

# Comparative Metabolomic And Transcriptomic Analysis Reveals A Coexpression Network of The Carotenoid Metabolism Pathway In The Panicle of *Setaria Italica*

**Hui Li**

College of Agriculture, Institute of Agricultural Bioengineering, Shanxi Agricultural University

**Shangling Han**

College of Agriculture, Institute of Agricultural Bioengineering, Shanxi Agricultural University

**Yiqiong Huo**

College of Agriculture, Institute of Agricultural Bioengineering, Shanxi Agricultural University

**Guifang Ma**

College of Agriculture, Institute of Agricultural Bioengineering, Shanxi Agricultural University

**Zhaoxia Sun**

Shanxi Key Laboratory of Germplasm Innovation and Molecular Breeding of Minor Crop

**Hongying Li**

Shanxi Key Laboratory of Germplasm Innovation and Molecular Breeding of Minor Crop

**Siyu Hou**

Shanxi Key Laboratory of Germplasm Innovation and Molecular Breeding of Minor Crop

**Yuanhuai Han (✉ [swgctd@163.com](mailto:swgctd@163.com))**

Shanxi Key Laboratory of Germplasm Innovation and Molecular Breeding of Minor Crop

---

## Research Article

**Keywords:** Foxtail millet, carotenoid, transcriptomics, metabolome, gene expression level

**Posted Date:** September 20th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-878056/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMC Plant Biology on March 8th, 2022. See the published version at <https://doi.org/10.1186/s12870-022-03467-2>.

# Abstract

**[Background]** The grains of foxtail millet are enriched in carotenoids, which endow this plant with a yellow color and extremely high nutritional value. However, the underlying molecular regulation mechanism and gene coexpression network remain unclear.

**[Methods]** The carotenoid species and content were detected by HPLC for two foxtail millet varieties at three panicle development stages. Based on a homologous sequence BLAST analysis, these genes related to carotenoid metabolism were identified from the foxtail millet genome database. The conserved protein domains, chromosome locations, gene structures and phylogenetic trees were analyzed using bioinformatics tools. RNA-seq was performed for these samples to identify differentially expressed genes (DEGs). A Pearson correlation analysis was performed between the expression of genes related to carotenoid metabolism and the content of carotenoid metabolites. Furthermore, the expression levels of the key DEGs were verified by qRT-PCR. The gene coexpression network was constructed by a weighted gene coexpression network analysis (WGCNA).

**[Result]** The major carotenoid metabolites in the panicles of DHD and JG21 were lutein and  $\beta$ -carotene. These carotenoid metabolite contents sharply decreased during the panicle development stage. The lutein and  $\beta$ -carotene contents were highest at the S1 stage of DHD, with values of 11.474  $\mu\text{g}/100\text{ mg}$  and 12.524  $\mu\text{g}/100\text{ mg}$ , respectively. Fifty-four genes related to carotenoid metabolism were identified in the foxtail millet genome. Cis-acting element analysis showed that these gene promoters mainly contain 'light-responsive' and 'ABA-responsive' elements. In the carotenoid metabolic pathways, *SiHDS*, *SiHMGS3*, *SiPDS* and *SiNCED1* were more highly expressed in the panicle of foxtail millet. The expression of *SiCMT*, *SiAACT3*, *SiPSY1*, *SiZEP1/2*, and *SiCCD8c/8d* was significantly correlated with the lutein content. The expression of *SiCMT*, *SiHDR*, *SiIDI2*, *SiAACT3*, *SiPSY1*, and *SiZEP1/2* was significantly correlated with the content of  $\beta$ -carotene. WGCNA showed that the coral module was highly correlated with lutein and  $\beta$ -carotene, and 13 structural genes from the carotenoid biosynthetic pathway were identified. Network visualization revealed 25 intramodular hub genes that putatively control carotenoid metabolism.

**[Conclusion]** Based on the integrative analysis of the transcriptomics and carotenoid metabolomics, we found that DEGs related to carotenoid metabolism had a stronger correlation with the key carotenoid metabolite content. The correlation analysis and WGCNA identified and predicted the gene regulation network related to carotenoid metabolism. These results lay the foundation for exploring the key target genes regulating carotenoid metabolism flux in the panicle of foxtail millet. We hope that these target genes could be used to genetically modify millet to enhance the carotenoid content in the future.

## Background

Foxtail millet (*Setaria italica* (L.) Beauv.] belongs to the *Setaria* genus of Poaceae grass family, and it is widely planted in Eurasian arid and semiarid areas as a kind of C4 crop that endures drought stress and infertile soils[1]. The ancestor of cultivated foxtail millet is green foxtail grass, which was domesticated early in China at approximately 16000 years ago according to archaeological evidence[2]. At present, four genome sequences of foxtail millet have been published: 'Yugu1', 'Zhanggu', 'Xiaom' and 'huagu11'[3–6]. The construction of a genome database and efficient genetic transformation system lay the foundation for investigating the functional genes, genetic evolution, plant architecture and physiology of foxtail millet, especially as a kind of C4 model and bioenergy grass system[1, 7, 8]. It is an important food crop in China and other Asian countries, and has small grains and higher nutritional value and contains edible fiber, protein, starch, vitamins and mineral elements[9]. Previous studies have shown that dehulled grains of foxtail millet enriched in carotenoids present a yellow color[10, 11]. Carotenoids, as the main source of vitamin A in the human body, present various functions, such as eyesight protection, antioxidation, and anticancer effects, and they also have preventive functions in a variety of cardiovascular diseases[12]. Biofortified carotenoid food could be essential for meeting the health requirements and reduce "hidden hunger" in developing areas. According to previous reports, the carotenoid content in grains of foxtail millet ranged from 189.1

µg/100 mg to 201.3 µg/100 mg, 7.2, 201.3 and 1.7 times that of wheat, brown rice and sorghum, respectively[13–15]. Hence, the grains of foxtail millet could be an excellent carotenoid food resource.

The carotenoid metabolism pathway in plants is well known. Carotenoids and their derivatives are composed of isopentenyl pyrophosphate (IPP) and its double bond isomer and dimethylallyl diphosphate (DMAPP). In plant cells, these IPP precursors are mainly synthesized by the mevalonate (MVA) and 2-C-methyl-D-erythritol 4-phosphate (MEP) pathways and the precursors of carotenoid synthesis mainly come from the MEP metabolic pathway[16, 17]. Finally, carotenoid synthesis begins with phytoene synthase (PSY) condensation of two geranylgeranyl pyrophosphate (GGPP) molecules to form a 15 cis isomer. In the next step, all trans-lycopene is produced by phytene saturase (PDS), zeta-e carotene desaturase (ZDS), carotenoid isomerase (Crtiso), and zeta-carotene isomerase (Z-ISO). β-carotene and α-carotene are synthesized by LCYB or LCYE catalysis, and then lutein and zeaxanthin are produced by hydroxylation of nonheme carotene hydroxylase (BCH1 and BCH2) and heme hydroxylase (cytochrome P450-type monooxygenase, CYP97A3 and CYP97C1). Under the catalysis of zeaxanthin cyclooxygenase (ZEP) and pansy xanthine decyclooxygenase (VDE), luteins are produced, including anther xanthine, viola xanthine and neoxanthine[18] (Fig. 1).

Many previous studies on genes involved in carotenoid biosynthesis and regulated networks in plants have been reported. Through introducing *PSY* and *CRTL* genes in rice, a new edible rice variety named “golden rice” was successfully developed[19]. In maize, they found that the genotype with favorable alleles of *crtRB1* and *lcyE* had a significant effect on the β-carotene content (7.9-fold and 2.1-fold higher) compared to the unfavorable genotype[20]. In tomato, two *PSY* isoforms are responsible for divergent functions of fruit-specific carotenoid accumulation[21]. For foxtail millet, although previous reports on carotenoid components and some key gene expression analyses have been reported, the global analysis of gene expression patterns involved in carotenoid metabolism during the panicle developmental stage has not yet been reported. Therefore, in this study, we proposed to explore the dynamic pattern of carotenoid accumulation and the relationship between carotenoid content and gene expression level during panicle development periods to deeply understand the underlying molecular mechanism of carotenoid metabolism and identify target genes for genetic modification in foxtail millet.

## Methods

### Plant materials

Foxtail millet varieties JG21 (Shanxi Academy of Agricultural Sciences, the Industrial Crop Institute) and DHD (Center for Crop Germplasm Resources, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences) were planted in the experimental field of Shanxi Agricultural University in April 2019 (N: 37°12', E: 112°28'). Samples were collected from the middle part of the panicle at S1 (grain-filling stage, 21 days after heading), S2 (middle and late filling stage, 39 days after heading) and S3 (grain maturation period, 53 days after heading) (Figure S1). At each stage, 10 individual panicles were collected, mixed and stored at -80 °C for the following analysis.

### Extraction and determination of carotenoids

The extraction of carotenoids followed Paul's method[22]. In brief, a fine powder generated from each sample (approximately 0.5 g) was added to 40 µl 50% KOH solution and 2 ml anhydrous ethanol with 0.1% BHT (butylated hydroxytoluene). After vortex mixing, the samples were bathed in water at 85 °C for 5 min. The sample was removed, and 1 ml cold water was added. Then, 1 ml n-hexane was added and centrifuged at 5000 g for 5 min after vortexing. The supernatant was removed and transferred to a new centrifuge tube. Then, 1 ml of ultrapure water was added to the final n-hexane phase, which was vortexed and centrifuged, and then the supernatant was dried with nitrogen. Subsequently, 200 µl 0.1% BHT-acetonitrile:methanol (20/80) was added and dissolved.

The chromatographic system was a DGLC dual ternary ultrahigh-performance liquid chromatography system (Thermo, USA). A YMC Carotenoid S-3 µm (150\*4.6 mm) liquid chromatographic column was used. The injection volume was 2 µl, the

column temperature was 40 °C, and the detection wavelength was 450 nm. The solvent systems contained mobile phase A (methanol/water (50/50)) and mobile phase B (acetonitrile (ACN)/ethyl acetate (75/25)). All solvents used were HPLC grade and filtered through a 0.2-mm filter prior to use. The gradient was 30% A:70% B for 0.5 min, then increased in intervals to 0.1% A:99.9% B for 5.5 min and then to 30% A:70% B for the last 2 min. UPLC–MS/MS was performed by Sanshu Biotechnology Co., Ltd. (ShangHai, China) (Figure S2).

