, and RT-qPCR Analysis

Total RNA was extracted from the samples using the Plant Total RNA Isolation Kit (Sangon, Shanghai, China). cDNA was generated using the MightyScript First Strand cDNA Synthesis Master Mix (Sangon, Shanghai). RT-qPCR was performed using the 2X SG Fast qPCR Master Mix (Sangon, Shanghai) according to the manufacturer's protocol. The *actin* gene was used as the reference. The reaction system of RT-qPCR was carried out as follows: 10 uL qPCR Master Mix, 0.5 µL of forward primer and reverse primer, 1 µL of diluted cDNA and 8 uL of ddH_2O. Then, the reaction system was performed on a Bio-Rad CFX96 (Hercules, CA, USA), and the program was used as follows: 95 °C for 60 s, followed by 40 cycles of 95 °C for 10 s, and 60 °C for 30 s. Three biological and three technical replicates were performed and the relative expression of *ClGLKs* was analyzed using the $2^{-\Delta\Delta CT}$ method [20]. The primer sequences used for RT-qPCR are listed in Table S4.

3. Results

3.1. Identification and Characterization of GLKs in Watermelon

By using a bioinformatics approach, a total of 48 GLKs were identified in the watermelon genome. These genes were annotated as *ClGLK1* to *ClGLK48* according to their chromosomal locations (Table S1). Furthermore, the amino acid lengths, MWs, pIs, signal peptides, and subcellular locations were determined to characterize the detailed physical and chemical properties of ClGLKs (Table S1). *ClGLKs* encoded proteins ranging from 138 (ClGLK42) to 710 (ClGLK20) amino acids in length with predicted molecular weights ranging from 16.26 kDa (ClGLK42) to 78.78 kDa (ClGLK20). The pI of ClGLK proteins ranged from 5.19 (ClGLK31) to 9.9 (ClGLK34). No ClGLKs were predicted to have signal peptides, indicating that they are all are likely to be non-secreted proteins. In addition, 44 GLKs were predicted to be localized in the nucleus. ClGLK30, ClGLK32, and ClGLK42 were predicted to be localized in the cytoplasm, with ClGLK37 in the chloroplast.

3.2. Phylogenetic Analysis, Gene structure, and Motif Analysis of ClGLKs

A single phylogenetic tree was built using the ClGLK protein sequences. Forty-eight ClGLKs were clustered into five separate groups based on bootstrap values (Figure 1). Group I harbored the largest number of ClGLKs with 13 members. Twelve ClGLKs were contained in groups II and V. Group III had seven ClGLKs, and group IV contained four members.

Gene structure analysis indicated that the gene structure varied among different groups, but in the same subfamily, most members showed great similarity. The *ClGLKs* in group I contained introns that varied from four to seven, with more than half of the members containing five introns (Figure 1). The introns in group II varied from three to six, with most genes containing five introns. Most genes in group III contained four introns, except for *ClGLK27* and *ClGLK47*, which contained three introns. No introns were identified in the *ClGLKs* of group IV. For group V, the intron number varied from four to twelve, with half of the members containing five introns.

Based on MEME analysis, the motif compositions were conserved within each group, though they differed between groups. Motifs 1 and 2 (Table S2), which corresponded to the myb SHAQKYF domain were shared in most ClGLKs except for ClGLK42 (Figure 1). Aside from these, a unique conserved motif composition was observed in each group. For example, motif 3 was only found in the group I proteins. Group II only contained motifs 1, 2, and 15, though 15 was absent in several members. Motifs 6, 9, and 10 were uniquely distributed in the proteins of group III. Motif 8 was present in the ClGLKs of group IV, and the group V ClGLKs contained motifs 4, 5, and 7, except for ClGLK8. The diverse motif compositions might lead to functional diversity among the different groups of ClGLKs.

The similar conserved motifs within each group indicated that their ClGLKs may have similar functions.

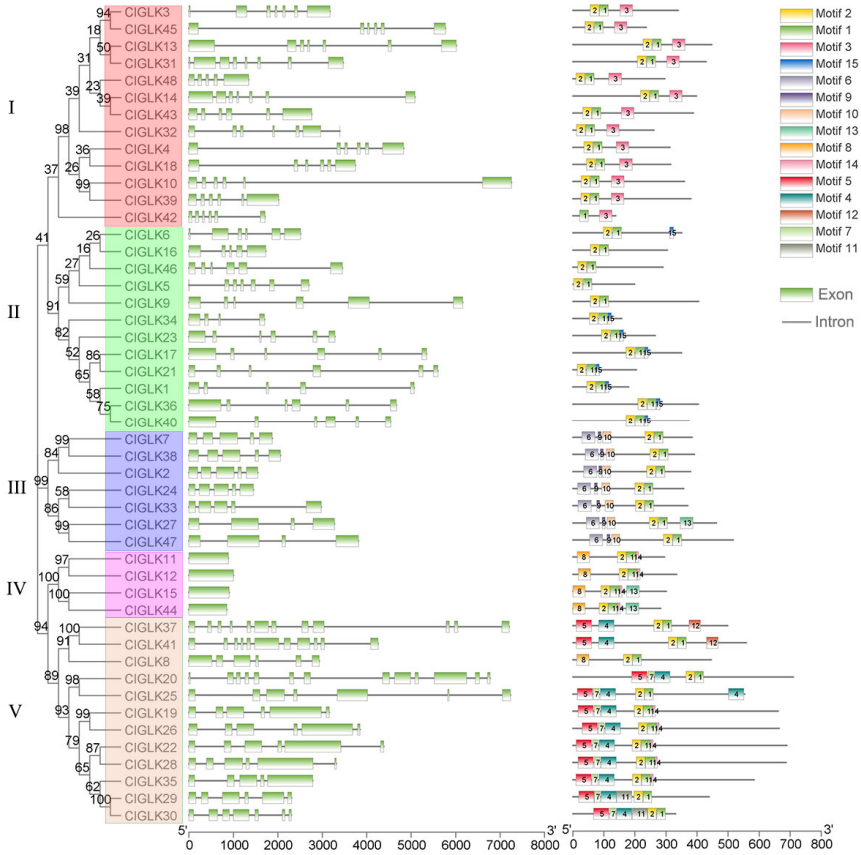

Figure 1. Analysis of the phylogenetic relationships, gene structures, and conserved motifs of GLKs in watermelon.

