Comparative Transcriptomics Provides Insight Into Rind Color Polymorphism in Watermelon

Dongming Liu
Henan Agricultural University
Huihui Yang
Henan Agricultural University
Dongling Sun
Henan Agricultural University
Minjuan Zhang
Henan Agricultural University
Junling Dou
Henan Agricultural University
Sen Yang
Henan University
Huayu Zhu
Henan Agricultural University
Jianbin Hu
Henan Academy of Agricultural Sciences
Shouru Sun
Henan Academy of Agricultural Sciences
Luming Yang (lumingyang@henau.edu.cn)
Henan Agricultural University https://orcid.org/0000-0002-4505-5414

Research article

Keywords: Watermelon, Rind color, Transcriptomics, Chlorophyll pathway, Carotenoid pathway

Posted Date: October 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-75546/v1

Abstract

Background

As an important agronomic trait affecting breeding and consumption choice, rind color of watermelon [*Citrullus lanatus* (Thunb.) Matsum. and Nakai] receives considerable attention from breeders. Although rind color occupies an important place in watermelon breeding work, the underlying molecular mechanism governing this complex trait has still not been determined.

Results

In the present study, chlorophyll and carotenoid contents were hypothesized to be the main factors determining rind color. Clearly different structures and numbers of chloroplasts and grana lamellae were found among materials with different rind colors. To study the genes involved in the coloration process of the yellow, dark green, light green, and light green-gray rinds, 84,516 unigenes were obtained from the rind transcriptome. Through DEGs screening and pathway analysis of chlorophyll and carotenoid mechanism, the genes with abnormal expression levels were identified. Then, the nucleotide sequences and expression levels of the putative genes responsible for different rind colors were studied.

Conclusions

This study provides a comprehensive insight on the factors affecting rind coloration and investigated expression levels of the critical genes concerning pigment metabolism in watermelon. These results would help to analyze the molecular mechanisms of rind coloration regulation and development of the related organelles.

Background

As one important vegetable crops around this world, watermelon (*Citrullus lanatus* (Thunb.) Matsum. & Nakai.) provides abundant carotenoids to worldwide consumers and plays important economical roles in many countries [1]. Various phenotypes including size and shape of fruit , color and pattern of rind, and flavour and color of flesh give the breeders chances to develop kinds of elite cultivars. Rind color of watermelon is an extremely important characteristic for markets and breeders choice. Although yellow and green are the basic colors of watermelon rind, different degrees of the basic colors existed in different plants [1, 2].

A number of genes and QTLs concerning rind color have been detected or identified in various horticultural crops, such as citrus, pummelo, chrysanthemum, and cherry [3, 4] [5, 6]. In cucurbitaceae plants, rind color is often controlled by one single gene. Orange rind color of cucumber is dominant to creamy color in mature fruit, and the *R2R3-MYB* transcription factor was identified as the candidate gene [7]. In wax gourd, the dark green fruit skin is controlled by the gene named *pc* [8]. In watermelon, sorts of inheritance patterns of fruit skin color have been raised. A general background color and various types of

mottling and stripping form different pictures on the watermelon rind. Take the background color 'green' for example, it can be sorted as light green, light green-gray, and dark green [9, 10]. With genetic analysis, Porter found that the light green is recessive to dark green [11]. Weetman proposed the hypothesis that there existing two genes responsible for the foreground stripe pattern and background color, respectively [12]. With help of GWAS analysis, Oren found that a 16-bp deletion in gene *ClCG09G012330* is the key factor for the light/dark green appearance [13]. With two constructed families, Kumar found that the solid dark green appearance follows a duplicate dominant epistasis fashion [14]. Being less complex than dark green rind, yellow rind in watermelon is controlled by one single gene, which is dominant to green [1]. Through BSA-seq and GWAS methods, the gene controlling yellow rind was delimited to a region of 91.42 kb on chromosome 4, but no candidate gene was obtained in this region [15].

Rind color in watermelon has always gained increasing attention from consumers and breeders, but still little knowledge of the molecular mechanism accounting for the fruit rind coloration was revealed until now. In the present study, a comparative transcriptome analysis was completed for DEGs screening and gene functional analysis of the dark green background rind material WM102, light green background rind material WT2, light green-gray background rind material WT13, and yellow background rind material WT4. In association with the pigment formation pathway, the key genes and pathways affecting rind coloration in different materials were studied. These results would provide new insight about the molecular mechanisms and key potential pathways involved in fruit rind coloration.

Results

Quantification of chlorophylls and carotenoids

Different color performance of plant organs and tissues is mainly caused by the pigment difference [16]. Content of chlorophylls and carotenoids in four watermelon materials was measured. In accordance with the visual appearance, material WM102 which shows a dark green skin owns the highest content of chlorophylls (chlorophyll a and chlorophyll b) (Fig. 2). The other ones, ranging from high to low chlorophyll content are WT2 (light green rind), WT13 (light green-gray rind), and WT4 (yellow rind) in turn. Compared with the other three materials, the chlorophyll content in WT4 was significantly reduced. In contrast to chlorophylls, content of carotenoids in WT4 is considerably higher than that in the other three materials which show a similar carotenoids content (Fig. 2). These result indicates that different rind color appearance in watermelon was due to the varying content of chlorophylls and carotenoids.

Ultrastructural changes in the rind of watermelon

TEM of chemically fixed rind samples could help to reveal the ultrastructural differences. The most noticeable difference focuses on the chloroplast number and structure. Material WM102 with dark green rind contains the most chloroplasts with approximately eight chloroplasts in each field of a microscope (Fig. 3A), and chloroplast granas of WM102 are composed of a set of compact grana lamellae (Fig. 3E). In WT2 and WT13, the number of grana and grana lamellae decreased slightly, but the structure of chloroplasts was not significantly changed compared with that in WM102 (Fig. 3B, 3C, 3F, 3G). In the

yellow-rind material WT4, the number of grana is significantly decreased, and the number of grana lamellae in each grana is also seriously decreased (Fig. 3D, 3H). Except for the number of grana and grana lamellae, the structure of grana also changed considerably. Because lacking grana lamellae, the size of the grana is observed to be smaller, and the shape appears to be a ball (Fig. 3H), which is obviously different from that in WM102, WT2, and WT13. The ultrastructural results illustrate that the difference in chloroplast number and structure may be the a reason for the different pigment content and rind color appearance.

De novo assembly of RNA-Seq reads

In the previous research, the gene responsible for yellow skin was delimited to the region of 1–91.42 kb on chromosome 4 of watermelon, but no candidate genes were found in this region [15]. We proposed that the target gene may not be assembled into the watermelon reference genome of cultivar 97103 [17] or cultivar Charleston Gray [18]. To provide possible reference information for the related gene study, the clean reads were *de novo* assembled and 542,314,676 clean reads were obtained after the 586,255,040 raw reads were screened (Additional file 1: Table S2). The 202,841 transcripts assembled with Trinity software [16] has a length of 325,191,335 bp (Additional file 1: Table S3) and 84,516 unigenes were obtained based on the assembled transcripts. Among the unigenes, 77.61% (65,595) are shorter than 1,099 bp, 20.90% (17,663) ranges from 1,099 to 4,999 bp, and only 1.49% (1,258) are longer than 4,000 bp (Additional file 1: Table S4). Compared with the 22,596 and 22,567 genes of the watermelon reference genome of cultivar 97103 and cultivar Charleston Gray, respectively, many more novel genes were obtained through *de novo* assembly. The newly assembled transcriptome would provide a reference resource for gene functional study in watermelon.

Functional annotation and classification of the unigenes

After assembly, unigenes were aligned to the six functional databases (Additional file 2: Fig. S1; Additional file 1: Table S5). To further evaluate the BLAST result, the E-value and similarity distributions to the Nr database were calculated. Results revealed that about 39.05% of the sequences were matched with the database (E-value<10−45) (Fig. 4a) and 67.92% of the sequences had similarities over 80% (Fig. 4b). Species distribution analysis showed that *Cucumis melo* and *Cucumis sativus* have a relatively high similarity score with 24.66% and 21.17% top BLASTx hits, respectively (Fig. 4c).

In GO classification, 28,095 unigenes (33.12%) were assigned to 65 GO term subcategories. The molecular function process contained 15 (23.08%) subcategories, and cellular component (32.31%) and category biological process (44.62%) are the three mostly enriched GO terms (Additional file 3: Fig. S2). The Cellular Component category contains 21 subcategories, and the Cellular Component category includes the most enriched subcategories. The top five enriched GO terms were cell, cell part, membrane, organelle, and membrane part (Additional file 3: Fig. S2). In the Biological Process Process, 29 subcategories were found, and the most enriched term was metabolic process followed by cellular process (Additional file 1: Table S6, Additional file 3: Fig. S2). In the eukaryote-specific COG (KOG) classification, 40,522 unigenes were classified into 26 KOG categories (Additional file 4: Fig. S3). General

function prediction accounted for the largest proportion (22.57%) followed by posttranslational modification, protein turnover, chaperones (9.80%) and signal transduction mechanisms (8.16%) (Additional file 4: Fig. S3; Additional file 1: Table S7). 21352 (25.23%) unigenes were mapped into 33 categories of pathway hierarchy level 2 and 5 categories of hierarchy level 1 (Additional file 5: Fig. S4). Among the 33 pathways, translation (10.84%) and carbohydrate metabolism (10.32%) were the most dominant groups followed by signal transduction (9.78%), overview (8.40%), energy metabolism (8.28%), and folding, sorting and degradation (7.52%) (Additional file 5: Fig. S4, Additional file 1: Table S8).