### Retrieval gene sequence, collinear analysis and chromosome location

The genes involved in the map00900 and map00906 metabolic pathways were analyzed and mapped to Kyoto Encyclopedia of Genes and Genomes Database (KEGG, <https://www.kegg.jp/>). The sequences of key genes in the carotenoid metabolism pathway of foxtail millet were obtained from homozygous genes in *Arabidopsis thaliana*. TBtools analysis tools were used for the gene sequence information analysis, collinearity analysis, and chromosome mapping analysis[23]. The online websites were used for protein domain prediction analysis (<http://pfam.xfam.org/search/>) and subcellular location prediction (<https://wolfsort.hgc.jp/>).

### RNA extraction and RNA-seq analysis

The developing panicle of DHD and JG21 were separated at the S1/S2/S3 filling stages, and total RNA was isolated using a Quick RNA Isolation kit (Takara Corporation, Dalian, China)[24]. RNA quality and concentration were assessed by 1% agarose electrophoresis (electrophoresis on a denaturing agarose gel) and a NanoPhotometer® spectrophotometer (IMPLEN, CA, USA). The library construction of qualified RNA samples was carried out with a target insert size of ~450 bp, and the quality of the RNA-seq libraries was evaluated by an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). After the library profile analysis, the RNA-seq libraries were sequenced on an Illumina HiSeq platform following standard methods by Novogene Life Sciences Pvt. Ltd. Beijing, China, with three biological replicates, and 150 bp paired-end reads were generated. After filtering the raw data, clean reads were mapped to the foxtail millet reference genome ([https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\\_Sitalica](https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Sitalica)) using Hisat2 (<http://ccb.jhu.edu/software/hisat2/index.shtml>)[25].

For gene expression quantification, HTSeq was used to count the read numbers mapped to each gene, and the FPKM (fragments per kilobase per million fragments) value was used to normalize the expression level of each gene. DEGs between different samples were identified using the R package DESeq with  $|\log_2\text{FoldChange}| > 1$  and  $P$  value  $< 0.05$  as the thresholds[26]. Volcano and MA plots of DEGs were drawn by the R package ggplot2. The KEGG (Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.jp/kegg>) and GO (Gene Ontology, <http://geneontology.org/>) annotations of DEGs were further performed with Goseq and KOBAS software, respectively[27]. The FPKM values of key genes involved in the carotenoid synthesis pathway at different developmental stages of 'DHD' and 'JG21' spikelets were obtained, and the correlation coefficients between the FPKM values of candidate genes and carotenoid content were calculated using SPSS 19.0 software (SPSS, Inc., Chicago, IL, USA). Heat maps of gene expression and correlation coefficients were drawn by TBtools[23]. A weighted gene coexpression network analysis of all genes was performed using the R package WGCNA.

### qRT-PCR analysis

qRT-PCR was performed to verify the expression patterns revealed by the RNA-seq study. Total RNA samples of three stages of foxtail millet panicles were extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). Purified RNA samples were reverse-transcribed using the PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Dalian, China) according to the manufacturer's protocol. Four transcripts were selected for the qRT-PCR assay. Gene-specific qRT-PCR primers were designed using Primer3 software (<http://primer3.ut.ee/>) (Table S1). qRT-PCR was carried out using a Bio–Rad CFX96 instrument (Bio–Rad Laboratories, USA). Each reaction mix was composed of 10  $\mu\text{l}$  2  $\times$  SYBR Green Master Mix Reagent (Vazyme Biotech, China), 2.0  $\mu\text{l}$  cDNA sample, and 400 nm gene-specific primers in a final volume of 20  $\mu\text{l}$  PCR conditions were as follows: 2 min at 95 °C, followed by 40 cycles of denaturation at 95 °C for 10 s and annealing at 60 °C for 40 s. The relative mRNA level for each gene was calculated using the  $2^{-\Delta\Delta\text{CT}}$  formula[28].

# Results

## Dynamic changes in carotenoid metabolites during panicle development stages

For two cultivars with different seed coat colors, five carotenoid metabolites, lutein,  $\beta$ -carotene, zeaxanthin, violaxanthin, and neoxanthin, were detected in the spikes at the three panicle development stages by HPLC. The results showed that all carotenoid metabolite contents were higher at the S1 stage than at the other stages of the two cultivars except zeaxanthin of DHD (Figure 2). The first major carotenoid metabolite was  $\beta$ -carotene, accounting for 43.28% to 58.18% of the total carotenoid content. The second major carotenoid metabolite was lutein, accounting for 27.82% to 50.36% of the total carotenoid content (Figure S3). Moreover, we found that the contents of two major carotenoid metabolites in JG21 were higher than those in DHD at the S3 stage. The highest lutein and  $\beta$ -carotene contents of 11.474  $\mu\text{g}/100\text{ mg}$  and 12.524  $\mu\text{g}/100\text{ mg}$ , respectively, were observed at the S1 stage of DHD.

## Sequence characteristics of genes related to the carotenoid metabolism pathway

Based on the homologous protein blast method, fifty-four genes related to carotenoid metabolism were identified from the genome database (Table 1). Of them, fifteen and fourteen genes encoded six and nine key enzymes involved in the MVA and MEP metabolism pathways, respectively. The remaining genes encoded fifteen enzymes involved in the carotenoid biosynthesis and degradation pathways. The protein lengths of these genes ranged from 89~752 amino acid residues. The protein molecular weights of these genes ranged from 26535~82255 Da. Moreover, 34 genes could be targeted to chloroplasts according to the protein prediction of subcellular localization. The remaining 2, 6, 4, 4 and 2 genes could be targeted to the cytoskeleton, endoplasmic reticulum, mitochondrion, nucleus and plasma membrane, respectively. The *SiGGPPS7b*, *SiLCYB* and *SiNNCED4* gene structures did not have any introns. The shortest gene genomic sequence length was *SiGGPPS7b* at 1098 bp. The longest gene genomic sequence length was *SiCYP97C1* at 18373 bp. These homologous genes for each gene family contained the same specific conserved protein domain (Figure 3). Fifty-four genes related to the carotenoid metabolism pathway were distributed on nine chromosomes of foxtail millet. Ten genes were located on the second chromosome, while only two genes were located on the sixth chromosome. Three gene clusters, *SiZEP1/2*, *SiCCD8b/c/d* and *SiID1/2*, were located on the second, fifth and seventh chromosomes, respectively. A genomic collinearity analysis showed 53, 46, 33, and 37 homologous genes in the *S. viridis*, *S. bicolor*, *Z. mays* and *O. sativa* genomes, respectively, compared with the *S. italica* genome (Figure 4).

Phylogenetic trees of homologous proteins related to carotenoid metabolism were constructed using the neighbor-joining method from *S. viridis*, *Z. mays*, *O. sativa* and *A. thaliana* (Figure S4). For the MEP pathway, *SiDXS1/2/3*, *SiID1/2* and *SiGGPPS7a/b/c* were clustered together into one group with orthologous proteins in *Z. mays* and *O. sativa*. For the MVA pathway, *SiAACT1/2/3* were more highly homologous proteins with *ZmAACT1/2*. *SiMVD1/2* were higher homologous with *OsMVD*. *SiHMGS1/2/3* were clustered into one group with *ZmHMGS* and *OsHMGS*. Moreover, *SiHMGR1/2/3* were more highly homologous with *ZmHMGR1* *OsHMGR3*-like and *OsHMGR3*. For the carotenoid biosynthesis and degradation pathway, *SiPSY1/2/3*, *SiBCH1/2*, *SiZEP1/2*, *SiNCED1a/1b/4/5*, and *SiCCD7/8a/8b/8c* were clustered into one group with orthologous proteins of *Z. mays* and *O. sativa*.

The promoters of these genes indicated that they were mainly involved with 'light responsiveness', 'plant hormone', 'drought stress resistance', 'MYB binding site', 'endosperm specific' and 'seed specific' cis-acting elements. Except for *SiMDC1* and *SiNCED1a*, the promoters of the remaining 52 genes contained a few ABA (abscisic acid) hormone cis-acting elements. Moreover, the promoters of all 54 genes had many MYB binding sites. Of them, the promoters of 51 and 14 genes had some 'drought stress resistance' and 'light responsiveness' cis-acting elements, respectively (Figure S5).

## Differentially expressed genes during the panicle development stage

Eighteen samples, including two cultivars at three panicle development stages, were sequenced and analyzed by RNA-seq. These samples showed higher repeatability and dependency for the same panicle development stage. In total, 125.69 Gb raw

read data were produced from 18 samples by RNA-seq, with Q30  $\geq$  92.31%. The range of raw read numbers mapped to the reference genome among these samples was from 83.09% to 94.63%. By comparing the gene expression levels of DHD\_S1 to that of DHD\_S2 and DHD\_S3, 2218, 3173 and 526 differentially expressed genes were identified, respectively. DHD had a total of 2218 genes with expression differences between S1 and S2 (D1), 526 genes with expression differences between S2 and S3 (D2), and 3173 genes with expression differences between S1 and S3 (D3). Similarly, JG21 had 2209 genes with expression differences between S1 and S2 (J1), 1544 genes between S2 and S3 (J2), and 3412 genes between S1 and S3 (J3). Among them, the number of DEGs between D1 and D2 was at least 24, and the number of DEGs between D1 and D3 was at most 1,784. D2 vs. D3, J1 vs. J2, J2 vs. J3, and J1 vs. J3 had 244, 187, 848 and 1184 differentially expressed genes, respectively. The number of differentially expressed genes shared by the D1, D2, and D3 stages was 209, while that of the J1, J2, and J3 stages was 175.

DEGs between DHD and JG21 at different stages of ear development were analyzed (Figure S6). There were 2690 DEGs between DHD and JG21 (C1) in the S1 period, 2149 DEGs between DHD and JG21 (C2) in the S2 period, and 604 DEGs between DHD and JG21 (C3) in the S3 period. In C1 vs. C2, C2 vs. C3, C2 vs. C3, there were 835, 66, and 61 genes were differentially expressed simultaneously and 227 genes were differentially expressed in all three periods.

The GO enrichment analysis showed that these DEGs mapped to “cellular processes”, “environmental information processing”, “genetic information processing”, “metabolism” and “organismal system”. The KEGG enrichment analysis showed that these DEGs mainly mapped to the ‘phenylpropanoid biosynthesis’, ‘flavonol and flavonol biosynthesis’ and ‘flavonoid biosynthesis’ pathways between DHD and JG21 at the three panicle development stages (Figure S7). However, in the S1 and S3 periods, the DEGs were mainly enriched in the ‘carotenoid biosynthesis’ pathway (Figure S8).