3.3. Phylogenetic Analysis of GLKs in Watermelon, Arabidopsis, and Tomato

An additional phylogenetic tree was generated using the GLKs of watermelon, *Arabidopsis* [21], and tomato [22] to infer the potential roles of *ClGLKs*. This phylogenetic tree was divided into five groups (Figure 2a). The placement of individual ClGLKs within each group was consistent with the results of the single phylogenetic tree of ClGLKs (Figures 1 and 2a). The potential biological functions of ClGLKs could be inferred from the GLKs that have demonstrated functions in *Arabidopsis* and tomato based on clustering. ClGLK8 clustered together with AtGLK1 (AT5G44190), AtGLK2 (AT2G20570), SlGLK1 (Solyc07g053630), and SlGLK2 (Solyc10g008160). Sequence alignments showed that ClGLK8 had the most similarity to SlGLK1 (59%) in tomato and AtGLK2 (50%) in *Arabidopsis* (Figure 2b), suggesting that it may have a similar function to these genes.

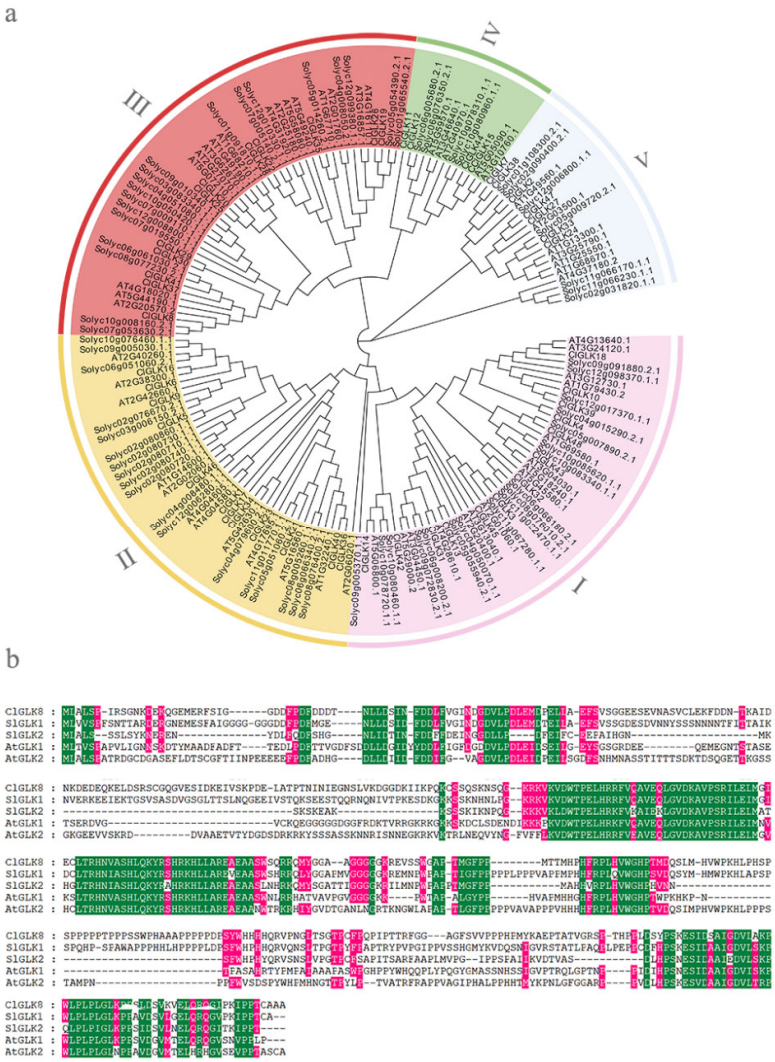

Figure 2. Phylogenetic analysis and sequence alignment of GLKs in watermelon, Arabidopsis, and tomato. (**a**) Phylogenetic tree of GLK proteins in watermelon, Arabidopsis, and tomato. (**b**) Sequence alignment of AtGLK1, AtGLK2, SlGLK1, SlGLK1, and ClGLK8.

3.4. Promoter Analysis of ClGLKs

Based on the cis-element analysis, 45 types of cis-elements were identified in the promoter regions of *ClGLKs* (Figure 3). Among these, cis-elements responsive to light were particularly abundant, accounting for 25 of the 45 cis-element types (Table S3). The promoters of all *ClGLKs* contained at least one type of light-responsive cis-element, indicating that light may play an important role in regulating the expression of *ClGLKs*. Aside from light, cis-elements related to stress response (drought, low temperature, and defense), hormone response (MeJA, abscisic acid, auxin, and gibberellin), circadian control, and endosperm expression were also identified in promoter regions (Figure 3, Table S3), indicating that *ClGLKs* may be involved in a variety of processes of plant development and stress response.

