


Article

# Transcriptomics Integrated with Metabolomics Unveil Carotenoids Accumulation and Correlated Gene Regulation in White and Yellow-Fleshed Turnip (*Brassica rapa ssp. rapa*)

Yanjing Ren <sup>1,2,†</sup> , Rui Han <sup>1,†</sup>, Yidong Ma <sup>1</sup>, Xiaojuan Li <sup>1</sup>, Changrong Deng <sup>1,2</sup>, Mengliang Zhao <sup>1,2</sup>, Jiang Li <sup>1,2</sup>, Quangang Hou <sup>1,2</sup>, Qiwen Zhong <sup>1,2,\*</sup> and Dengkui Shao <sup>1,2,3,\*</sup>

<sup>1</sup> Academy of Agriculture and Forestry Sciences, Qinghai University, Xining 810016, China; renyan0202@163.com (Y.R.); hanrui201@126.com (R.H.); mydwn0709@163.com (Y.M.); l889989l@163.com (X.L.); dengchang\_rong@126.com (C.D.); 8304269@163.com (M.Z.); 2004990019@qhu.edu.cn (J.L.); hqgboy@163.com (Q.H.)

<sup>2</sup> Qinghai Key Laboratory of Vegetable Genetics and Physiology, Xining 810016, China

<sup>3</sup> State Key Laboratory of Crop Stress Biology for Arid Area, College of Horticulture, Northwest A&F University, Yangling, Xianyang 712100, China

\* Correspondence: zhongqiwen@qhu.edu.cn (Q.Z.); 2006990015@qhu.edu.cn (D.S.); Tel.: +86-0971-531-1167 (Q.Z. & D.S.)

† These authors contributed equally to this work.

**Abstract:** Turnip (*Brassica rapa ssp. rapa*) is considered to be a highly nutritious and health-promoting vegetable crop, whose flesh color can be divided into yellow and white. It is widely accepted that yellow-fleshed turnips have higher nutritional value. However, reports about flesh color formation is lacking. Here, the white-fleshed inbred line, W21, and yellow-fleshed inbred line, W25, were profiled from the swollen root of the turnip at three developmental periods to elucidate the yellow color formation. Transcriptomics integrated with metabolomics analysis showed that the *PSY* gene was the key gene affecting the carotenoids formation in W25. The coding sequence of *BrrPSY-W25* was 1278 bp and that of *BrrPSY-W21* was 1275 bp, and *BrrPSY* was more highly expressed in swollen roots in W25 than in W21. Transient transgenic tobacco leaf over-expressing *BrrPSY-W* and *BrrPSY-Y* showed higher transcript levels and carotenoids contents. Results revealed that yellow turnip formation is due to high expression of the *PSY* gene rather than mutations in the *PSY* gene, indicating that a post-transcriptional regulatory mechanism may affect carotenoids formation. Results obtained in this study will be helpful for explaining the carotenoids accumulation of turnips.

**Keywords:** *Brassica rapa ssp. rapa*; carotenoids accumulation; carotenoid biosynthesis pathway; candidate gene analysis; gene function validation



**Citation:** Ren, Y.; Han, R.; Ma, Y.; Li, X.; Deng, C.; Zhao, M.; Li, J.; Hou, Q.; Zhong, Q.; Shao, D. Transcriptomics Integrated with Metabolomics Unveil Carotenoids Accumulation and Correlated Gene Regulation in White and Yellow-Fleshed Turnip (*Brassica rapa ssp. rapa*). *Genes* **2022**, *13*, 953. <https://doi.org/10.3390/genes13060953>

Received: 20 April 2022

Accepted: 23 May 2022

Published: 26 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Turnip (*B. rapa ssp. rapa* 2n = 20) is one of the most important Curciferaceae leaf and root vegetable crops in China and throughout East Asia, which is used for human consumption, Tibetan medicine and animal fodder [1]. Turnip is a rich source of glucosinolates [2,3], dietary phenolics [4], dietary fiber, vitamin C [5] and other bioactive compounds [6]. It was also found that flavonoid compounds in Tibetan turnip, p-coumaric acid and glucoside (p coumaric acid-β-D-glucopyranoside), were closely related to an anti-hypoxia effect [7]. Color is an important commodity trait for fruits and vegetables [8]. Yellow-fleshed turnips are a large part of the turnip species, which are highly popular with consumers. Generally, yellow-fleshed turnips are caused by the accumulation of carotenoids.

Carotenoids are a class of important natural pigments with 40 carbons in their backbone, which provide the precursors for vitamin A synthesis [9]. In addition, carotenoids can reduce various chronic diseases due to their antioxidant properties [10]. Carotenoids can absorb light within 400–500 nm. Therefore, the high accumulation of carotenoids

makes many plants appear yellow, orange and red [11]. In carrot, red is attributed to the massive accumulation of lycopene and  $\beta$ -carotene, orange is due to the enrichment of  $\beta$ -carotene, while yellow is thanks to the concentration of lutein [12]. In tomato, lycopene is predominant in the red type [13],  $\delta$ -carotene and lycopene are major components in the orange/red type [14],  $\beta$ -carotene and lycopene are predominant in the orange type [15]. In pepper, capsanthin is the main pigment in the red type; zeaxanthin, capsanthin and lutein are predominant in the orange type [16]; lutein and  $\beta$ -carotene are the main two compounds in the yellow type [17]. In orange cauliflower and orange heading Chinese cabbage,  $\beta$ -carotene is predominant [18,19].

Many genes have been reported to control carotenoids' regulation in vegetable crops. Carotenoids synthesis begins with geranylgeranyl diphosphate (GGPP) which is formed by the condensation of isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP) [20]. GGPP is condensed into colorless phytoene by the activity of phytoene synthase (PSY) [21]. Then, the colorless phytoene forms multiple compounds, such as  $\beta$ -carotene,  $\delta$ -carotene, zeaxanthin, lutein, violaxanthin and neoxanthin, through a series of reactions [22–24]. In orange carrot, the *CYP97A3* gene was reported to be responsible for the high levels of  $\alpha$ -carotene accumulation [25]. In tomato, the *SISGR1* gene was reported to interact with *PSY1* and regulate lycopene enrichment [26]. In pepper, the *CCS* gene was considered to be the main gene that controls orange formation [16]. In heading Chinese cabbage, a large insertion in the *CRTISO* gene leads to the orange formation [27]. These studies suggested that most yellow/orange traits are controlled by the structure genes. In addition, many transcription factors have been proved to be involved in carotenoids biosynthesis. Wu et al. [28] reported that *SIMYB72*, a R2R3-MYB subfamily, directly bound to phytoene synthase, *z*-carotene isomerase and lycopene-cyclase genes and regulated carotenoids biosynthesis. Shi et al. [29] also published that *SIZHD17* is involved in the control of chlorophyll and carotenoids metabolism in tomato fruit. To date, the carotenoid component and correlated gene analyses in white and yellow-fleshed turnip have been infrequent.

In this study, we identified the carotenoid components and correlated genes in two cultivars at three developmental periods by transcriptome and metabolome methods. The data from this study enhanced our understanding of the carotenoid accumulation and correlated gene expression in turnips, and provided an insight for the improvement of yellow turnips or other yellow/orange-fleshed root vegetable crops.

## 2. Materials and Methods

### 2.1. Plant Materials

Two *B. rapa* ssp. *rapa* cultivars (white-fleshed cultivar, W21, and yellow-fleshed cultivar, W25) were planted in a randomized field plot according to standard agricultural practices in a field at the Xining experimental station (36°42' N; 101°45' E) of the Academy of Agriculture and Forestry Sciences, Qinghai University. They were planted in June and swollen root tissue samples were collected at 30 d (the beginning of swelling of the root), 50 d (swelling period of the root) and 70 d (maturation period of the root as a commercial organ) after planting. The swollen root was washed with tap water and the peel was removed. Then, the swollen root tissues were cut into pieces and stored at  $-80^{\circ}\text{C}$  for further analysis. According to the sampling period, samples were marked as W1, W2, W3, Y1, Y2 and Y3, separately. Each cultivar included five individuals and each stage contained three samples. In total, 18 samples were quickly frozen and stored.

### 2.2. RNA Extraction, Quantification and Transcriptome Sequencing

Eighteen libraries representing the six flesh samples and the three replicates were constructed for transcriptome sequencing. The RNA extraction, quantification and transcriptome sequencing were conducted by Illumina HiSeq high-throughput sequencing platform (Illumina, San Diego, CA, USA) at Wuhan MetWare Biotechnology Co., Ltd. (Wuhan, China) following their standard procedures [30] and the raw data were obtained. After removing the adapter sequences, reads with unknown nucleotides (more than 10%

ambiguous residues N) and reads with low quality (containing >50% bases with quality score < 20), the clean reads were held. The high-quality reads were de novo assembled into transcripts using Trinity (Version 2.6.6 [31]) by employing a paired-end method [32]. The dataset is available from the NCBI (<https://www.ncbi.nlm.nih.gov/>, accessed on 27 July 2020) Short Read Archive (SRA) under accession number PRJNA645708.

### 2.3. Unigene Annotations and Differentially Expressed Unigenes Identification

All assembled unigenes were annotated based on BLAST searches [33] and searched against NCBI non-redundant protein sequences (Nr) [34], Swiss-Prot (<http://www.ebi.ac.uk/uniprot/>, accessed on 27 July 2020) [35], TrEMBL, Clusters of Orthologous Groups/euKaryotic Ortholog Groups (COG/KOG) (<http://www.ncbi.nlm.nih.gov/COG/>, accessed on 27 July 2020) [36,37], Gene Ontology (GO) (<http://www.geneontology.org/>, accessed on 27 July 2020) [38] and Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>, accessed on 27 July 2020) [39] with a threshold of the e-value <  $1.0 \times 10^{-5}$ . After predicting the amino acid sequence of unigenes, the annotation information of unigenes was obtained by using HMMER v3.0 software (Robert D. Finn, EMBL-EBI European Bioinformatics Institute, Cambridge, UK) against the Pfam database (<http://pfam.sanger.ac.uk/>, accessed on 27 July 2020) with a threshold of the e-value < 0.01. Differential expression genes (DEGs) were compared between two varieties at three developmental period and determined by the DESeq2 [40,41] with the  $|\log_2\text{Fold Change}| \geq 1$  and false discovery rate (FDR) correction set at  $p < 0.05$ . GO enrichment and KEGG pathway enrichment of the DEGs were measured by using the topGO method and KOBAS2.0 [42,43].