### Gene expression characteristics related to carotenoid metabolism

To investigate the differential expression levels of genes related to the carotenoid metabolism pathway, we analyzed 54 gene expression values from the RNA-seq of JG21 and DHD at three panicle development stages (Figure S9). In the MEP pathway, the *SiHDS* gene was more highly expressed at the S1 stage of DHD and JG21 than the other genes, with TPM values of 60.3 and 60.4, respectively. However, the *SiIDI3* gene had nearly no expression at the three panicle development stages of DHD and JG21. In the MVA pathway, the highest expression level was observed for the *SiHMGS3* gene during panicle development of DHD and JG21, with TPM values ranging from 41.0 to 70.5. However, *SiMVK* had the lowest expression level compared to other genes. In the carotenoid biosynthesis and degradation pathway, only *SiPDS* had a higher expression level, with TPM values ranging from 54.3 to 70.7. The remaining genes had both lower expression levels, with TPM values ranging from 0.0 to 25.9. Moreover, only *SiNCED1a* had a higher expression level than the other genes during the panicle development stage of the two cultivars. The others had lower expression levels for the two cultivars.

### Relationship between carotenoid metabolites and gene expression levels

Based on Pearson’s correlation coefficient analysis, the relationship between the expression levels of genes related to carotenoid metabolism and major carotenoid metabolite contents during the panicle development stage of the two cultivars was investigated (Figure S10). *SiDXS3*, *SiMCT*, *SiHDR*, *SiIDI2*, *SiAACT1/3*, *SiMVK-like*, *SipMVK*, *SiPSY1*, *SiZDS* and *SiZEP1/2* were significantly positively correlated with the neoxanthin content at  $P < 0.05$ . However, *SiCCD8b* was significantly negatively correlated with the neoxanthin content at  $P < 0.05$ . Moreover, *SiDXS3*, *SiMCT*, *SiAACT3*, *SiZEP1/2*, and *SiCCD8a/8c/8d* were significantly positively correlated with the violaxanthin content at  $P < 0.05$ . *SiMCT*, *SiAACT3*, *SiPSY1*, *SiZEP2*, and *SiCc8a/8d* were significantly positively correlated with the lutein content at  $P < 0.05$ . Remarkably, *SiZEP1/2* was significantly positively correlated with four carotenoid metabolites except zeaxanthin ( $r = 0.882, 0.903, 0.974, 0.894, 0.899, 0.969, 0.914$  and  $0.909$ , respectively). However, *SiPSY2* and *SiNCED1b/4* were significantly positively correlated with the zeaxanthin content. *SiCMT*, *SiHDR*, *SiIDI2*, *SiAACT3*, *SiPSY1* and *SiZEP12* were significantly positively correlated with the  $\beta$ -carotene content. Moreover, *SiBCH1* was significantly negatively correlated with the  $\beta$ -carotene content ( $r = -0.818$ , at  $P < 0.05$ ). Furthermore, *SiPSY1* was significantly positively correlated with the contents of neoxanthin, lutein, and  $\beta$ -carotene at  $P < 0.05$ .

( $r=0.921$ ,  $0.818$  and  $0.833$ , respectively). *SiCCD8b* was negatively correlated with all carotenoids except zeaxanthin ( $P<0.05$ ,  $r=-0.903$  and  $-0.811$ ;  $P>0.05$ ,  $r=-0.792$  and  $-0.804$ ).

### Carotenoid metabolism related to the gene coexpression network

The total gene expression value obtained from the RNA-seq data was used to construct a coexpression network. All genes were divided into 22 coexpression modules based on the WGCNA method (Figure S11). The major carotenoid metabolites correlated with the coexpression network module were investigated and analyzed. Of the 22 modules, the MEcoral module was more highly correlated with lutein and  $\beta$ -carotene. Interestingly, the module contained 13 key genes involved in the carotenoid metabolism pathway, which were composed of *SiMCT*, *SiIDI2*, *SiGGPPS7a*, *SipMVK-like*, *SipMVK*, *SiAACT1/3*, *SiPDS*, *SiPSY1*, *SiBCH1*, *SiZEP1/2* and *SiCD8b* (Figure 5). Moreover, we found that 25 transcription factors had higher weight values than the other genes related to carotenoid metabolism (Figure 6). Based on the DNA binding site analysis, these transcription factors potentially interacted with 13 carotenoid metabolism genes. Finally, we found that there were interaction relationships between *SiMADS8* and *SiMCT*, *SiSPL18* and *SiPSY1*, *SiSPT* and *SiZEP2*. Specifically, the regulated *SPL* gene could be a conserved repressor to regulate *PSY* and carotenoid metabolism flux according to a previously reported reference. These results suggested that the gene coexpression network related to the carotenoid metabolism pathway was available and receivable.

### qRT-PCR verified the expression of key genes related to carotenoid metabolism

The relative expression levels of key genes related to carotenoid metabolism during the panicle development stage of DHD were determined by qRT-PCR (Figure S12). We found that *SiDXS1/3*, *SiPSY1/2* and *SiGGPP7a/b/c*, as major rate-limiting enzymes during the panicle developmental stages, showed a consistent downregulation trend. Among them, the expression levels of *SiDXS1*, *SiPSY1* and *SiGGPP7* were 1.6-, 2.04-, 2.35- and 4.91-times higher at the S1 stage compared with the S3 stage, respectively. The expression levels of *SiDXS1*, *SiPSY2/3* and *SiGGPP7b* in JG21 cells were higher than those in DHD cells at the S1 stage. The *SiDXS1* expression level in JG21 was 9.58-times higher than that in DHD at the S1 stage.

For the above coexpression gene pairs, the expression levels of *SiIDI2*, *SiZEP2*, *SiMCT*, *SiSPL18* and *SiMADS8* sharply declined during the panicle development stage. The expression level of *SiSPL18* was 3.55-times higher at the S1 stage than at the S3 stage in DHD. The expression level of *SiMADS8* in DHD was 1.97-times higher than that of JG21 at the S1 stage. The expression levels of *SiZEP2*, *SiSPL18* and *SiMADS8* at the S1 stage in DHD were 1.12-, 1.15- and 1.97-times higher than those in JG21. Moreover, *SiIDI2* expression was 1.23-fold higher in JG21 cells than in DHD cells. Based on the correlation analysis, there was a stronger correlation between the relative expression value and TPM value, with  $R^2=0.9982$ . This finding suggests that these analysis results were reliable.

## Discussion

### Characteristics of carotenoid content variations in plants

With an increasing number of people on our planet, food security issues and the energy crisis are increasingly outstanding and currently face challenges for developing areas[29]. Specifically, many people have presented worsening health statuses because of a lack of dietary vitamin intake for the human body. Carotenoids, as essential vitamins, cannot be synthesized by humans and animals, which play an important role in antioxidation of reactive oxygen and reducing the risk of modern civilization diseases, i.e., cancer, cardiovascular or photosensitivity disorders[12]. To date, more than 750 different carotenoid metabolites have been detected and annotated from the natural world from bacteria, algae and higher plants[30-32]. Although golden rice can help effectively alleviate vitamin A deficiency via the food supply for people living in developing areas, natural food enriched in vitamin A is more easily accepted than GMF (genetically modified food). Previously, reports showed that carotenoid metabolites were enriched in the grains of foxtail millet, although the dynamic changes in the carotenoid content and species are not well understood at present. For *Arabidopsis* leaves, the carotenoid metabolites are mainly composed of lutein,  $\beta$ -carotene, neoxanthin and violaxanthin. Lutein is a major carotenoid metabolite that accounts

for 46.55% of the total carotenoid content[32]. In corn, lutein and zeaxanthin are the most abundant carotenoid metabolites in all immature and mature grains[33]. Among 201 corn inbred lines with different grain colors, significant differences in carotenoid content occurred, with the total carotenoid content ranging from 95.5 to 629.6  $\mu\text{g}/100\text{ mg}$ . Moreover, the most abundant carotenoids in corn kernels were zeaxanthin (14.43  $\mu\text{g}/100\text{ mg}$ ) and lutein (12.32  $\mu\text{g}/100\text{ mg}$ )[34]. In wheat and wheat bread, lutein is the main carotenoid component (0.72–3.07  $\mu\text{g}/100\text{ mg}$ ), followed by zeaxanthin (0.88–1.84  $\mu\text{g}/100\text{ mg}$ ) and  $\beta$ -carotene (0.07–0.33  $\mu\text{g}/100\text{ mg}$ )[30, 31]. The main carotenoid metabolites in brown rice are  $\beta$ -carotene and lutein, up to 1.50  $\mu\text{g}/100\text{ mg}$  and 1.09  $\mu\text{g}/100\text{ mg}$ , respectively, while the content of zeaxanthin was as low as 0.37  $\mu\text{g}/100\text{ mg}$ [14]. The total carotenoid content in mature grains of foxtail millet (192.3  $\mu\text{g}/100\text{ mg}$ ) was 10- to 100-times higher than that in the above cereal crops[11]. In our study, the major carotenoid metabolites were lutein (47.35%) and  $\beta$ -carotene (44.62%). Moreover, we found a significant difference in the total carotenoid content in the grains of foxtail millet between the two varieties at  $P < 0.05$ . A comparison between DHD and JG21, which have white color grains and yellow color grains, respectively, at the early stage of grain showed that the content of lutein and  $\beta$ -carotene in DHD (11.51  $\mu\text{g}/100\text{ mg}$  and 12.56  $\mu\text{g}/100\text{ mg}$ , respectively) was significantly higher than that in JG21 (8.94  $\mu\text{g}/100\text{ mg}$  and 9.86  $\mu\text{g}/100\text{ mg}$ , respectively). However, at the end of development, the contents of the two types of carotene in the white-grained variety DHD were 5.43  $\mu\text{g}/100\text{ mg}$  and 2.59  $\mu\text{g}/100\text{ mg}$ , respectively, which were lower than that of the yellow-grained variety JG21 (5.61  $\mu\text{g}/100\text{ mg}$  and 4.21  $\mu\text{g}/100\text{ mg}$ , respectively) during this period.

### Key genes related to carotenoid metabolism controlled carotenoid metabolism flux

A total of 54 genes related to carotenoid metabolism were identified and analyzed in foxtail millet. Furthermore, we found that these genes had stronger genomic collinearity relationships among cereal crops. We also found that the genome distribution characteristics of these genes were similar to the results of *Brassica napus*[35], indicating that they were essential for maintaining plant development, physiology and biochemical processes in plants.