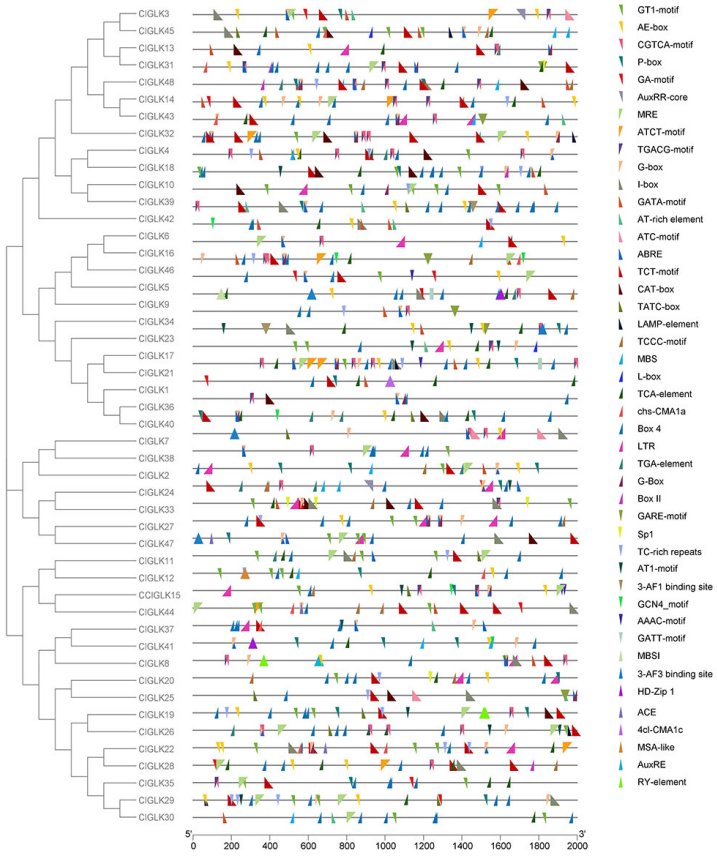

Figure 3. Cis-elements analysis of the promoters of ClGLKs. The classification and annotation of the cis-elements are listed in Table S3.

3.5. Expression Analysis of Response of ClGLKs to Biotic and Abiotic stresses

To gain insights into the potential roles of *ClGLKs* in response to abiotic and biotic stresses, gene expression was assayed via RNA-Seq after subjecting watermelons to the cucumber green mottle mosaic virus and drought stress (Figure 4a). For CGMMV treatment, the expression of 11 *ClGLKs* (*ClGLK23*, *ClGLK25*, *ClGLK29*, *ClGLK30*, *ClGLK34*, *ClGLK46*, *ClGLK1*, *ClGLK6*, *ClGLK9*, *ClGLK16*, and *ClGLK31*) were at low levels and not affected by viral infection. Among the responsive *ClGLKs*, 30 genes showed increased the expression compared to the control. Among these, 11 genes (*ClGLK21*, *ClGLK41*, *ClGLK42*, *ClGLK10*, *ClGLK27*, *ClGLK43*, *ClGLK14*, *ClGLK22*, *ClGLK47*, *ClGLK37*, and *ClGLK35*) increased the expression greater than two-fold. Of *ClGLKs* that decreased in expression, *ClGLK18*, *ClGLK33*, and *ClGLK44* were down-regulated more than two-fold. These data suggest that most *ClGLKs* responded to *CGMMV* infection, though the variability in responses suggest different mechanisms in response to CGMMV infection.

Under drought stress, a total of 46 *ClGLKs* showed changes in expression compared to the control. Among these, 26 genes were up-regulated while 20 were down-regulated (Figure 4b). Further analysis found that 20 genes changed more than two-fold compared to the control, suggesting that these genes may be significantly regulated by osmotic stress treatment. Twelve *ClGLKs* (*ClGLK8*, *ClGLK29*, *ClGLK30*, *ClGLK5*, *ClGLK37*, *ClGLK16*, *ClGLK46*, *ClGLK24*, *ClGLK26*, *ClGLK13*, *ClGLK33*, and *ClGLK11*) were significantly up-regulated, suggesting that these genes may play a positive role in response to osmotic stress. Eight *ClGLKs*, including ClGLK1, *ClGLK32*, *ClGLK20*, *ClGLK23*, *ClGLK42*, *ClGLK36*, *ClGLK18*, and *ClGLK2* were significantly down-regulated. Our data suggest that *ClGLKs* have varied expression responses to abiotic stress.

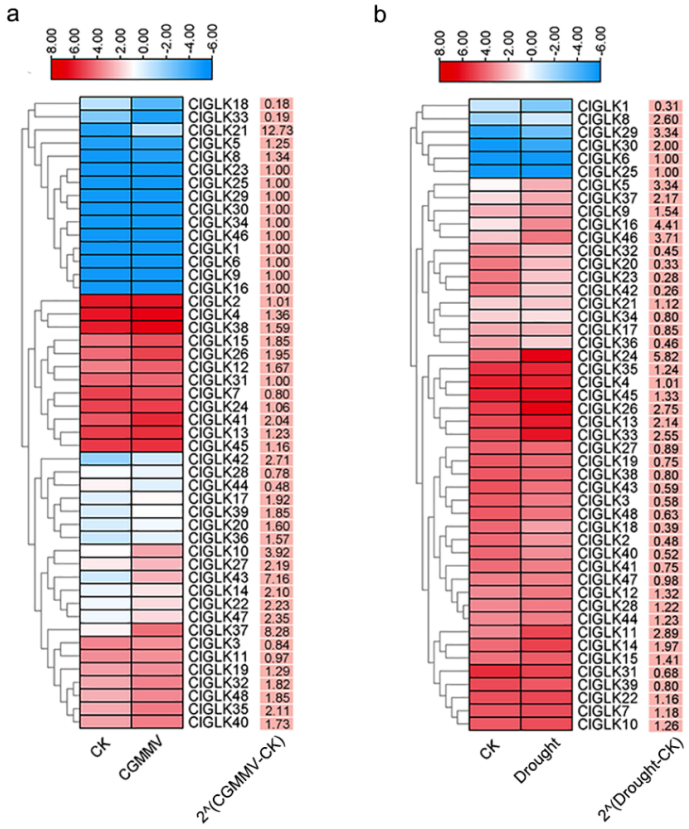

Figure 4. Expression analysis of ClGLKs under biotic and abiotic stress treatment. (**a**) The expression heatmap of ClGLKs under CGMMV infection. The data of $2^{\hat{}(CGMMV-CK)}$ indicated the fold-change of the expression level of ClGLKs under CGMMV stress compared to the control. (**b**) The expression heatmap of ClGLKs under drought stress. The data of $2^{\hat{}(CGMMV-CK)}$ indicated the fold-change of the expression level of ClGLKs under drought stress compared to the control.