Identification of DEGs and validation of RNA-Seq Data

Gene expression levels of the samples were estimated with the FPKM value and the correlation coefficients of the twelve samples were also analyzed and are listed in Additional file 1: Table S9. As shown in Fig. 5, comparison of gene expression between WT13 and WM102 showed that 5,905 genes were significantly differentially expressed, with 3,690 up-regulated and 2,215 down-regulated. For WT13 vs WT2, there were 5,811 DEGs, with 3,101 being up-regulated and 2.710 down-regulated. For WT13 vs WT4, 6179 genes were significantly differentially expressed, with 3486 being up-regulated and 2693 down-regulated. In the comparison group of WT2 and WM102, 3,954 up-regulated and 2,627 down-regulated DEGs were identified. For WT2 vs WT4, 8149 genes were significantly differentially expressed, with 4313 up-regulated and 3836 down-regulated. In the comparison group of WT4 and WM102, 1961 u-pregulated and 1111 down-regulated DEGs were identified (Fig. 5). In all the comparison groups, only 62 DEGs in common were found (Fig. 6). A total of 14 unigenes were selected to validate the RNA-Seq data using qRT-PCR with three biological replicates. The qRT-PCR results showed that the gene expression patterns of the 14 unigenes had similar trends with the RNA-Seq data with a positive correlation coefficient of 1.124 (Additional file 6: Fig. S5), implying the accuracy of the data.

Expression of genes involved in chlorophyll metabolism

In higher plants, chlorophyll metabolism, including chlorophyll biosynthesis, chlorophyll cycle, and chlorophyll degradation, is catalyzed by a series of enzyme complexes. Changes in the expression patterns of any of these reactions may affect rind color appearance. In this experiment, chlorophyll and carotenoid content also varied in WM102, WT2, WT13, and WT4 in accordance with the different color appearances. Accordingly, the core genes that encoded the enzymes involved in the chlorophyll metabolic pathway were investigated. A total of 50 unigenes encoding 23 chlorophyll metabolism-related enzymes were identified (Fig. 7; Additional file 1: Table S10).

Among these unigenes, 9 enzymes (i.e., GSA, HEMC, HEMG, CHLH/D/I, CHLG, CAO, and RCCR) were found to belong to a single gene family, and the remaining ones belong to multigene families. As shown in Fig. 7, almost all the genes related to chlorophyll metabolism were highest expressed in WM102, which has the darkest rind and the highest chlorophyll content and number of chloroplasts. In WT4, expression of *CHLD* is considerably down-regulated compared with materials WM102, WT2, and WT13, with a similar expression pattern as other material(s). In WT13, the expression of the *HEMD* gene was down-regulated in the chlorophyll biosynthesis part but up-regulated in the chlorophyll degradation part of the *SGR* gene

compared with materials WM102, WT2, and WT4. In WT2, the gene expression of *HEME* was up-regulated compared with the other three materials. Except for the above particular genes, the expression levels of all the related genes in these four materials appear to be similar (Fig. 7; Additional file 1: Table S10).

Expression of genes involved in carotenoid metabolism

Carotenoids can be biosynthesized by many living organisms, such as photosynthetic bacteria, cyanobacteria, algae, and higher plants. Carotenoids usually play vital roles in light harvesting, photoprotection, photomorphogenesis, nonphotochemical quenching, and lipid peroxidation [19, 20]. The various kinds and contents of carotenoids impart different colors to organisms, ranging from colorless to yellow, orange, and red. Because of the different coloration expression and carotenoids content in the four watermelon materials, the expression levels of genes involved in carotenoid metabolism were investigated.

A total of 80 unigenes encoding 18 carotenoid metabolism-related enzymes were identified (Fig. 8; Additional file 1: Table S11). Among these unigenes, only one gene (*CCS*) was found to belong to a single gene family, and the remaining genes were demonstrated to be from multigene families. Carotenoids are a diverse class of lipid-soluble isoprenoids built from the 5-carbon compound IPP [21, 22]. Biosynthesis of isoprenoid is limited by the first committed step catalyzed by DXS [23]. Six members were found in the *DXS* family. *DXS1* and *DXS5* were up-regulated in WT13 and WM102, respectively. Gene *DXR* catalyzes the first committed step of the MEP pathway for isoprenoid biosynthesis in plants [24]. The expression levels of *DXR1* and *DXR2* in WT13 are considerably higher than those of the other three materials. As the terminal gene of the MEP pathway, *LytB* can enhance the production of carotenoids [25]. Both *LytB* (*LytB*1 and *LytB2*) were down-regulated compared with the other three materials. Four members exist in the IPI family. A varied expression level was exhibited by materials WT4, WM102, WT2, and WT13 in genes *IPI1* and *IPI2*. As a key enzyme in carotenoid biosynthesis, *PSY* belongs to the branching enzyme that determines the flux of carbon source toward carotenoids [26, 27]. The expression level of PSY1 in WT13 is up-regulated. Xanthophylls are produced from β-carotene with reactions of epoxidation-catalyzed ZEP [28]. In contrast to the similar expression level in WT13, the expression of the *ZEP1* gene was significantly up-regulated (Fig. 8; Additional file 1: Table S11). Except for the mentioned genes with different expression levels above, similar expression levels were presented by most of the materials. Although the rind color of WT4 turns yellow and carotenoid content is higher than those of the other three green materials, no gene with significantly different expression levels was found.

Discussion

As an important visual characteristic of watermelon, rind color generally affects consumption customs. Pigment kind and content are the two primary elements determining rind color expression. In this study, we attempted to elucidate the mechanism of differential pigmentation of watermelon rinds. The expression profiles of four materials named WT4, WM102, WT2, and WT13 with different background

rind colors were studied. Our findings suggest that the different rind colors are the result of a combination of multiple pigments.

Chlorophyll and carotenoid are the main pigments affecting watermelon rind color. Varying the chlorophyll and carotenoid contents would change the color expression in many plants. For example, Cui found that in the Arabidopsis mutant gdc1-3, whose leaves turned pale green, the total chlorophyll content was reduced by approximately 82% compared with wild-type plants [29]. Similar to the *Arabidopsis* gdc1-3 mutant, the content of both Chl a and Chl b in the wheat yellow leaf color mutant is considerably lower than that in the wild type [30]. In this study, we also measured the chlorophyll and carotenoid content. Results showed that the content of both Chl a and Chl b decreased considerably compared with the other three green rind materials, especially compared with WM102, a material with dark green rind. However, the carotenoid content was higher in WT4 than in WM102, WT2, and WT13. Light green and light green-gray colors are exhibited by WT2 and WT13, not as dark green in WM102. Accordingly, the contents of Chl a and Chl b are lower than WM102 but higher than WT4. However, a similar carotenoid content was observed for materials WM102, WT2, and WT13. These results show that in accordance with other yellow-mutant plants, the chlorophyll and carotenoid contents may affect the color expression of the green tissues of plants.

In material WT4, we found that the number of grana and grana lamellae decreased significantly and that the structure of grana changed notably. A similar abnormal grana structure could be found in wheat mutant Y [30], indicating the important role for rind coloration of grana structure. In WT2 and WT13, although chlorophyll content also decreased, the structure of the chloroplast did not change as much as that of WM102. These results indicate that abnormal chloroplast development occurs in WT4 and that the decreased chlorophyll number may be a result of abnormal chloroplast development.

