Integrated metabolome and transcriptome reveals the mechanism of the flower coloration in cashew (Anacardium occidentale Linn.)

Zhongrun Zhang (✉ zhrzhang@126.com)  
Chinese Academy of Tropical Agricultural Sciences

Weijian Huang  
Chinese Academy of Tropical Agricultural Sciences

Liyan Xiao  
Chinese Academy of Tropical Agricultural Sciences

Haijie Huang  
Chinese Academy of Tropical Agricultural Sciences

Research Article

Keywords: cashew, flower coloration, anthocyanin, transcriptome, metabolome, structural genes, metabolites

Posted Date: September 26th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2039851/v1

License: ☑️ ☐ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background

Cashew is a widely attractive food in the world with high nutritional value for human health. The quality of cashew is closely associated with the flower coloration. However, the underlying mechanism of flower coloration of cashew was poorly understood. Here, we identified two cashew cultivars with variable flower coloration. Then, integrated analysis of metabolome and transcriptome was performed to explore the potential mechanisms.

Results

The results showed that 8 flavonoid compounds were markedly accumulated in red flower cashew, including taxifolin, procyanidin B2, delphinidin-3-O-glucoside, peonidin-3-glucoside, L-epicatechin, chalconaringenin, luteolin, naringenin and chrysin. In particular, delphinidin-3-O-glucoside, peonidin-3-glucoside and procyanidin B2 were key metabolites that promoted the formation of red flower coloration. Transcriptome analysis revealed that genes involved in flavonoid biosynthesis were also varied among these two cultivars. The flavonoid biosynthesis direction was reprogrammed to synthesize delphinidin-3-O-glucoside, peonidin-3-glucoside and procyanidin B2. Typically, four core genes encoding anthocyanin synthesis were also highly expressed in cashew with red flower, including CFI (TRINITY_DN4346_c0_g1), CHS2 (TRINITY_DN16129_c0_g1), CHS1 (TRINITY_DN2623_c0_g1) and F3H (TRINITY_DN8780_c0_g1).

Conclusions

This study sheds light on flavonoid metabolic pathways and candidate genes varied in two cashew cultivars and lays a foundation for the breeding programs of cashew.

Background

Cashew (Anacardium occidentale Linn) is an evergreen tree of Anacardium genus in the family of Anacardiaceae, which is main economic crops in tropical areas [1, 2]. The cashew nut is one of the four famous dried fruits and widely attractive in the world. It possesses abundant nutrients including unsaturated fatty acid, microelements, polysaccharose and amino acids, which are beneficial to human health such as anti-aging, antioxidation and antifatigue [3]. Besides, other cashew organisms such as roots, leaves could be made into tea beverages [4], and some studies have shown that the extract of cashew conferred a certain therapeutic effect on diabetes by promoting cells uptake blood sugar. [5, 6]. Thus, cashew planting industry has great economic, social and medical benefits. To date, the planting areas of cashew was mainly distributed in Hainan and Yunnan province in China, especially Ledong county in Hainan is the major agricultural region with cultivation area accounting for more than 60% [7]. However, cashew planting industry was limited by numerous abiotic and biotic stress, including drought,
low temperature, pathogens attack and insects herbivory [1, 8, 9]. Much effort in germplasm resources introduction, high-yield cultivation and varieties breeding and promotion has been made to deal with these stress. The identification of cashew cultivars was always the focus. Of note, flower coloration was a conspicuous phenotype to distinguish varieties and we found two cashew varieties with different colors of bisexual flowers in our cashew germplasm bank. Hence our work aimed to reveal the mechanisms of flower coloration in cashew plants, which was important to elucidate the characterization of cashew cultivars contributing to the development of cashew industry of China.

