Unveiling molecular mechanisms of pigment synthesis in gardenia (*Gardenia jasminoides*) fruits through integrative transcriptomics and metabolomics analysis

Kangqin Li [a,b,1], Lixin Yu [a,b,1], Liqin Gao [a,b], lingzhi Zhu [a,b], Xiaotao Feng [c], Shaoyong Deng [a,b,*]

[a] *Jiangxi Academy of Forestry, Nanchang 330032, China*
[b] *Engineering Research Center for Gardenia of National Forestry and Grassland Administration, Nanchang 330032, China*
[c] *College of Forestry, Jiangxi Agricultural University, Jiangxi, Nanchang 330045, China*

ARTICLE INFO

ABSTRACT

This study conducted a combined transcriptomics and metabolomics analysis in premature and mature developmental stages of *Gardenia jasminoides* Ellis fruits to identify the molecular mechanisms of pigment synthesis. The transcriptomics data produced high-quality clean data amounting to 46.98 gigabytes, exhibiting a mapping ratio of 86.36% to 91.43%. Transcriptomics analysis successfully identified about 3,914 differentially expressed genes which are associated with pivotal biological processes, including photosynthesis, chlorophyll, biosynthetic processes, and protein-chromophore linkage pathways. Functional diversity was clarified by the Clusters of Orthologous Groups (COG) classification, which focused mainly on pigment synthesis functions. Pathways analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) revealed critical pathways affecting pigment development. Metabolomics studies were carried out utilizing Ultra Performance Liquid Chromatography and mass spectrometry (UPLC-MS). About 480 metabolites were detected via metabolomics investigation, the majority of that were significantly involved in pigment synthesis. Cluster and pathway analyses revealed the importance of pathways such as plant secondary metabolite biosynthesis, biosynthesis of phenylpropanoids and plant hormone signal transduction in pigment synthesis. Current research advances our comprehension of the underlying mechanisms at the molecular level governing pigment synthesis in gardenia fruits, furnishing valuable insights for subsequent investigations.

1. Introduction

Gardenia, a perennial shrub from the *Rubiaceae* family, renowned for its therapeutic and nutritional characteristics, establishing a significant role in the concept of food homology in medicine (Wenping et al., 2017). In Asian countries, the fruit of the gardenia is utilized as a natural dye and in traditional Chinese herbal medicine. Gardenia boasts a rich medicinal background and is among the initial set of plants recognized for both therapeutic and ingestible purposes by China's Ministry of Health. Gardenia fruits exhibit a spectrum of therapeutic effects, including anti-inflammatory, pain-relieving, hepatoprotective properties and properties that stimulate bile production, lower lipid levels, prevent blood clotting, and protect the nervous system (Liu et al., 2014). Due to its widespread usage and high demand, gardenia is extensively cultivated

in China, particularly in the provinces of Jiangxi, Hunan, Hubei, Fujian, Zhejiang, Anhui, Sichuan, Guizhou, Jiangsu, Henan, and Shandong (Liu et al., 2014). Currently, the artificially cultivated area of Gardenia in China spans approximately 17,000 ha, yielding an annual output nearing 40,000 metric tons (Shahid and Mohammad, 2013).

For almost two hundred years, natural dye extracts have been crucial in providing color, enhancing properties, and acting as antioxidants in the food, pharmaceutical, and textile industries. Synthetic chemical dyes have mostly taken over natural dyes because they last longer, are faster to use and are more convenient (Bechtold et al., 2006). However, synthetic dyes can harm the environment and human health, a "green" movement has started, leading to restrictions like bans on products with certain synthetic dyes (Boo et al., 2012; El-Shishtawy et al., 2009). Plant-based colorants are well known for being good for the environment and

they also have antioxidant and antimicrobial effects. This makes them useful in various things like food, beauty products, and medications (Pyle, 2003; Park et al., 2013; Siwawej, and Jarayapun, 1998). Moreover, the increased use of chemical dyes has underscored the need to preserve local species and their connection with humans (Zhou et al., 2014). Gardenia pigments are used as natural colorants in textiles, medicine and foods industries. They are also used as natural photosensitizers in solar cell manufacturing industries (Shen et al., 2014; Choi et al., 2001; Tangpao et al., 2018; Sangare et al., 2023). In addition to the Gardenia fruits, its foliage and flower tissues also contain absorbing pigments. The primary substances that give Gardenia and saffron yellow color are crocins, glycosides of C-20 carotenoid aglycone crocetins (Nagata et al., 2016; Peng et al., 2013).

The two important bioactive substances found in gardenia fruits are geniposide and crocins (Wang, 2015). Because of the presence of crocetin, crocin-1, crocin-2, crocin-3 and other crocin components, gardenia is widely used as a food coloring agent. Due to their water solubility, these yellow pigments have been employed as natural food colorants in Japan for an extended period, particularly in juices, jellies, candies, and noodles (Liu, 2010). The iridoid compounds found in gardenia fruits, notably geniposide and gardenoside, are recognized for their tranquilizing properties (Pan et al., 2019).

The transformation of the white mesocarp to orange or red during fruit maturation is attributed to the synthesis and accumulation of apocarotenoids, particularly crocin-1 which are derived from carotenoids. Flavonoids, a subgroup of secondary metabolites known as plant polyphenols, play a crucial part in the synthesis of plant pigments (Tsanakas et al., 2014). Beyond their involvement in aesthetic aspects, flavonoids serve as crucial signaling molecules when plants encounter stress, activating defense related signaling pathways and regulatory mechanisms. The development of fruit color involves complex processes which are influenced by various factors, including plant hormones. Different plant hormones play vital roles in signal transduction pathways that control the regulation of genes implicated in pigment synthesis. The phenylpropanoid pathway marks the starting point of flavonoid synthesis, transforming phenylalanine into 4-coumaroyl-CoA via enzymatic processes. Consequently, an in-depth analysis of metabolites at different developmental stages becomes instrumental in comprehending the pigment formation process and advancing the development of health products derived from gardenia fruits. In gardenia fruit, flavonoids represent a crucial class of metabolic components influencing fruit color and providing antioxidant properties. Concurrently, genocides, another significant constituent in Gardenia, contribute to the fruit's medicinal value. The synergy between flavonoids and genocides is believed to underlie the advantageous properties associated with gardenia fruits (Kieffer et al., 2016). HPLC analysis of gardenia fruits showed that, mature fruits have the most crocin, while immature ones have the most genepisode. Carotenoids, such as beta-carotene, are vital pigments responsible for fruits' red, orange, and yellow colors. They are accessory pigments in photosynthesis, capturing light energy and protecting chlorophyll from damage (Havaux and Niyogi, 1999). The sugars produced during photosynthesis influence the formation and accumulation of these pigments. The break down products of chlorophyll contribute to color changes during fruit ripening. Chlorophyll degradation reveals other pigments, influencing the visual appearance of ripe fruits (Barry et al., 2008).

The current study aims to identify genes and metabolites and their regulatory networks that determine the color properties of gardenia fruits and establish a theoretical foundation for enhancing pigment synthesis in gardenia breeding efforts. Several genes and their regulated pathways were identified through transcriptomic sequencing of gardenia fruits at premature and mature growth stages. In addition, the UPLC-MS method was employed to evaluate the metabolites during the premature and mature growth stages. This study will offer insightful viewpoints on the molecular processes controlling the pigment synthesis, laying the groundwork for focused regulation of essential metabolic components in gardenia.

2. Materials and methods

2.1. Plant materials

Gardenia fruits at both early and ripe stages were collected from various locations, including Jiangyi Town (31°50′20″N 120°17′42″E), Gongqing City (31.321821°N 121.545769°E), Yushui District, Xinyu City (27°50′N 115°00′E), China. Each sample was collected in 5 g portions for further investigations and three biological duplicates were prepared for every sample. After collection, all samples were stored in liquid nitrogen and maintained at −80 °C until they were ready for further analysis. Metabolite analysis and transcriptome sequencing (RNA-seq) were performed on these samples. The premature samples were defined as "H293," while mature samples were defined as "H1830" in the case of transcriptomics analysis. In contrast, premature samples were defined as "H1," and mature samples were defined as "H2" in the case of metabolomics study.