### 2.4. Extraction of Carotenoids, Metabolite Identification and Quantification

The sample preparation, extract analysis, metabolite identification and quantification were performed at Wuhan MetWare Biotechnology Co., Ltd. (Wuhan, China) following their standard procedures previously fully described by Inbaraj et al. [44], Petry and Mercadante [45], and Zhou et al. [26]. The standard curves of eighteen carotenoid components (phytoene, phytofluene, lycopene,  $\delta$ -carotene,  $\alpha$ -carotene, zeinoxanthin, lutein,  $\gamma$ -carotene,  $\beta$ -carotene,  $\beta$ -cryptoxanthin, zeaxanthin, antheraxanthin, violaxanthin, neoxanthin, apocrotinal, astaxanthin, capsanthin, capsorubin) were plotted and linear equations of the standard curves were calculated. Thus, carotenoid component content was calculated accurately. Differential expression metabolites were compared between two varieties at three developmental period and determined by the DESeq2 [40,41] with the  $|\text{Fold Change}| \geq 1$  and false discovery rate (FDR) correction set at  $p < 0.05$ . Significance level analysis was performed by SPSS 26.0.

The chemical reagents used in this study, including methanol (MeOH), ethanol (EtOH), acetone, methyl tert-butyl ether and BHT, purchased from Merck (Darmstadt, Germany), and Milli-Q water (Millipore, Bradford, MA, USA), were used in all experiments. All standards and formic acid were purchased from Sigma (St. Louis, MO, USA). The stock solutions of standards were prepared at a concentration of 1 mg/mL and were stored at  $-20^\circ\text{C}$ .

### 2.5. Weighted Gene Co-Expression Network Analysis

The searched DEGs were used to generate co-expression network modules by the weighted gene co-expression network analysis (WGCNA) package and the obtained co-expression modules were merged on eigengenes [46]. Eigengene value was calculated for each module, which was used to search the association with carotenoid metabolites. The co-expression network between structural genes and transcription factors was also created based on the Pearson correlation coefficient ( $\text{PCC} > 0.8$ ).

### 2.6. DNA/RNA Extraction, Candidate Gene Prediction and Cloning

The extractions of the genomic DNA from the leaf sample and total RNA from the different tissues were performed as described by Ren et al. [47]. The candidate gene was predicted according to the association analysis of DEGs and DEMs. The specific primers for the candidate gene (Table S1) were designed using Primer Premier 6.0 on the basis of the reference genome of *B. rapa* [48,49] and the sequences of the candidate gene were cloned from 'W21' and 'W25'. The PCR products were purified and used for TA cloning, and the recombinant plasmids were transformed into *E. coli* strain DH5 $\alpha$ . The recombinant plasmids were sequenced by Sunny Biotechnology Co. Ltd. (Shanghai, China) and sequence alignment was performed with DANMAN software (Lynnon BioSoft, San Ramon, CA, USA).

### 2.7. Sequence Analysis and Gene Expression Analysis of BrrPSY Gene

Gene sequences obtained from 'W21' and 'W25' were called *PSY-W* and *PSY-Y*, respectively. Gene structure of exons and introns was predicted based on the *PSY* gene in *B. rapa*. Sequence alignments were performed by the multiple sequence alignment module of DNAMAN software (Lynnon BioSoft, USA). The physicochemical properties of the proteins encoded by the BrrPSY gene were predicted using on-line software tools SOPMA and ExPaSy. The phylogenetic relationship tree of PSY from multiple species was constructed with MEGA 6.0 with the neighbor-joining method.

The analysis of gene expression was carried out by using quantitative real-time PCR (RT-qPCR) and the actin gene was used as the reference gene [49]. The specific primers of the BrrPSY gene are shown in Table S1. The qPCR reactions were performed in a Roche LightCycler 480 Detection System and all qPCR reactions of each sample were performed with three biological repetitions and three technical repetitions. The relative expression levels of genes were measured by the  $2^{-\Delta\Delta C_t}$  method [50].

### 2.8. Transformation of Protoplasts and Fluorescence Observation

To clarify the location of the BrrPSY protein in the cells, the GFP gene was fused with BrrPSY as a reporter. The recombinant vector PBI211GFP-BrrPSY was transferred into Arabidopsis protoplasts using a polyethylene glycol-mediated approach. The transformed protoplasts were incubated at 23 °C for 16 h with low light level and then fluorescence was observed with a laser confocal microscope. PBI211GFP was used as the negative control.

### 2.9. Transformation of Tobacco

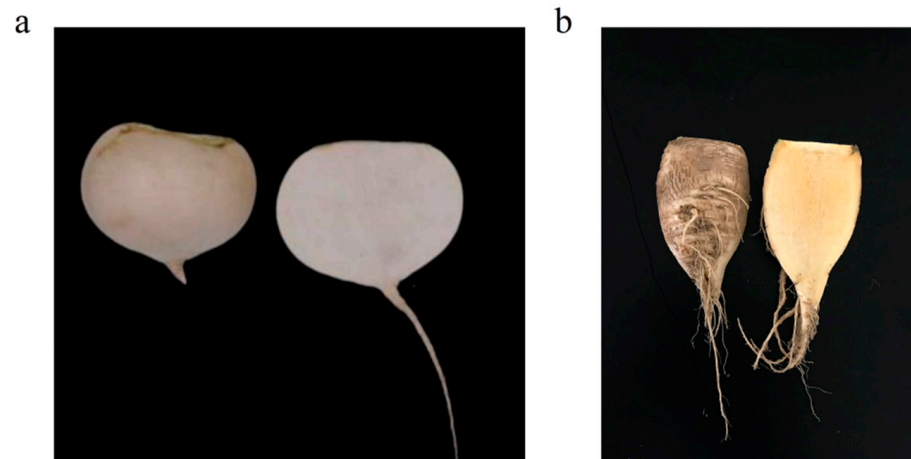
To validate the function of the BrrPSY gene, the *PSY-Y* and *PSY-W* genes with *ScaI* and *XbaI* restriction enzyme cutting sites were cloned into a PCAMBIA2300 vector, respectively. Then, *Agrobacterium tumefaciens* GV3101 cell lines carrying PCAMBIA2300, PCAMBIA2300-*PSY-Y* and PCAMBIA2300-*PSY-W* were transferred into a tobacco line using the injection infiltration method, separately. Three day later, the infiltration areas were collected for RNA extraction, RT-qPCR and carotenoid content detection. The primers used in these experiments are listed in Table S1.

## 3. Results and Discussion

### 3.1. De Novo Transcriptome Assembly in the Two Turnips at Three Development Stages

Swollen root tissue samples of two *B. rapa* ssp. *rapa* varieties (white-fleshed cultivar, W21, and yellow-fleshed cultivar, W25) were collected for the RNA-seq (Figure 1a,b). Each sample included five individuals and each stage contained three samples. A total of 173.82G clean bases with an average of 9.65 Gb sequencing data, 94.09% of bases scoring Q30 and 45.21% GC content for each sample, were obtained (Table S2). Compared with Zhuang et al. [1], we obtained more sequencing data in swollen root tissue samples of turnips than in root skin samples (6.20 Gb). After split joint by Trinity software, a total of 298,926 transcripts were generated with an N50 length of 1365 bp, an N90 of 362 bp and a mean length of 879 bp. The obtained transcripts were used as reference sequences for

the unigene analysis based on the corset hierarchical clustering. The results showed that a total of 253,720 unigenes were obtained with an N50 length of 1424 bp, an N90 of 460 bp and a mean length of 989 bp. In other studies, Zhuang et al. [1] obtained 76,152 unigenes after assembly in turnip root skin samples, and Lin et al. [51] generated 84,132 unigenes in turnip floral buds, which suggest that more unigenes are expressed in turnip swollen root tissue than in root skin samples and floral bud samples.