Carotenoid precursor metabolites are formed by the condensation of the 5-carbon precursors IPP and DMAPP, which are produced through the MEP pathway in plastids[36]. DXS and DXR, as important rate-limiting enzymes in the MEP pathway, play an important role in the regulation of carotenoid metabolism flux[37]. For foxtail millet, the results of the phylogenetic tree showed that the three *SiDXS* genes were divided into three independent branches, suggesting that each *SiDXS* gene could play a different role in the biosynthesis of terpenoids. Similar results existed in other plants[38-40]. Moreover, our results showed that *SiDXS3* was specifically expressed in the panicle of foxtail millet, which indicated that this gene could be the key gene for the carotenoid accumulation at the early panicle development stage. The *IDI* gene was responsible for regulating the ratio of IPP and DMAPP products in the MEP and MVA pathways[41]. Overexpression of different exogenous *IDI* genes in *E. coli* could promote the biosynthesis of  $\beta$ -carotene, lycopene, astaxanthin and zeaxanthin *in vivo*[42-44]. Our results showed that the *IDI2* gene was more highly expressed at the panicle development stage among the three *IDI* gene members correlated with the content of major carotenoid metabolites.

The cytoplasmic MVA pathway also contributes to the synthesis of IPP and DMAPP, which provide precursors for the biosynthesis of sesquiterpenes, polyterpenes, sterols, and glycols and the formation of ubiquinone in mitochondria. In Arabidopsis, the expression level of *AtAACT2* was six times that of *AtAACT1*[45]. In our study, the expression of *SiAACT1* was also significantly higher than that of *SiAACT2/3*, which indicated that *SiAACT1* plays a more important role in cytoplasmic isoprenoid biosynthesis during the panicle development stage of foxtail millet. PSY, as a rate-limiting enzyme in the carotenoid biosynthesis pathway, is easily regarded as the main bottleneck of carotenoid circulation. Ectopic expression of the *PSY* gene and *PaCRT1* gene in rice created the first- and second-generation 'golden rice', which had higher carotenoid contents of 16  $\mu\text{g}/100\text{ mg}$  and 367  $\mu\text{g}/100\text{ mg}$ , respectively[19, 46]. Moreover, overexpression of *PSY1* in tomato plants significantly increased the carotenoid content in tomato fruit[47, 48]. Our results showed that the three *PSY* genes have differentially expressed patterns during the panicle development stage, suggesting that they have differential functions in regulating carotenoid metabolism flux[49]. Additionally, many studies have demonstrated that *ZEP* is an important node for fine-tuning carotenoid metabolism in Arabidopsis[50, 51]. The SNP variants of *ZEP* in sorghum and Arabidopsis were significantly correlated with the zeaxanthin content and lutein/zeaxanthin ratio[52, 53]. Meanwhile, we found that two

members of the *ZEP* gene in foxtail millet were highly correlated with the content of carotenoid metabolites. However, the expression of *SiZEP2* was 72-times higher than that of *SiZEP1*, suggesting that *SiZEP2* could be a major gene for regulating carotenoid metabolism.

The function of CCDs as nonhaem iron-dependent enzymes was to cleave carotenoids by catalyzing the oxidative cleavage of a double bond to form either a ketone or an aldehyde, which reduced the carotenoid content in *Arabidopsis thaliana*, *Chrysanthemum morifolium*, *Fragaria × ananassa* and *Solanum tuberosum*[54-58]. In foxtail millet, *SiCCD8a*, *SiCCD8b* and *SiCCD8d* were tandem repeats distributed on chromosome 5. However, their expression patterns were significantly different. Moreover, we found that *SiCCD8a* and *SiCCD8d* were positively correlated with carotenoid metabolites while *SiCCD8b* was negatively correlated with them. Taken together, these results indicate that the molecular mechanism underlying carotenoid metabolism that maintains the balance of carotenoid synthesis and degradation during the panicle development stage in foxtail millet is relatively complex and challenging.

### Transcription regulation network related to carotenoid metabolism

In recent years, a number of studies have shown that some transcription factors directly or indirectly regulate the expression of carotenoid metabolism and metabolism genes and further control carotenoid metabolism flux in plants. To date, research reports on the interactive relationship between some transcription factors and genes related to carotenoid metabolism have been identified and proven in *Arabidopsis*. Nevertheless, the regulatory network related to the carotenoid metabolism pathway is still unknown and lacks in-depth investigation. In our coexpression network, we found that 25 transcription factors had a stronger interaction with five genes related to carotenoid metabolism. However, only the interaction relationship between SPT and *ZEP2*, SPL18 and *PSY1*, MADS8 and *MCT* could be identified and predicted through the hTFtarget tool. A well-known gene named PHYTOCHROME INTERACTING FACTOR 1 (PIF1), which is a key transcription regulator of carotenoid biosynthesis, could be activated by phytochrome under red light signal treatment during the process of seedling deyellowing[59]. PIF1 is phosphorylated by phytochrome when activated by light and subsequently degraded by the proteasome by inhibiting the expression of *AtPSY* in the dark[60, 61]. PIF1 also initiates genes related to chlorophyll biosynthesis and chloroplast development[62]. Another important bZIP transcription factor, LONG HYPOCOTYL 5 (HY5), could antagonize PIF1 during the process of photomorphogenesis[63, 64]. Spatula (SPT) annotated a bHLH transcription factor that regulates ABA metabolism to control the gynoecium and promote vegetative growth and seed dormancy[65-68]. Hence, we speculated that SPT, as a key regulator of carotenoid metabolism to ABA, had a stronger interaction relationship with *SiZEP2* in the panicle of foxtail millet[69].

In tomato, an SPL gene mutant with a colorless and immature tomato could not produce lycopene because *PSY1* is not expressed[70]. Overexpression of *AtmiR156b*, which repressed *AtSPL3* expression[71], enhanced the content of lutein and  $\beta$ -carotene in rape seeds[72, 73]. Hence, we suggest that the *SiPSY1* gene interacts with the SPL18 gene as a potential regulator of carotenoid metabolism.

Several MADS-box regulators affect the expression of the tomato *CBP* gene, including tomato AGAMOUS-like 1 (TAGL1), RIPENING INHIBITOR (RIN), and FRUITFULL1/2 (FUL1/2)[74]. These MADS-box proteins directly or indirectly positively regulated the expression of *SIPSY1*, *SIPSY2*, *SIZDS*, *SIZ-ISO*, *SICRTISO* and *SIBCH* while negatively regulating the expression of *SILCYB* and *SILCYE*[75-80]. Our results showed that the *SiMADS8* and *SiMCT* genes had a stronger coexpression relationship.

Based on these results, we obtained a gene coexpression network related to carotenoid metabolism and found that three gene pairs had a stronger interaction relationship. These results lay the foundation for exploring the underlying molecular regulation mechanism of carotenoid metabolism in the panicle of foxtail millet.

## Declarations

### Ethical approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

All datasets supporting the results of this article are included within the article and its supplementary information.

### Competing interests

The authors declare that they have no conflicts of interest in this paper.

### Funding

This work was supported by the National Natural Science Foundation of China (31771810, 32070366); Cultivate Scientific Research Excellence Programs of Higher Education Institutions in Shanxi (2019KJ020); Shanxi Agricultural University Science and Technology Innovation Fund Project (2020BQ12).

### Acknowledgments

We would like to thank Professor Donald Grierson, University of Nottingham, UK, for engaging in discussions and helping with the manuscript.

### Author's contribution

YH and SH conceived and designed the experiments. HL and SH wrote the manuscript. GM performed the experiments and analyzed the data. QH, YH, and SH administrated the project. ZS, YH and HL revised the manuscript. All authors have read and gave final approval for publication.

## References

1. Doust, A.N., et al., *Foxtail millet: a sequence-driven grass model system*. Plant Physiol, 2009. **149**(1): p. 137-41.
2. Diao, X. and G. Jia, *Origin and domestication of foxtail millet*, in *genetics and genomics of setaria*, A. Doust and X. Diao, Editors. 2017, Springer International Publishing: Cham. p. 61-72.
3. Zhang, G., et al., *Genome sequence of foxtail millet (Setaria italica) provides insights into grass evolution and biofuel potential*. Nat Biotechnol, 2012. **30**(6): p. 549-54.
4. Bennetzen, J.L., et al., *Reference genome sequence of the model plant Setaria*. Nat Biotechnol, 2012. **30**(6): p. 555-61.
5. Yang, Z., et al., *A mini foxtail millet with an Arabidopsis-like life cycle as a C4 model system*. Nat Plants, 2020. **6**(9): p. 1167-1178.
6. Wang, J., et al., *De novo genome assembly of a foxtail millet cultivar Huagu11 uncovered the genetic difference to the cultivar Yugu1, and the genetic mechanism of imazethapyr tolerance*. BMC Plant Biol, 2021. **21**(1): p. 271.
7. Peng, R. and B. Zhang, *Foxtail millet: a new model for C4 plants*. Trends Plant Sci, 2021. **26**(3): p. 199-201.
8. Lata, C., S. Gupta, and M. Prasad, *Foxtail millet: a model crop for genetic and genomic studies in bioenergy grasses*. Crit Rev Biotechnol, 2013. **33**(3): p. 328-43.

9. Okarter, N. and R.H. Liu, *Health benefits of whole grain phytochemicals*. Crit Rev Food Sci Nutr, 2010. **50**(3): p. 193-208.
10. Shen, R., et al., *Identification of carotenoids in foxtail millet (*Setaria italica*) and the effects of cooking methods on carotenoid content*. Journal of Cereal Science, 2015. **61**: p. 86-93.
11. Zhang, B., et al., *Carotenoid composition and expression of biosynthetic genes in yellow and white foxtail millet [*Setaria italica* (L.) Beauv.]*. Journal of Cereal Science, 2019. **85**: p. 84-90.
12. Fiedor, J. and K. Burda, *Potential role of carotenoids as antioxidants in human health and disease*. Nutrients, 2014. **6**(2): p. 466-88.
13. Moreau, R.A., et al., *A comparison of the levels of oil, carotenoids, and lipolytic enzyme activities in modern lines and hybrids of grain sorghum*. Journal of the American Oil Chemists' Society, 2016. **93**(4): p. 569-573.
14. Lamberts, L. and J.A. Delcour, *Carotenoids in raw and parboiled brown and milled rice*. Journal of Agricultural and Food Chemistry, 2008. **56**(24): p. 11914-11919.
15. Qin, X., et al., *Distinct expression and function of carotenoid metabolic genes and homoeologs in developing wheat grains*. BMC plant biology, 2016. **16**(1): p. 155-155.
16. Rodriguez-Concepcion, M., *Supply of precursors for carotenoid biosynthesis in plants*. Arch Biochem Biophys, 2010. **504**(1): p. 118-22.
17. Vranova, E., D. Coman, and W. Gruissem, *Network analysis of the MVA and MEP pathways for isoprenoid synthesis*. Annu Rev Plant Biol, 2013. **64**: p. 665-700.
18. Watkins, J.L. and B.J. Pogson, *Prospects for carotenoid biofortification targeting retention and catabolism*. Trends Plant Sci, 2020. **25**(5): p. 501-512.
19. Paine, J.A., et al., *Improving the nutritional value of Golden Rice through increased pro-vitamin A content*. Nature Biotechnology, 2005. **23**(4): p. 482-487.
20. Zunjare, R.U., et al., *Influence of rare alleles of  $\beta$ -carotene hydroxylase and lycopene epsilon cyclase genes on accumulation of provitamin A carotenoids in maize kernels*. Plant Breeding, 2017. **136**(6): p. 872-880.
21. Cao, H., et al., *A neighboring aromatic-aromatic amino acid combination governs activity divergence between tomato phytoene synthases*. Plant Physiol, 2019. **180**(4): p. 1988-2003.
22. Nogueira, M., et al., *Subchromoplast sequestration of carotenoids affects regulatory mechanisms in tomato lines expressing different carotenoid gene combinations*. Plant Cell, 2013. **25**(11): p. 4560-79.
23. Chen, C., et al., *TBtools: an integrative toolkit developed for interactive analyses of big biological data*. Molecular Plant, 2020. **13**(8): p. 1194-1202.
24. Nicolaides, N.C. and C.J. Stoeckert, Jr., *A simple, efficient method for the separate isolation of RNA and DNA from the same cells*. Biotechniques, 1990. **8**(2): p. 154-6.
25. Perteza, M., et al., *Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown*. Nat Protoc, 2016. **11**(9): p. 1650-67.
26. Wang, L., et al., *DEGseq: an R package for identifying differentially expressed genes from RNA-seq data*. Bioinformatics, 2010. **26**(1): p. 136-8.