2.2. RNA extraction, library preparation and clustering

The RNAprep Pure Plant Kit (Tiangen, Beijing, China) was utilized to extract total RNA from gardenia fruits according to instructions provided by the manufacturer. The RNA purity and concentration was assessed using The Nanodrop 2000 (Thermo Fisher Scientific, USA). Sequencing libraries were constructed using qualified RNA samples. We employed 1 μg of RNA for each sample to generate sequencing libraries using the NEB Next® Ultra™ RNA Library Prep Kit. The index-coded samples were clustered using a cBot Cluster Generation System with the TruSeq PE Cluster Kit v4-cBot-HS (Illumina, USA), following the manufacturer's guidelines. Post clustering, paired-end reads were produced by sequencing the prepared library on an Illumina system.

2.3. Quality control and comparative analysis

Custom Perl scripts were used to process the raw reads in fastq. The adapter-containing, ploy-N-containing, and low-quality reads were removed from the raw data. Sequence duplication levels, Q20, Q30, and GC-content analyses were all performed concurrently on the clean data. Afterwards, adaptor sequences and low-quality sequencing reads were removed from the data sets. The raw sequences were processed and then transformed into clean reads. These clean reads were then mapped to the genomic sequence of the reference. Only readings displaying a perfect match or a single mismatch were subjected to further analysis and annotation based on the reference genome. Software called Hisat2 Tools was utilized to map to the reference genome.

2.4. Gene functional annotation and differential expression analysis

Several databases, including NCBI non-redundant nucleotide sequences, COG, KEGG and GO databases, were utilized for annotating gene functionalities. COG is a database collecting phylogenetic classification of proteins, which can supply information on the orthologous classification of gene products. FPKM values were employed to evaluate the levels of gene expression. Analysis of differential expression between two sample groups was performed using the DESeq2. The calculated P-values underwent adjustment through the Benjamini and Hochberg method to manage the false discovery rate. Genes identified as exhibiting differential expression by DESeq2 were those with an adjusted P-value of < 0.05. Significantly differential expression was determined by a False Discovery Rate (FDR) < 0.05 and Fold Change ≥ 2.0 thresholds. The GO enrichment analysis for these differentially expressed genes (DEGs) utilized the GOseq R packages, employing the Wallenius non-central hypergeometric distribution method (Mao et al., 2005), which adjusts for potential biases in gene length within the DEGs. Statistical

enrichment of differential expression genes in KEGG pathways was assessed using KOBAS software (Wang et al., 2021).

2.5. Extraction and analysis of metabolites

Metabolites were extracted from gardenia fruits for subsequent analysis. At first, freeze-dried samples were grounded using zirconia beads for 1.5 min in a stirring mill (MM 400, Germany). Each 100 mg sample was dissolved in a volume of 1.0 mL of the 70 % methanol solution, and later, the sample was extracted using a rotator at 4℃ for the whole night. In the next step, the samples were centrifuged at $10,000 \times g$ for 10 min and absorbed using a Carbon-GCB SPE Cartridge. Finally, the extracts were passed through filters with a pore size of 0.45 μm prior to UPLC-MS analysis. For the chromatographic analysis analytical grade chemicals with higher purity were used. The UPLC columns performance was assessed using a described protocol (Zelena et al., 2009). Metabolites were analyzed using a UPLC system (Waters, USA) couples with a mass spectrometer (Waters, USA) with higher sensitivity.

2.6. Data processing

At first, the obtained data was converter to mzXML using MS Convert in Proteo Wizard software (Smith et al., 2006) and later, using XCMS, the data was processed for further clarification of obtained data (Navarro-Reig et al., 2015). Finally, metabolites identification was according to accuracy mass (<30 ppm) and MS/MS results that were compared with HMDB (Wishart et al., 2007) and KEGG (Ogata et al., 1999). To ensure precise identification of metabolites, only ion peaks exhibiting relative standard deviations (RSDs) of less than 30 % in quality control (QC) samples were preserved following normalization. Various metabolites were identified using pre-determined P-value and VIP threshold in the statistical test (Kieffer et al., 2016).

2.7. Pathway analysis of identified metabolites

Metabo Analyst software was employed for analyzing the pathways of divergent metabolites (Xia and Wishart, 2011). Subsequently, the KEGG pathway was utilized to map the identified metabolites and interpret higher level systemic functions biologically. The KEGG Mapper tool was then used to visualize the pathways linked with the metabolites.

3. Results

3.1. Library construction for RNA-sequencing and data alignment to reference genome

The quality control of the samples successfully generated about 46.98 Gb clean. An impressive performance was observed, with more than 92.84 % of bases in each sample achieving a Q-score no less than Q30. The sequencing data statistics are thoroughly outlined in Table 1. The table provides data on sequencing quality metrics for six samples, labeled H1830-1 through H293-3. Each sample is evaluated for clean

Table 1
Sequencing data statistics.

Samples	Clean reads	Clean bases	GC Content	% ≥ Q30
H1830-1	29,282,311	8,784,693,300	45.07 %	93.48 %
H1830-2	25,602,902	7,680,870,600	44.74 %	93.07 %
H1830-3	26,723,302	8,016,990,600	44.75 %	92.84 %
H293-1	26,167,704	7,850,311,200	44.17 %	93.61 %
H293-2	23,959,855	7,187,956,500	44.20 %	93.12 %
H293-3	24,874,013	7,462,203,900	44.59 %	93.35 %

*Note: (1) Samples: Sample name; (2) Clean reads: Counts of clean PE reads; (3) Clean bases: total base number of Clean Data; (4) GC content: Percentage of G, C in clean data; (5) ≥ Q30%: Percentage of bases with Q-score no less than Q30.

reads, clean bases, GC content, and percentage of bases with a quality score of Q30 or higher. Overall, these metrics indicate high quality sequencing data across all samples, with a consistent GC content around 44–45 % and a high percentage of bases (above 92 %) meeting the Q30 threshold. Notably, for sample H1830-1, 29,282,311 clean reads were generated, accounting for a total of 8,784,693,300 clean bases with a GC content of 45.07 % and an outstanding 93.48 % of the bases had a Q-score equal to or exceeding Q30. Similarly, H1830-2, H1830-3, H293-1, H293-2, and H293-3 exhibited robust sequencing outcomes, with consistently high percentages of bases meeting or surpassing the Q30 threshold. These sequencing data statistics underscore the reliability and high quality of the obtained data, laying a solid foundation for subsequent bioinformatics analyses and interpretation of the experimental results. The HISAT2 system was utilized for aligning RNA-seq data to the predetermined reference genome (Gardenia_jasminoides. Gardenia_v2. genome. a). Subsequently, String Tie was employed to assemble the mapped reads, enabling the assembly and quantification of transcripts, capturing multiple spliced variants for each gene locus. Across the samples, the mapping ratio ranged from 86.36 % to 91.43 %. The mapping results indicated that H293 samples were comprised of 68.15 % exon, 8.09 % intron, and 23.76 % intergenic region, while H1830 samples were comprised of 69.06 % exon, 6.86 % intron and 24.08 % intergenic region (Shown in Fig. 1).

Table 2 presents the performance of sample mappings, encompassing total reads, mapped reads, uniquely mapped reads, multiply mapped reads, reads mapped to the sense chain ('+'), and reads mapped to the antisense chain ('-'). In this table, the statistics on data mapping for six samples, labeled H1830-1 to H293-3 are presented. This data provides insights into the quality and efficiency of the sequencing and mapping processes for each sample, highlighting the proportion of reads that are accurately aligned to the genome and their strand distribution.

3.2. Gene structure optimization

About 3,890 genes were subjected to structure optimization in this study, as detailed in Table 3. The table provides comprehensive information, including gene ID, gene locus, strand orientation, site of optimization (on 3′ UTR or 5′ UTR), original region (starting and ending positions of original annotated genes) and optimized region (starting and ending positions of optimized genes) of *Gj1A102T38, Gj1A10T65, Gj1A119T21, Gj1A11T101, Gj1A122T86, Gj1A129T31, Gj1A129T33, Gj1A131T65* and *Gj1A131T66* genes.