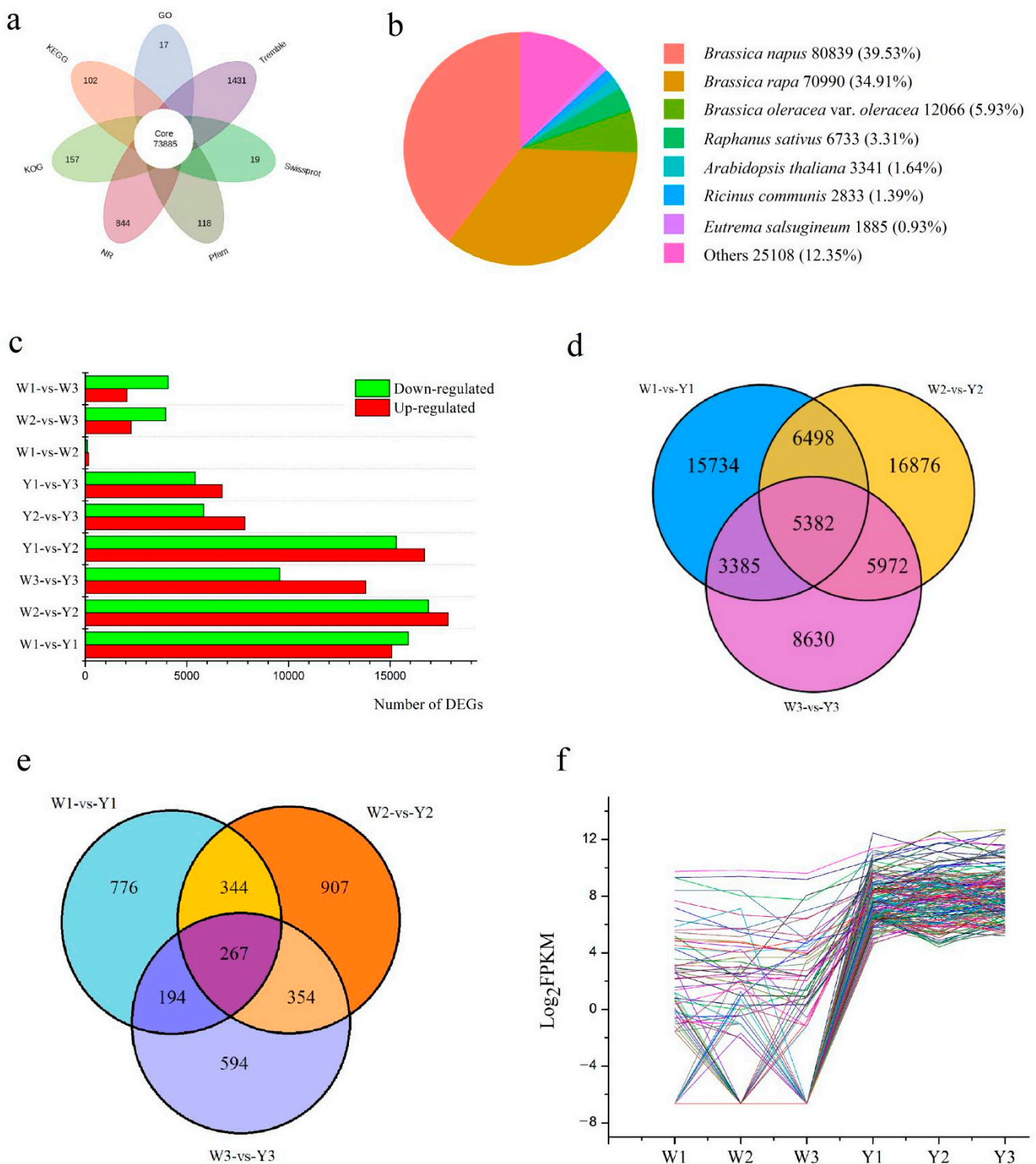


**Figure 1.** The phenotypes of the turnips at 70 days after planting. (a) W21. (b) W25.

The size and distribution of transcripts and unigene length analysis showed that 133,252 (44.58%) transcripts were less than 500 bp in length and 26,755 (8.95%) transcripts were more than 2000 bp in length, 88,523 (34.89%) unigenes were less than 500 bp in length and 26,755 (10.55%) unigenes were more than 2000 bp in length (Figure S1a, Table S3). Coding sequence (CDS) length and predicted protein sequence were also analyzed. A total of 193,914 CDSs were searched and 99,798 (51.47%) CDSs were less than 500 bp and 6869 (3.54%) CDSs were greater than 2000 bp in length (Figure S1b, Table S4). The results of predicted protein sequence analysis revealed that 152,628 (76.40%) protein sequences were less than 300 amino acids and 1711 (0.86%) protein sequences were greater than 1000 amino acids in length (Figure S1c).

### 3.2. Unigene Annotation

Based on Kyoto Encyclopedia of Genes and Genomes (KEGG), NCBI non-redundant protein sequences (Nr), Gene Ontology (GO), Clusters of Orthologous Groups/euKaryotic Ortholog Groups (COG/KOG), Swiss-Prot, Trembl and Pfam databases, there were 151,257 (69.54%), 230,345 (80.15%), 172,078 (67.82%), 118,324 (46.64%), 143,975 (56.75%), 204,633 (80.65%) and 143,807 (56.68%) functional unigene annotations, respectively, and 207,651 (81.84%) unigenes were assigned to at least one database (Table S5). After the integrated analysis, 73,885 (29.12%) unigenes were found to be annotated in all seven databases (Figure 2a, Table S6). Compared with the Nr database, an interesting result observed was that the functional information of the homologous sequences was similar to that of *B. napus* (80,389, 39.53%) than that of *B. rapa* (70,990, 34.91%) (Figure 2b). In fact, many researchers consider turnip to be a subspecies of Chinese cabbage. At the genomic level, the results of the phylogenetic relationship of the chloroplast and mitochondrial genome of turnip showed that turnip is closely related to *B. rapa* [52,53]. Results of homologous sequence alignment in this study may provide a new insight into the root development of *B. napus*, *B. rapa* and turnip.



**Figure 2.** Unigene annotation and identification of DEGs in two turnips. (a) The petal diagram of number of unigenes annotated in seven databases. (b) The top BLAST hit species based on the unigene annotation of turnip fresh root. (c) DEGs in different pairwise comparisons. (d) Venn diagram of DEGs in W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3. (e) Venn diagram of differentially expressed transcription factors in W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3. (f) The trend change map of 163 transcription factors shows higher expression level in W25 than in W21 in all three pairwise comparisons of W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3.

Gene Ontology annotation results showed that 38,256 unigenes were classified as biological process, 55,166 unigenes were classified into cellular component and 33,987 unigenes were annotated as molecular function (Table S7). In terms of KOG classification, 24,147 unigenes were general function prediction only, 13,277 unigenes were categorized into posttranslational modification, protein turnover and chaperones and 12,696 unigenes were categorized into signal transduction mechanisms (Table S8).

### 3.3. Identification of Differentially Expressed Unigenes

To identify the DEGs between W21 and W25, gene expression levels were estimated with the fragments per kilobase of exon per million fragments mapped (FPKM) values. The FPKM values in W21 and W25 at the different developmental stages were compared and the DEGs were selected with the  $|\log_2\text{Fold Change}| \geq 1$  and false discovery rate (FDR) correction set at  $p < 0.05$ . Compared with Y1, 15,089 unigenes were up-regulated and 15,910 unigenes were down-regulated in W1. Compared with Y2, 17,844 unigenes were up-regulated and 16,884 unigenes were down-regulated in W2. Compared with Y3, 13,798 unigenes were up-regulated and 9571 unigenes were down-regulated in W3 (Figure 2c). After searching the DEGs in pairwise comparisons of W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3, the common DEGs in all three pairwise comparisons were then searched; 5382 common unigenes were identified in all three pairwise expressions and are shown in a Venn diagram (Figure 2d).

Differentially expressed transcription factors (TFs) were also identified in the two turnips at three development stages. There were 1581, 1872 and 1409 differentially expressed TFs and 267 common differentially expressed TFs identified in the pairwise comparisons of W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3 (Figure 2e, Table S9). Among these TFs, 18, 18, 14, 11, 11 and 10 unigenes were classified into *ethylene-responsive transcription factor (ERF)*, *C3H*, *bZIP*, *bHLH*, *WRKY* and *nuclear transcription factor Y subunit A (NF-YA)* TF family, and they were considered as predominant common differentially expressed TFs. Among these TF gene families, 163 TFs exhibited a higher expression level in W25 than in W21 in all three pairwise comparisons of W1-vs-Y1, W2-vs-Y2 and W3-vs-Y3, and the predominant TFs were the same as the common differentially expressed TFs. An interesting phenomenon was that most of the above highly expressed TFs were mostly not expressed in W21 and there were 107/163 (65.64%). Based on the  $\text{Log}_2\text{FPKM}$  values, the trend change map of these 163 TFs was constructed (Figure 2f).

### 3.4. Carotenoid Metabolites in Turnip at Three Developmental Stages

Though two turnip cultivars were planted simultaneously and grown in the same field and under the same conditions, the colors of the swollen root differed distinctly. The flesh color was white in W21 and yellow in W25. To elucidate the carotenoid metabolites in turnip, W21 and W25 were chosen for carotenoid metabolome profiling analysis. A total of 18 carotenoid compounds were detected and 10/18 were detected in swollen root tissues of the two turnips. Ten carotenoid compounds (phytoene, lycopene,  $\alpha$ -carotene,  $\gamma$ -carotene,  $\beta$ -carotene, lutein, zeaxanthin, violaxanthin, neoxanthin and apocarotenal) were detected in W25 and four compounds (phytoene, lycopene,  $\gamma$ -carotene and apocarotenal) were unique. Seven carotenoid compounds ( $\alpha$ -carotene,  $\beta$ -carotene, lutein, zeaxanthin, violaxanthin, neoxanthin and  $\beta$ -cryptoxanthin) were detected in W21 and only one compound ( $\beta$ -cryptoxanthin) was unique. At the same time, all detected carotenoid components were quantified and the contents are shown in Table 1. Among these carotenoids, phytoene was colorless, thus, the high contents of lycopene and  $\gamma$ -carotene made W25 develop a yellow swollen root.