Flower coloration was a complicated and interesting phenomenon in nature, which could contribute to entomophily pollination and biotic and abiotic endurance for plants [10, 11, 12]. Much progress have been made in deciphering the color formation in recent years. It is well-known that three major classes of pigments including flavonoids/anthocyanins, carotenoids and chlorophylls were universally distributed in plants. Among them, flavonoids/anthocyanins and carotenoids were mainly involved in yellow, blue and red flower coloration, whereas chlorophylls were a class of essential photosynthetic compounds that exist nearly all plants and were mainly involved in green coloration in plants [13, 14, 15]. Of note, as one of main subgroups of flavonoids, anthocyanins were mainly responsible for the coloration of flower of red, blue, purple and deepening colors. Anthocyanins are water-soluble pigments which are mainly deposited in vacuoles of cells. The main anthocyanins in plants include geranium, centaurea, paeoniflorin, petunia, delphinium and mallow, and their proportion in cell determined the final coloration of flower [15]. Further studies have shown that anthocyanins are synthesized through the flavonoid pathways and are regulated by many structural genes, including chalcone synthase (CHS), chalcone isomerase (CHI), dihydroflavonol 4-reductase (DFR), flavanone 3-hydroxylase (F3H), flavonoid 3-O-glucosyltransferase (UFGT), anthocyanidin-O-methyltransferase (AOMT), flavonol synthase (FLS), flavone synthase (FNS) and anthocyanidin synthase (ANS) [16–18]. Furthermore, some transcription factors (TFs) were extensively reported to regulate above structural genes by binding their promoters, such as basic leucine zipper domain transcription factor bZIP, MYB, bHLH and MADS-box [19–21]. Besides, increasing evidence showed that the protein complex co-regulated the biosynthesis of anthocyanins that was relatively authentic regulation network in plants, such as MBW (MYB-bHLH-WD) and WMBW (WD-MYB-bHLH-WRKY) complex were discovered in multiple plants [22–24].

Substantial documents have stressed the importance of anthocyanins in flower coloration, yet the anthocyanins in flower of cashew were still poorly understood. Importantly, the flower coloration of cashew considered to be one of the pivotal indexes to evaluate characterizations of cashew. Meanwhile, the accumulation of anthocyanins and related flavonoids were also closely implicated with the capacity of stress resistance and ornamental value [25, 26]. In the present study, we identified two typical cashew cultivars with different flower colors and occurrence in color transformation. Based on the analysis of metabolome and transcriptome, we identified the candidate metabolites and genes involved in the formation of flower coloration, thus providing a new understanding of the molecular mechanism of cashew flower coloration. These results expected to contribute to breeding and promotion of cashew.
Results

The variation in flower color of B2_5 and B7 cultivars in cashew altered metabolome and transcriptome patterns

We obtained two cashew cultivars designated as B2_5 and B7 in germplasm resource in Ledong county. Interestingly, we found the flower colors of B2_5 and B7 were white when they initially bloomed, but the colors of mature flowers varied among both cultivars. During the mature stage, the flower color of B2_5 cultivar was always white, whereas the flower color of B7 turned red in its mature stage (Fig. 1A). Considering the function of flavonoids and carotenoids in flower coloration, it is possible that the contents of both kinds of metabolites were differentially accumulated among four stages of cashew flowers. Thus metabolomics was performed on both kinds of cultivars at mature and bloom stages. PCA plots showed conspicuously separate groups among four samples that verified the accumulation of different metabolites that might eventually the distinction in flower coloration (Fig. 1B). Further, we exhibited the expression of all metabolites in a heatmap based on unsupervised hierarchical clustering analysis, indicating that authentic variation of metabolites between two cashew cultivars with different flower colors (Fig. 1D). Meanwhile, we also observed the transcripts fluctuation among these samples. The four clear separations among samples from transcriptomic profiles showed the 35% variations and 25% variations represented by PC1 and PC2, respectively, suggesting that the difference in flower coloration between two cashew cultivars also caused significantly variable expression in transcriptomic profile among these groups (Fig. 1C). Consistently, all upregulated and downregulated transcripts among these samples were depicted in a heatmap (Fig. 1E). Overall, these results supported that the variation of flower colors in B2_5 and B7 markedly affected metabolome and transcriptome during different blooming stage in cashew.