27. Kanehisa, M., et al., *KEGG for linking genomes to life and the environment*. Nucleic Acids Res, 2008. **36**(Database issue): p. D480-4.
28. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
29. Brinkman, H.-J., et al., *High food prices and the global financial crisis have reduced access to nutritious food and worsened nutritional status and health*. The Journal of Nutrition, 2010. **140**(1): p. 153S-161S.
30. Digesù, A.M., et al., *Genetic variability in yellow pigment components in cultivated and wild tetraploid wheats*. Journal of Cereal Science, 2009. **50**(2): p. 210-218.
31. Leenhardt, F., et al., *Genetic variability of carotenoid concentration, and lipoxygenase and peroxidase activities among cultivated wheat species and bread wheat varieties*. European Journal of Agronomy, 2006. **25**(2): p. 170-176.
32. Taylor, K.L., et al., *High-performance liquid chromatography profiling of the major carotenoids in Arabidopsis thaliana leaf tissue*. J Chromatogr A, 2006. **1121**(1): p. 83-91.
33. Song, J., et al., *Comparison of carotenoid composition in immature and mature grains of corn (Zea Mays L.) varieties*. International Journal of Food Properties, 2016. **19**(2): p. 351-358.
34. Owens, B.F., et al., *A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels*. Genetics, 2014. **198**(4): p. 1699-1716.
35. Li, P., et al., *Carotenoid biosynthetic genes in Brassica rapa: comparative genomic analysis, phylogenetic analysis, and expression profiling*. BMC genomics, 2015. **16**(1): p. 492-492.
36. Blanc, G. and K.H. Wolfe, *Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution[W]*. The Plant Cell, 2004. **16**(7): p. 1679-1691.
37. Carretero-Paulet, L., et al., *Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase*. Plant Molecular Biology, 2006. **62**(4): p. 683-695.
38. Walter, M.H., J. Hans, and D. Strack, *Two distantly related genes encoding 1-deoxy-d-xylulose 5-phosphate synthases: differential regulation in shoots and apocarotenoid-accumulating mycorrhizal roots*. The Plant Journal, 2002. **31**(3): p. 243-254.
39. Phillips, M.A., et al., *Functional identification and differential expression of 1-deoxy-d-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (Picea abies)*. Plant Molecular Biology, 2007. **65**(3): p. 243-257.
40. Cordoba, E., et al., *Functional characterization of the three genes encoding 1-deoxy-D-xylulose 5-phosphate synthase in maize*. Journal of Experimental Botany, 2011. **62**(6): p. 2023-2038.
41. Berthelot, K., et al., *Isopentenyl diphosphate isomerase: A checkpoint to isoprenoid biosynthesis*. Biochimie, 2012. **94**(8): p. 1621-1634.
42. Albrecht, M. and G. Sandmann, *Light-stimulated carotenoid biosynthesis during transformation of maize etioplasts is regulated by increased activity of isopentenyl pyrophosphate isomerase*. Plant physiology, 1994. **105**(2): p. 529-534.
43. Sun, J., et al., *A novel cytoplasmic isopentenyl diphosphate isomerase gene from tomato (Solanum lycopersicum): cloning, expression, and color complementation*. Plant Molecular Biology Reporter, 2010. **28**(3): p. 473-480.

44. Gallagher, C., M. Cervantes-Cervantes, and E. Wurtzel, *Surrogate biochemistry: use of Escherichia coli to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation*. Applied Microbiology and Biotechnology, 2003. **60**(6): p. 713-719.
45. Jin, H., Z. Song, and B.J. Nikolau, *Reverse genetic characterization of two paralogous acetoacetyl CoA thiolase genes in Arabidopsis reveals their importance in plant growth and development*. The Plant Journal, 2012. **70**(6): p. 1015-1032.
46. Ye, X., et al., *Engineering the provitamin a ( $\beta$ -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm*. Science, 2000. **287**(5451): p. 303.
47. Fraser, P.D., et al., *Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism*. The Plant Cell, 2007. **19**(10): p. 3194-3211.
48. Tomato Genome, C., *The tomato genome sequence provides insights into fleshy fruit evolution*. Nature, 2012. **485**(7400): p. 635-641.
49. Li, F., C. Murillo, and E.T. Wurtzel, *Maize Y9 encodes a product essential for 15-cis-zeta-carotene isomerization*. Plant Physiol, 2007. **144**(2): p. 1181-9.
50. Xi, W., et al., *The genes of CYP, ZEP, and CCD1/4 play an important role in controlling carotenoid and aroma volatile apocarotenoid accumulation of apricot fruit*. Frontiers in Plant Science, 2020. **11**: p. 2105.
51. Suematsu, K., et al., *Comparative transcriptome analysis implied a ZEP paralog was a key gene involved in carotenoid accumulation in yellow-fleshed sweetpotato*. Scientific reports, 2020. **10**(1): p. 20607-20607.
52. Gonzalez-Jorge, S., et al., *ZEAXANTHIN EPOXIDASE activity potentiates carotenoid degradation in maturing seed* Plant Physiology, 2016. **171**(3): p. 1837-1851.
53. Cruet-Burgos, C., et al., *Advancing provitamin A biofortification in sorghum: Genome-wide association studies of grain carotenoids in global germplasm*. The Plant Genome, 2020. **13**(1): p. e20013.
54. Schaub, P., et al., *Nonenzymatic  $\beta$ -carotene degradation in provitamin a-biofortified crop plants*. Journal of Agricultural and Food Chemistry, 2017. **65**(31): p. 6588-6598.
55. Bruno, M., et al., *Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites*. Journal of Experimental Botany, 2016. **67**(21): p. 5993-6005.
56. Ohmiya, A., et al., *Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals*. Plant Physiology, 2006. **142**(3): p. 1193-1201.
57. García-Limones, C., et al., *Functional characterization of FaCCD1: a carotenoid cleavage dioxygenase from strawberry involved in lutein degradation during fruit ripening*. Journal of Agricultural and Food Chemistry, 2008. **56**(19): p. 9277-9285.
58. Campbell, R., et al., *The metabolic and developmental roles of carotenoid cleavage dioxygenase4 from potato*. Plant Physiology, 2010. **154**(2): p. 656-664.
59. Shin, J., et al., *Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors*. Proceedings of the National Academy of Sciences, 2009. **106**(18): p. 7660.
60. Bae, G. and G. Choi, *Decoding of light signals by plant phytochromes and their interacting proteins*. Annual Review of Plant Biology, 2008. **59**(1): p. 281-311.

61. Shen, H., et al., *Light-induced phosphorylation and degradation of the negative regulator PHYTOCHROME-INTERACTING FACTOR1 from Arabidopsis depend upon its direct physical interactions with photoactivated phytochromes*. The Plant Cell, 2008. **20**(6): p. 1586-1602.
62. Toledo-Ortiz, G., E. Huq, and M. Rodríguez-Concepción, *Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors*. Proceedings of the National Academy of Sciences, 2010. **107**(25): p. 11626.
63. Shi, H., et al., *DET1 degrades HFR1 but stabilizes PIF1 to precisely regulate seed germination*. Proceedings of the National Academy of Sciences, 2015. **112**(12): p. 3817.
64. Zhu, L., et al., *CUL4 forms an E3 ligase with COP1 and SPA to promote light-induced degradation of PIF1*. Nature Communications, 2015. **6**(1): p. 7245.
65. Alvarez, J. and D.R. Smyth, *CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS*. Development, 1999. **126**(11): p. 2377-2386.
66. Tsiantis, M., *Plant development: multiple strategies for breaking seed dormancy*. Current Biology, 2006. **16**(1): p. R25-R27.
67. Josse, E.-M., et al., *A DELLA in disguise: SPATULA restrains the growth of the developing arabidopsis seedling* The Plant Cell, 2011. **23**(4): p. 1337-1351.
68. Penfield, S., et al., *Cold and light control seed germination through the bHLH transcription factor SPATULA*. Current Biology, 2005. **15**(22): p. 1998-2006.
69. Vaistij, F.E., et al., *MOTHER-OF-FT-AND-TFL represses seed germination under far-red light by modulating phytohormone responses in Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, 2018. **115**(33): p. 8442.
70. Eriksson, E.M., et al., *Effect of the colorless non-ripening mutation on cell wall biochemistry and gene expression during tomato fruit development and ripening*. Plant Physiology, 2004. **136**(4): p. 4184-4197.
71. Wang, J.-W., B. Czech, and D. Weigel, *miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in Arabidopsis thaliana*. Cell, 2009. **138**(4): p. 738-749.
72. Gandikota, M., et al., *The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings*. The Plant Journal, 2007. **49**(4): p. 683-693.
73. Wei, S., et al., *Enhanced seed carotenoid levels and branching in transgenic Brassica napus expressing the Arabidopsis miR156b gene*. Journal of Agricultural and Food Chemistry, 2010. **58**(17): p. 9572-9578.
74. Stanley, L. and Y.-W. Yuan, *Transcriptional regulation of carotenoid biosynthesis in plants: so many regulators, so little consensus*. Frontiers in Plant Science, 2019. **10**: p. 1017.
75. Fujisawa, M., T. Nakano, and Y. Ito, *Identification of potential target genes for the tomato fruit-ripening regulator RIN by chromatin immunoprecipitation*. BMC Plant Biology, 2011. **11**(1): p. 26.
76. Fujisawa, M., et al., *A large-scale identification of direct targets of the tomato MADS box transcription factor RIPENING INHIBITOR reveals the regulation of fruit ripening*. The Plant Cell, 2013. **25**(2): p. 371-386.
77. Giménez, E., et al., *Functional analysis of the Arlequin mutant corroborates the essential role of the Arlequin/TAGL1 gene during reproductive development of tomato*. PloS one, 2010. **5**(12): p. e14427-e14427.