**Table 1.** All detected carotenoid components and their contents (mg/100g DW) in two turnips (*B. rapa* ssp. *rapa*).

Compounds Detected	W25			W21		
	Y1	Y2	Y3	W1	W2	W3
Phytoene	37.050 ± 0.920 <sup>aA</sup>	23.800 ± 0.658 <sup>bA</sup>	18.77 ± 0.491 <sup>cA</sup>	ND <sup>1</sup>	ND	ND
Phytofluene	ND	ND	ND	ND	ND	ND
Lycopene	21.800 ± 0.170 <sup>aB</sup>	13.950 ± 0.210 <sup>bB</sup>	11.050 ± 0.780 <sup>cB</sup>	ND	ND	ND
δ-carotene	ND	ND	ND	ND	ND	ND
α-carotene	0.032 ± 0.003 <sup>cD</sup>	0.070 ± 0.010 <sup>bEF</sup>	0.201 ± 0.012 <sup>aE</sup>	0.023 ± 0.003 <sup>cF</sup>	0.013 ± 0.001 <sup>dE</sup>	0.011 ± 0.001 <sup>dF</sup>
Zeinoxanthin	ND	ND	ND	ND	ND	ND
Lutein	0.214 ± 0.002 <sup>dD</sup>	0.287 ± 0.004 <sup>cDEF</sup>	0.710 ± 0.004 <sup>aD</sup>	0.314 ± 0.019 <sup>bC</sup>	0.138 ± 0.001 <sup>eC</sup>	0.120 ± 0.009 <sup>fC</sup>
γ-carotene	1.905 ± 0.035 <sup>bC</sup>	4.720 ± 0.227 <sup>aC</sup>	1.192 ± 0.006 <sup>cC</sup>	ND	ND	ND
β-carotene	0.179 ± 0.012 <sup>cD</sup>	0.537 ± 0.001 <sup>bD</sup>	0.712 ± 0.005 <sup>aD</sup>	0.153 ± 0.012 <sup>dE</sup>	0.138 ± 0.002 <sup>eC</sup>	0.074 ± 0.005 <sup>fD</sup>
β-Cryptoxanthin	ND	ND	ND	0.134 ± 0.005 <sup>aE</sup>	0.068 ± 0.002 <sup>bD</sup>	0.041 ± 0.001 <sup>cE</sup>
Zeaxanthin	0.160 ± 0.003 <sup>dD</sup>	0.179 ± 0.006 <sup>cDEF</sup>	0.268 ± 0.003 <sup>aE</sup>	0.207 ± 0.009 <sup>bD</sup>	0.127 ± 0.010 <sup>eC</sup>	0.079 ± 0.003 <sup>fD</sup>
Antheraxanthin	ND	ND	ND	ND	ND	ND
Violaxanthin	0.325 ± 0.010 <sup>eD</sup>	0.459 ± 0.011 <sup>dDE</sup>	0.973 ± 0.008 <sup>bCD</sup>	1.450 ± 0.133 <sup>aA</sup>	0.809 ± 0.016 <sup>bB</sup>	0.468 ± 0.046 <sup>bB</sup>
Neoxanthin	0.175 ± 0.017 <sup>eD</sup>	0.377 ± 0.047 <sup>dDEF</sup>	1.270 ± 0.099 <sup>aC</sup>	1.205 ± 0.091 <sup>aB</sup>	0.849 ± 0.019 <sup>bA</sup>	0.740 ± 0.007 <sup>cA</sup>
Apocarotenal	0.003 ± 0.000 <sup>cD</sup>	0.012 ± 0.001 <sup>aF</sup>	0.008 ± 0.000 <sup>bE</sup>	ND	ND	ND
Astaxanthin	ND	ND	ND	ND	ND	ND
Capsanthin	ND	ND	ND	ND	ND	ND
Capsorubin	ND	ND	ND	ND	ND	ND

ND: not detected. Lowercase letters (a, b, c, d, e, f) indicate significant level of carotenoid components in different samples. Upper case letters (A, B, C, D, E, F) indicate significant level of different carotenoid components in the same sample.

In the study of carotenoid metabolites, components of the carotenoids in most species are already well understood, such as in carrot [12], tomato [13–15], pepper [16,17] and orange cauliflower and orange heading Chinese cabbage [18,19]. Additionally, Hadjipieri et al. [21] also indicated that trans-lutein and trans-β-carotene were the major carotenoids in the peel of loquat fruit, and trans-β-cryptoxanthin, followed by trans-β-carotene and 5,8-epoxy-β-carotene, to be the most predominant carotenoids in the flesh of loquat fruit. Xu et al. [24] reported significant accumulation of antheraxanthin, zeaxanthin, neoxanthin and β-cryptoxanthin in orange zucchini. Our findings were different from previous studies in most vegetable crops, as we found that the high accumulation of lycopene and γ-carotene made the turnip develop a yellow-fleshed root.

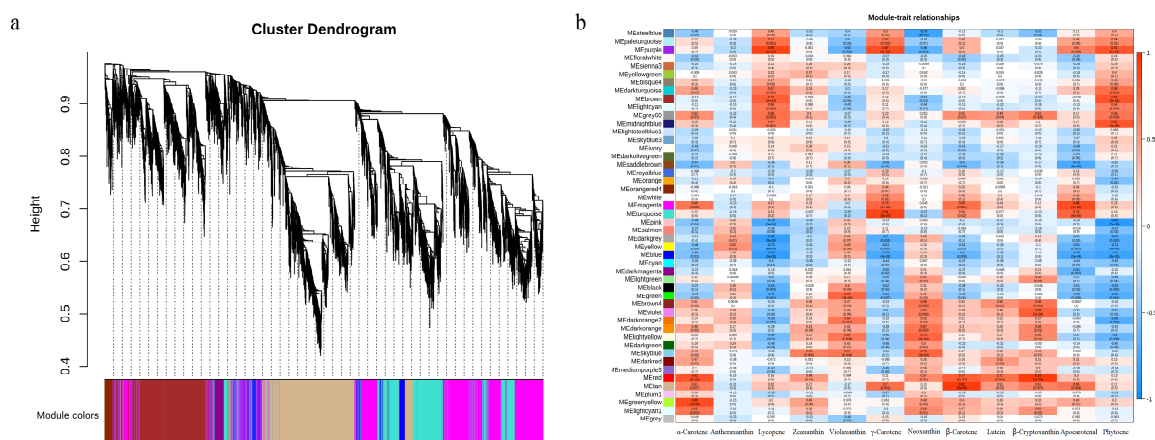
The trends of the carotenoid component content changes were different. There were three trends of the content changes of carotenoid compounds detected from stage 1 to stage 3 in W25. The content of phytoene and lycopene decreased, the content of α-carotene, lutein, β-carotene, zeaxanthin, violaxanthin and neoxanthin increased and the content of γ-carotene and apocarotenal increased from Y1 to Y2 and decreased from Y2 to Y3. Of these carotenoid compounds, the content of phytoene at stage one (Y1) was higher than others, reaching 37.050 ± 0.920 mg/100 g DW, which was predominant. The content of apocarotenal at stage one (Y1) was lower than others, reaching 0.003 ± 0.000 mg/100 g DW. β-Cryptoxanthin was detected only in W21 and all carotenoid compounds detected in W21 maintained a similar change trend, continuously declining.

### 3.5. Weighted Gene Co-Expression Network and Module-Carotenoid Metabolite Correlation Analysis

In order to identify the genes related to the carotenoid components, the weighted gene co-expression network analysis (WGCNA) was performed to investigate the co-expression networks of DEGs and a total of 48 co-expression modules were identified based on their similar expression patterns (Figure 3a). The heatmap of module-carotenoid correlations was also created and the results showed that the accumulation of transcripts for the purple module was correlated with carotenoid metabolites, including phytoene, lycopene and γ-carotene ( $r > 0.8$ ). The accumulation of transcripts for the brown module was correlated with phytoene ( $r = 0.82$ ) and the accumulation of transcripts for the turquoise module was



correlated with  $\gamma$ -carotene ( $r = 0.91$ ) (Figure 3b), which suggest that unigenes clustering in these modules may be closely related to the formation of yellow root in W25.

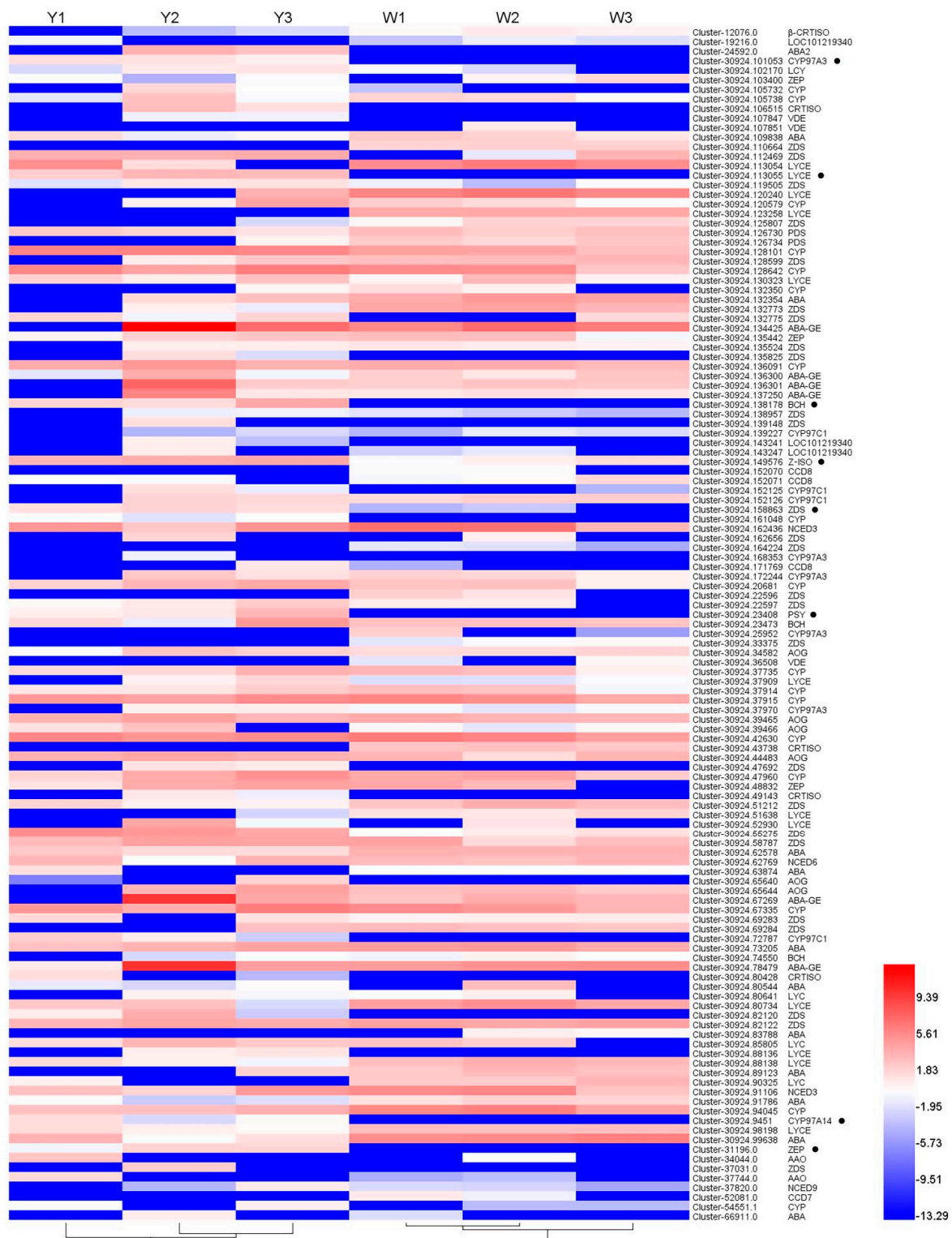


**Figure 3.** Transcriptomic and metabolic correlation analysis in turnip. (a) Forty-eight co-expression modules identified by weighted gene co-expression network analysis. (b) The heatmap of module-carotenoid correlations.