The accumulation of flavonoid compounds promoted flower color change in B7 cultivar.

To investigate the phenomenon that the flower color of B7 cultivar turned red during its growing development, integrated analysis of transcriptome and metabolome for initial stage (N1B7) and mature stage (M1B7) was performed. We examined the metabolites accumulation of N1B7 and M1B7, the result showed that metabolites related to flavonoid biosynthesis, flavone and flavonol biosynthesis and phenylalanine metabolism were markedly enriched in comparison of M1B7 vs. N1B7 (Fig. 2A). Then we abstracted these metabolites and plotted them referring to their content in two groups with a heatmap (Fig. 2B), we found many flavonoid metabolites including taxifolin, isoquercitrin, delphinidin-3-O-glucoside, L-epicatechin, chalconaringenin, luteolin, naringenin, hesperetin, rutin, chrysirin, hesperetin 7-O-glucoside, kaempferol, cianidanol, fisetin and peonidin-3-glucoside were significantly upregulated as well as N-((S)-jasmonoyl)-S-isoleucine and kaurenoic acid belonging to phytohormone and diterpenoids respectively in M1B7. Increasing evidence supported that accumulation of flavonoids contributed to deepen the color of fruits or leaves [21, 27]. In the contrast, we noticed that multiple phytohormones including gibberellin, salicylic acid and abscisic acid as well as phenylalanine compounds containing phenylalanine isomerides and phenethylamine were marked enriched in N1B7. We reasoned many plant
hormones molecules probably participated in regulating the transformation in flower coloration in B7 cashew cultivar. Subsequently, we further analyzed the transcriptomic patterns of M1B7 vs. N1B7 to explore potential transcripts changes caused variable content of metabolites relevant to flavonoid and phenylalanine compounds (Fig. 2D). KEGG pathway enrichment analysis of M1B7 vs N1B7 indicated that DEGs involved in phytohormone signal transduction, amino acid metabolism, carbon source metabolism and secondary metabolites biosynthesis were enriched (Fig. 2C). Especially we found the biosynthesis pathways involved in flavonoid and phenylalanine compounds were conspicuously enriched. Then synthetic pathway flow plots were depicted based on DEMs and DEGs associated with flavonoid and phenylalanine biosynthesis (Fig. 2E and S1). We found gene TRINITY_DN4346_c0_g1 annotated as Chalcone-flavanone isomerase was positively related with accumulation of chrysin, naringenin, hesperetin and hesperetin 7-O-glucoside; gene TRINITY_DN8652_c0_g1 annotated as Flavonoid monoxygenase was negatively related with accumulation of luteolin, isoquercitrin and rutin; two TIFY gene annotated as TRINITY_DN6010_c0_g1 and TRINITY_DN89_c0_g1 were negatively associated with accumulation of JA (Fig. S1).

The accumulation of flavonoid and phenylalanine compounds caused the difference of flower color in M1B7 and M1B2_5.