3.3. Functional analysis of identified genes during developmental stages of gardenia

The transcriptomic analysis identified several genes, including *Gj1A102T38, Gj1A10T65, Gj1A119T21, Gj1A11T101, Gj1A122T86, Gj1A129T31, Gj1A129T33, Gj1A131T65, Gj1A131T66, Gj9P574T13, Gj3P320T22, Gj3A320T89, Gj3A320T88, Gj11X432T33, Gj9X715T20, Gj9A119T113, Gj4P41T15, Gj4A41T73, Gj3X308T65, Gj9P121T25, Gj9A987T39, Gj9A902T126, Gj9A1T81, Gj9A1014T60,* etc. with diverse biological processes including photosynthesis, chlorophyll biosynthetic process, amine metabolic process, protein-chromophore linkage, etc. About 225 genes associated with KEGG and 302 genes with GO analysis. These annotations provided valuable insights into the potential functions and pathways associated with the newly identified genes. In the comparison between H1830_vs_H293, approximately 3,914 differentially expressed genes (DEGs) were detected, with 2,412 genes showing up-regulation and 1,502 exhibiting down-regulation.

The expression profile of these DEGs is illustrated in Fig. 2. MA plot shows the overall distribution of gene expression and fold change of expression level between H1830 and H293 gardenia fruit samples (Fig. 2**A**). The MA-plot illustrates the distribution of genes exhibiting differential expression. On the y-axis, the $\log_2$ FC in expression between premature and mature gardenia fruits is depicted, while the x-axis
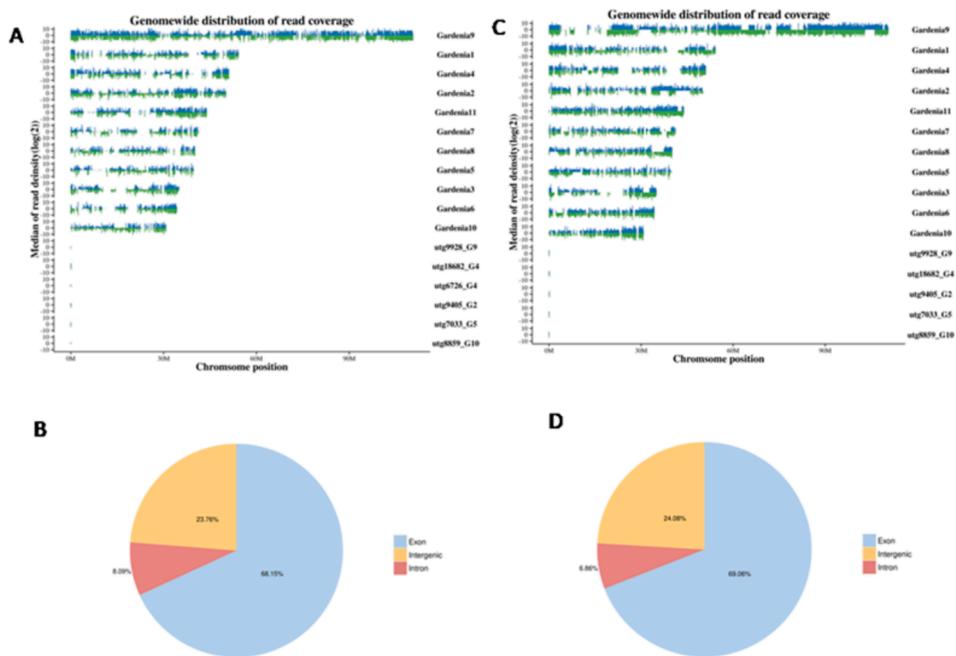
Fig. 1. RNA-seq data against *G. jasminoides* genome. RNA-seq data was aligned to a reference genome using the HISAT2 system and then assembled using String Tie, enabling transcript assembly and quantification. The mapping results revealed the composition of different genomic regions within the samples.

Table 2
Statistics on data mapping.

Sample	Total Reads	Mapped Reads	Uniq Mapped Reads	Multiple Map Reads	Reads Map to '+'	Reads Map to '-'
H1830-1	58,564,622	52,511,277 (89.66 %)	49,620,146 (84.73 %)	2,891,131 (4.94 %)	28,569,481 (48.78 %)	28,551,563 (48.75 %)
H1830-2	51,205,804	46,214,500 (90.25 %)	43,888,547 (85.71 %)	2,325,953 (4.54 %)	24,739,133 (48.31 %)	24,689,100 (48.22 %)
H1830-3	53,446,604	48,414,574 (90.58 %)	46,146,145 (86.34 %)	2,268,429 (4.24 %)	25,772,950 (48.22 %)	25,777,955 (48.23 %)
H293-1	52,335,408	47,033,573 (89.87 %)	45,860,284 (87.63 %)	1,173,289 (2.24 %)	24,245,564 (46.33 %)	24,312,044 (46.45 %)
H293-2	47,919,710	43,814,861 (91.43 %)	42,693,497 (89.09 %)	1,121,364 (2.34 %)	22,627,598 (47.22 %)	22,679,114 (47.33 %)
H293-3	49,748,026	42,961,669 (86.36 %)	41,890,686 (84.21 %)	1,070,983 (2.15 %)	22,162,661 (44.55 %)	22,210,847 (44.65 %)

*Note: Sample: sample ID in system; Total Reads: Counts of Clean Reads, counted as single end; Mapped Reads: Counts of mapped reads and the proportion of that in clean data; Uniq Mapped Reads: Counts of reads mapped to a unique position on reference genome and proportion of that in clean data; Multiple Mapped Reads: Counts of reads mapped to multiple positions on reference genome and proportion of that in clean data; Reads Map to '+': Counts of reads mapped to the sense chain and the proportion of that in clean data.

Table 3
Gene structure optimization.

Gene ID	Locus	Strand	Site	Original Region	Optimize Region
Gj1A102T38	Gardenia1:10231591–10232564	−	5′	10232523–10232523	10232523–10232564
Gj1A10T65	Gardenia1:1052281–1066880	+	3′	1066520–1066520	1066520–1066880
Gj1A119T21	Gardenia1:11922369–11926894	−	5′	11926838–11926838	11926838–11926894
Gj1A11T101	Gardenia1:1148929–1150402	−	5′	1150237–1150244	1150237–1150402
Gj1A122T86	Gardenia1:12222175–12223003	+	3′	12222898–12222898	12222898–12223003
Gj1A129T31	Gardenia1:12892424–12897092	+	3′	12896437–12896437	12896437–12897092
Gj1A129T33	Gardenia1:12959626–13042352	−	3′	12960106–12960106	12959626–12960106
Gj1A131T65	Gardenia1:13095159–13104227	+	3′	13103964–13103964	13103964–13104227
Gj1A131T66	Gardenia1:13133949–13138180	+	3′	13137981–13137981	13137981–13138180

represents the log$_2$-transformed average expression level for each gene across all samples (FPKM). The up-regulated, downregulated, and normal genes are presented as red, blue, and ash grey colors respectively. The volcano plot illustrating gene expression in "H1830" and "H293" is depicted in Fig. 2**B**. The discernible segregation of significantly expressed genes from non-significantly expressed is facilitated through distinct color codes. The upregulated, downregulated, and normal genes are presented as red, blue, and ash-grey colors. This volcano plot depicts clusters of data points near the center and a spreading pattern radiating outward. Because P-values undergo negative transformation, higher positions along the y-axis indicate smaller P-values.