Chlorophyll metabolism is comprised of three steps: chlorophyll biosynthesis, chlorophyll cycle, and chlorophyll degradation. Comparisons of the expression levels of chlorophyll metabolism-related genes among the four materials were carried out. Some genes with particularly different expression levels were identified. With the F2 population, Liu confined the gene for yellow rind to a region of 91.42 kb on chromosome 4 and found three genes, named *Cla97C04G068450*, *Cla97C04G068460*, and *Cla97C04G068470* in this region. However, the RT-qPCR result showed that the expression of the three genes was not significantly regulated [15]. Another gene, termed *CHLH* (*Cla97C04G068530*/ *Cla002769*), near this region was speculated to be correlated with rind color in watermelon through GWAS analysis [31]. The DNA sequence and gene expression level were studied. The results showed that no sequence changes were found in the coding or promoter region of the *Cla97C04G068530*/ *Cla002769* gene in WT4 compared with WM102 and 97103. However, the expression level was considerably lower in WT4 than in WT2, WT13, and WM102 (Fig. 9 B), and the relationship between the *CHLH* gene and the yellow rind phenotype warrants further research. The RNA-Seq results show that the expression of *CHLD* (*Cla97C02G049100*/ *Cla008566*), a gene in the chlorophyll metabolism pathway, is notably down-regulated in WT4 compared with materials WM102, WT2, and WT13. However, the *CHLD* (*Cla97C02G049100*/ *Cla008566*) gene was located on chromosome 2, not within the 91.42 kb region of

chromosome 4. Regarding the relationship between the yellow rind trait and the *CHLD* and *CHLH* genes, further study is required.

In a previous study, the gene for dark green rind in inbred line material '9904' was narrowed to a candidate region to 31.4 kb on chromosome 8, which contains 4 genes. One nonsynonymous SNP mutation (C→G) of gene *ClCG08G017810* (*Cla000814*) at 29,880,100 bp and a 16-bp insertion from the position of − 207 bp was identified. Therefore, the gene numbered *ClCG08G017810* (*Cla000814*) was selected as the candidate gene for the dark green rind [32]. The transcriptome analysis results showed that the expression level of the *ClCG08G017810* gene (*Cla000814*) in the dark-green material WM102 was also higher than that in the other three materials (Fig. 9 C). Sequencing results show that the same SNP (C→G) at 27994761 and the same insertion of 16 bp in the promoter region as in material '9904' was also found in dark green material WM102, but none of the above mutation styles was found in materials WT2, WT13, and WT4. Because the mutation style is found in both the dark rind materials 9904 and WM102, the factors causing this trait are still unknown. Determining whether the SNP or the insertion of 16 bp in the promoter region is the key factor responsible for darkened rind coloration requires further work.

The important roles of transcription factors in plant pigment formation have been stressed recently. For example, the R2R3-*MYB* transcription factor in kiwifruit (*Actinidia deliciosa*) was observed to modulate chlorophyll and carotenoid accumulation by activating the promoter of the kiwifruit lycopene beta-cyclase (*AdLCY-β*) gene, which plays a key role in the carotenoid biosynthetic pathway [33]. In birch, the *GLK1* transcription factor directly interacts with the promoter of genes related to antenna protein, chlorophyll biosynthesis, and photosystem subunit synthesis and regulates their expression, thereby further affecting chloroplast development [34]. In the present study, the main reason for carrying out a *de novo* assembly of the transcriptome is to screen for the gene responsible for the yellow rind of watermelon that may not be assembled into the watermelon reference genome of cultivar 97103. The sequences of all the unigenes were blasted with those of the genes of the watermelon reference genome of cultivar 97103. According to the BLAST results, a unigene (TRINITY_DN39858_c0_g4) belonging to the *ERF* (ethylene response factor) gene family with a higher expression level in material WT4 than in WT2, WT13, or WM102 was identified (Fig. 9 D). The expression results showed that a low and stable expression level of the unigene TRINITY_DN39858_c0_g4 was detected in WT4, but no expression was found in WT2, WT13, or WM102 (Fig. 9 D). Genes from the ERF gene family encode transcriptional regulators concerning a variety of functions that participate in plant developmental and physiological processes. Yin found that transient and stable overexpression of CitERF13 resulted in rapid chlorophyll degradation in *Nicotiana tabacum* leaves and made leaves yellow. Expression of the gene *CitERF13* sharply increases, and the fruit color also changes to yellow with ethylene treatment [34]. Except for the unigene TRINITY_DN39858_c0_g4, another transcription factor (TRINITY_DN41795_c0_g5) belonging to the *MYB* gene family with higher expression levels in material WT4 was also found (Fig. 9 E). A number of *MYB*-type transcription factors participate in plant pigment development and rind coloration in cucumber, sweet cherry, tomato, apple, and rice [35, 36] [37, 38] [39] [40, 41]. Transcription factors are also important elements involved in the light green phenotype in Cucurbitaceae plants. With a GWAS study, Oren found that the APRR2 gene is associated with fruit pigment accumulation in both melon and

watermelon, and the gene *ClCG09G012330* in watermelon was selected as the candidate gene for the light green phenotype. Alternative splicing was speculated to be responsible for a 16-bp deletion in the mRNA of the light rind parent [13]. In the present study, the transcriptional level in the light green material WT2 was also considerably different from that in the dark green material WM102 (Fig. 9 F).

Since various rind colors are exhibited in different watermelon materials, a complex functional and regulatory mechanism may affect the coloration process in watermelon rinds. Several gene related to rind color variety was mapped recently, but the complex regulatory network and metabolic mechanism is still unknown. Concerning the important roles rind color plays in affecting consumption choice and plant coloration, future research should investigate the key genes and pathways participating in pigment metabolism and formation.

Conclusions

In this study, the different content of chlorophyll and carotenoids was speculated to be the key factors affecting watermelon rind color appearance, and the abnormal chloroplast development and decreased chloroplast number should be the reasons for the decreased chlorophyll content. 202,841 transcripts with a total length of 325,191,335 bp were obtained with the de novo assemble result. In all the comparison groups, only 62 DEGs in common were found among thousand of the DEGs, indicating the core function the 62 DEGs may play in watermelon rind coloration. Some key genes with different expression level in the pigment metabolic pathways were identified. All these genes may be involved in the complex regulatory network and metabolic mechanism during watermelon rind coloration.

Methods

Pigment content determination and RNA extraction

Four watermelon inbred lines with different background rind colors were used as the materials in this study. WM102 is an inbred line after eight generations of self-cross from bush sugar baby with dark green background rind. WT2, WT4 and WT13 are three inbred lines with light green background rind, yellow background and light green-gray background rind, respectively (Fig. 1); and they are generated from local watermelon germplasm with at least six generations of self-crossing. Chlorophylls and total carotenoid content determination and RNA extraction was completed at 14 DAP [15].

Ultrastructural analysis

The rind was cut into 1 mm × 1 mm × 3 mm pieces and fixed with 2.5% glutaraldehyde (v/v) at 4 °C until the tissues sank in the solution. All the tissues were dehydrated using a graded series of ethanol (30%, 50%, 70%, 80%, 90%, and 95% for 1 hour each) followed by 100% ethanol solution (twice for 1 hour each). Next, samples were soaked in a graded series of mixtures of ethanol and acetone (3:1, 1:1, 1:3) for three 30-min cycles followed by 100% acetone solution for 1 hour. Finally, the samples were immersed overnight in embedding medium. Ultrathin slices (60 nm) were cut using an ultramicrotome (Leica UC7)

and collected on copper grids. Subsequently, the slices were stained with uranyl acetate and lead citrate for 15 minutes. After air drying overnight, the samples were observed using a transmission electron microscope (HT7700; Hitachi).

RNA-Seq library construction, sequencing, and read mapping

For RNA-Seq library construction, the extracted RNA samples were sent to Personal Technologies Co. (Shanghai). After RNA concentration and integrity assessment, mRNA was purified. RNA-Seq library construction was completed followed the steps previous described [15].

The library preparations were sequenced on an Illumina HiSeq2000 platform. Clean reading were obtained after low-quality reads and reads containing the adapter poly-N were removed from raw reads. The clean reads were subsequently assembled *de novo* using Trinity software [43], and the constructed unigenes were referred to as genes in this study. All unigenes were functionally annotated with the NR [44], Pfam [45], Swiss-Prot [46], COG [47], KEGG [48], and GO [49] databases.

Screening and expression level validation of DEGs

FPKM was used to normalize the mapped read counts for the gene expression level estimation. DEGs were obtained with DEGSeq2 program [50] and P-value ≤ 0.05, with the Benjamini and Hochberg method [15]. To further study the biological significance of the DEGs, Go and KEGG enrichment analysis was completed with GOseq R package [52] and KOBAS software [53], respectively. Accuracy of the expression levels of some DEGs was determined [15] and sequences of primers for qRT-PCR are listed in Additional file 1: Table S1.

Abbreviations

QTLs: quantitative trait loci; DEGs: differentially expressed genes: GWAS: Genome-Wide Association Studies; BSA: bulk segregant analysis; DAP: days after pollination; TEM: Transmission electron microscopy; ChlH: Mg-chelatase H subunit; IPP: isopentenyl diphosphate; DXS: Deoxy-d-xylulose-5-phosphate synthase; ChlD: Mg-chelatase D subunit; CCS: Capsanthin/Capsorubin synthase; MEP: 2-C-methyl-d-erythritol 4-phosphate; Psy: phytoene synthase gene; DXR1: 1-deoxy-D-xylulose-5-phosphate reductoisomerase; PSY: phytoene synthase gene; ZEP: zeaxanthin epoxidase.