We also noticed that M1B7 and M1B2_5 exhibited different flower colors in the period of mature stage. First, we compared the metabolites accumulation between M1B7 and M1B2_5, indicating that 155 metabolites were increased and 79 metabolites were reduced in the comparison of M1B7 vs. M1B2_5 (Fig. 3A). Then we performed a KEGG pathway enrichment analysis to inquire the distribution of these metabolites. We found increased metabolites were mainly enriched in alpha-linolenic acid metabolism, arginine and proline metabolism, aminoacyl-tRNA biosynthesis and flavonoid biosynthesis, and metabolites related to sesquiterpenoid and triterpenoid biosynthesis, folate biosynthesis and tryptophan metabolism were reduced in M1B7 vs. M1B2_5 (Fig. 3C). We abstract these metabolites to construct a heatmap showing specific relative content of metabolites (Fig. 3B), we found that many flavonoid compounds such as quercitrin, genistein, chrysin, L-epicatechin, chalcone-arigenin, naringenin, taxifolin, procyanidin B2, tiliroside and kaempferol-3-O-rutinoside were significantly increased in M1B7, which were closely implicated with flower color formation. Meanwhile, we noticed that some phenols were increased compared with M1B2_5, including p-anisic acid, 2-(4-hydroxyphenyl) ethanol, trimethoprim, 3,4-dihydroxybenzaldehyde, 3,4,5-trimethoxycinnamyl alcohol, phenethyl alcohol, 2-phenylethanol, caffeic acid, benzaldehyde, n-(p-hydroxyphenethyl) actinidine, phloretic acid;3-(2-hydroxyphenyl) propanoic acid and ethylparaben, which were reported to facilitate red, blue and purple color of flower upon the accumulation of these metabolites [28, 29]. As we expected, transcriptome analysis reached a consensus with metabolome that phenylpropanoid and flavonoid biosynthesis pathways exhibited more active than other pathways in KEGG enrichment analysis of M1B7 vs M1B2_5 (Fig. 3D). We next abstracted a total of 87 gene involved in two pathways to plot a heatmap, showing 45 genes with marked upregulation were separated into two groups which was implicated with phenylpropanoid and flavonoid biosynthesis (Fig. 3E). And metabolic pathway maps showed that genes associated with K00487 and K01859 were positively related to the accumulation of chrysin, naringenin chalcone, naringenin, dihydroquercetin and
epicaterchin and genes associated with K05280 and K13464 were negatively related to the accumulation of epicaterchin and quercitrin (Fig. S2). Overall, these results suggested that the change of genes and metabolites relevant to phenylpropanoid and flavonoid metabolic pathways led to the flower color difference between M1B7 and M1B2_5.

The accumulation of flavonoids facilitated formation of red flower.

To explore the potential mechanism for the formation of red flower in M1B7, we compared the gene and metabolite components between M1B7 vs M1B2_5 and M1B7 vs N1B7 with venn diagrams (Fig. 4A, C). these results indicated that 6 metabolites and 46 genes were shared in both groups, we therefore reasoned these metabolites such as chrysin, L-epicatechin, chalcone-aringenin, taxifolin, naringenin and ferulic acid were key metabolites that resulted in formation of red flower. Interestingly, we noticed that these metabolites belonged to flavonoids except for ferulic acid derived from phenylpropanoids. Then we abstracted specific content of core metabolites in metabolome, suggesting 5 flavonoid compounds were significantly accumulated in M1B7 whereas the content of ferulic acid and its isomers were reduced in M1B7 (Fig. 4B). Obviously, it was the accumulation of the 5 flavonoids that promoted the formation of red flower in mature B7 cultivar. To further investigate detailed transcriptional regulation associated with these metabolites, we plotted these 46 genes with a heatmap showing 27 upregulated and 19 downregulated genes were mainly related with flavanone synthesis (Fig. 4D). Furthermore, key metabolites and relevant genes were integrated in the metabolic pathway depicted in Fig. 5. It is speculated that variable gene expression of eight genes (Table 1) facilitated the accumulation of flavonoid compounds. Overall, we concluded that increased flavonoid components facilitated the transformation of flower color in B7 cultivar.
Table 1
Core genes involved in the accumulation of flavonoid components