It's customary for volcano plots to include threshold indicators for adjusted P-values, highlighting genes deemed statistically differentially expressed based on their adjusted P-values between premature and mature stages. The x-axis, representing log$_2$ fold-change, emphasizes substantial differences at extreme values, with points closer to 0 indicating genes with similar mean expression levels. Greater dispersion indicates greater disparity in gene expression between premature and mature stages. In Fig. 2**C, a** depiction is provided of the hierarchical clustering analysis conducted on the DEGs, focusing on genes exhibiting comparable or identical expression patterns in two different developmental stages of gardenia fruits. The identified DEGs from the
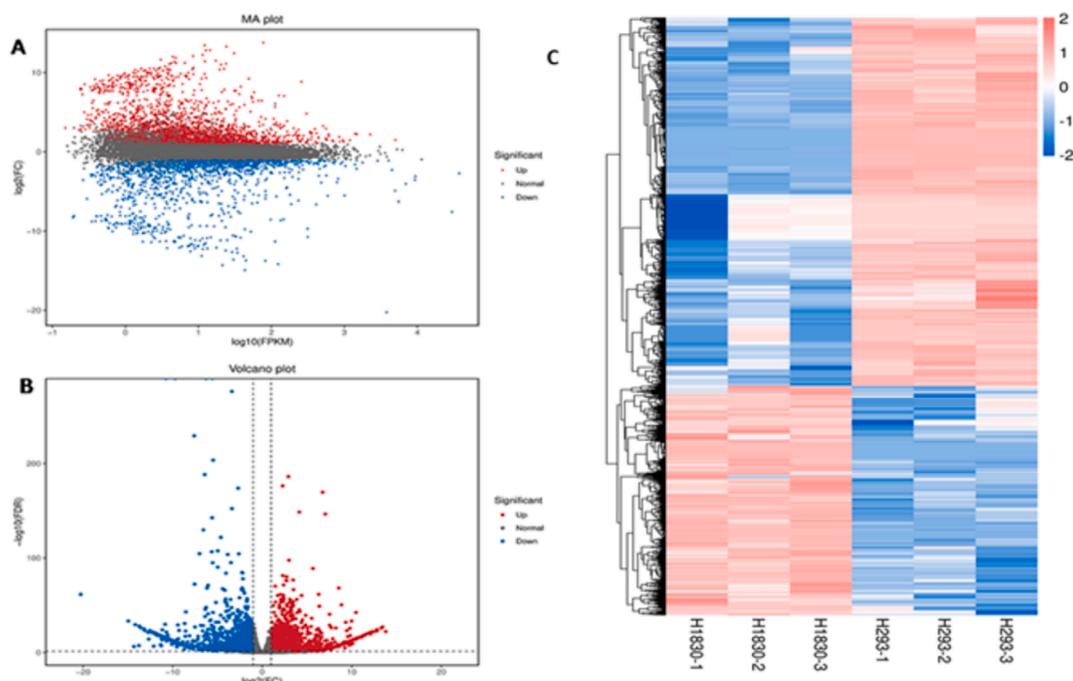
Fig. 2. A. MA plot of differentially expressed genes; **B.** Volcano plot on differential expression; (Each dot represents a single gene. the dots colored in red and blue stand for significant up-regulated and down-regulated genes, respectively. Ash grey dots stand for the genes without significant difference in expression between two samples); **C.** Hierarchical clustering was conducted on the differentially expressed genes, with each column representing a single sample and rows indicating genes. Gene expression levels (FPKM) were $\log_{10}$-normalized and visualized with varying colors according to the scale bar.

differential expression analysis underwent comprehensive annotation. In comparing H1830_vs_H293, 3,575 DEGs were annotated, providing a wealth of information on their functional attributes. The DEGs exhibited functional diversity, with 2,964 genes annotated in the GO database and 2,460 in KEGG. From these annotations a deeper understanding of the biological processes, pathways, and molecular functions associated with the differentially expressed genes are obtained.

3.4. COG Classification, gene ontology and KEGG pathway analysis revealed several key pathways that are involved in pigment formation in gardenia fruits

The COG functional classification demonstrated that the DEGs spanned 26 COG categories, as illustrated in Fig. 3 **A**. Notably, most DEGs fell under the predicting of overall function category, followed by

Fig. 3. A. Summary of COG classifications on DEGs; X-axis: COG classification terms; Y-axis: Gene numbers. **B.** Classification of DEGs based on GO.

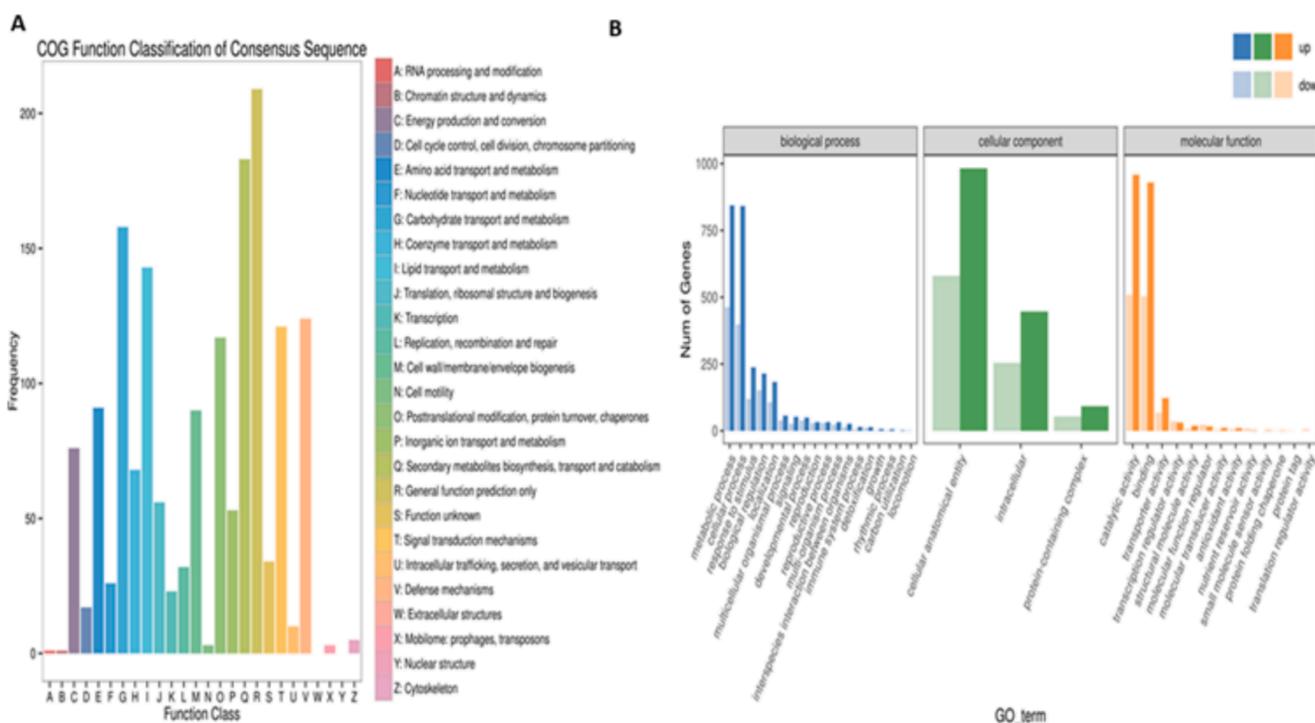secondary metabolites biosynthesis, transport and metabolism, and carbohydrate transport and metabolism. Additionally, DEGs were found to be involved in lipid transport and metabolism, defense mechanisms, and cell wall/membrane/envelope biogenesis, among others. These findings underscore the significant roles genes play within these categories in pigment production. To clarify the functional roles of DEGs, a comprehensive analysis was conducted utilizing a set of 3,914 DEGs through GO analysis. The DEGs were systematically categorized into three principal classifications: cellular component, molecular function, and biological process, as depicted in Fig. 3 **B**. The results of the GO annotation revealed prominent terms within each category specifically, within the cellular component category, terms such as cellular, anatomical entity, intracellular, and protein-containing complex exhibited dominance. Binding and catalytic activity were dominant in term of molecular function while terms associated with metabolic process, cellular process, response to stimulus, and biological regulation in the biological process category, emerged as the foremost contributors. The analysis of GO terms revealed a diverse array of regulators and proteins encoded by genes associated with pigment production in gardenia fruits. We conducted a statistical enrichment analysis using the KOBAS software on DEGs within KEGG pathways to gain deeper insights into the physiological processes governing pigment production. The results indicated that DEGs were mapped to 130 KEGG pathways (Fig. 4 **A and 4B**). Numerous genes exhibited differential expression in pathways such as plant hormone signal transduction, carbon metabolism, ABC transporters, photosynthesis, glycolysis, and carotenoid biosynthesis, emphasizing their relevance to the intricate mechanisms underlying pigment production in gardenia fruits.

The highest representation of unigenes was observed in specific metabolic pathways where plant pathogen interaction leading with 165 genes (12.33 %), followed by plant hormone signal transduction with 114 genes (8.52 %), phenylpropanoid biosynthesis with 73 genes (5.46 %), carbon metabolism with 67 genes (5.01 %), MAPK signaling pathway with 67 genes (5.01 %) and starch and sucrose metabolism with 52 genes (3.89 %). This distribution highlights the prominence of these pathways in the transcriptomic landscape, shedding light on the molecular mechanisms and regulatory networks associated with these vital biological processes. The assigned pathways offer crucial insights into pigment formation processes, providing valuable information for

further investigations. Our study indicates that plant hormone signal transduction, phenylpropanoid biosynthesis, and carbon metabolism are implicated in the regulation of pigment formation.