Declarations

Authors' contributions

DL, YY, SY and LY designed the experiment and revised the manuscript. HY and DS carried out the pigment content and structure analysis, HZ, JHB, SS, and CM performed most experiment and wrote the manuscript. XW, MZ and JH collected and prepared RNA samples for RNA-seq and performed some RT-PCR. All authors read and approved the final manuscript.

Acknowledgements

We thank Professor Yuxiang Yuan and Dr. Xiaochun Wei from the Biochemistry and Molecular Platform of Henan Academy of agricultural sciences, for their help in providing experimental instruments and instructions.

Funding

The transcriptomic data analysis in this study were financially supported by the China Postdoctoral Science Foundation (2018M630823), and the experimental material was supported by National Natural Science Foundation of China (31902041) and Key Scientific Research Projects of the Higher Education Institutions of Henan Province (19A210002). The funding bodies played no role in not only the design of the study but also the collection and analysis of data and manuscript writing.

Availability of data

Genbank accession for the RNA-sequencing dataset of WT4 and WM102 is PRJNA549842 and PRJNA597690 is the accession number of WT2 and WT13.

Ethics approval and consent to participate

Not application.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Dou J, Lu X, Ali A, Zhao S, Zhang L, He N, Liu W: **Genetic mapping reveals a marker for yellow skin in watermelon (Citrullus lanatus L.)**. *Plos One* 2018, **13**(9):e0200617.
2. Rhodes B, Zhang X: **Gene list for watermelon**. *Cucurbit Genet Coop* 1995.
3. Lacerna M, Bayogan E: **Rind color change and granulation in pummelo [Citrus maxima (Burm. ex Rumph.) Merr.] fruit as influenced by 1-methylcyclopropene**. *International Food Research Journal* 2018, **25**(4).
4. Zheng X, Zhu K, Sun Q, Zhang W, Wang X, Cao H, Tan M, Xie Z, Zeng Y, Ye J: **Natural Variation in CCD4 Promoter Underpins Species-specific Evolution of Red Coloration in Citrus Peel**. *Molecular Plant* 2019.

5. Ohmiya A, Kishimoto S, Aida R, Yoshioka S, Sumitomo K: **Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals**. *Plant Physiology* 2006, **142**(3):1193-1201.
6. Sooriyapathirana SS, Khan A, Sebolt AM, Wang D, Bushakra JM, Lin-Wang K, Allan AC, Gardiner SE, Chagné D, Iezzoni AF: **QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (Prunus avium L.)**. *Tree genetics & genomes* 2010, **6**(6):821-832.
7. Li Y, Wen C, Weng Y: **Fine mapping of the pleiotropic locus for black spine and orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB transcription factor**. *Theoretical & Applied Genetics* 2013, **126**(8):2187-2196.
8. Jiang B, Liu W, Xie D, Peng Q, He X, Lin YE, Liang Z: **High-density genetic map construction and gene mapping of pericarp color in wax gourd using specific-locus amplified fragment (SLAF) sequencing**. *Bmc Genomics* 2015, **16**(1):1035.
9. Wehner TC: **Watermelon**. In: *Vegetables I.* Springer; 2008: 381-418.
10. Gusmini G, Wehner TC: **Qualitative Inheritance of Rind Pattern and Flesh Color in Watermelon**. *Journal of Heredity*, **97**(2):177-185.
11. Porter D: **Inheritance of certain fruit and seed characters in watermelons**. *Hilgardia* 1937, **10**(12):489-509.
12. Weetman LM: **Inheritance and correlation of shape size, and color in the watermelon**. *Iowa Agric Exp Stn Res Bull* 1937, **228**:222-256.
13. Oren E, Tzuri G, Vexler L, Dafna A, Meir A, Saar U, Faigenboim A, Kenigswald M, Portnoy V, Schaffer AAJb: **Multi-allelic APRR2 Gene is Associated with Fruit Pigment Accumulation in Melon and Watermelon**. 2019:542282.
14. Kumar R, Wehner TC: **Discovery of second gene for solid dark green versus light green rind pattern in watermelon**. *Journal of Heredity* 2011, **102**(4):489-493.
15. Liu D, Yang H, Zhang M, Yang S, Wang X, Zhu H, Wei X, Yuan Y, Yang L: **Comparative Transcriptome Analysis Provides Insights into Yellow Rind Formation and Fine Mapping of the clyr (yellow rind) Gene in Watermelon**. *Frontiers in Plant Science* 2020, **11**:192.
16. Luo F, Cheng S-C, Cai J-H, Wei B-D, Zhou X, Zhou Q, Zhao Y-B, Ji S-J: **Chlorophyll degradation and carotenoid biosynthetic pathways: Gene expression and pigment content in broccoli during yellowing**. *Food Chemistry* 2019:124964.
17. Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z: **The draft genome of watermelon (Citrullus lanatus) and resequencing of 20 diverse accessions**. *Nature Genetics* 2013, **45**(1):51-U82.
18. Levi A, Hernandez L, Thimmapuram J, Donthu R, Wright C, Ali C, Wechter W, Reddy U, Mikel M: **Sequencing the genome of the heirloom watermelon cultivar Charleston Gray**. In: *XX Plant and Animal Genome Conference, San Diego pp: 2011*; 2011: P047.
19. DellaPenna D, Pogson BJJARPB: **Vitamin synthesis in plants: tocopherols and carotenoids**. 2006, **57**:711-738.

20. Grotewold EJARPB: **The genetics and biochemistry of floral pigments**. 2006, **57**:761-780.
21. Lichtenthaler HK, Schwender J, Disch A, Rohmer MJFl: **Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway**. 1997, **400**(3):271-274.
22. Lichtenthaler HK, Rohmer M, Schwender JJPp: **Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants**. 1997, **101**(3):643-652.
23. Carretero-Paulet L, Cairo A, Botella-Pavía P, Besumbes O, Campos N, Boronat A, Rodríguez-Concepción MJPmb: **Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase**. 2006, **62**(4-5):683-695.
24. Yan X, Zhang L, Wang J, Liao P, Zhang Y, Zhang R, Kai GJApp: **Molecular characterization and expression of 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) gene from Salvia miltiorrhiza**. 2009, **31**(5):1015-1022.
25. Cunningham FX, Lafond TP, Gantt EJJob: **Evidence of a role for LytB in the nonmevalonate pathway of isoprenoid biosynthesis**. 2000, **182**(20):5841-5848.
26. Bramley P, Teulieres C, Blain I, Bird C, Schuch W: **Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5**. *The Plant Journal* 1992, **2**(3):343-349.
27. Shewmaker CK, Sheehy JA, Daley M, Colburn S, Ke DY: **Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects**. *The Plant Journal* 1999, **20**(4):401-412.
28. Fanciullino A-L, Dhuique-Mayer C, Luro F, Casanova J, Morillon R, Ollitrault PJJoA, Chemistry F: **Carotenoid diversity in cultivated citrus is highly influenced by genetic factors**. 2006, **54**(12):4397-4406.
29. Cui Y-L, Jia Q-S, Yin Q-Q, Lin G-N, Kong M-M, Yang Z-NJPp: **The GDC1 gene encodes a novel ankyrin domain-containing protein that is essential for grana formation in Arabidopsis**. 2011, **155**(1):130-141.
30. Zhang L, Liu C, An X, Wu H, Feng Y, Wang H, Sun D: **Identification and genetic mapping of a novel incompletely dominant yellow leaf color gene,Y1718, on chromosome 2BS in wheat**. *Euphytica*, **213**(7).
31. Guo S, Zhao S, Sun H, Wang X, Wu S, Lin T, Ren Y, Gao L, Deng Y, Zhang J: **Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits**. *Nature Genetics* 2019:1-8.
32. Li B, Zhao S, Dou J, Ali A, Gebremeskel H, Gao L, He N, Lu X, Liu W: **Genetic mapping and development of molecular markers for a candidate gene locus controlling rind color in watermelon**. *Theoretical and Applied Genetics* 2019, **132**(10):2741-2753.
33. Ampomah-Dwamena C, Thrimawithana AH, Dejnoprat S, Lewis D, Espley RV, Allan ACJNP: **A kiwifruit (Actinidia deliciosa) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation**. 2019, **221**(1):309-325.