<table>
<thead>
<tr>
<th>Numeration</th>
<th>ID</th>
<th>function</th>
<th>P-Value (M1B7 vs. N1B7)</th>
<th>P-Value (M1B7 vs. M1B2_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>TRINITY_DN4346_c0_g1</td>
<td>Chalcone–flavanone isomerase (CFI)</td>
<td>0.019*</td>
<td>0.10</td>
</tr>
<tr>
<td>B1</td>
<td>TRINITY_DN16129_c0_g1</td>
<td>Chalcone synthase 2 (CHS2)</td>
<td>0.008**</td>
<td>0.03*</td>
</tr>
<tr>
<td>B2</td>
<td>TRINITY_DN2623_c0_g1</td>
<td>Chalcone synthase 1 (CHS1)</td>
<td>0.03*</td>
<td>0.40</td>
</tr>
<tr>
<td>C</td>
<td>TRINITY_DN8780_c0_g1</td>
<td>Flavonol synthase/flavanone 3-hydroxylase (F3H)</td>
<td>0.04*</td>
<td>0.08</td>
</tr>
<tr>
<td>D</td>
<td>TRINITY_DN2918_c0_g1</td>
<td>Leucoanthocyanidin dioxygenase (LDOX)</td>
<td>0.02*</td>
<td>0.09</td>
</tr>
<tr>
<td>E</td>
<td>TRINITY_DN8652_c0_g1</td>
<td>Flavanone 3'-hydroxylase (F3H)</td>
<td>0.01*</td>
<td>0.02*</td>
</tr>
<tr>
<td>F</td>
<td>TRINITY_DN28797_c0_g1</td>
<td>Cytochrome P450</td>
<td>0.14</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*, P < 0.05; **, P < 0.01

Discussion

Plant shows variable colors attributes to distinct pigments deposition [14]. Elucidating the mechanisms of flower coloration was important for breeding excellent variety. Flavonoids consider to be essential components can facilitate plants exhibit different colors (ref). In particular, anthocyanidin as a popular pigment, is closely implicated with familiar colors such as red, blue and purple shown in leaves, flowers and fruits [30]. Thus, deciphering the function of both kinds of metabolites in flower coloration will contribute to our understanding of the mechanisms of flower coloration. As illustrated by our study, we identified two cashew cultivars named B7 and B2_5 exhibiting distinct flower colors in mature stage and the flower color of B7 transformed from white to red during its growing development. Of note, we found that the changes in flavonoids were closely associated with the formation of flower coloration in cashew that numerous flavoids highly accumulated in M1B7 comparing to N1B7 and M1B2_5. Despite some researches mentioned that carotenoid metabolites also facilitated the formation of red flower, our data showed no significance in the accumulation of carotenoids among pairwise comparisons. Further transcriptomic analysis manifested that genes involved in flavonoids were differentially expressed among these groups, which was consistant with metabolomics. Overall, our results provided the landscape of metabolic changes during the stage of flower coloration in cashew, which further
contributed to elucidate the mechanisms of flower coloration and provide theoretical basis for breeding excellent variety.

Increasing evidence confirmed that cyanidin, delphinidin, peonidin and their derivants with glucosyl participated in the formation of red coloration in flowers [31, 32]. Given the color transformation of B7 in the growing period, we found 6 flavonoid compounds were significantly accumulated with the growing development of B7. These metabolites including naringenin, luteolin, taxifolin, isoquercitrin, which were important substrates for anthocyanidin biosynthesis. Importantly, delphinidin 3-glucoside (Dp3g) and peonidin-3-glucoside (Pn3g) were highly accumulated in M1B7 and both were verified to facilitate the formation of red color in several plants [28]. Thus, the accumulation of Dp3g and Pn3g promoted transformation of red color in mature B7 cultivar.

Phytohormone serves as a crucial messenger to regulate multiple biological process in plants [33, 34]. Much progress has been made in deciphering plant-pathogen interaction, leaf senescence, temperature sensing, fruit coloring and so on [35]. Numerous studies have reported that ABA, ETH, GA and IAA influenced plant coloration by means of activating or suppressing transcription factors to regulate gene expression in anthocyanin synthesis pathway [36] (Liu et al., 2021). Interestingly, we noticed that JA was markedly accumulated compared primary stage of B7 and M1B2_5. We speculated that JA might affect the component type of flavonoid metabolism and eventually affect the flower coloration. Although a few studies manifested that exogenous JA application could promote anthocyanin accumulation especially delphinidin, but the detailed mechanism need to be further deciphered [37].