3.5. Analysis of differentially expressed metabolites during gardenia fruits developmental stages

In the metabolites analysis, many differential metabolites were identified from chromatographic analysis (Table 4). Based on differential metabolite screening of H1 and H2, about 480 metabolites were identified, and 243 metabolites were differentially expressed, among which 169 metabolites were upregulated and 74 metabolites were downregulated (Fig. 5 **A**). Some of the metabolites, including 2-Phenylethanol, o-Xylene, Catechol, 1,3-Benzenediol, Oxalacetic acid, Creatinine, Uracil, 3-Methyl-2-oxovaleric acid, L-Proline, Succinic acid, Betaine, L-Valine, 2-Methylserine, L-Leucine, D-Malic acid, L-Malic acid etc. were identified. Several metabolites, including Lariciresinol, Chlorogenic acid, Norethindrone, 4-methyl benzaldehyde, etc., were upregulated, while metabolites, including mannitol, O-succinyl-L-homoserine, adenosine, etc., were down regulated (Table 5).

3.6. Cluster analysis of metabolites

According to the metabolic profile, the developmental stages of the gardenia fruits "H1″ and "H2," were clustered together (Fig. 5 **B**). An apparent phenotypic variation was observed between the premature and ripening group stages. The signal intensity of all metabolites generated from MS was proportionate to their abundance. By observing accumulation patterns, metabolites could be distinctly categorized into two primary clusters, each containing several sub-clusters. The figure illustrates the relative content using a color gradient, with increasing expression levels depicted in progressively redder shades and decreasing expression levels in bluer shades. Each column corresponds to a distinct sample, while rows represent metabolite names. The clustering tree depicted on the left side of the figure is specifically tailored for the differential metabolites, providing a visual depiction of their interrelationships.
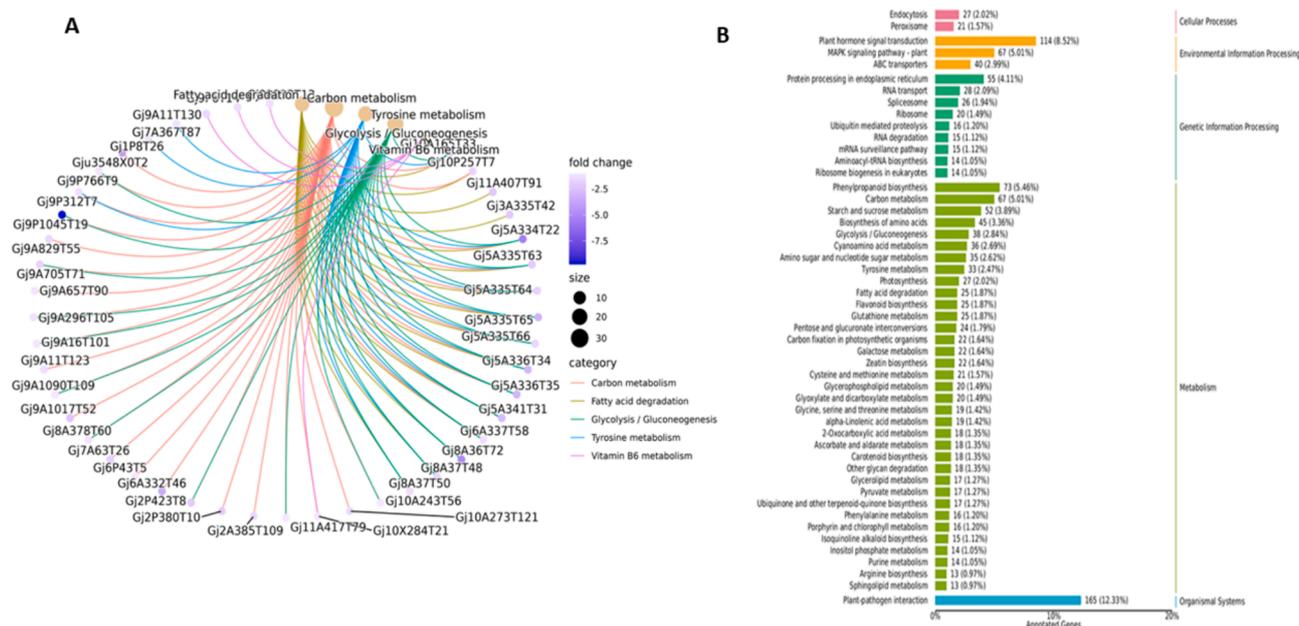

Fig. 4. KEGG Classification on DEGs; **A.** Genes, their pathways, fold change and size; **B.** KEGG pathways (X- axis); X-axis: Quantity of genes annotated to the KEGG pathway.

Table 4
List of some identified metabolites.

ID	name	mz	formula	class	CAS
M105T250	2-Phenylethanol	105.0698	C8H10O	Benzene and substituted derivatives	60-12-8
M107T234	o-Xylene	107.0854	C8H10	Benzene and substituted derivatives	95-47-6
M110T34_2	Catechol	110.0197	C6H6O2		120-80-9 / 12385-08-9
M111T234	1,3-Benzenediol	111.0443	C6H6O2	Benzene and substituted derivatives	108-46-3
M113T603	Oxalacetic acid	112.986	C4H4O5	Keto acids and derivatives	328-42-7 / 149-63-3
M113T578	Creatinine	112.986	C4H7N3O	Carboxylic acids and derivatives	60-27-5
M113T55	Uracil	113.0351	C4H4N2O2	Diazines	66-22-8
M113T264	3-Methyl-2-oxovaleric acid	113.06	C6H10O3	Keto acids and derivatives	1460-34-0
M116T51	L-Proline	116.0708	C5H9NO2	Carboxylic acids and derivatives	147-85-3
M117T44	Succinic acid	117.0189	C4H6O4	Carboxylic acids and derivatives	110-15-6
M118T74	Betaine	118.0865	C5H11NO2	Carboxylic acids and derivatives	107-43-7 / 590-46-5
M118T139	L-Valine	118.0865	C5H11NO2	Carboxylic acids and derivatives	72-18-4
M119T150	2-Methylserine	119.0495	C4H9NO3		5424-29-3
M132T60	L-Leucine	132.1024	C6H13NO2	Carboxylic acids and derivatives	61-90-5
M133T151	D-Malic acid	133.0147	C4H6O5	Fatty Acyls	636-61-3
M133T263	L-Malic acid	133.0149	C4H6O5	Fatty Acyls	636-61-3
M133T54	L-Ribulose	133.0499	C5H10O5	Organooxygen compounds	2042-27-5/551-84-8
M133T66	Glutaric acid	133.0499	C5H8O4	Carboxylic acids and derivatives	110-94-1
M133T177	L-Asparagine	133.1015	C4H8N2O3	Carboxylic acids and derivatives	2058-58-4
M135T351	Pulegone	135.1166	C10H16O	Prenol lipids	89-82-7
M137T188	Salicylic acid	137.0253	C7H6O3	Benzene and substituted derivatives	69-72-7
M138T335	4-Nitrophenol	138.0202	C6H5NO3	Phenols	100-02-7
M145T687	Anabasine	144.9823	C10H14N2		13078-04-1 / 494-52-0
M147T77	L-Glutamic acid	147.0447	C5H9NO4	Carboxylic acids and derivatives	6893-26-1
M147T43	L-Lysine	147.1135	C6H14N2O2	Carboxylic acids and derivatives	56-87-1
M149T550	3-Methyladenine	149.024	C6H7N5	Imidazopyrimidines	5142-23-4
M151T151	2-Methoxy-4-vinylphenol	151.0757	C9H10O2	Phenols	7786-61-0
M152T53	Guanine	152.0569	C5H5N5O	Imidazopyrimidines	73-40-5
M156T49	L-Histidine	156.0427	C6H9N3O2	Carboxylic acids and derivatives	71-00-1
M159T53	Gluconolactone	159.0304	C6H10O6	Organooxygen compounds	90-80-2
M163T34	Acetylcysteine	162.9977	C5H9NO3S	Carboxylic acids and derivatives	616-91-1
M163T329	Caffeic acid	163.0396	C9H8O4	Cinnamic acids and derivatives	331-39-5/501-16-6
M163T194	Phenylpyruvic acid	163.0411	C9H8O3	Benzene and substituted derivatives	156-06-9
M164T90	L-Phenylalanine	164.0725	C9H11NO2		673-06-3
M167T343	Vanillic acid	167.0356	C8H8O4	Benzene and substituted derivatives	121-34-6
M167T250	Caffeyl alcohol	167.0709	C9H10O3	Benzene and substituted derivatives	
M169T89	Gallic acid	169.0149	C7H6O5	Benzene and substituted derivatives	149-91-7
M169T259	Uric acid	169.0501	C5H4N4O3	Imidazopyrimidines	69-93-2
M170T558	8-Amino-7-oxononanoate	170.1182	C9H17NO3		4707-58-8
M185T252	Sebacic acid	185.1176	C10H18O4	Fatty Acyls	111-20-6
M191T44	Citric acid	191.0194	C6H8O7	Carboxylic acids and derivatives	77-92-9 / 126-44-3
M193T127	*trans*-Ferulic acid	193.051	C10H10O4	Cinnamic acids and derivatives	537-98-4 / 1135-24-6
M195T518	Neocnidilide	195.1389	C12H18O2	Lactones	4567-33-3
M204T75_1	N-Acetylmannosamine	204.0853	C8H15NO6	Organooxygen compounds	7772-94-3/3615-17-6
M209T260	(−)-Jasmonic acid	209.1191	C12H18O3	Fatty Acyls	c("6894-38-8″, "59366-47-1″)
M217T310	N-a-Acetylcitrulline	217.1075	C8H15N3O4	Carboxylic acids and derivatives	33965-42-3
M221T276	6-Acetyl-D-glucose	221.0672	C8H14O7	Carboxylic acids and derivatives	
M222T100	Carbofuran	222.1133	C12H15NO3	Coumarans	1563-66-2
M227T537	Deoxycytidine	227.165	C9H13N3O4	Pyrimidine nucleosides	951-77-9
M237T332	Capsidiol	237.1852	C15H24O2	Prenol lipids	37208-05-2
M244T69	Cytidine	244.094	C9H13N3O5	Pyrimidine nucleosides	65-46-3/147-94-4