34. Gang H, Li R, Zhao Y, Liu G, Chen S, Jiang JJJoeb: **The birch GLK1 transcription factor mutant reveals new insights in chlorophyll biosynthesis and chloroplast development**. 2019.
35. Jin W, Wang H, Li M, Wang J, Yang Y, Zhang X, Yan G, Zhang H, Liu J, Zhang KJPbj: **The R 2 R 3 MYB transcription factor P av MYB 10.1 involves in anthocyanin biosynthesis and determines fruit skin colour in sweet cherry (P runus avium L.)**. 2016, **14**(11):2120-2133.
36. Furukawa T, Maekawa M, Oki T, Suda I, Iida S, Shimada H, Takamure I, Kadowaki KiJTPJ: **The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp**. 2007, **49**(1):91-102.
37. Meng R, Zhang J, An L, Zhang B, Jiang X, Yang Y, Zhao ZJJopgr: **Expression profiling of several gene families involved in anthocyanin biosynthesis in apple (Malus domestica Borkh.) skin during fruit development**. 2016, **35**(2):449-464.
38. Li Y, Wen C, Weng YJT, genetics a: **Fine mapping of the pleiotropic locus B for black spine and orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB transcription factor**. 2013, **126**(8):2187-2196.
39. Lun Y, Wang X, Zhang C, Yang L, Gao D, Chen H, Huang SJE: **A CsYcf54 variant conferring light green coloration in cucumber**. 2016, **208**(3):509-517.
40. Adato A, Mandel T, Mintz-Oron S, Venger I, Levy D, Yativ M, Domínguez E, Wang Z, De Vos RC, Jetter RJPG: **Fruit-surface flavonoid accumulation in tomato is controlled by a SlMYB12-regulated transcriptional network**. 2009, **5**(12):e1000777.
41. Ballester A-R, Molthoff J, de Vos R, te Lintel Hekkert B, Orzaez D, Fernández-Moreno J-P, Tripodi P, Grandillo S, Martin C, Heldens JJPp: **Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color**. 2010, **152**(1):71-84.
42. Harmut A: **Chlorophylls and carotenoids: pigments of photosynthetic membranes**. *Methods Enzymol* 1987, **148**:350-383.
43. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q: **Full-length transcriptome assembly from RNA-Seq data without a reference genome**. *Nature Biotechnology* 2011, **29**(7):644.
44. Pruitt KD, Tatusova T, Maglott DR: **NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins**. *Nucleic acids research* 2005, **33**(suppl_1):D501-D504.
45. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J: **Pfam: the protein families database**. *Nucleic Acids Research* 2014, **42**(Database issue):222-230.
46. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer ELL: **The Pfam Protein Families Database**. *Nucleic Acids Research* 2008, **32**(1):D138.
47. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN: **The COG database: an updated version includes eukaryotes**. *Bmc Bioinformatics* 2003, **4**(1):41-41.

48. Kanehisa M: **The KEGG database**. *Novartis Foundation Symposium* 2002, **247**(247):91.
49. Ashburner M, Ball CA, Blake JA, Botstein D, Cherry JM: **Gene ontology: Tool for the unification of biology**. *Nature Genetics* 2000, **25**(1):25-29.
50. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome Biology* 2014, **15**(12):550.
51. Benjamini Y, Hochberg Y: **On the adaptive control of the false discovery rate in multiple testing with independent statistics**. *Journal of educational and Behavioral Statistics* 2000, **25**(1):60-83.
52. Young MD, Wakefield MJ, Smyth GK, Oshlack A: **goseq: Gene Ontology testing for RNA-seq datasets**. *R Bioconductor* 2012.
53. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes**. *Nucleic acids research* 2000, **28**(1):27-30.

Figures

Figure 1

Fruit of watermelon cultivars at 14 DAP. (A) WM102; (B) WT2; (C) WT13; (D) WT4.

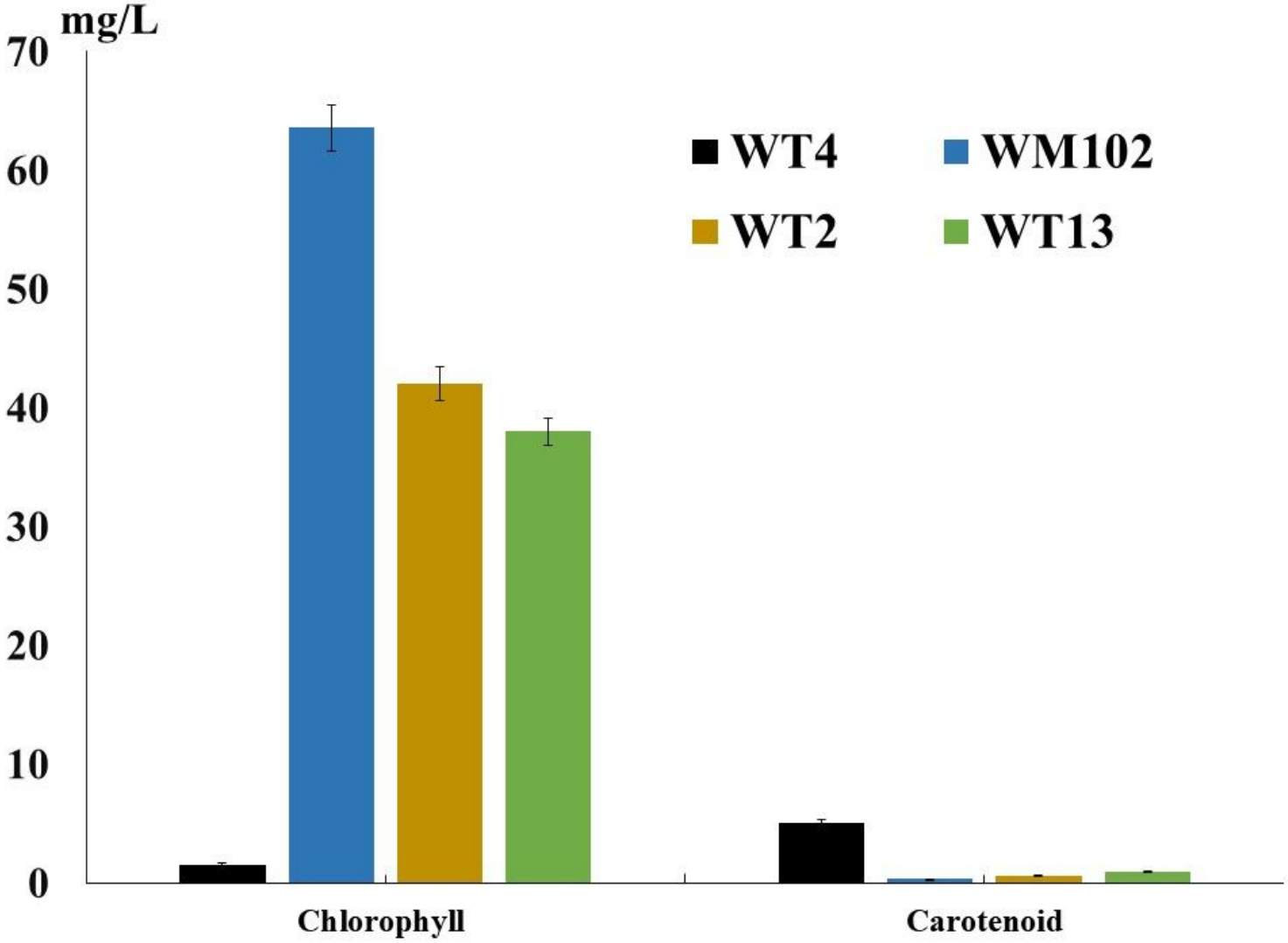

Figure 2

Chlorophyll and carotenoid content in rind of watermelon material WM102, WT2, WT13, and WT4.