3.6. Characterization of Pericarp Phenotype and Chloroplast Development

In order to analyze the color difference between the two cultivars (Figure 5a), the Chl content of the fruit pericarp at three development stages in these two cultivars was determined. Chla, Chlb, and total Chl were found to increase consistently throughout development in cultivar "W-21-72" (Figure 5b). Notably, the contents of all three chlorophyll types were significantly higher at S3, about three-fold higher than at S2, indicating that the dark green pericarp may be caused by significantly increased Chl content. For cultivar "W-21-301", there was no obvious change in the content of Chla, Chlb, and total Chl between S1 and S2. However, the Chl content at S3 was significantly higher than at S2, about a four-fold increase. In addition, the Chla, Chlb, and total Chl contents in cultivar "W-21-301" were lower than in cultivar "W-21-72" at all three developmental stages. At the S3 stage, Chla, Chlb, and total Chl contents in cultivar "W-21-72" were about two-fold higher than in cultivar "W-21-301", indicating that the color difference between the two cultivars could be due to higher Chl accumulation.

The chloroplast structure was analyzed to uncover the mechanistic basis for the color difference between the two cultivars. Chloroplast development was found to be accelerated in the pericarp of the cultivar "W-21-72", in which the chloroplast size and thylakoid density increased following fruit growth (Figure 5c). Compared to the cultivar "W-21-301", the chloroplast size and thylakoid density were higher in "W-21-72" (Figure 5c). These results suggest that the dark green formation of the pericarp was promoted by chloroplast development and Chl accumulation.

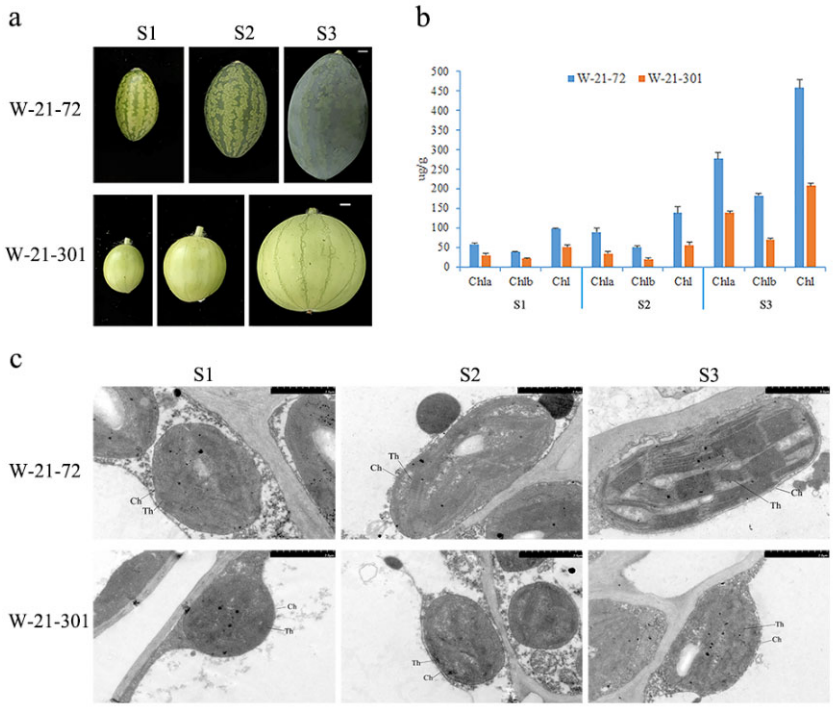

Figure 5. Phenotype analysis of the fruit in two cultivars of watermelon. (**a**) Fruit phenotype of two cultivars at three stages. S1, S2 and S3 represent the stages before, during, and after the change in pericarp color in the cultivar "W-21-72", respectively. (**b**) Chl content in the pericarp of the fruit at three stages. (**c**) Chloroplast ultrastructure (bars, 0.2 μm) in the pericarp of the fruit at three stages. Ch, chloroplast. Th, thylakoid.

3.7. Organ-Specific Expression Analysis of ClGLKs in the Pericarp of the Fruit

To gain insights into the potential function of *ClGLKs* in fruit development, their expression patterns were characterized using RNA-Seq data obtained from the pericarp of the two aforementioned cultivars. *ClGLKs* were found to have both conserved and divergent patterns of expression in these two watermelon cultivars (Figure 6). For example, nine genes (*ClGLK6*, *ClGLK9*, *ClGLK16*, *ClGLK23*, *ClGLK25*, *ClGLK29*, *ClGLK30*, *ClGLK34*, and *ClGLK46*) were either lowly expressed or not detected during fruit development in cultivars "W-21-72" and "W-21-301". Some genes, such as *ClGLK1*, *ClGLK4*, *ClGLK8*, *ClGLK13*, and *ClGLK15*, had high expression levels in both cultivars. In addition, *ClGLK8* and *ClGLK24* had a similar expression pattern, significantly increasing from S2 to S3 in both cultivars.

Some *ClGLKs*, however, had different expression patterns between the two cultivars (Figure 6). The expression of *ClGLK12*, *ClGLK32*, and *ClGLK44* decreased throughout development in "W-21-72", while no obvious changes were observed in "W-21-301". *ClGLK21* and *ClGLK26* increased expression significantly from S2 to S3 in "W-21-301", while no changes were observed in "W-21-72".