Consistently, we tried to explain the difference of flower color between M1B7 and M1B2_5. Interestingly, B2_5 cultivar seemed to be totally different from B7 in metabolites distribution. We noticed that procyanidin B2_5 as well as flavonoid compounds were markedly accumulated in M1B7. Besides, naringenin and naringenin chalcone were still highly enriched. Increasing evidence supported that naringin and dihydroflavonol were important branch points of flavonoid metabolism, and the formation of different color was caused by different proportion of flavonoid metabolites [38]. As we expected, we found chalcone-flavanone isomerase (TRINITY_DN4346_c0_g1) was markedly upregulated to promote the biosynthesis of naringenin. Similar study was documented in Arabidopsis thaliana that verified that chalcone-flavanone isomerase catalyzes the intramolecular cyclization of bicyclic chalcones into tricyclic (S)-flavanones is responsible for the isomerization of chalcone into naringenin [39]. Interestingly, flavanone 3-hydroxylase (F3H) (TRINITY_DN8652_c0_g1) and two chalcone synthase (CHS) TRINITY_DN16129_c0_g1 and TRINITY_DN2623_c0_g1 were negatively related with biosynthesis of naringenin and dihydroquercetin, respectively. Substantial documents have reported that CHS was implicated with multiple biological process including pathogen resistance, UV damage and formation of flower color, especially CHS behaved more active and higher expression level in response to pathogen challenge [40]. Our data indicated that CHS genes were downregulated and downstream naringenin was accumulated that was positively related with flower of red degree. In consequence, the activation or suppression of CHS could alter the flow and distribution of downstream metabolites thereby cause different physiological change in plants.
It was worth noting that we identified 5 core metabolites associated with red coloration such as taxifolin, chalcone aringenin, L-epicatechin, naringenin and chrysin, which possess the ability of antimicrobial and insect resistant. Flavonoids have been confirmed to function in response to multiple biotic and abiotic stress in plants and most of them exerted promising application in prevention of harmful microbes [41, 42]. We therefore reasoned that B7 cultivar may behave enhanced resistance or tolerance to deal with unpredictable disadvantages than B2_5 cultivar. Overall, our results revealed that the active flavonoid biosynthesis pathway facilitated abundant accumulation of anthocyanin and eventually formed the red coloration in B7 cultivar. These results expected to provide a valuable guidance for cashew breeding programs.

**Conclusion**

In the present study, metabolome and transcriptome analysis were applied to identify the key metabolites and genes responsible for the flower coloration of two cashew plants. Eventually, we found that 17 increased metabolites (15 flavonoids, 1 phytohormone and 1 diterpenoid) were implicated with the transformation of flower coloration. Especially two pivotal flavonoids including delphinidin 3-glucoside (Dp3g) and peonidin-3-glucoside (Pn3g) were highly accumulated in mature B7 cultivar, which was directly related to red flower coloration. Typically, we found that it was the six metabolites including chrysin, L-epicatechin, chalcone-aringenin, taxifolin, naringenin and ferulic acid that caused the difference in flower coloration between B7 (red) and B2_5 (white) cultivars. Further transcriptome analysis showed 7 DEGs including CFI (TRINITY_DN4346_c0_g1), two CHS (TRINITY_DN16129_c0_g1 and TRINITY_DN2623_c0_g1), two F3H (TRINITY_DN8780_c0_g1 and TRINITY_DN8652_c0_g1), LDOX (TRINITY_DN2918_c0_g1) and cytochrome P450 (TRINITY_DN28797_c0_g1), which were identified the candidate regulators contributing to form variable flower coloration between two cashew cultivars. These results provide valuable information and new insights for further selective breeding of cashew plants.

**Materials And Methods**

**Plant material extraction**

Two *Anacardium occidentale* L. cultivars, designated as B2_5 and B7, were collected from Ledong germplasm nursery in Hainan province in China. The changeable coloration of bisexual flower were observed in both varieties, then we chosed them for research materials. We harvested young and mature bisexual flowers in two cultivars and each set contained three replicates, designated as N1B2_5 and N1B7 (immature flowers), M1B2_5 and M1B7 (mature flowers). All samples were collected, immediately frozen in liquid nitrogen, and then stored at -80°C.