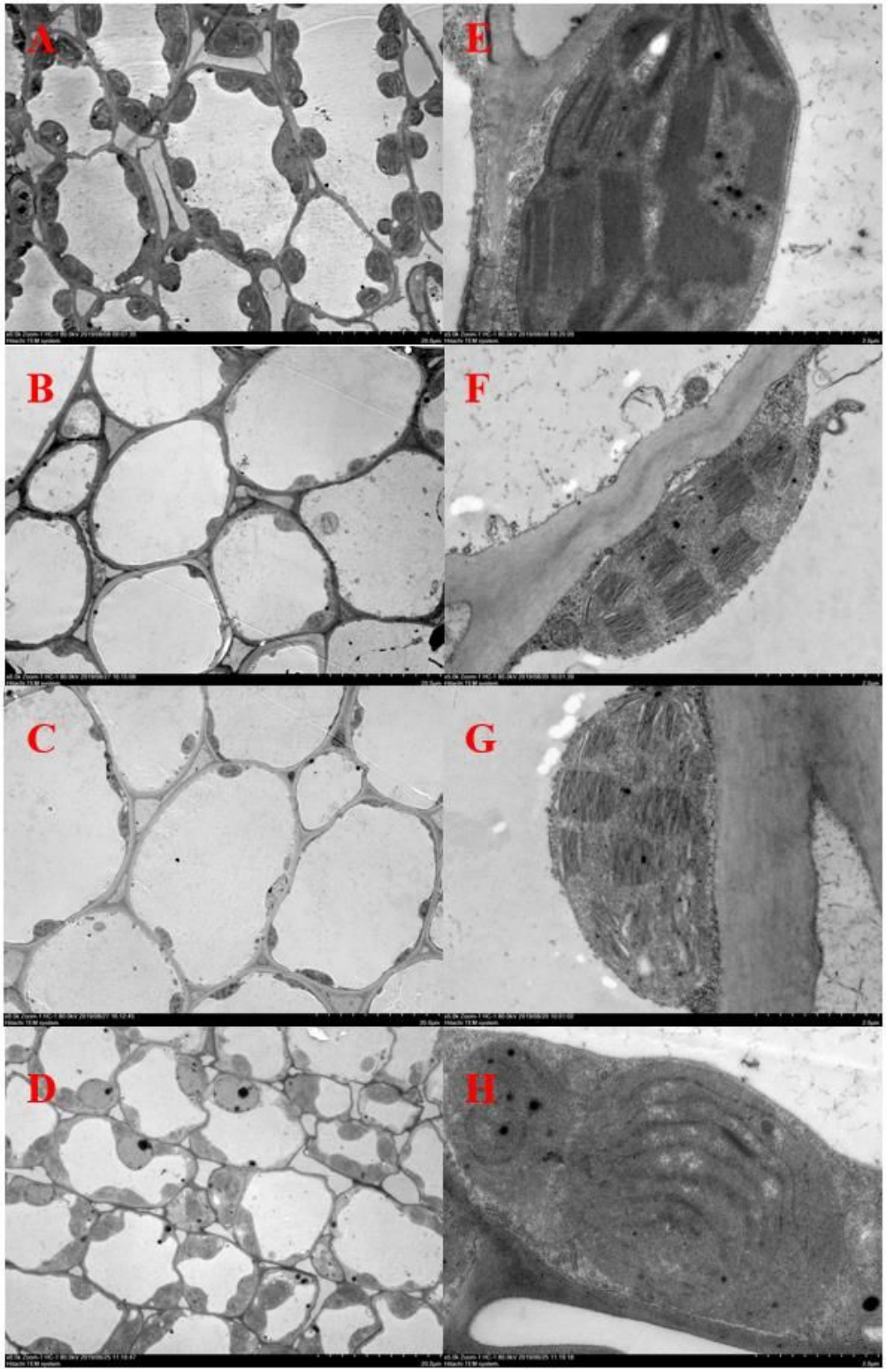

Figure 3

TEM scope of the rind of four materials. A, B, C, and D shows the low-magnification TEM images of material WM102, WT2, WT13, and WT4, respectively. Scale bar=20 μm. E, F, G, and H shows the high-magnification TEM images of material WM102, WT2, WT13, and WT4, respectively. Scale bar=2 μm.

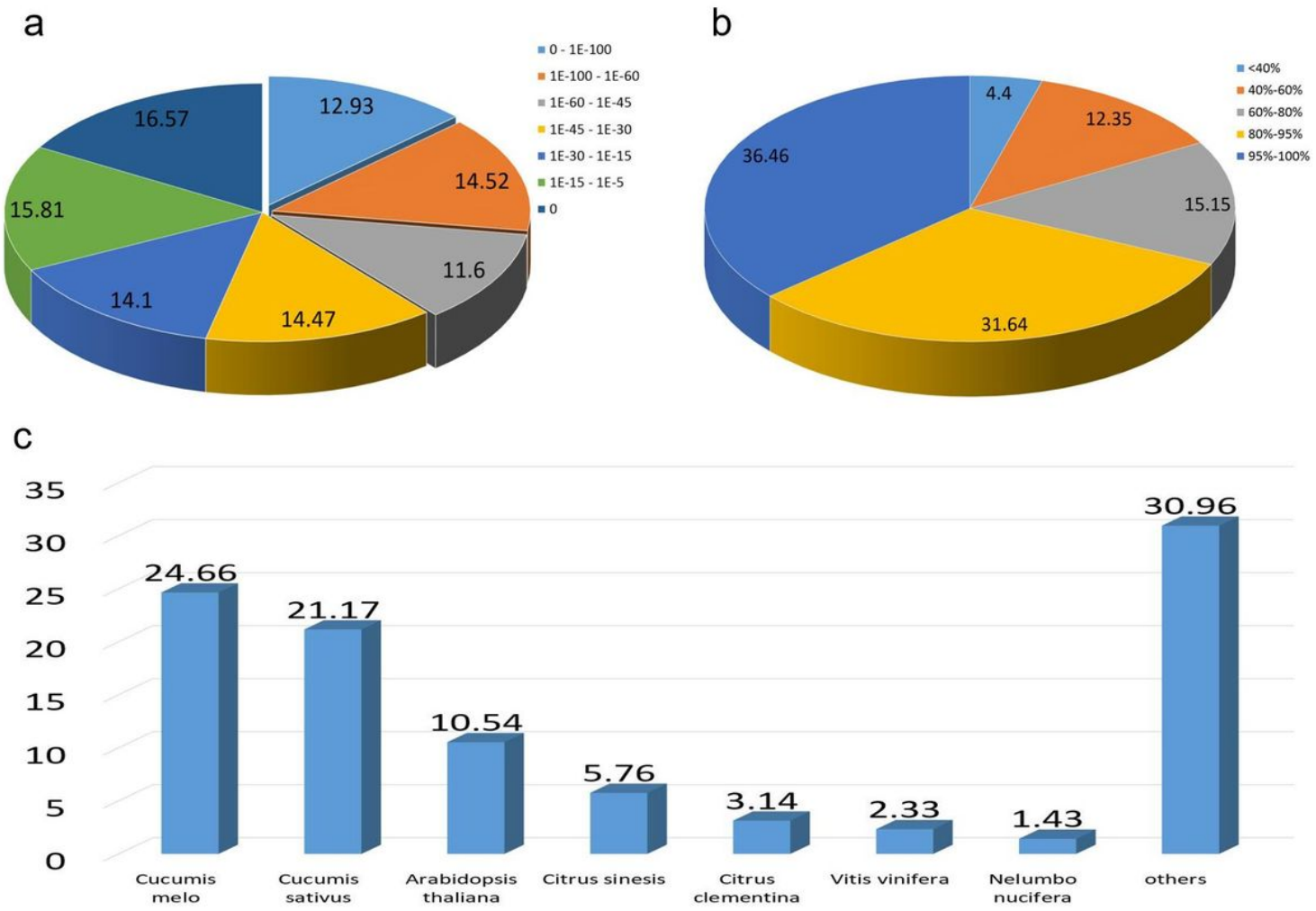

Figure 4

Characteristics of homology analysis for C. lanatus unigenes against the Nr database with an E-value of 10−5. (a) Te E-value distribution of BLASTx hits for each assembled unigene (%); (b) Te similarity distribution of BLASTx hits for each assembled unigenes (%); (c) Species-based distribution of the top BLASTx hits for each assembled unigenes (%).