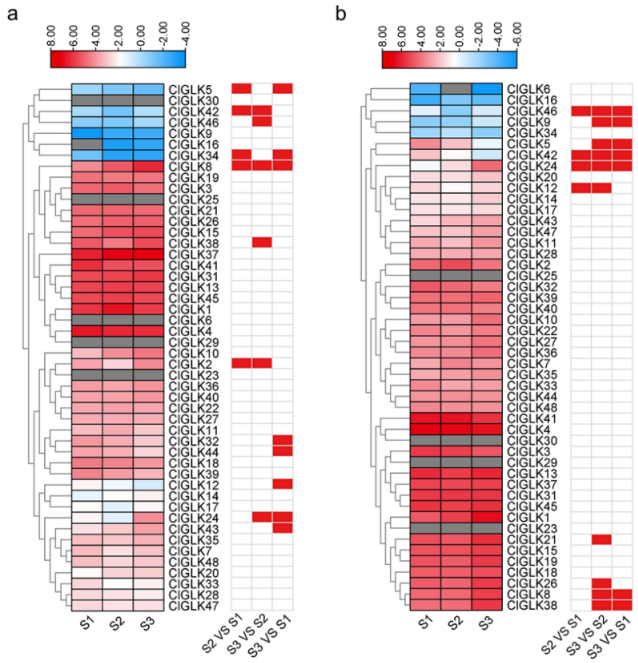

Figure 6. Expression analysis of ClGLKs in the pericarp of two cultivars of watermelon. (**a**) The expression heatmap of ClGLKs at three stages in the cultivar "W-21-72". (**b**) The expression heatmap of ClGLKs at three stages in the cultivar "W-21-301". The red color in the right table represents the DEGs in the transcriptome analysis.

3.8. Characterization of the Role of ClGLK8 in the Color Change of the Fruit Pericarp

ClGLK8, a homolog of *GLK1-2*, may also play key roles in chloroplast development and Chl accumulation. From the expression data, *ClGLK8* had high expression levels in watermelon fruit (Figure 6). In addition, it was found to significantly increase in expression in the fruit pericarp of "W-21-72" with each growth stage (Figure 6a). In the "W-21-301" cultivar, *ClGLK8* significantly increased in expression only from S2 to S3 (Figure 6b), which correspond to the increase in Chl content from S2 to S3 (Figure 5b). These results suggest that *ClGLK8* may play important roles in the Chl biosynthesis and chloroplast development of the fruit pericarp.

To further explore the role of *ClGLK8* in the color formation of fruit pericarp, its expression level was compared between the two cultivars using both RNA-Seq and RT-qPCR data (Figure 7). The RT-qPCR data for *ClGLK8* were consistent with the RNA-Seq data, and the results showed that the relative expression level of *ClGLK8* in "W-21-72" was more than two-fold higher than "W-21-301" at the S2 and S3 stages. These results suggest that the differences in *ClGLK8* expression may play a role in contributing to the color difference between the two cultivars.

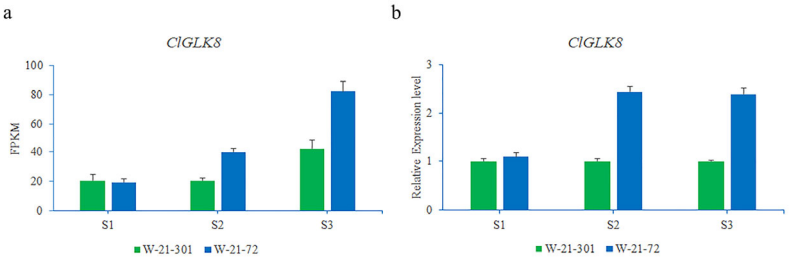

Figure 7. Expression analysis of ClGLK8 in the pericarp of two cultivars of watermelon. (**a**) The FPKM expression values of ClGLK8 at three stages in two cultivars. (**b**) The RT–qPCR data of ClGLK8 at three stages in two cultivars.

4. Discussion

Watermelon is an important horticultural crop widely cultivated worldwide. Its sweetness, sugar ratios, and skin color, among other traits, are important qualities of the fruit [23], and improving these traits is a principal goal for breeding. GLKs are a class of transcription factors widely found in plants that have been implicated in the regulation of photosynthesis-related gene expression, chloroplast development, differentiation, and functional maintenance. The potential of GLKs to increase crop yields and improve fruit quality has been noted [13,15,16], and they may also be involved in leaf senescence and stress response [24]. Therefore, a comprehensive identification and analysis of GLKs in watermelon is necessary to identify candidate GLKs for the purpose of breeding watermelons to improve fruit quality and stress defense.

In this study, 48 *GLKs* were identified in the watermelon genome. The number of GLKs in watermelon was less than that in *Arabidopsis* (55) [21], tomato (66) [22], and maize (59) [25], which might be related to the relatively small genome of watermelon. The protein sequence length, MW, and pI were found to be quite variable among the ClGLKs. Phylogenetic analysis showed that the GLK gene of watermelon can be divided into five groups, while *Arabidopsis* and tomato were divided into seven and five groups, respectively. The results of the gene structure and motif composition analysis of GLKs in these species showed similar exon-intron organization and conserved motif compositions in each group, suggesting that the phylogenetic grouping of GLKs may have a significant correlation with the gene structure and motif composition. In addition, motifs 1 and 2, which corresponded to the myb SHAQKYF domain were identified in GLKs of *Arabidopsis*, tomato, and watermelon, indicating that these motifs may play a critical role in maintaining the function of GLKs. Most GLKs in *Arabidopsis*, tomato, and maize were predicted to localize in the nucleus and some members were predicted as the chloroplast proteins. In watermelon, most ClGLKs were also predicted as the nucleus proteins while ClGLK37 was predicted to localize in the chloroplast, suggesting that it may be related to the growth and development of the chloroplast.