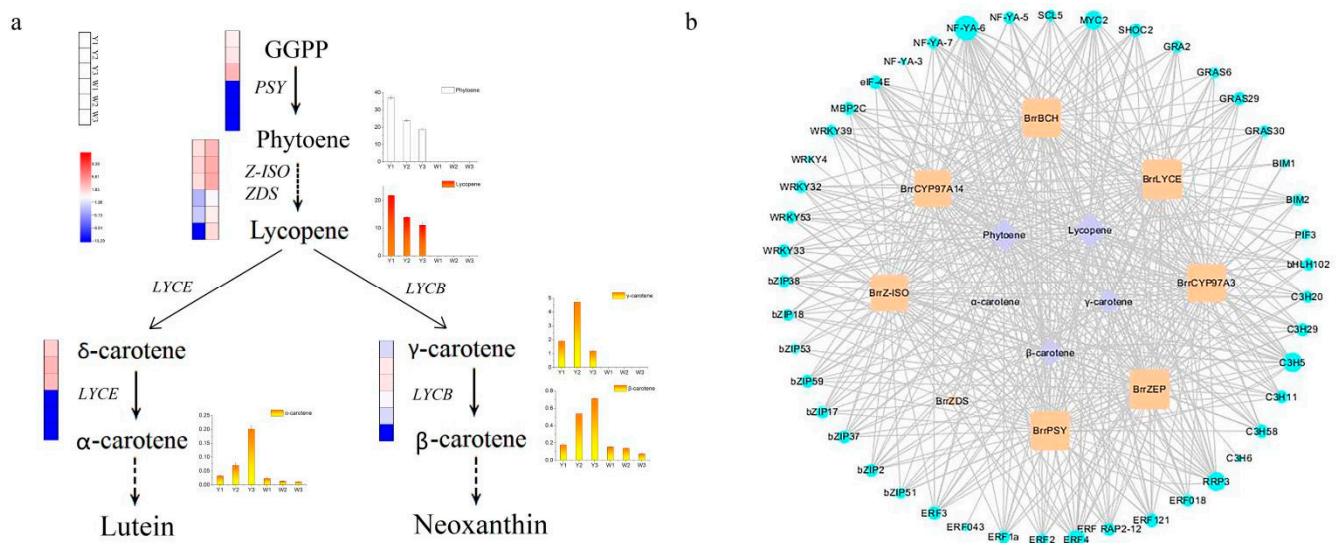
### 3.6. Genetic Basic of Carotenoid Metabolites

Combining analysis of DEGs and carotenoid metabolites,  $|\text{correlation coefficient}| > 0.8$  was considered to represent the gene that regulated the metabolites. In order to identify the genes related to the carotenoid components, the genes involved in carotenoid biosynthesis pathways were analyzed. In total, 125 DEGs (Table S10) were annotated as involved in the carotenoid biosynthesis pathways and the expression levels of these unigenes are shown in Figure 4. Unigene expression level results showed that eight unigenes, *CYP97A3* (Cluster-30924.101053), *Lycopene  $\epsilon$ -cyclase (LYCE)* (Cluster-30924.113055),  *$\beta$ -carotene 3-hydroxylase (BCH)* (Cluster-30924.138178),  *$\zeta$ -carotene isomerase (Z-ISO)* (Cluster-30924.149576),  *$\zeta$ -carotene desaturase (ZDS)* (Cluster-30924.158863), *phytoene synthase (PSY)* (Cluster-30924.23408), *CYP97A14* (Cluster-30924.9451) and *ZEP* (Cluster-31196.0), showed higher expression levels in W25, while they were mostly not expressed in W21, at three developmental stages.

Based on these differentially expressed structural genes and differential accumulation of carotenoid metabolites in the carotenoid biosynthesis pathway, we reconstructed the carotenoid biosynthesis pathway and predicted the molecular mechanisms resulting in the yellow-fleshed root coloration in turnip (Figure 5a). Five carotenoid metabolites, phytoene, lycopene,  $\alpha$ -carotene,  $\beta$ -carotene and  $\gamma$ -carotene, were highly accumulated in W25 and were used to reconstruct the carotenoid biosynthesis pathway. Of the above differentially expressed structural unigenes, *PSY* is the first enzyme to produce the colorless carotenoid 15-cis-phytoene in the carotenoid biosynthetic pathway [54,55]. Cluster-30924.23408, annotated as *PSY*, was highly expressed in W25, while it was mostly not expressed in W21. Similarity, phytoene accumulated in W25, while it was not detected in W21. The expression pattern of the *PSY* gene was consistent with the accumulation of 15-cis-phytoene in the two turnips, which explained the lack of phytoene in W21. The colorless phytoene was converted to the red-colored all-trans-lycopene via a series of enzymes, including phytoene desaturase (*PDS*),  $\zeta$ -carotene isomerase (*Z-ISO*),  $\zeta$ -carotene desaturase (*ZDS*) and carotenoid isomerase (*CRTISO*) [8,56]. However, the deficiency of the 15-cis-phytoene in W21 meant that the downstream lycopene and  $\gamma$ -carotene were also not detected in W21. The contents of downstream carotenoid metabolites were very interesting. Eight downstream carotenoid metabolites were detected in W21 except apocarotenal, which suggests that metabolites in other biosynthetic pathways may influence the formation of downstream carotenoid metabolites.



**Figure 4.** The  $\text{Log}_2\text{FPKM}$  heatmap of 125 differentially expressed unigenes involved in the carotenoid biosynthesis pathways. ● indicates differentially expressed unigenes involved in the carotenoid biosynthesis pathways with higher expression level in yellow-fleshed turnip.



**Figure 5.** The regulatory network of key differentially expressed genes in yellow-fleshed turnip. (a) The regulatory network of carotenoid accumulation in yellow-fleshed turnip. *PSY*, phytoene synthase; *Z-ISO*,  $\zeta$ -carotene isomerase; *ZDS*,  $\zeta$ -carotene desaturase; *LYCE*, lycopene  $\epsilon$ -cyclase; *LYCB*, lycopene  $\beta$ -cyclase. The differentially expressed gene changes were represented by the  $\log_2$ FPKM. The differential carotenoid metabolite changes are shown in the column diagram. (b) Purple circles represent carotenoid metabolites. Yellow blocks represent structural genes involved in carotenoid biosynthesis pathway. Blue circles represent different families of transcription factors whose expression was highly correlated with the above five carotenoid metabolites. The size of the circle and block represents the number of correlations.

### 3.7. Identification of Key Transcription Factors Regulating Sugar Metabolism

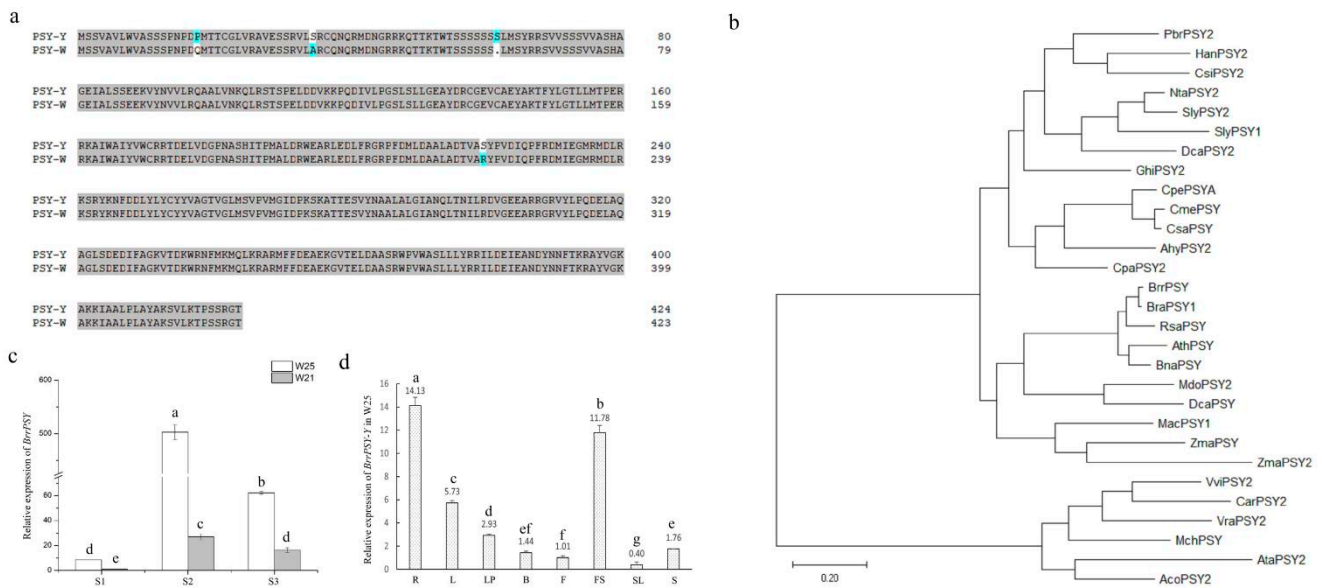
In order to predict the transcription factors involved in the carotenoids biosynthetic pathways, 45 differentially expressed TFs whose expressions were highly correlated with the above five carotenoid metabolites and eight carotenoid biosynthetic pathway structural unigenes were shown to form a correlated network (Figure 5b). All these transcription factors showed lower expression in W21 than in W25, including four *NF-YA* genes, five *WRKY* genes, eight *bZIP* genes, eight *ERF* genes, six *zinc finger CCCH domain-containing protein (C3H)* genes and 14 other TFs. Results suggested that these TFs may correspond to the putative regulators controlling carotenoids biosynthesis.