78. Itkin, M., et al., *TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network*. The Plant Journal, 2009. **60**(6): p. 1081-1095.

79. Martel, C., et al., *The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner* Plant Physiology, 2011. **157**(3): p. 1568-1579.

80. Shima, Y., et al., *Tomato FRUITFULL homologues act in fruit ripening via forming MADS-box transcription factor complexes with RIN*. Plant Molecular Biology, 2013. **82**(4): p. 427-438.

## Tables

**Table 1.** Carotenoid metabolism gene name, ID, and sequence information.

Gene	E.C.	NCBI ID	Xiaomi ID	yugu1 ID	isoelectric point	molecule weight	PSORT	protein length
SiDXS1	2.2.1.7	1.02E+08	Si3g23880	Seita.3G245500	6.8	77109.14	cyto	717
SiDXS2	2.2.1.7	1.02E+08	Si2g07510	Seita.2G068200	6.85	76986.96	mito	721
SiDXS3	2.2.1.7	1.02E+08	Si4g03680	Seita.4G031100	6.01	78989.42	chlo	721
SiDXR	1.1.1.267	1.02E+08	Si5g07910	Seita.5G071800	6.44	51245	cysk	472
SiMCT	2.7.7.60	1.02E+08	Si5g40650	Seita.5G407200	8.2	32333.52	chlo	297
SiCMK	2.7.1.148	1.02E+08	Si5g35430	Seita.5G351100	6.04	43879.51	chlo	404
SiMDS	4.6.1.12	1.02E+08	Si1g27830	Seita.1G272600	9.52	65443.67	nucl	605
SiHDS	1.17.7.1	1.02E+08	Si1g23200	Seita.1G223900	5.63	82255.08	cyto	746
SiHDR	1.17.7.4	1.02E+08	Si9g10620	Seita.9G105600	5.64	51719.51	chlo	466
SiIDI1	5.3.3.2	1.02E+08	Si3g23600	Seita.3G241900	5.94	32873.71	chlo	297
SiIDI2	5.3.3.2	1.02E+08	Si2g34540	Seita.2G344000	6.28	37523.76	mito	334
SiIDI3	5.3.3.2	1.02E+08	Si2g34530	Seita.2G343900	5.19	26535.25	cyto	233
SiGGPPS7a	2.5.1.1	1.02E+08	Si2g36780	Seita.2G368100	6.18	38272.95	mito	359
SiGGPPS7b	2.5.1.1	1.02E+08	Si5g07040	Seita.5G062900	6.46	37432.52	chlo	365
SiGGPPS7c	2.5.1.1	1.02E+08	Si1g27280	Seita.1G266500	6.15	35656.98	chlo	340
SiAACT1	2.3.1.9	1.02E+08	Si7g28520	Seita.7G280000	6.02	41059.23	chlo	401
SiAACT2	2.3.1.9	1.02E+08	Si5g08110	Seita.5G074000	8.11	43122.41	chlo	421
SiAACT3	2.3.1.9	1.02E+08	Si5g31350	Seita.5G308600	5.83	41231.32	chlo	401
SiHMGS1	2.3.3.10	1.02E+08	Si9g54750	Seita.9G564900	6.1	70931.78	nucl	649
SiHMGS2	2.3.3.10	1.02E+08	Si6g23300	Seita.6G234800	6.11	48769.71	nucl	436
SiHMGS3	2.3.3.10	1.02E+08	Si2g28300	Seita.2G279400	5.82	52344.33	nucl	470
SiHMGR1	1.1.1.34	1.02E+08	Si1g31130	Seita.1G294900	9.02	56290.87	plas	540
SiHMGR2	1.1.1.34	1.02E+08	Si6g20840	Seita.6G208200	7.94	60576.86	E.R.	574
SiHMGR3	1.1.1.34	1.02E+08	Si2g26210	Seita.2G257000	8.35	62101.51	E.R.	584
SiMVK	2.7.1.36	1.02E+08	Si3g26580	Seita.3G273700	5.42	40498.61	plas	387
SiMVK-like	2.7.1.36	1.02E+08	Si2g35580	Seita.2G354500	5.34	40499.55	cyto	387
SipMVKp	2.7.4.2	1.02E+08	Si9g45700	Seita.9G467300	6.19	54814.23	cyto	512
SiMDC1	4.1.1.33	1.02E+08	Si1g35390	Seita.1G351000	5.96	46029.43	chlo	420
SiMDC2	4.1.1.33	1.02E+08	Si3g38700	Seita.3G395300	6.03	46365.95	cyto	423
SiPSY1	2.5.1.32	1.02E+08	Si4g27520	Seita.4G288600	8.97	46899.94	chlo	415
SiPSY2	2.5.1.32	1.02E+08	Si2g30580	Seita.2G303000	9.04	48655.63	chlo	440
SiPSY3	2.5.1.32	1.02E+08	Si3g38930	Seita.3G397800	8.78	45494.92	chlo	409
SiPDS	1.3.5.5	1.02E+08	Si9g50120	Seita.9G515900	8.38	69490.07	chlo	619

SiZDS	1.3.5.6	1.02E+08	Si2g08440	Seita.2G077800	7.98	63378.75	chlo	575
SiZ-ISO	5.2.1.12	1.02E+08	Si3g30280	Seita.3G304800	9.3	40735.41	chlo	373
SiCRTISO	5.2.1.13	1.02E+08	Si8g16170	Seita.8G158400	6.3	64075.75	chlo	592
SiLYCE	5.5.1.18	1.02E+08	Si5g21910	Seita.5G258300	6.24	50077.69	cyto	442
SiLYCB	5.5.1.19	1.02E+08	Si1g06300	Seita.1G055200	7.18	53665.03	chlo	495
SiCYP97C1	1.14.14.158	1.02E+08	Si9g33310	Seita.9G336100	5.92	65129.85	chlo	583
SiCYP97A3	1.14.-.-	1.02E+08	Si1g36810	Seita.1G367300	5.88	70177.46	chlo	645
SiBCH1	1.14.15.24	1.02E+08	Si9g54300	Seita.9G559200	9.12	33335.42	chlo	309
SiBCH2	1.14.15.24	1.02E+08	Si7g21990	Seita.7G209000	9.83	33887.3	chlo	311
SiZEP1	1.14.15.21	1.02E+08	Si7g13140	Seita.7G116800	7.16	69731.49	mito	635
SiZEP2	1.14.15.21	1.02E+08	Si7g13100	Seita.7G116400	8.82	82101.33	chlo	751
SiVDE	1.23.5.1	1.02E+08	Si7g08440	Seita.7G067200	5.32	51047.11	chlo	450
SiNCED1a	1.13.11.51	1.02E+08	Si1g31780	Seita.1G288400	6.33	70862.24	chlo	659
SiNCED1b	1.13.11.51	1.02E+08	Si9g15380	Seita.9G156500	6.08	65918.88	chlo	607
SiNCED4	1.13.11.51	1.02E+08	Si2g04470	Seita.2G035400	6.3	62697.22	chlo	582
SiNCED5	1.13.11.51	1.02E+08	Si3g38270	Seita.3G391000	6.56	70850.96	chlo	659
SiCCD7	1.13.11.68	1.02E+08	Si7g20330	Seita.7G189300	8.9	67763.96	chlo	619
SiCCD8a	1.13.11.69	1.02E+08	Si8g10360	Seita.8G101900	7.91	59808.24	chlo	542
SiCCD8b	1.13.11.69	1.02E+08	Si5g32040	Seita.5G315800	6.48	62634.96	chlo	577
SiCCD8c	1.13.11.69	1.02E+08	Si5g32020	Seita.5G315600	6.04	61291.42	chlo	560
SiCCD8d	1.13.11.69	1.02E+08	Si5g32030	Seita.5G315700	7.66	60585.91	chlo	553