**RNA Extraction and Sequencing**

Total RNA was extracted from plant tissue using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol and digested with DNase I (Takara, Dalian, China). The quality and purity of
total RNA were evaluated by stringent RNA quality control. The mRNA was enriched by Oligo(dT) beads and fragmented into short fragments using fragmentation buffer. cDNA library construction and sequencing were performed by Annoroad Gene Technology (Beijing, China). All products were sequenced on the Illumina HiSeq2500 platform of Gene Denovo Biotechnology Co. (Guangzhou, China).

RNA-seq Data Analysis and Annotation

The raw reads were filtered to remove adapter-polluted reads, low-quality reads, and reads with more than 5% ambiguous nucleotides. These clean reads with high quality were subjected to the following analyses. Trinity software [43] was used to perform the de novo transcriptome assembly with default parameter values. All transcripts were annotated from databases, including the Gene Ontology (GO) database, Kyoto Encyclopedia of Genes and Genomes (KEGG) database, NCBI non-redundant (Nr) database, Swiss-Prot protein database, and Pfam database.

Differentially expressed gene (DEG) identification and analysis The expression level of each unigene was calculated by reads per kilobase million mapped reads (RPKM) to assess the length and depth of sequencing. Then, the differences in the expression abundance of each gene between each pair of compared samples were calculated by DESeq 2 software (version 1.4.5). Each resulting p-value was adjusted to a q-value, following the Benjamini-Hochberg procedure for controlling the false discovery rate. The DEGs were identified with $q \leq 0.05$ and $|\log_2(\text{fold-change})| \geq 1$ as thresholds.

Extraction and separation of metabolites

We collected petal samples from B2_5 and B7 lyophilized samples were used in our experiment for metabolome analysis. The weight of 100 mg powder was weighted and extracted overnight at 4°C with 1.0 mL 70% aqueous methanol. Then, the extracts were analyzed via a LC-MS/MS system (HPLC, Shim-pack UFLC SHIMADZU CBM30A system, Shimadzu, Kyoto, Japan; MS, Applied Biosystems 6500 Q TRAP, Applied Biosystems, AB Sciex, Waltham, MA, USA). The analytical conditions were as follows: Waters ACQUITY UPLC HSS T3 C18 column (1.8 µm, 2.1 mm × 100 mm); water (0.04% acetic acid) to acetonitrile (0.04% acetic acid) mobile phase; gradient program: 100:0 (A/B) at 0 min, 5:95 (A/B) at 11.0 min, 5:95 (A/B) at 12.0 min, 95:5 (A/B) from 12.1 min to 15.0 min; flow rate of 0.40 mL/min; column temperature of 40°C; injection volume of 2 µL. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (Q TRAP)-MS.

Analysis of metabolomics profiles

Data changes are performed by the R software package xcms software. Metaboanalyst 3.0 was used for principal component analysis (PCA) and path enrichment analysis. Principal component analysis showed metabolic changes in different groups. Principal component analysis was performed on three-dimensional metabolic data involving metabolite names, sample names and normalized peak areas. The data are further processed by means of mean center and unit variance scale. Generate principal
component analysis diagram to explain cluster separation. Treatment with index $p < 0.05$ and $\log(FC) > 2$ was used to identify metabolites with significant differences.

**Abbreviations**

CFI: Chalcone–flavanone isomerase; CHS: chalcone synthase; CHI: chalcone isomerase; DFR: dihydroflavonol 4-reductase; F3H: flavanone 3-hydroxylase; UFGT: flavonoid 3-$\alpha$-glucosyltransferase; AOMT: anthocyanin $\alpha$-methyltransferase; FLS: flavonol synthase; FNS: flavone synthase; ANS: anthocyanidin synthase; TFs: transcription factors; PCA: principal component analysis; Dp3g: delphinidin 3-glucoside; Pn3g: peonidin-3-glucoside; ABA: abscisic acid; ETH: ethephon; GA: gibberellin; IAA: indole-3-acetic acid; JA: jasmonic acid; UV: Ultraviolet

**Declarations**

**Acknowledgments**

Not applicable.