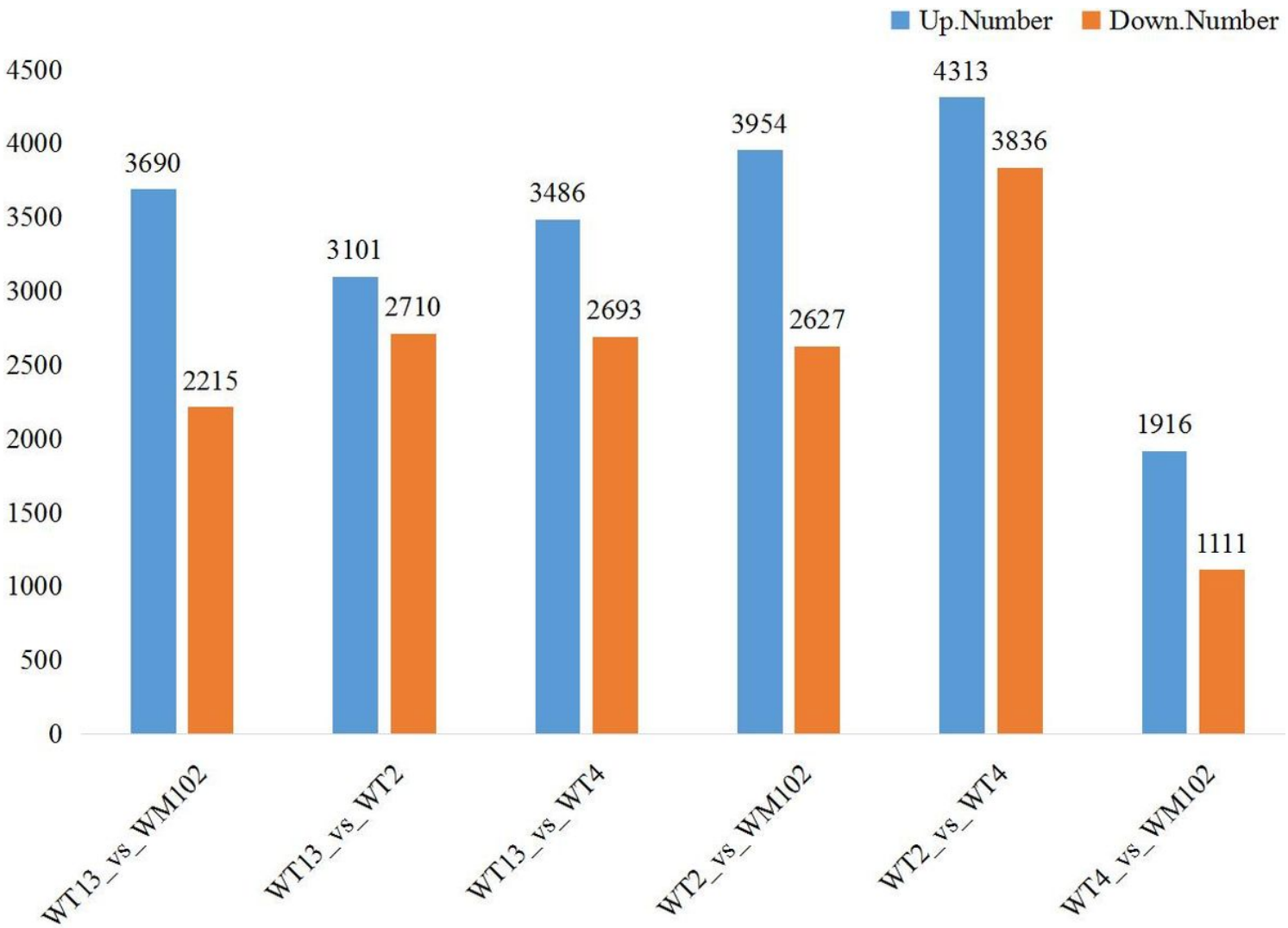

Figure 5

Number of DEGs among four watermelon lines.

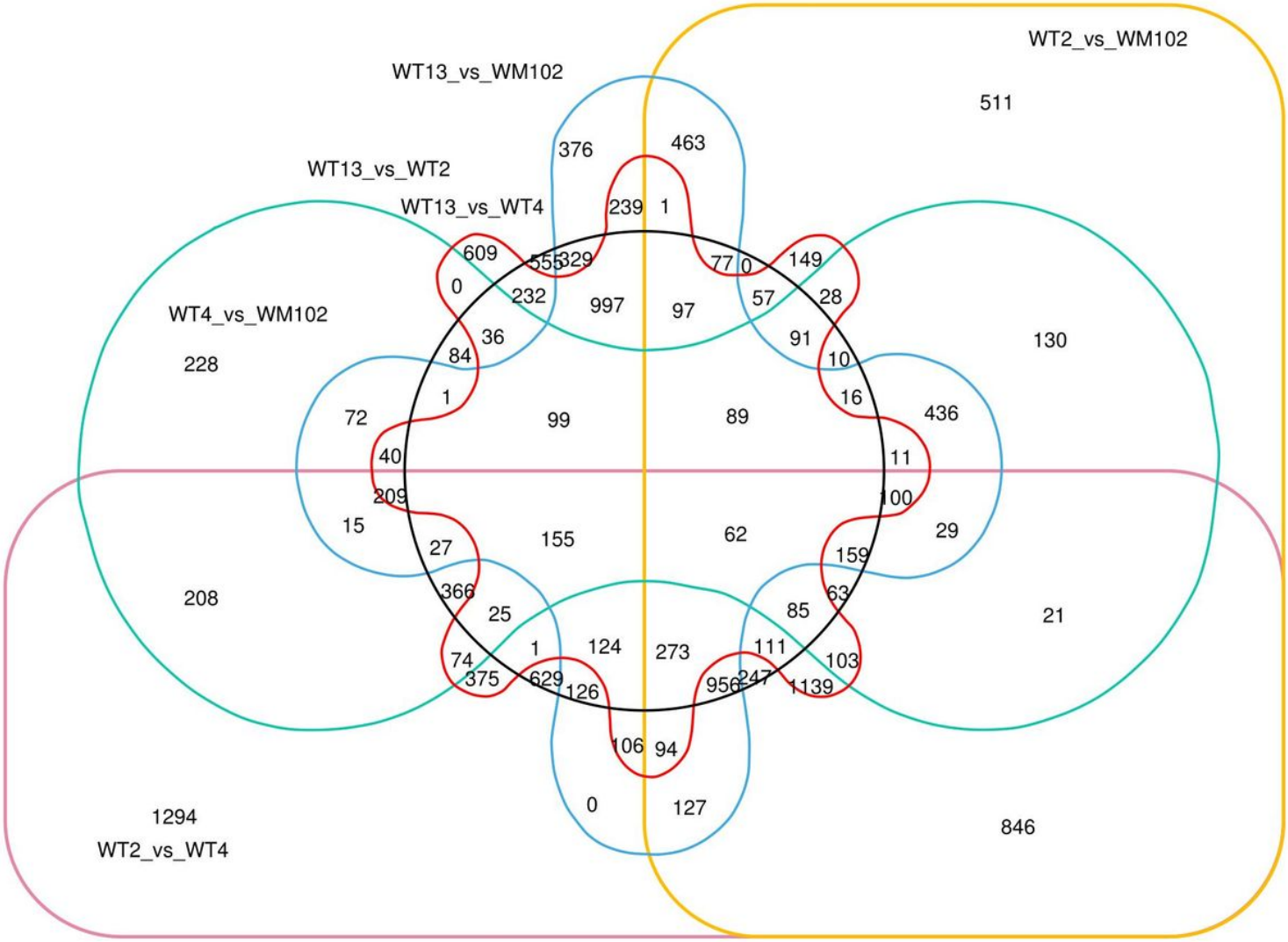

Figure 6

Venn maps of DEGs.

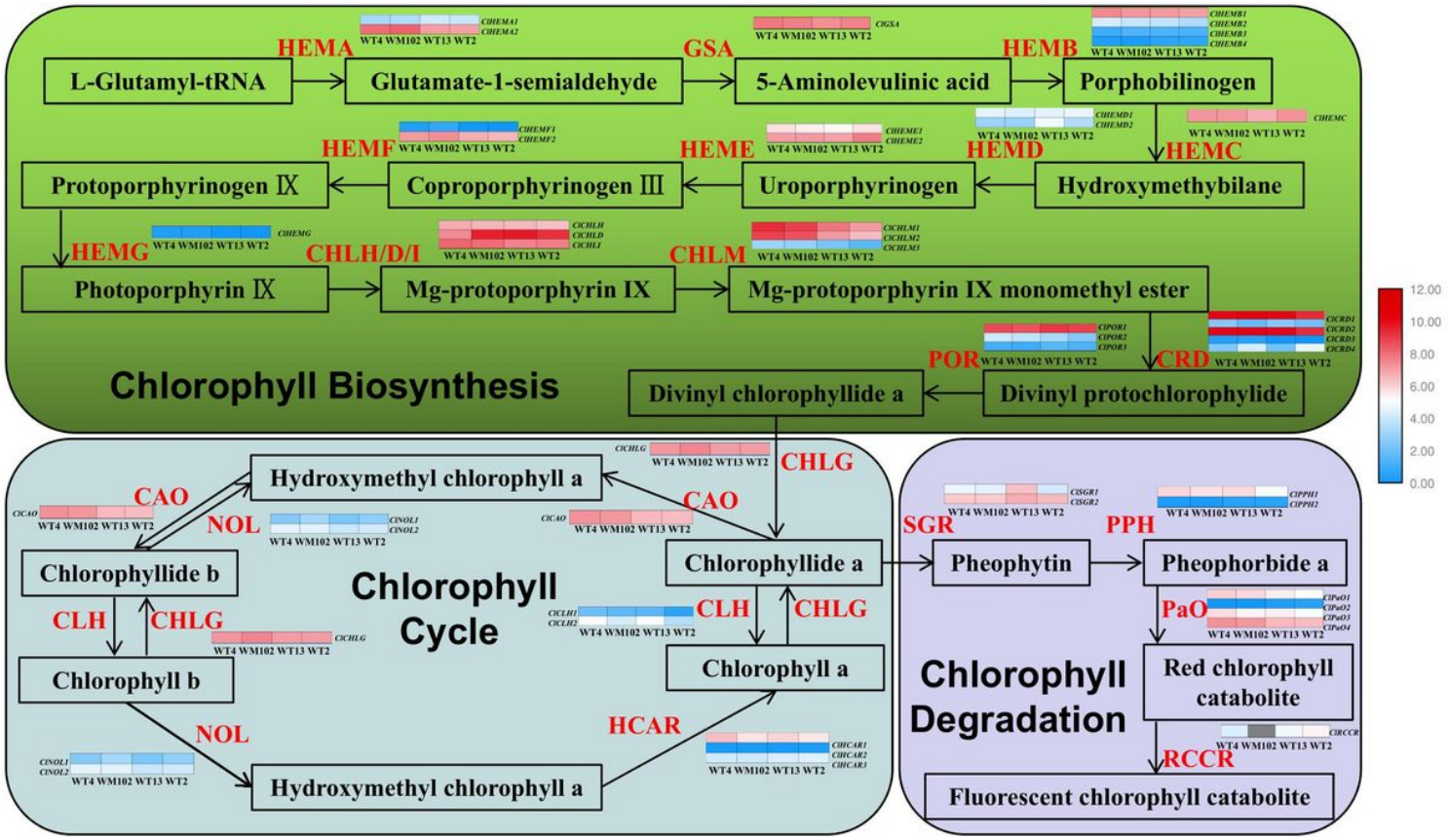

Figure 7

The Chl metabolic pathway in higher plants.