Previous studies have found that different signals such as light, hormones, and stress regulate the expression of *GLKs*. The expression of *GLKs* in maize, rice, and *Arabidopsis* were found to be induced by light [5,10]. In tomato, IAA can directly repress the expression of *SlGLK2* [26]. The binding site of ARFs was identified in the promoter region of *SlGLK1*, and *SlARF10* was shown to bind to the promoter to activate the expression of *SlGLK1* [27]. In addition, *GLKs* may also enhance plant defenses against environmental stress [28]. Cis-elements responsive to light, hormones (auxin, SA, ABA), and biotic stress were identified in the promoter regions of *ClGLKs*, suggesting that these signals may play important roles in regulating the *ClGLKs* involved in different developmental and growth processes.

Drought is one of the main stressors that limits watermelon growth and crop yields. *GLKs* may regulate the plant defenses to abiotic stressors such as cold and drought. Overexpression of the cotton *GhGLK1* in *Arabidopsis* significantly improved plant resistance to drought and low temperature [28]. Ectopic expression of the peanut *AhGLK1* gene in *atglk1/atglk2* double mutants can promote chlorophyll biosynthesis and photosynthesis and promote plant recovery after drought-induced damage [29]. After PEG6000 treatment, the expression of 20 *ClGLKs* was found to be significantly changed by drought (twelve up regulated, eight down regulated), suggesting that these genes may play a variety of roles in response to drought stress.

GLKs have also shown positive resistance to biotic stress. Overexpression of *AtGLK1* can enhance the plant resistance to *Fusarium graminearum* and *Botrytis cinerea* while the *atglk1/atglk2* double mutant is more susceptible to cucumber mosaic virus and *Botrytis cinerea* [30–32]. Hetero-expression of peanut *AhGLK1b* in *Arabidopsis* resulted in strong resistance to *Pst DC3000* and *Ralstonia solanacearum* [33]. *GLKs* may induce the expression of basic defense genes that significantly up-regulate or bind key factors of viral immunity, thus initiating an immune response [31,32]. CGMMV is an important viral pathogen harmful to Cucurbitaceae crops, causing severe disease symptoms in infected plants. For

watermelon, it results in the transformation of the inner pulp to a water-soaked dirty red and even flesh acidulation in the fruit, a phenomenon known as 'blood flesh' [34], leading to serious economic losses. After CGMMV infection, the expression of 11 *ClGLKs* (*ClGLK21*, *ClGLK41*, *ClGLK42*, *ClGLK10*, *ClGLK27*, *ClGLK43*, *ClGLK14*, *ClGLK22*, *ClGLK47*, *ClGLK37*, and *ClGLK35*) showed a significant increase while *ClGLK18*, *ClGLK33*, and *ClGLK44* decreased, indicating that these genes may participate in response to CGMMV.

One pair of homolog *GLKs*, *GLK1* and *GLK2*, have been shown to promote chloroplast development in plants. For example, both *AtGLK1* and *AtGLK2* promote increases in the number and volume of chloroplasts and thus the accumulation of Chl in *Arabidopsis* [5]. In tomato, *SlGLK1* and *SlGLK2* have similar functions as *AtGLK1* and *AtGLK2* [13]. In this study, only a single *ClGLK*, *ClGLK8*, was clustered together with *AtGLK1-2* and *SlGLK1-2* in the phylogeny, suggesting that *ClGLK8* may have a similar function as these proteins, and another homolog *GLK* in watermelon may have been lost during the evolution of the *ClGLKs*.

Though *GLK1* and *GLK2* have similar functions, they do have different expression patterns. Like *Arabidopsis,* tomato, and pepper, studies have revealed that *GLK1* is often highly expressed in the leaves while *GLK2* is highly expressed in the fruit. In *Arabidopsis*, only *AtGLK2* was expressed in the siliques, and an *AtGLK2* mutant resulted in a light green silique [5]. In tomato, both *SlGLK1* and *SlGLK2* are expressed in the leaves, though only *SlGLK2* is expressed in fruits and plays a key role in regulating the green shoulder fruit of the U phenotype [13]. In pepper, the expression patterns of *CaGLK1* and *CaGLK2* were similar to *SlGLK1* and *SlGLK2*, and *CaGLK2* is also a key regulator of chloroplast development in immature fruits [35]. From our data, *ClGLK8* had high expression in the fruit pericarp of the two watermelon cultivars at different stages, suggesting that *ClGLK8* has a similar expression pattern to *GLK2*. Phenotypic analysis indicated that increased chloroplast volume, thylakoid density, and Chl content were correlated with the color change in the "dark green" watermelon cultivar "W-21-72". In the "light green" watermelon cultivar "W-21-301", the Chl content also significantly increased from the S2 to S3 stage. Transcriptome analysis indicated that the expression of *ClGLK8* significantly increased in accordance with the changes in chloroplasts, thylakoid development, and Chl content, suggesting that *ClGLK8* may play a key role in promoting the color change in watermelon fruit. In addition, chloroplast size, thylakoid density, and Chl content were increased in the "dark green" fruit compared to the "light green" cultivar, and *ClGLK8* also had a relatively high expression level in the "dark green" fruit pericarp. This result suggests that *ClGLK8* may play an important role in contributing to the color difference between the two cultivars. Overexpression of *AtGLK2* and *SlGLK2* has been shown to increase the soluble solids, sugar, and carotenoid content in fruits, indicating the potential role of *ClGLK8* in the improvement in fruit quality. Further functional analyses are necessary to confirm this.

5. Conclusions

In this study, 48 *GLKs* were identified in the watermelon genome. Aside from the physical and chemical properties, we also analyzed their phylogenetic relationship, gene structure, and motif composition. Cis-elements and expression analysis suggest that most *ClGLKs* were induced by abiotic and biotic stress, and some members may play important roles in the modulation of drought stress and CGMMV infection. Furthermore, phenotype analysis of the "dark green" and "green" skin color of the watermelon fruit indicated that the difference in skin color was due to the larger chloroplast and high content of Chl. Analysis of the gene expression profiles suggests that *ClGLK8* may play a key role in contributing to this difference in phenotype. Our findings provide a theoretical reference for the further study of the role of the *ClGLKs* in stress defense, and the suggested genes could be used for breeding to improve the fruit quality of watermelon.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae9020229/s1, Table S1: Characteristics of *ClGLKs* in watermelon; Table S2: Motif sequence identified in ClGLKs; Table S3: Cis-elements identified in the promoter of *ClGLKs*; Table S4: Primers used for RT-qPCR.