Based on carotenoid metabolism in tomato fruit, Wu et al. [28] and Shi et al. [29] identified an R2R3-MYB transcription factor, *SIMYB72*, and demonstrated that the *SIMYB72*-interacting protein *SIZHD17*, which belongs to the zinc-finger homeodomain transcription factor family, also functions in carotenoid metabolism. They also found that the expressions of *PSY1* and *Z-ISO* were suppressed due to the regulation by *SIZHD17*, which decreased the lycopene content in fruits. In addition, Mannen et al. [57] reported that *phytochrome-interacting factor 5 (PIF5)*, a basic helix–loop–helix family transcription factor, positively regulates expression of key enzymatic genes in the MEP pathway and improves the accumulation of chlorophyll and carotenoids in cultured cells. We also identified a basic helix–loop–helix family transcription factor, *PIF3*, which was highly correlation with  $\gamma$ -carotene ( $R = 0.974$ ). The possible regulation mechanism of *PIF3* involved in the control of carotenoid metabolism in turnip root will investigated in the future. All these findings provided new insight into deep exploration of the formation of yellow turnips.

### 3.8. Cloning and Analysis of Candidate Gene Sequence

Based on the reconstructed carotenoid biosynthesis pathway in the yellow-fleshed root coloration of turnip, the *PSY* gene was considered as the candidate gene controlling the yellow-fleshed root coloration. In order to compare the *PSY* gene sequence between

W21 and W25, the sequences of *PSY* were cloned, sequenced and named as *PSY-W* and *PSY-Y* for W21 and W25, respectively. The full length gDNA sequence of *PSY-W* and *PSY-Y* was 2086 bp and 2106 bp, cDNA sequence was 1275 bp and 1278 bp, respectively. Referring to the *PSY* gene sequence of Chinese cabbage in the BRAD database, the *BrrPSY* gene in turnip contains six exons and five introns and *PSY-W* in turnip was totally consistent with the *PSY* in Chinese cabbage. Sequence alignment revealed four mutations in the *PSY-Y* amino acid sequence compared with *PSY-W*, including mutations of proline to glutamine (P-Q) at the 18th amino acid position, serine to alanine (S-A) at the 35th amino acid position, a serine (S) insertion at the 62nd amino acid position and serine to arginine (S-R) at the 220th amino acid position (Figure 6a).



**Figure 6.** The analysis of candidate gene controlling the yellow-fleshed root coloration in turnip. (a) Sequence alignment of *PSY-Y* and *PSY-W* amino acid sequence. (b) Phylogenetic relationship between *BrrPSY* and other *PSY* genes. *Vigna angularis* (*VraPSY2*: XM\_017575246.1); *B. rapa* subsp. *pekinensis* (*BraPSY1*: FJ227935.2); *Nicotiana tabacum* (*NtaPSY2*: NM\_001325140.1); *Ananas comosus* (*AcoPSY2*: XM\_020224223.1); *Musa acuminata* (*MacPSY1*: JX195664.1); *Arachis duranensis* (*AhyPSY2*: XM\_016076234.2); *Helianthus annuus* (*HanPSY2*: AJ304825.1); *Zea mays* (*ZmaPSY1*: NM\_001114652.2; *ZmaPSY2*: EU958083.1); *Cucurbita pepo* (*CpePSYA*: JX912285.1); *Momordica charantia* (*MchPSY*: XM\_022285988); *Aegilops tauschii* subsp. *Tauschii* (*AtaPSY2*: XM\_020291471); *Coffea arabica* (*CarPSY2*: XM\_027218762); *Cucumis sativus* (*CsaPSY*: XM\_004148859); *Daucus carota* subsp. *Sativus* (*DcaPSY1*: NM\_001329177.1; *DcaPSY2*: DQ192187); *Gossypium hirsutum* (*GhiPSY2*: XM\_016825070.1); *B. napus* (rape) (*BnaPSY*: NM\_001316292.1); *Solanum lycopersicum* (*SlyPSY1*: NM\_001347838.1; *SlyPSY2*: NM\_001247742.2); *Carica papaya* (*CpaPSY2*: DQ666828.1); *Camellia sinensis* (*CsiPSY2*: KM519981.1); *Vitis riparia* (*VviPSY2*: XM\_034827956.1); *Raphanus sativus* (*RsaPSY*: XM\_018588986.1); *Arabidopsis thaliana* (*AthPSY*: BT000450.1); *Cucumis melo* (*CmePSY*: Z37543.1); *Malus domestica* (*MdoPSY2*: NM\_001294092.1); *Pyrus bretschneideri* (*PbrPSY2*: XM\_009343584.2). (c) Relative expression of *BrrPSY* gene at different development stages in turnip root. (d) Relative expression of *BrrPSY-Y* gene in different tissues of turnip in yellow-fleshed W25. Lowercase letters in (c,d) indicate the significance levels.

To clarify the evolutionary relationship between PSY protein sequences in turnip and in other species, twenty-eight PSY proteins sequences were downloaded from the National Center for Biotechnology Information and a phylogenetic tree was constructed by MEGA 6.0 software. Results revealed that PSY in turnip (*BrrPSY*) and PSY in Brassicaceae species were clustered into the same clade, and there was a closer genetic relationship with *BraPSY* in Chinese cabbage and *RsaPSY* in *Raphanus sativus* rather than *AthPSY* in *Arabidopsis thaliana* and *BnaPSY* in *B. napus* (Figure 6b).

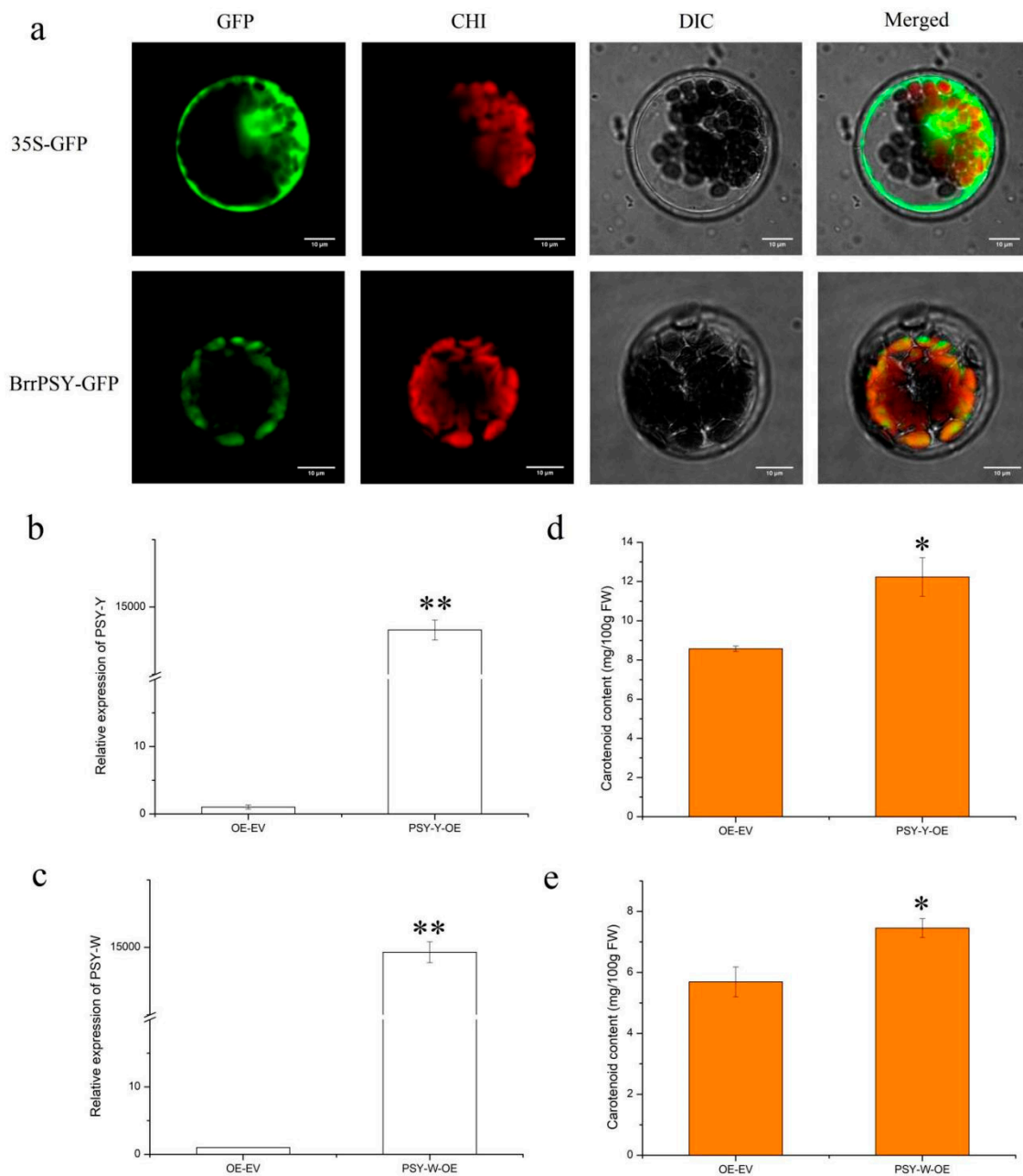
Relative expression levels of the *BrrPSY* gene in turnips at different root development stages in W21 and W25 and the *BrrPSY-Y* gene in W25 in the vegetative growth period (leaf, leaf petiole, fleshed root) and reproductive growth period (bud, flower, stem leaf, flower stem and seed) were measured by RT-qPCR. Results of relative expression levels of the *BrrPSY* gene in turnips at different root development stages showed that the expression levels of *BrrPSY* in W25 were evidently higher than in W21 at three stages, expression levels at S2 and S3 stages were higher than in S1 and *BrrPSY* expression levels at the S2 stage in W25 were the highest (Figure 6c). This result was consistent with transcript accumulation by RNA-seq. Results of the *BrrPSY-Y* gene in W25 in different tissues revealed that the *BrrPSY-Y* gene was expressed in various tissues at different levels. The *BrrPSY-Y* gene had the highest expression in the root, and the relative expression of each tissue was as follows: root > flower stem > leaf > petiole > seeds > bud > flower > stem leaf (Figure 6d).