3.7. KEGG pathway analysis of metabolites at premature and mature stages of gardenia fruits

From this analysis, the major KEGG pathways associated with metabolites showing differential expression are displayed in Fig. 5 **C**. The KEGG analysis of differentially expressed metabolites from premature and mature stages of gardenia fruits indicated that highest number of metabolites was involved in the biosynthesis of plant secondary metabolites (map01060) which was followed by biosynthesis of amino acids (map01230), biosynthesis of phenylpropanoids (map01061), phenylpropanoid biosynthesis (map00940), 2-Oxocarboxylic acid metabolism (map01210), biosynthesis of plant hormones (map01070), ABC transporters (map02010), alpha-Linolenic acid metabolism (map00592), phosphotransferase system (PTS) (map02060), protein digestion and absorption (map04974), carbon metabolism (map01200), metabolic pathways (map01100) etc. The network diagram of pathways is presented in Fig. 6. "Amino acid biosynthesis pathway" illustrates a structured design of the biosynthetic routes for twenty amino acids. The central component is the KEGG module responsible for converting three-carbon compounds from glyceraldehyde-3P to pyruvate, along with the pathways involving serine and glycine. In case of "biosynthesis of phenylpropanoids" and "phenylpropanoid biosynthesis", phenylpropanoids constitute a cluster of plant secondary compounds originating from phenylalanine, serving diverse roles as both structural components and signaling agents. Phenylalanine undergoes initial deamination to form cinnamic acid, which is subsequently subjected to hydroxylation and frequent methylation, yielding coumaric acid and other phenylpropane (C6-C3) unit-containing acids. The carboxyl groups of these acids, when CoA-activated, undergo reduction to yield the corresponding aldehydes and alcohols. These alcohols, known as monolignols, serve as the foundational molecules for lignin biosynthesis. "Biosynthesis of plant hormones" outlines the biosynthesis of various plant hormones, which are crucial for plant growth and development. This pathway includes the synthesis processes for key hormones such as auxins, cytokinins, gibberellins, abscisic acid, ethylene, brassinosteroids, jasmonates, and salicylic acid. Each hormone is synthesized through distinct biochemical pathways involving multiple enzymatic reactions. These pathways are interconnected with other metabolic routes, such as glycolysis and the
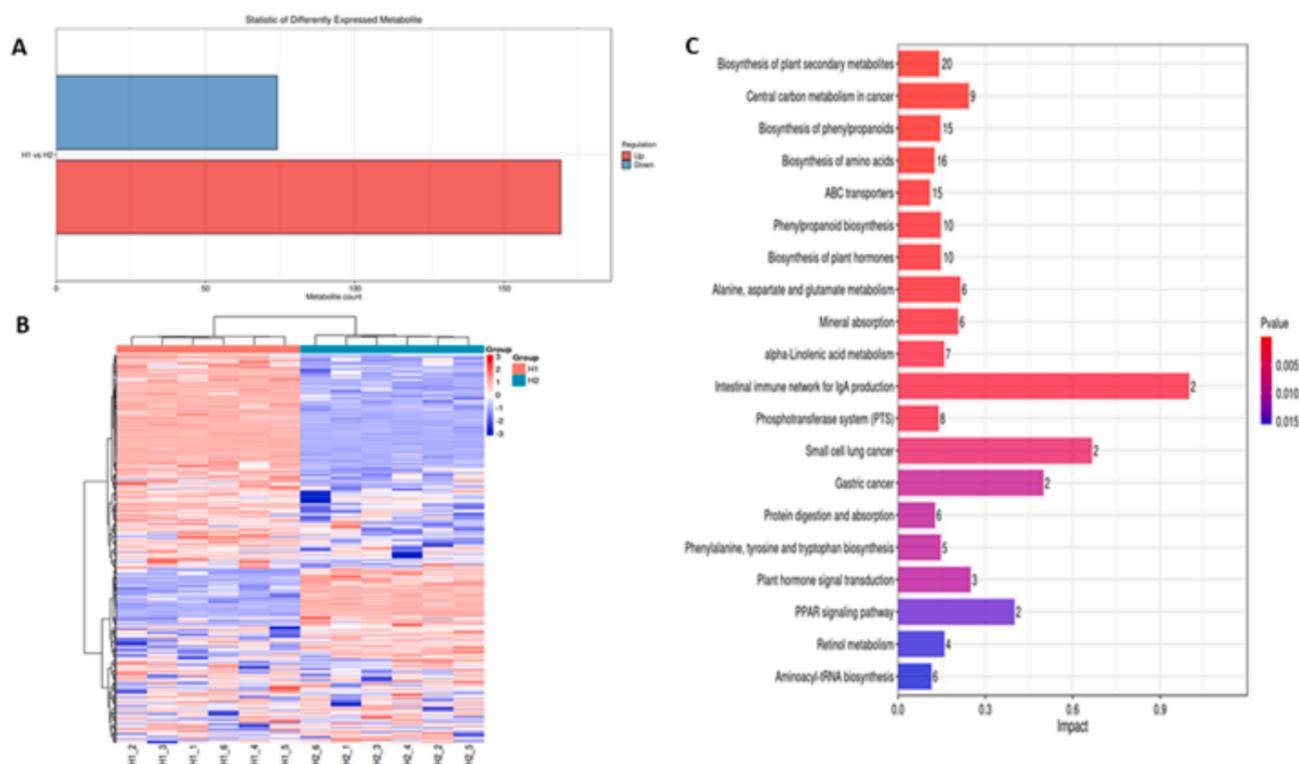
Fig. 5. A. Statistical histogram of differential metabolites; **B**. Heat map visualization of detected metabolites in gardenia fruits at studied developmental stages; Various colors in the figure represent the relative content. The metabolite names are shown in the rows, while the samples are represented in the columns; darker shades of blue indicate lower expression levels, whereas darker shades of red indicate higher expression levels. **C.** Histogram of metabolic pathway influencing factors. A deeper red color indicates a lower P-value, while a deeper blue color means a more significant P-value.

pentose phosphate pathway, reflecting the complexity and integration of hormone biosynthesis with overall plant metabolism. The KEGG pathway, "Metabolic pathways" offers a comprehensive summary of metabolic processes across various biological systems. It encompasses a broad spectrum of metabolic functions, including the breakdown and creation of carbohydrates, lipids, amino acids, nucleotides, and energy metabolism. This pathway diagram consolidates specific pathways such as glycolysis, the citric acid cycle, oxidative phosphorylation, and the pentose phosphate pathway, demonstrating their interconnections and regulatory mechanisms. It also includes pathways for the synthesis of vital molecules like amino acids and nucleotides, alongside secondary metabolites and vitamins. This comprehensive perspective aids in understanding how cells uphold balance, adapt to environmental shifts, and manage their growth and development.