Author Contributions: Conceptualization, Q.D. and L.J.; Investigation, L.J.; Software, X.J.; Visualization, M.W. and Y.W.; Supervision, H.J.; Writing—original draft preparation, Q.D.; Writing—review and editing, L.J., F.Y. and C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Important Science & Technology Specific Projects of Anhui Province (202203a06020018); the Youth Development Fund from Anhui Academy of Agricultural Sciences (QNYC-201906); the National Modern Agricultural Industrial Technology System Project (CARS-25); and the Anhui Province Vegetable industry Technology System (2021-711).

Data Availability Statement: Data are available upon request.

Acknowledgments: We would like to thank Joseph Elliot at the University of Kansas for his assistance with the English language and grammatical editing of the manuscript.

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

References

1. Llorente, B.; Segretin, M.E.; Giannini, E.; Lobais, C.; Juárez, M.E.; Paulsen, I.T.; Blanco, N.E. Homecoming: Rewinding the reductive evolution of the chloroplast genome for increasing crop yields. *Nat. Commun.* **2021**, *12*, 6734. [CrossRef]
2. Cocaliadis, M.F.; Fernández-Muñoz, R.; Pons, C.; Orzaez, D.; Granell, A. Increasing tomato fruit quality by enhancing fruit chloroplast function. A double-edged sword? *J. Exp. Bot.* **2013**, *65*, 4589–4598. [CrossRef]
3. Zhao, M.-H.; Li, X.; Zhang, X.-X.; Zhang, H.; Zhao, X.-Y. Mutation mechanism of leaf color in plants: A review. *Forests* **2020**, *11*, 851. [CrossRef]
4. McFadden, G.I. Chloroplast origin and integration. *Plant Physiol.* **2001**, *125*, 50–53. [CrossRef]
5. Fitter, D.W.; Martin, D.J.; Copley, M.J.; Scotland, R.W.; Langdale, J.A. GLK gene pairs regulate chloroplast development in diverse plant species. *Plant J.* **2002**, *31*, 713–727. [CrossRef]
6. Oda-Yamamizo, C.; Mitsuda, N.; Sakamoto, S.; Ogawa, D.; Ohme-Takagi, M.; Ohmiya, A. The NAC transcription factor ANAC046 is a positive regulator of chlorophyll degradation and senescence in Arabidopsis leaves. *Sci. Rep.* **2016**, *6*, 23609. [CrossRef]
7. Nadakuduti, S.S.; Holdsworth, W.L.; Klein, C.L.; Barry, C.S. KNOX genes influence a gradient of fruit chloroplast development through regulation of GOLDEN 2-LIKE expression in tomato. *Plant J.* **2014**, *78*, 1022–1033. [CrossRef]
8. Yan, F.; Gao, Y.; Pang, X.; Xu, X.; Zhu, N.; Chan, H.; Hu, G.; Wu, M.; Yuan, Y.; Li, Z. BEL1-LIKE HOMEODOMAIN4 regulates chlorophyll accumulation, chloroplast development, and cell wall metabolism in tomato fruit. *J. Exp. Bot.* **2020**, *71*, 5549–5561. [CrossRef]
9. Riechmann, J.L.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C.-Z.; Keddie, J.; Adam, L.; Pineda, O.; Ratcliffe, O.; Samaha, R.J.; et al. Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. *Science* **2000**, *290*, 2105–2110. [CrossRef]
10. Rossini, L.; Cribb, L.; Martin, D.J.; Langdale, J.A. The maize golden2 gene defines a novel class of transcriptional regulators in plants. *Plant Cell* **2001**, *13*, 1231–1244. [CrossRef]
11. Nakamura, H.; Muramatsu, M.; Hakata, M.; Ueno, O.; Nagamura, Y.; Hirochika, H.; Takano, M.; Ichikawa, H. Ectopic overexpression of the transcription factor OsGLK1 induces chloroplast development in non-green rice cells. *Plant Cell Physiol.* **2009**, *50*, 1933–1949. [CrossRef]
12. Powell, A.L.; Nguyen, C.V.; Hill, T.; Cheng, K.L.; Figueroa-Balderas, R.; Aktas, H.; Ashrafi, H.; Pons, C.; Fernández-Muñoz, R.; Vicente, A.; et al. Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science* **2012**, *336*, 1711–1715. [CrossRef]
13. Nguyen, C.V.; Vrebalov, J.T.; Gapper, N.E.; Zheng, Y.; Zhong, S.; Fei, Z.; Giovannoni, J.J. Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. *Plant Cell* **2014**, *26*, 585–601. [CrossRef]
14. Waters, M.T.; Wang, P.; Korkaric, M.; Capper, R.G.; Saunders, N.J.; Langdale, J. A GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. *Plant Cell* **2009**, *21*, 1109–1128. [CrossRef]
15. Li, X.; Wang, P.; Li, J.; Wei, S.; Yan, Y.; Yang, J.; Zhao, M.; Langdale, J.A.; Zhou, W. Maize GOLDEN2-LIKE genes enhance biomass and grain yields in rice by improving photosynthesis and reducing photoinhibition. *Commun. Biol.* **2020**, *3*, 151. [CrossRef]
16. Yeh, S.-Y.; Lin, H.-H.; Chang, Y.-M.; Chang, Y.-L.; Chang, C.-K.; Huang, Y.-C.; Ho, Y.-W.; Lin, C.-Y.; Zheng, J.-Z.; Jane, W.-N.; et al. Maize Golden2-like transcription factors boost rice chloroplast development, photosynthesis, and grain yield. *Plant Physiol.* **2022**, *188*, 442–459. [CrossRef]
17. Brown, A.C.; Summers, W. Carbohydrate accumulation and color development in watermelon. *J. Am. Soc. Hortic. Sci.* **1985**, *110*, 683–687. [CrossRef]