**Authors’ contributions**

Z.Z. and H.H. conceived and designed research. W.H. and L.X. conducted experiments. Z.Z., W.H. and H.H. analyzed data. Z.Z., W.H., L.X. and H.H. wrote the manuscript. All authors contributed to the revision of this manuscript and approved the final manuscript.

**Funding**

This work was supported by Central Public-interest Scientific Institution Basal Research Fund(NO. 1630032022018), National Tropical Plants Germplasm Resource Center Projects and Agricultural germplasm resources protection project of Department of Agriculture and Rural Affairs of Hainan Province.

**Availability of data and materials**

All data generated or analysed during this study are included in this published article and its Supplementary information files. The raw RNA-seq data are freely available at: http://www.ncbi.nlm.nih.gov/bioproject/PRJNA860659.

**Ethics approval and consent to participate**

The authors confirm that all methods comply with relevant institutional, national, and international guidelines and legislations. The cashew cultivars (B2_5 and B7) used is from our laboratory and was planted in the Ledong cashew germplasm nursery of Tropical Crops Genetic Resources Institute, Chinese
Academy of Tropical Agricultural Sciences, which is the authors’ subordinate unit. Thus, the authors have the permission to collect the cashew material.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Supplementary Information

Additional file 1: Figure S1. The metabolic pathway of JA biosynthesis in M1B7 vs. N1B7 comparison.

Additional file 2: Figure S2. The metabolic pathway of quercitrin and JA biosynthesis in M1B7 vs. M1B2_5 comparison.

References


Figures
Figure 1

The variation of flower colors in B2_5 and B7 cultivars affected metabolome and transcriptome patterns. A. Phenotype of flower colors of B2_5 and B7 cultivars in blooming stages. B and C. PCA plots exhibited distinctly separate groups among four experimental samples. D and E. The heatmaps showed significant up-/down regulation of all DEMs and DEGs among four groups.
Figure 2

The accumulation of flavonoid compounds promoted the flower color change of B7 cultivar in cashew. A. Pathway enrichment analysis of M1B7 vs N1B7. Parameter of “-ln P value” assessed the significance of metabolites enrichment and parameter of “impact” assessed the importance of metabolic pathways. B and D. heatmaps showed DEMs and DEGs involved in flavonoid and phenylalanine biosynthesis. C. KEGG pathway enrichment analysis of M1B7 vs. N1B7. Top 20 KEGG pathways were exhibited according
to P-value. E. The Metabolic pathway maps showed related DEMs and DEGs involved in flavonoid and phenylalanine biosynthesis. Up or down regulation transcripts and metabolites marked yellow-gray and red-green.
The variation of flavonoid and phenylpropanoid biosynthesis caused the distinction of flower color in mature B7 and B2_5 cultivars in cashew. A. Volcano plot showed up or down regulated metabolites in comparison of M1B7 vs M1B2_5. B and E. Heatmaps exhibited DEMs and DEGs involved in flavonoid and phenylpropanoid biosynthesis respectively. C and D. KEGG pathway enrichment analysis showed the distribution of DEMs and DEGs relevant to flavonoid and phenylpropanoid biosynthesis in M1B7 vs M1B2_5.

The accumulation of flavonoid compounds promoted the formation of red flower in B7 cultivar. A and C. Venn diagram showed the distribution of DEMs and DEGs in comparison of M1B7 vs M1B2_5 and M1B7 vs N1B7. B. The relative expression of core metabolites that facilitated the formation of red flower. *, p<0.05; **, p<0.01. D. Heatmap exhibited the core genes that facilitated the formation of red flower.
Figure 5

Reconstruction of the flavonoid biosynthetic pathway with the differentially expressed structural genes.

Supplementary Files

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