Fruit colour changes are characterized by reducing chlorophyll content within chloroplasts and accumulating carotenoids and pigments resembling flavonoids or anthocyanins in various cellular compartments. Complex biosynthetic pathways in gardenia fruits, which entail multiple enzymes, are responsible for the production of carotenoids, anthocyanins, and chlorophylls. In the meantime, variations in the constituents and quantities of different pigments accumulated in the fruits of diverse resources are frequently caused by their variation related genes. In addition to different factors, the endogenous hormones of each pigment and the external environment influence the synthesis and metabolism of these pigments. When the gardenia fruits reach their mature and immature stages, they finally display distinct colour phenotypes. Anthocyanins, a significant category of water-soluble plant pigments within the flavonoid family, play a crucial role in providing vibrant colours to plants, flowers, and fruits, spanning the spectrum from pink to scarlet, purple, and blue. Photosynthesis, a fundamental process in plants, is intricately linked to the formation of pigments in plant fruits.

4. Discussion

This study undertakes a thorough examination of the transcriptome and metabolome of gardenia fruits at premature and mature growth phases, aiming to identify critical regulatory networks governing color development. Fruit pigment formation is mainly associated with the biosynthesis of pigments like anthocyanins, carotenoids, and chlorophyll derived from different metabolic pathways. These processes drive the distinct color transformation observed during fruit maturation (Yang et al., 2016; Zhang et al., 2021; Zhang et al., 2023).

In the current study, about 3,914 genes were identified by transcriptome analysis. Prominent genes were linked to a variety of biological processes such as photosynthesis, protein-chromophore linkage, chlorophyll biosynthesis and amine metabolic process etc. The transcriptome analysis of 'Sanbianhong' jujube peels at three growth stages identified DEGs enriched in chlorophyll, carotenoid, and anthocyanin metabolic pathways, respectively (Li et al., 2023). Crocetin, crocin-1, crocin-2, crocin-3 and other crocin components from gardenia are used as natural colorant in the food industry. Transcriptome sequencing identified critical enzymes involved in geniposide and crocin biosynthesis pathways in a study (Wang et al., 2021). As crocin components are derived from carotenoids therefore carotenoids are one of the key ingredients in pigment synthesis in gardenia fruits. The multi-allelic *APRR2* gene has been identified as a critical factor influencing accumulation of fruit pigments in both melon and watermelon, as demonstrated in a study (Oren et al., 2019). Another study identified *GjPSY*, a gene involved in carotenoid biosynthesis, in gardenia fruits during different developmental stages (Ichi et al., 1995). Transcriptome sequencing of two developmental phases of gardenia fruit 29 unigenes were anticipated to take part in the manufacture of carotenoids (Yang et al., 2016). Both structural and regulatory genes oversee the production of anthocyanins. Structural genes are responsible for encoding biosynthetic enzymes, which actively participate in the synthesis of

Table 5

The KEGG pathways of differentially expressed metabolites.

Pathway_id	Pathway name	Total	Hits	Pvalue	$-\text{Log}_{10}$ (P-value)	FDR	Impact
map01060	Biosynthesis of plant secondary metabolites	141	20	4.81253E-07	6.31762636905604	0.000260357981596094	0.1418
map05230	Central carbon metabolism in cancer	37	9	9.09834E-06	5.04103771299551	0.00177111773177837	0.2432
map01061	Biosynthesis of phenylpropanoids	103	15	9.82136E-06	5.00782857931414	0.00177111773177837	0.1456
map01230	Biosynthesis of amino acids	128	16	3.55532E-05	4.44912072094932	0.00480857675653194	0.125
map02010	ABC transporters	137	15	0.000282702	3.54867067286405	0.0218906753540853	0.1095
map00940	Phenylpropanoid biosynthesis	68	10	0.000283243	3.54784006480222	0.0218906753540853	0.1471
map01070	Biosynthesis of plant hormones	68	10	0.000283243	3.54784006480222	0.0218906753540853	0.1471
map00250	Alanine, aspartate and glutamate metabolism	28	6	0.000627476	3.20240269662689	0.042433082478245	0.2143
map04978	Mineral absorption	29	6	0.000765519	3.11604408310328	0.0460161915892994	0.2069
map00592	alpha-Linolenic acid metabolism	44	7	0.001459876	2.83568395894888	0.0709157546548616	0.1591
map04672	Intestinal immune network for IgA production	2	2	0.001536497	2.81346839210842	0.0709157546548616	1
map02060	Phosphotransferase system (PTS)	57	8	0.001572993	2.80327329022867	0.0709157546548616	0.1404
map05222	Small cell lung cancer	3	2	0.00448975	2.34777784748027	0.186842670236113	0.6667
map05226	Gastric cancer	4	2	0.008746991	2.05814133700304	0.321968332334041	0.5
map04974	Protein digestion and absorption	47	6	0.009601709	2.01765145450495	0.321968332334041	0.1277
map00400	Phenylalanine, tyrosine and tryptophan biosynthesis	34	5	0.009822267	2.00778826302965	0.321968332334041	0.1471
map04075	Plant hormone signal transduction	12	3	0.010117304	1.99493518560051	0.321968332334041	0.25
map03320	PPAR signaling pathway	5	2	0.014202061	1.84764862950295	0.409044235924079	0.4
map00830	Retinol metabolism	25	4	0.015325886	1.8145743957176	0.409044235924079	0.16
map00970	Aminoacyl-tRNA biosynthesis	52	6	0.015480339	1.81021953820757	0.409044235924079	0.1154
map01210	2-Oxocarboxylic acid metabolism	134	11	0.015877872	1.79920769322378	0.409044235924079	0.0821
map00254	Aflatoxin biosynthesis	27	4	0.020003349	1.69889728060748	0.488195391743588	0.1481
map05223	Non-small cell lung cancer	6	2	0.020755072	1.6828757534543	0.488195391743588	0.3333
map05143	African trypanosomiasis	8	2	0.036784773	1.43433191682709	0.829190096780435	0.25
map00020	Citrate cycle (TCA cycle)	20	3	0.041635411	1.38053714128365	0.900990299442688	0.15
map01065	Biosynthesis of alkaloids derived from histidine and purine	35	4	0.046836795	1.32941283029449	0.93890378730775	0.1143
map01064	Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid	67	6	0.046858415	1.32921241004693	0.93890378730775	0.0896
map04212	Longevity regulating pathway − worm	10	2	0.056150393	1.25064719794141	1	0.2
map04979	Cholesterol metabolism	10	2	0.056150393	1.25064719794141	1	0.2
map00622	Xylene degradation	38	4	0.060340437	1.21939154625845	1	0.1053
map00945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	25	3	0.072881175	1.13738463138507	1	0.12
map00981	Insect hormone biosynthesis	25	3	0.072881175	1.13738463138507	1	0.12
map00590	Arachidonic acid metabolism	75	6	0.073319346	1.13478141653904	1	0.08
map04922	Glucagon signaling pathway	26	3	0.080095993	1.09638920864765	1	0.1154
map00360	Phenylalanine metabolism	60	5	0.085479009	1.06814052175322	1	0.0833
map00073	Cutin, suberine and wax biosynthesis	27	3	0.087607574	1.05745834710409	1	0.1111
map00720	Carbon fixation pathways in prokaryotes	44	4	0.09279823	1.03246030794466	1	0.0909
map00591	Linoleic acid metabolism	28	3	0.095404009	1.02043337540651	1	0.1071
map00460	Cyanoamino acid metabolism	45	4	0.098878255	1.00489920846063	1	0.0889
map00270	Cysteine and methionine metabolism	63	5	0.100324254	0.998594062483045	1	0.0794
map04923	Regulation of lipolysis in adipocytes	14	2	0.102543515	0.989091798615287	1	0.1429
map00052	Galactose metabolism	46	4	0.105139079	0.978235830768682	1	0.087
map00903	Limonene and pinene degradation	64	5	0.105542752	0.976571584123864	1	0.0781
map05322	Systemic lupus erythematosus	3	1	0.113278083	0.945854107758379	1	0.3333
map04071	Sphingolipid signaling pathway	15	2	0.115368166	0.937914012731409	1	0.1333
map00627	Aminobenzoate degradation	86	6	0.121105199	0.916837211694484	1	0.0698
map00960	Tropane, piperidine and pyridine alkaloid biosynthesis	68	5	0.127707967	0.893782008232087	1	0.0735
map00310	Lysine degradation	50	4	0.131893737	0.879775827095494	1	0.08
map04142	Lysosome	4	1	0.148121781	0.829381074901365	1	0.25