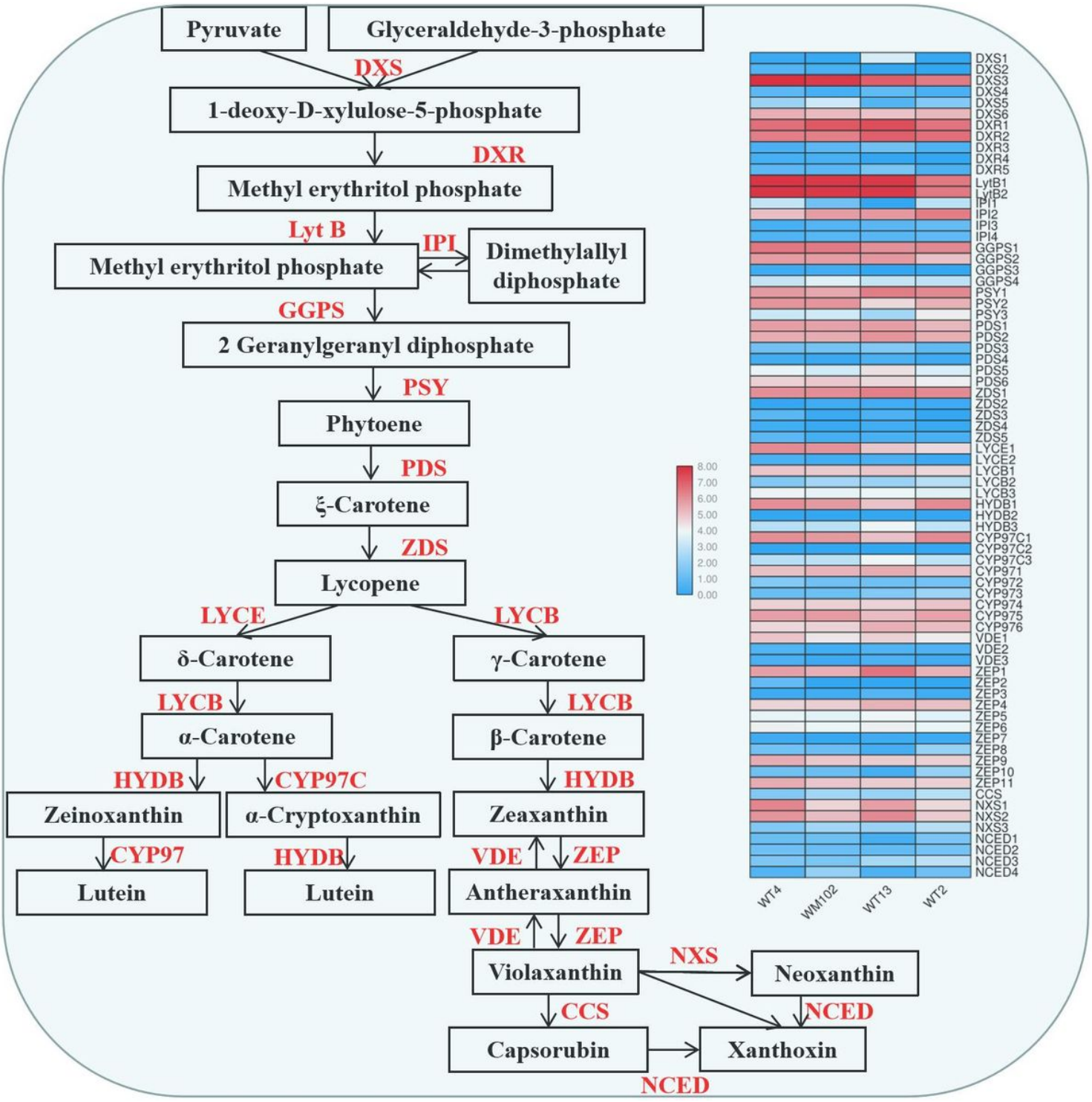

Figure 8

Carotenoid biosynthetic pathway in higher plants.

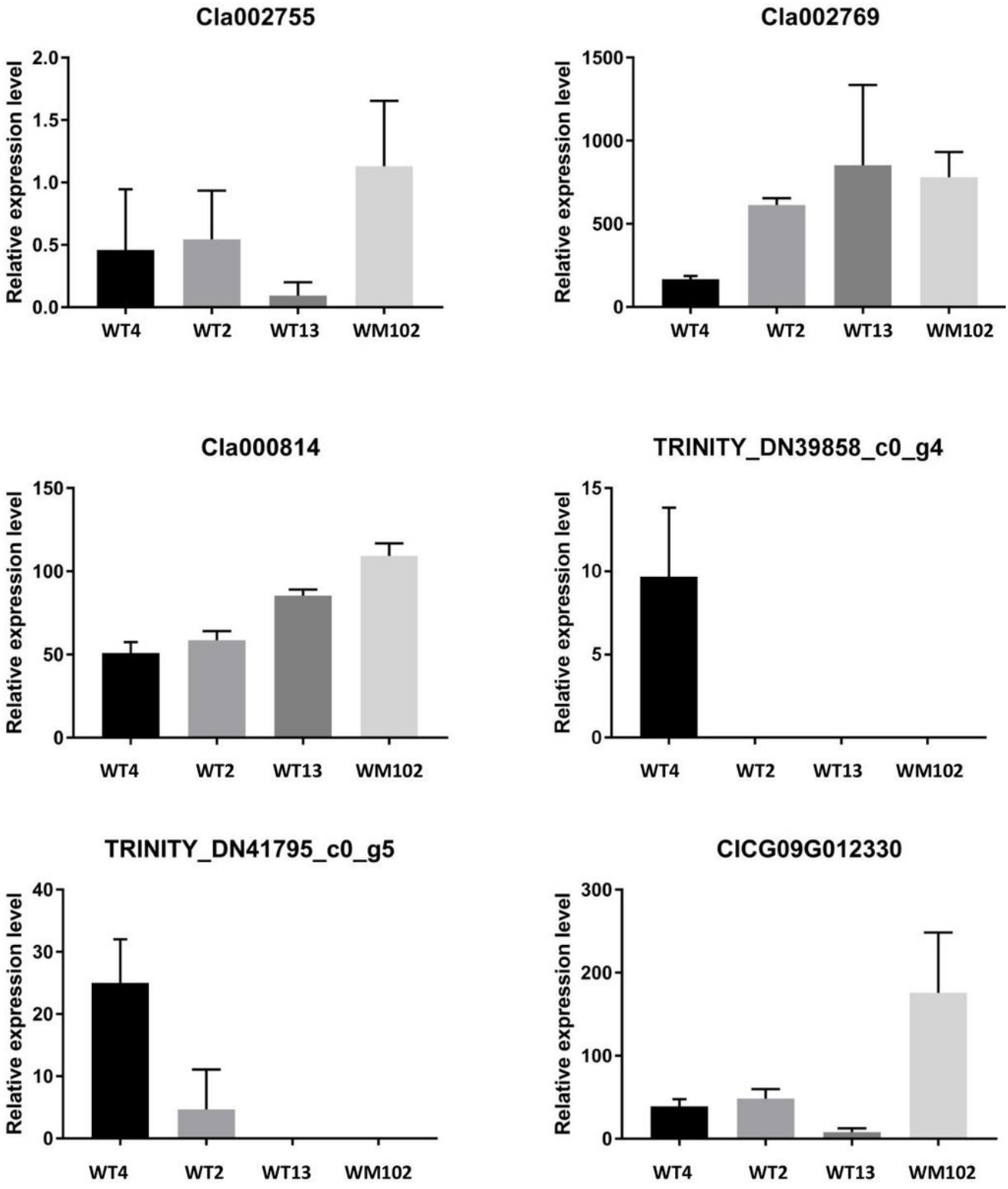

Figure 9

Expression of watermelon rind coloration related genes detected in the transcriptome data.

Supplementary Files

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

- Additionalfile6.doc
- Additionalfile5.doc
- Additionalfile4.doc
- Additionalfile3.doc
- Additionalfile2.doc
- Additionalfile1.xls