## Figures

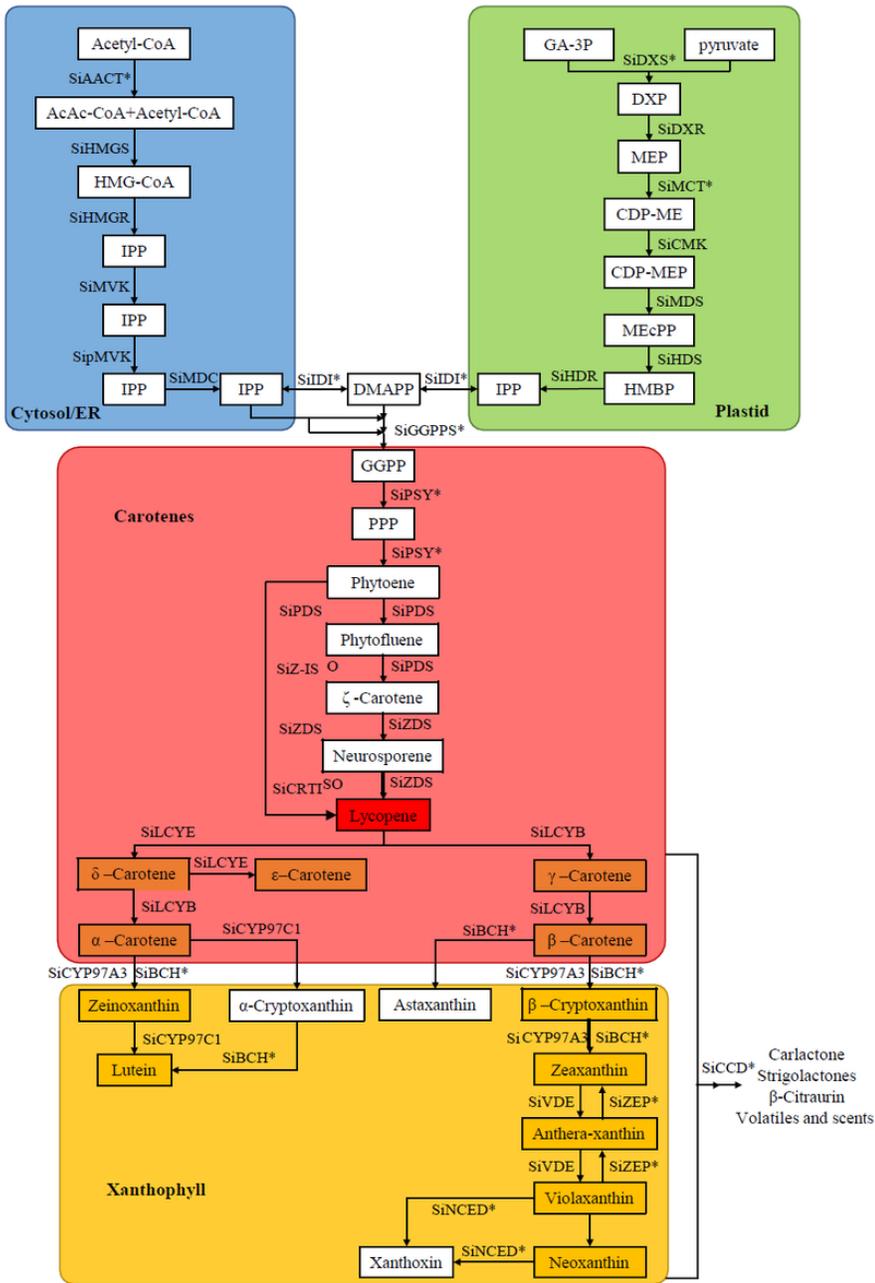
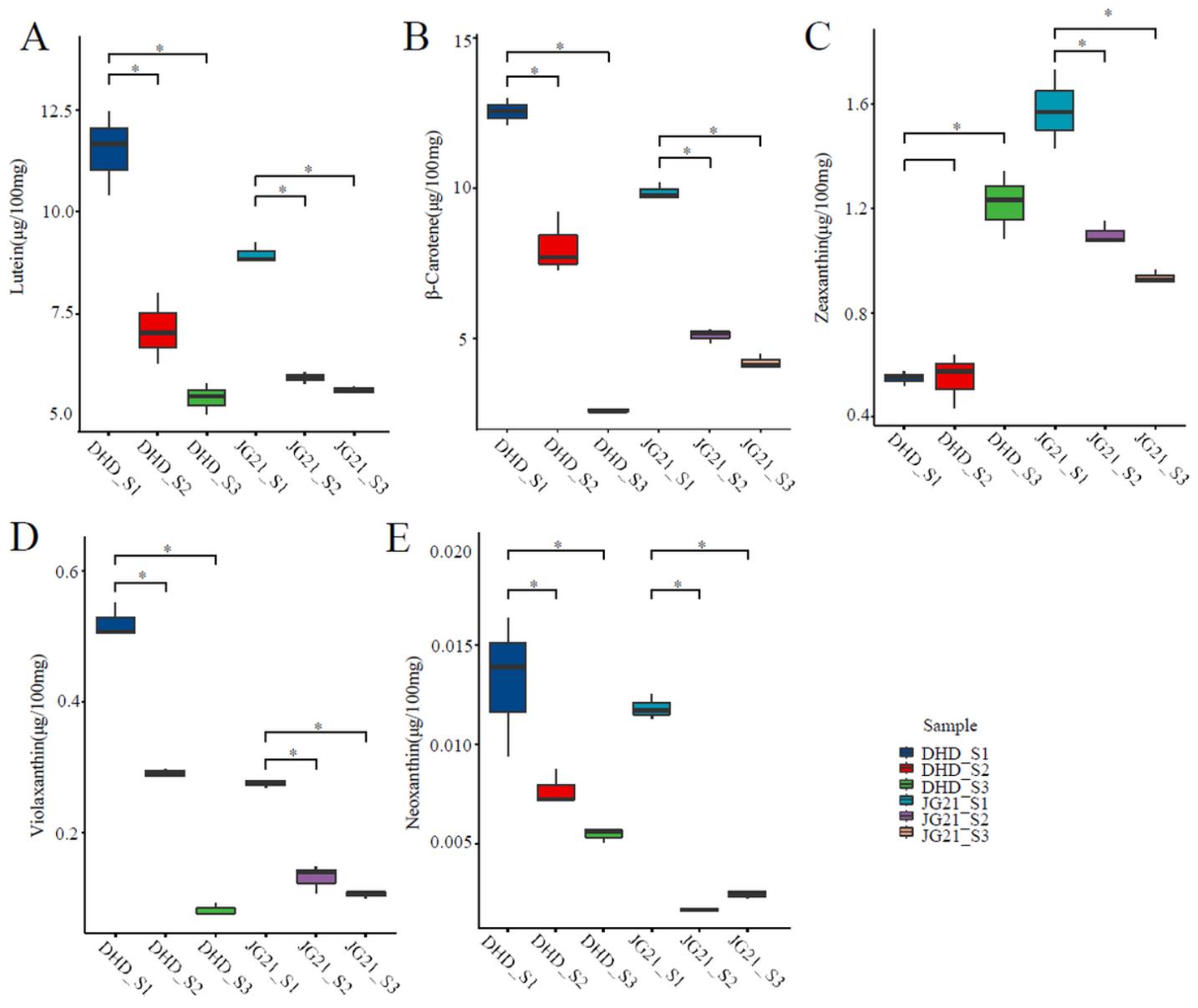


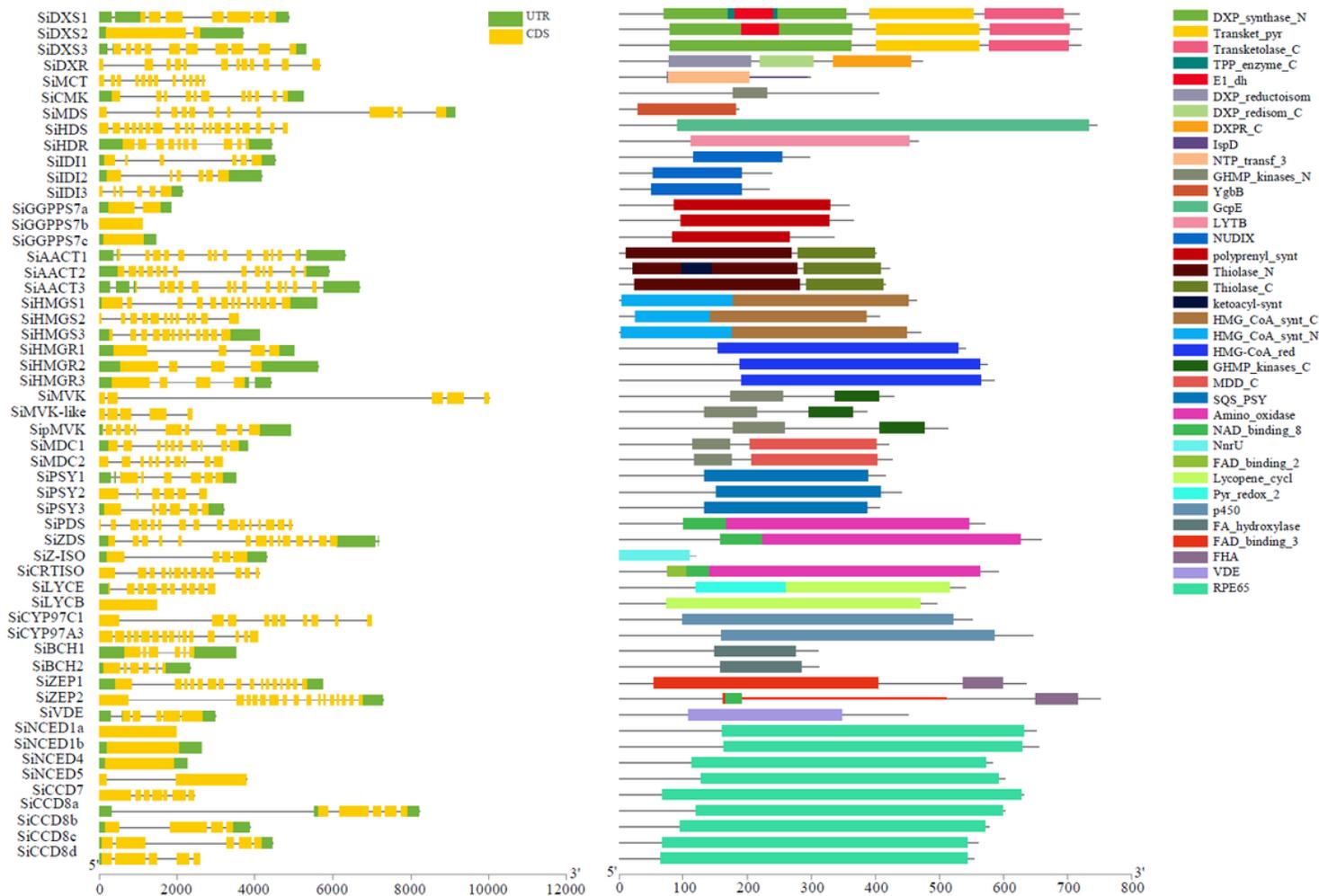
Figure 1

Foxtail millet carotenoid metabolism pathway.