### 3.9. Subcellular Localization and Functional Analysis of *BrrPSY* Gene in Turnip

In order to clear the *BrrPSY* protein localization in cells, we performed assays for the transient transformation of *Arabidopsis* protoplasts with PBI211GFP-*BrrPSY* recombinant plasmids and a PBI211GFP plasmid was used as control. Fluorescence of transformed *Arabidopsis* protoplasts showed that the *BrrPSY* protein appeared as green fluorescence at 488 nm, and the chloroplast in *Arabidopsis* protoplasts appeared as red fluorescence at 555 nm. The merged field showed green and red fluorescence presenting the superimposed orange, indicating that *BrrPSY* protein was localized in the chloroplasts (Figure 7a).

To test the ability of *BrrPSY* to impact the accumulation of carotenoid content in vivo, we generated over-expressing constructs targeting the *PSY-Y* and *PSY-W* genes. For this purpose, *Agrobacterium tumefaciens* strain GV3101 containing recombinant plasmid 35S::*PSY-Y* or 35S::*PSY-W* was injected into a tobacco line using the injection infiltration method, and corresponding empty plasmids were used as control. The relative expression of *PSY-Y* and *PSY-W* was measured by qRT-PCR analysis and the results showed that *PSY-Y* and *PSY-W* were both significantly up-regulated in the transient expression tobacco (Figure 7b,c). Similarly, the carotenoid contents were increased in both the *PSY-Y* and *PSY-W* over-expressing tobacco (Figure 7d,e), confirming a direct regulation of *BrrPSY* on carotenoid content in turnip. Results revealed that yellow turnips are due to high expression of the *PSY* gene rather than mutations in the *PSY* gene, indicating that a posttranscriptional regulatory mechanism may affect carotenoid formation.

Since PSY catalyzes the first committed reaction, it is considered to be the major rate-limiting enzyme regulating the flux into carotenoid production [37,38], while there are few reports on the function of a mutated *PSY* gene. Our research confirmed the mutated *BrrPSY* gene in carotenoid accumulation and validated the function of *PSY-Y* and *PSY-W* genes. All these findings provide new insight into deep exploration of the formation of yellow turnips.



**Figure 7.** Subcellular localization and functional analysis of *BrrPSY* gene. (a) The subcellular localization of BrrPSY protein. (b,c) The relative expression levels of *PSY-Y* and *PSY-W* genes in the over-expressing tobacco leaves and control. (d,e) The carotenoid contents in the over-expressing tobacco leaves and control. \* indicates a significant difference when  $p < 0.05$ , \*\* indicates a significant difference when  $p < 0.01$ .

#### 4. Conclusions

In this study, we identified eight unigenes involved in carotenoid biosynthesis and that were significantly upregulated in W25 and barely expressed in W21. Metabolomics profiling in turnip revealed that four carotenoid components, phytoene, lycopene,  $\gamma$ -carotene and apocarotenal, were unique in W25. Thus, we hypothesized high accumulation of lycopene and  $\gamma$ -carotene in W25 to cause the yellow-fleshed root coloration in turnip. A carotenoid biosynthesis pathway was reconstructed and forty-five transcription factors were identified to predict the yellow-fleshed root coloration in turnip. The *PSY* gene was the key gene affecting the carotenoid formation in W25. The coding sequence of BrrPSY-W25 was 1278 bp

and BrrPSY-W21 was 1275 bp, and BrrPSY was more highly expressed in swollen roots in W25 than in W21. Transient transgenic tobacco leaf overexpressing BrrPSY-W and BrrPSY-Y showed higher transcript levels and carotenoid contents. Results revealed that yellow turnip formation is due to high expression of the *PSY* gene rather than mutations in the *PSY* gene, indicating that a posttranscriptional regulatory mechanism may affect carotenoid formation.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13060953/s1>, Table S1: All primers used in this paper. Table S2: The summary of the transcriptome sequencing dataset and quality control; Table S3: The size and distribution of transcripts and unigene length data; Table S4: Coding sequence length and predicted protein sequence analysis data; Table S5: Functional annotation of assembled unigenes in turnip; Table S6: Number of unigenes annotated in seven databases; Table S7: Gene Ontology classification of unigenes in turnips; Table S8: EuKaryotic Ortholog Groups classification of unigenes in turnips; Table S9: The identified differentially expressed TFs in two turnips at three development stages; Table S10: The FPKM values of 125 differentially expressed unigenes involved in the carotenoid biosynthesis pathways; Figure S1: The column diagram of transcripts, unigene, coding sequences and predicted protein sequence size and distribution in turnips.

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

**Funding:** This research was funded by the project of key research and transformation in Qinghai Province (2021-NK-124), and Qinghai Key Laboratory of Vegetable Genetics and Physiology (2020-ZJ-Y02).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The dataset is available from the NCBI Short Read Archive (SRA) under accession number PRJNA645708.

**Acknowledgments:** The authors would like to thank Wuhan Metware Biotechnology Co., Ltd. for transcript sequencing and carotenoid metabolite detection.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

1. Zhuang, H.; Lou, Q.; Liu, H.; Han, H.; Wang, Q.; Tang, Z.; Ma, Y.; Wang, H. Differential Regulation of Anthocyanins in Green and Purple Turnips Revealed by Combined De Novo Transcriptome and Metabolome Analysis. *Int. J. Mol. Sci.* **2019**, *20*, 4387. [[CrossRef](#)] [[PubMed](#)]
2. Rochfort, S.J.; Imsic, M.; Jones, R.; Trenerry, V.C.; Tomkins, B. Characterization of flavonol conjugates in immature leaves of pak choi (*Brassica rapa* L. ssp. *chinensis* L. (Hanelt.)) by HPLC-DAD and LC-MS/MS. *J. Agric. Food Chem.* **2006**, *54*, 4855–4860. [[CrossRef](#)]
3. Zhang, H.; Schonhof, I.; Krumbein, A.; Gutezeit, B.; Li, L.; Stützel, H.; Schreiner, M. Water supply and growing season influence glucosinolate concentration and composition in turnip root (*Brassica rapa* ssp. *rapifera* L.). *J. Plant Nutr. Soil Sci.* **2008**, *171*, 255–265. [[CrossRef](#)]
4. Chung, I.M.; Rekha, K.; Rajakumar, G.; Thiruvengadam, M. Production of glucosinolates, phenolic compounds and associated gene expression profiles of hairy root cultures in turnip (*Brassica rapa* ssp. *rapa*). *3 Biotech* **2016**, *6*, 175. [[CrossRef](#)] [[PubMed](#)]
5. Fahey, J.W.; Talalay, P.; Gustine, D.I.; Flores, H.E. The role of crucifers in cancer chemoprotection. In *Phytochemicals and Health*; American Society of Plant Physiologist: Rochville, MD, USA, 1995; pp. 87–93.
6. Thiruvengadam, M.; Chung, I.M. Selenium, putrescine, and cadmium influence health-promoting phytochemicals and molecular-level effects on turnip (*Brassica rapa* ssp. *rapa*). *Food Chem.* **2015**, *173*, 185–193. [[CrossRef](#)] [[PubMed](#)]
7. Li, Y.; Han, J.; Zhang, Y.; Chen, Y.; Zhang, Y. Prophylactic effect and mechanism of p-coumaric acid against hypoxic cerebral edema in mice. *Resp. Physiol. Neurobi.* **2019**, *260*, 95–104. [[CrossRef](#)] [[PubMed](#)]
8. Yuan, H.; Zhang, J.; Nageswaran, D.; Li, L. Carotenoid metabolism and regulation in horticultural crops. *Hortic. Res.* **2015**, *2*, 15036. [[CrossRef](#)]