18. Sagar, M.; Chervin, C.; Mila, I.; Hao, Y.; Roustan, J.-P.; Benichou, M.; Gibon, Y.; Biais, B.; Maury, P.; Latché, A. SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiol.* **2013**, *161*, 1362–1374. [CrossRef] [PubMed]
19. 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]
20. 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]
21. Alam, I.; Wu, X.; Yu, Q.; Ge, L. Comprehensive Genomic Analysis of G2-like Transcription Factor Genes and Their Role in Development and Abiotic Stresses in Arabidopsis. *Diversity* **2022**, *14*, 228. [CrossRef]
22. Wang, Z.-Y.; Zhao, S.; Liu, J.-F.; Zhao, H.-Y.; Sun, X.-Y.; Wu, T.-R.; Pei, T.; Wang, Y.; Liu, Q.-F.; Yang, H.-H.; et al. Genome-wide identification of Tomato Golden 2-Like transcription factors and abiotic stress related members screening. *BMC Plant Biol.* **2022**, *22*, 82. [CrossRef]
23. Kyriacou, M.C.; Leskovar, D.I.; Colla, G.; Rouphael, Y. Watermelon and melon fruit quality: The genotypic and agro-environmental factors implicated. *Sci. Hortic.* **2018**, *234*, 393–408. [CrossRef]
24. Chen, M.; Ji, M.; Wen, B.; Liu, L.; Li, S.; Chen, X.; Gao, D.; Li, L. GOLDEN 2-LIKE transcription factors of plants. *Front. Plant Sci.* **2016**, *7*, 1509. [CrossRef]
25. Liu, F.; Xu, Y.; Han, G.; Zhou, L.; Ali, A.; Zhu, S.; Li, X. Molecular evolution and genetic variation of G2-like transcription factor genes in maize. *PLoS ONE* **2016**, *11*, e0161763. [CrossRef]
26. Lupi, A.C.D.; Lira, B.S.; Gramegna, G.; Trench, B.; Alves, F.R.R.; Demarco, D.; Peres, L.E.P.; Purgatto, E.; Freschi, L.; Rossi, M. Solanum lycopersicum GOLDEN 2-LIKE 2 transcription factor affects fruit quality in a light-and auxin-dependent manner. *PLoS ONE* **2019**, *14*, e0212224. [CrossRef]
27. Yuan, Y.; Mei, L.; Wu, M.; Wei, W.; Shan, W.; Gong, Z.; Zhang, Q.; Yang, F.; Yan, F.; Zhang, Q.; et al. SlARF10, an auxin response factor, is involved in chlorophyll and sugar accumulation during tomato fruit development. *J. Exp. Bot.* **2018**, *69*, 5507–5518. [CrossRef]
28. Liu, J.; Mehari, T.G.; Xu, Y.; Umer, M.J.; Hou, Y.; Wang, Y.; Peng, R.; Wang, K.; Cai, X.; Zhou, Z.J. GhGLK1 a key candidate gene from GARP family enhances cold and drought stress tolerance in cotton. *Front. Plant Sci.* **2021**, *12*, 759312. [CrossRef]
29. Liu, X.; Li, L.; Li, M.; Su, L.; Lian, S.; Zhang, B.; Li, X.; Ge, K.; Li, L. AhGLK1 affects chlorophyll biosynthesis and photosynthesis in peanut leaves during recovery from drought. *Sci. Rep.* **2018**, *8*, 2250. [CrossRef]
30. Han, X.-Y.; Li, P.-X.; Zou, L.-J.; Tan, W.-R.; Zheng, T.; Zhang, D.-W.; Lin, H.-H. GOLDEN2-LIKE transcription factors coordinate the tolerance to Cucumber mosaic virus in Arabidopsis. *Biochem. Biophys. Res. Commun.* **2016**, *477*, 626–632. [CrossRef]
31. Murmu, J.; Wilton, M.; Allard, G.; Pandeya, R.; Desveaux, D.; Singh, J.; Subramaniam, R.J. A rabidopsis GOLDEN2-LIKE (GLK) transcription factors activate jasmonic acid (JA)-dependent disease susceptibility to the biotrophic pathogen H yaloperonospora arabidopsidis, as well as JA-independent plant immunity against the necrotrophic pathogen B otrytis cinerea. *Mol. Plant Pathol.* **2014**, *15*, 174–184.
32. Savitch, L.V.; Subramaniam, R.; Allard, G.C.; Singh, J. The GLK1 'regulon'encodes disease defense related proteins and confers resistance to Fusarium graminearum in Arabidopsis. *Biochem. Biophys. Res. Commun.* **2007**, *359*, 234–238. [CrossRef]
33. Ali, N.; Chen, H.; Zhang, C.; Khan, S.A.; Gandeka, M.; Xie, D.; Zhuang, W. Ectopic expression of AhGLK1b (GOLDEN2-like transcription factor) in Arabidopsis confers dual resistance to fungal and Bacterial Pathogens. *Plant Cell Physiol.* **2020**, *11*, 343. [CrossRef]
34. Li, X.; An, M.; Xia, Z.; Bai, X.; Wu, Y. Transcriptome analysis of watermelon (*Citrullus lanatus*) fruits in response to Cucumber green mottle mosaic virus (CGMMV) infection. *Sci. Rep.* **2017**, *7*, 16747. [CrossRef]
35. Brand, A.; Borovsky, Y.; Hill, T.; Rahman, K.A.A.; Bellalou, A.; Van Deynze, A.; Paran, I. CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit. *Theor. Appl. Genet.* **2014**, *127*, 2139–2148. [CrossRef]