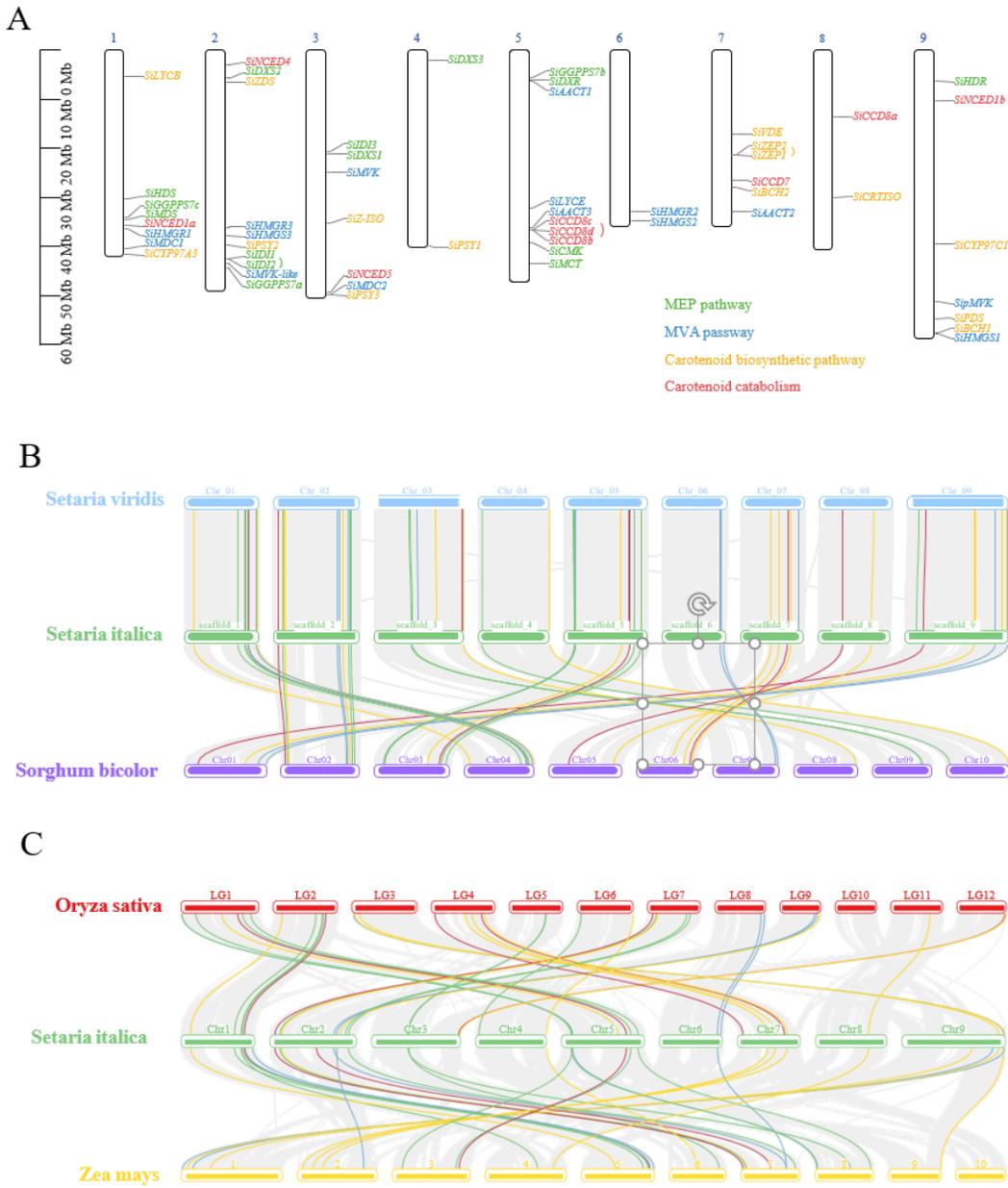
**Figure 2**

DHD and JG21 carotenoid content at the three developmental stages of panicle. (A) Lutein. (B)  $\beta$ -carotene. (C) Zeaxanthin. (D) Violaxanthin. (E) Neoxanthin.



**Figure 3**

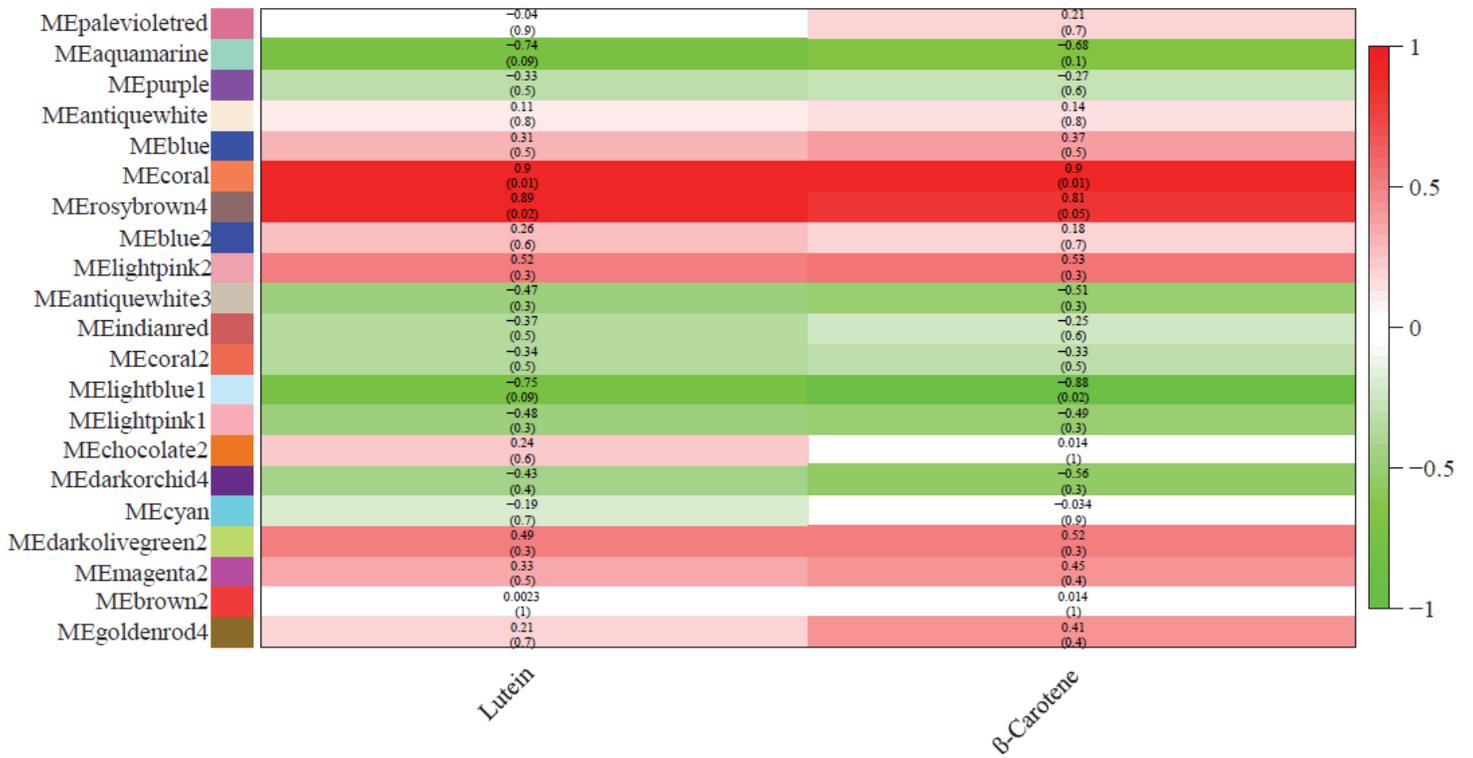
Gene structure and protein structure of the carotenoid metabolism pathway in foxtail millet.



**Figure 4**

(A) Subcellular location of carotenoid metabolism genes. (B) Collinearity analysis of millet carotenoid metabolism genes in close-source species. (C) Collinearity analysis of millet carotenoid metabolism genes in the main monocot crops.

### Module-trait relationships



**Figure 5**

Correlation analysis of the same carotene content in the gene coexpression network module.

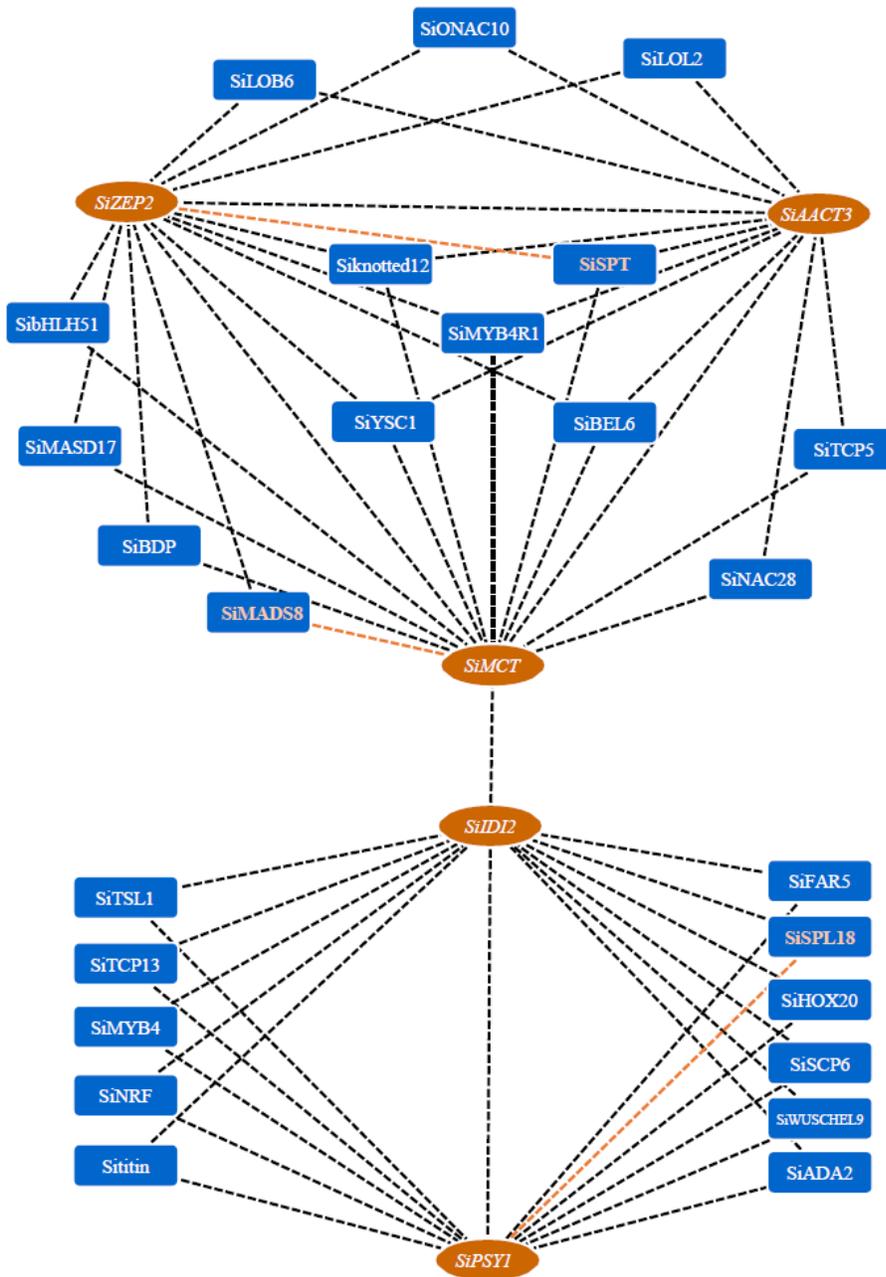


Figure 6

Coexpression network of important carotenoid metabolism genes and related transcription factors.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [genomicscarotenoid9.5s.pptx](#)