9. Fraser, P.D.; Bramley, P.M. The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* **2004**, *43*, 228–265. [[CrossRef](#)]
10. Fiedor, J.; Burda, K. Potential role of carotenoids as antioxidants in human health and disease. *Nutrients* **2014**, *6*, 466–488. [[CrossRef](#)]
11. Hirschberg, J. Carotenoid biosynthesis in flowering plants. *Curr. Opin. Plant Biol.* **2001**, *4*, 210–218. [[CrossRef](#)]
12. Rodriguez-Concepcion, M.; Stange, C. Biosynthesis of carotenoids in carrot: An underground story comes to light. *Arch. Biochem. Biophys.* **2013**, *539*, 110–116. [[CrossRef](#)]
13. Tadmor, Y.; King, S.; Levi, A.; Davis, A.; Meir, A.; Wasserman, B.; Hirschberg, J.; Lewinsohn, E. Comparative fruit colouration in watermelon and tomato. *Food Res. Int.* **2005**, *38*, 837–841. [[CrossRef](#)]
14. Ronen, G.; Cohen, M.; Zamir, D.; Hirschberg, J. Regulation of carotenoid biosynthesis during tomato fruit development: Expression of the gene for *lycopene epsiloncyclase* is down-regulated during ripening and is elevated in the mutant Delta. *Plant J.* **1999**, *17*, 341–351. [[CrossRef](#)]
15. Ronen, G.; Carmel-Goren, L.; Zamir, D.; Hirschberg, J. An alternative pathway to betacarotene formation in plant chromoplasts discovered by map-based cloning of beta and old-gold color mutations in tomato. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11102–11107. [[CrossRef](#)]
16. Rodriguez-Urbe, L.; Guzman, I.; Rajapakse, W.; Richins, R.D.; O’Connell, M.A. Carotenoid accumulation in orange-pigmented *Capsicum annuum* fruit, regulated at multiple levels. *J. Exp. Bot.* **2012**, *63*, 517–526. [[CrossRef](#)]
17. Ha, S.H.; Kim, J.B.; Park, J.S.; Lee, S.W.; Cho, K.J. A comparison of the carotenoid accumulation in *capsicum* varieties that show different ripening colours: Deletion of the *capsanthin-capsorubin synthase* gene is not a prerequisite for the formation of a yellow pepper. *J. Exp. Bot.* **2007**, *58*, 3135–3144. [[CrossRef](#)]
18. Li, L.; Paolillo, D.J.; Parthasarathy, M.V.; Dimuzio, E.M.; Garvin, D.F. A novel gene mutation that confers abnormal patterns of b-carotene accumulation in cauliflower (*Brassica oleracea* var. *botrytis*). *Plant J.* **2001**, *26*, 59–67. [[CrossRef](#)]
19. Su, T.; Yu, S.; Zhang, J.; Yu, Y.; Zhang, D.; Zhao, X.; Wang, W. Loss of function of the carotenoid isomerase gene BrCRTISO confers orange color to the inner leaves of Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Plant Mol. Biol. Rep.* **2014**, *33*, 648–659. [[CrossRef](#)]
20. Beck, G.; Coman, D.; Herren, E.; Ruiz-Sola, M.A.; Rodriguez-Concepcion, M.; Gruissem, W.; Vranova, E. Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* **2013**, *82*, 393–416. [[CrossRef](#)]
21. Hadjipieri, M.; Georgiadou, E.C.; Marin, A.; Diaz-Mula, H.M.; Goulas, V.; Fotopoulos, V.; Tomás-Barberán, F.A.; Manganaris, G.A. Metabolic and transcriptional elucidation of the carotenoid biosynthesis pathway in peel and flesh tissue of loquat fruit during on-tree development. *BMC Plant Biol.* **2017**, *17*, 102. [[CrossRef](#)]
22. Rubio, A.; Rambla, J.L.; Santaella, M.; Gómez, M.D.; Orzaez, D.; Granell, A.; Gómez-Gómez, L. Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in b-ionone release. *J. Biol. Chem.* **2008**, *283*, 24816–24825. [[CrossRef](#)] [[PubMed](#)]
23. Frusciante, S.; Diretto, G.; Bruno, M.; Ferrante, P.; Pietrella, M.; Prado-Cabrero, A.; Rubio-Moraga, A.; Beyer, P.; Gomez-Gomez, L.; Al-Babili, S.; et al. Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12246–12251. [[CrossRef](#)]
24. Xu, X.; Lu, X.; Tang, Z.; Zhang, X.; Lei, F.; Hou, L.; Li, M. Combined analysis of carotenoid metabolites and the transcriptome to reveal the molecular mechanism underlying fruit colouration in zucchini (*Cucurbita pepo* L.). *Food Chem. Mol. Sci.* **2021**, *2*, 100021. [[CrossRef](#)] [[PubMed](#)]
25. Arango, J.; Jourdan, M.; Geoffriau, E.; Beyer, P.; Welsch, R. Carotenoid hydroxylase activity determines the levels of both a-carotene and total carotenoids in orange carrots. *Plant Cell* **2014**, *26*, 2223–2233. [[CrossRef](#)] [[PubMed](#)]
26. Zhou, X.; Welsch, R.; Yang, Y.; Álvarez, D.; Riediger, M.; Yuan, H.; Fish, T.; Liu, J.; Thannhauser, T.W.; Li, L. Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3558–3563. [[CrossRef](#)] [[PubMed](#)]
27. Zhang, J.; Yuan, H.; Fei, Z.; Pogson, B.J.; Zhang, L.; Li, L. Molecular characterization and transcriptome analysis of orange head Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Planta* **2015**, *241*, 1381–1394. [[CrossRef](#)] [[PubMed](#)]
28. Wu, M.; Xu, X.; Hu, X.; Liu, Y.; Cao, H.; Chan, H.; Gong, Z.; Yuan, Y.; Luo, Y.; Feng, B.; et al. SIMYB72 Regulates the Metabolism of Chlorophylls, Carotenoids, and Flavonoids in Tomato Fruit. *Plant Physiol.* **2020**, *183*, 854–868. [[CrossRef](#)] [[PubMed](#)]
29. Shi, Y.; Pang, X.; Liu, W.; Wang, R.; Su, D.; Gao, Y.; Wu, M.; Deng, W.; Liu, Y.; Li, Z. SIZHD17 is involved in the control of chlorophyll and carotenoid metabolism in tomato fruit. *Hortic. Res.* **2021**, *8*, 259. [[CrossRef](#)]
30. Conesa, A.; Madrigal, P.; Tarazona, S.; Gomez-Cabrero, D.; Cervera, A.; McPherson, A.; Szczesniak, M.W.; Gaffney, D.J.; Elo, L.L.; Zhang, X.; et al. A Survey of Best Practices for RNA-Seq Data Analysis. *Genome Biol.* **2016**, *17*, 13. [[CrossRef](#)]
31. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full length transcriptome assembly from RNA Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
32. Yu, R.; Liang, X.; Wei, Z.; Wang, Y.; Luo, X.; Wang, R.; Zhu, X.; Xie, Y.; Karanja, B.; Liu, L. De novo taproot transcriptome sequencing and analysis of major genes involved in sucrose metabolism in Radish (*Raphanus sativus* L.). *Front. Plant Sci.* **2016**, *7*, 585. [[CrossRef](#)]
33. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]



34. Deng, Y.Y.; Li, J.Q.; Wu, S.F.; Zhu, Y.P.; Chen, Y.W.; He, F.C. Integrated nr Database in Protein Annotation System and Its Localization. *Comput. Eng.* **2006**, *32*, 71–74.
35. Apweiler, R.; Bairoch, A.; Wu, C.H.; Barker, W.C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; et al. UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115–D119. [[CrossRef](#)]
36. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A. The COG database: A tool for genome scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [[CrossRef](#)]
37. Koonin, E.V.; Fedorova, N.D.; Jackson, J.D.; Jacobs, A.R.; Krylov, D.M.; Makarova, K.S.; Mazumder, R.; Mekhedov, S.L.; Nikolskaya, A.N.; Rao, B.S.; et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **2004**, *5*, R7. [[CrossRef](#)]
38. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
39. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)]
40. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
41. Varet, H.; Brillet-Guéguen, L.; Coppée, J.Y.; Dillies, M.A. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS ONE* **2016**, *11*, e0157022. [[CrossRef](#)]
42. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **2010**, *11*, R14. [[CrossRef](#)]
43. Xie, C.; Mao, X.; Huang, J.; Ding, Y.; Wu, J.; Dong, S.; Kong, L.; Gao, G.; Li, C.; Wei, L. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **2011**, *39*, W316–W322. [[CrossRef](#)]
44. Inbaraj, H.B.S.; Lu, C.F.; Hung, W.B.; Wu, C.L.; Lin, B.H. Chen Determination of carotenoids and their esters in fruits of *Lycium barbarum* Linnaeus by HPLC–DAD–APCI–MS. *J. Pharmaceut. Biomed.* **2018**, *47*, 812–818. [[CrossRef](#)] [[PubMed](#)]
45. Petry, F.C.; Mercadante, A.Z. New method for carotenoid extraction and analysis by HPLC–DAD–MS/MS in freeze-dried citrus and mango pulps. *J. Brazil Chem. Soc.* **2018**, *29*, 205–215. [[CrossRef](#)]
46. Wang, R.; Shu, P.; Zhang, C.; Zhang, J.; Chen, Y.; Zhang, Y.; Du, K.; Xie, Y.; Li, M.; Ma, T.; et al. Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytol.* **2022**, *233*, 373–389. [[CrossRef](#)] [[PubMed](#)]
47. Ren, Y.J.; He, Q.; Ma, X.; Zhang, L. Characteristics of Color Development in seeds of Brown- and Yellow-seeded Heading Chinese Cabbage and Molecular Analysis of *Brsc* the Candidate Gene Controlling Seed Coat Color. *Front. Plant Sci.* **2017**, *8*, 1410. [[CrossRef](#)] [[PubMed](#)]
48. Wang, X.; Wang, H.; Wang, J.; Sun, R.; Wu, J.; Liu, S.Y.; Bai, Y.; Mun, J.; Bancroft, I.; Cheng, F.; et al. The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **2011**, *43*, 1035–1039. [[CrossRef](#)] [[PubMed](#)]
49. Zhang, L.; Cai, X.; Wu, J.; Liu, M.; Gro, S.; Cheng, F.; Liang, J.; Cai, C.; Liu, Z.; Liu, B.; et al. Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Hortic. Res.* **2019**, *5*, 50. [[CrossRef](#)]
50. Livak, K.; 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](#)]
51. Lin, S.; Miao, Y.; Su, S.; Xu, J.; Jin, L.; Sun, D.; Peng, R.; Huang, L.; Cao, J. Comprehensive analysis of *Ogura* cytoplasmic male sterility-related genes in turnip (*Brassica rapa* ssp. *rapifera*) using RNA sequencing analysis and bioinformatics. *PLoS ONE* **2019**, *14*, e0218029. [[CrossRef](#)]
52. Han, R.; Tian, M.; Zhang, G.; Shao, D.; Ren, Y. Complete chloroplast genome sequence of turnip (*Brassica rapa* ssp. *rapa*): Genome structure and phylogenetic analysis. *Mitochondrial DNA B* **2020**, *5*, 3573–3575. [[CrossRef](#)]
53. Ren, Y. The complete mitochondrial genome of turnip (*Brassica rapa* ssp. *rapa*). *Mitochondrial DNA B* **2021**, *6*, 1566–1567. [[CrossRef](#)]
54. Fantini, E.; Falcone, G.; Frusciante, S.; Giliberto, L.; Giuliano, G. Dissection of tomato lycopene biosynthesis through virus-induced gene silencing. *Plant Physiol.* **2013**, *163*, 986–998. [[CrossRef](#)]
55. Peng, G.; Wang, C.; Song, S.; Fu, X.; Azam, M.; Grierson, D.; Xu, C. The role of 1-deoxy-d-xylulose-5-phosphate synthase and phytoene synthase gene family in citrus carotenoid accumulation. *Plant Physiol. Biochem.* **2013**, *71*, 67–76. [[CrossRef](#)]
56. Fanciullino, A.L.; Dhuique-Mayer, C.; Luro, F.; Morillon, R.; Ollitrault, P. Carotenoid biosynthetic pathway in the Citrus genus: Number of copies and phylogenetic diversity of seven genes. *J. Agric. Food Chem.* **2007**, *55*, 7405–7417. [[CrossRef](#)]
57. Mannen, K.; Matsumoto, T.; Takahashi, S.; Yamaguchi, Y.; Tsukagoshi, M.; Sano, R.; Nakayama, T. Coordinated transcriptional regulation of isopentenyl diphosphate biosynthetic pathway enzymes in plastids by phytochrome-interacting factor 5. *Biochem. Biophys. Res. Commun.* **2014**, *443*, 768–774. [[CrossRef](#)]