anthocyanins (Li et al., 2015). Three genes: *CkCHS-1 (chalcone synthase), CkDFR (dihydroflavonol 4-reductase),* and *CkANS (anthocyanidin synthase)*, were identified to be involved in pigment formation in orchid flowers. These genes participate in the anthocyanin biosynthesis pathway (Zhou et al., 2021). In soybean flowers, flavonoids primarily anthocyanins are the most prevalent compounds that contribute to pigment synthesis (Sundaramoorthy et al., 2015). In apple it was found that, *MdIAA26* overexpression induced anthocyanin biosynthesis (Wang et al., 2020). Carotenoids, a crucial group of plant pigments normally appears as yellow, orange, and red colors (Delgado-Vargas et al., 2000). They lipids soluble pigments, synthesized in chloroplasts, where they play a vital role in photosynthesis (Tanaka et al., 2008; Sun et al., 2020; Sun et al., 2022). In *Citrus sinensis*, transcription factor CsMADS6 stimulates the expression of *LCYb1, PSY* and *PDS*, leading to increased carotenoid levels (Lu et al., 2018). In Papaya, CpbHLH1 and CpbHLH2 facilitate in the activation of *LCYb1* and *PSY* expression, thereby stimulating the accumulation of carotenoids. (Zhou et al., 2019). MYB7 regulates carotenoid accumulation by activating the lycopene *beta-cyclase (LCYB)* gene and influencing chlorophyll biosynthesis in

Kiwifruit (Ampomah-Dwamena et al., 2019). In pepper fruits, chloroplasts become chromoplasts where carotenoids are synthesized and stored. The color of ripe peppers mostly comes from these carotenoids (Tian et al., 2015; Kuai et al., 2018).

The characteristics of gardenia fruits varied notably across different growth stages, particularly in color, size, and shape. The KEGG pathways of both transcriptomics and metabolomics analysis identified several significant pathways, including plant hormone Signal transduction, phenylpropanoid biosynthesis, carbon metabolism, biosynthesis of plant secondary metabolites, biosynthesis of plant hormones, etc., during the developmental stages. The diagram of plant hormone signal transduction pathway is presented in Fig. 7. In Arabidopsis, Abscisic acid (ABA), a crucial plant hormone derived from carotenoids, plays a vital role in plant stress factors and participates in various developmental processes. Moreover, this hormone governs seed ripening (Tuan et al., 2018). This pathway comprises three primary constituents: the PYR (pyrabactin resistance)/PYL; PP2C (protein phosphatase 2C; acting as a suppressor), and SnRK2 (sucrose non-fermenting 1–related protein kinase 2; acting as an enhancer) (Todaka
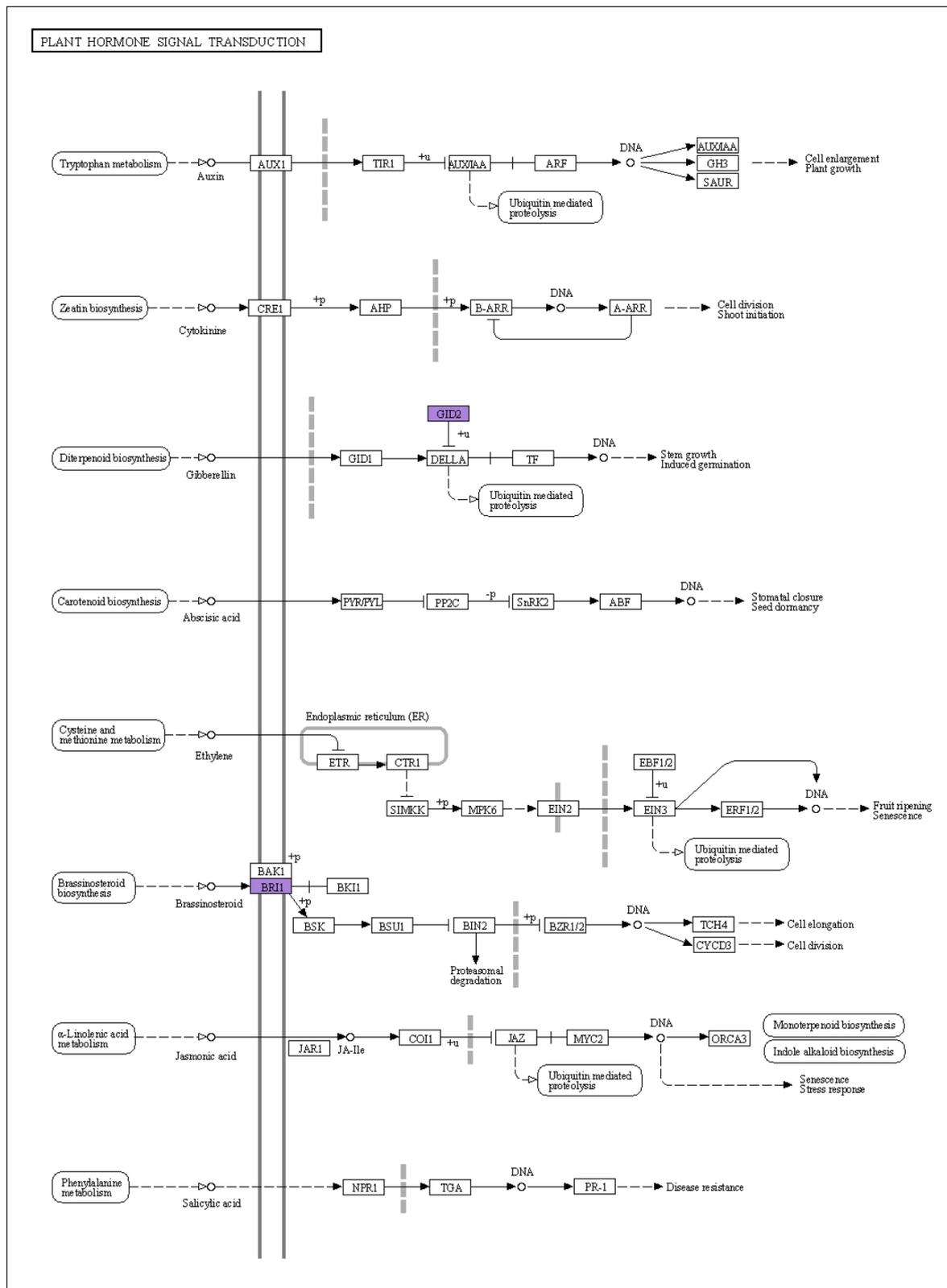
Fig. 6. Network diagram; Pathways were indicated by blue dots and metabolites were indicated by other dots. The size of a pathway dot indicates the number of associated metabolites. The color gradient of metabolite dots reflects the magnitude of the log$_2$ (fold change) value.

et al., 2015). These elements collectively constitute an ABA perception and signal transduction complex (Wang et al., 2017; Xie et al., 2019). The pigment synthesis during maturity of gardenia fruits might follow the similar mechanism.

Further pathway classification and statistical analysis provided insights into the metabolite changes during gardenia fruit development. The metabolomics analysis detected several metabolites, including 2-phenylethanol, o-Xylene, Catechol, 1, 3-benzenediol, etc. Among the detected metabolites, 2-phenylethanol is derived from the phenylpropanoid biosynthesis pathway. Therefore, 2-phenylethanol may be involved in pigment formation in gardenia fruits. O-xylene may interact with secondary metabolites in pigment synthesis and fruit coloration. Phenolic compounds, like Catechol, from the phenylpropanoid pathway is required to synthesize pigments such as anthocyanins and flavonoids, affecting fruits color. 1,3-Benzenediol belongs to the class of phenolic compounds and is categorized within the larger group of plant secondary metabolites originating from the phenylpropanoid pathway and involved in the biosynthesis of pigments like anthocyanins and flavonoids, influencing fruit pigment synthesis (Jaakola, 2013).

The biosynthesis of fruit pigments, particularly anthocyanins, is critically linked to the phenylpropanoid pathway. Phenylpropanoids serve as source materials for diverse secondary metabolites, such as flavonoids, which contribute to the vibrant colors of fruits. The critical steps in the phenylpropanoid biosynthesis pathway and its connection to fruit pigment formation are described as follows: The phenylpropanoid pathway begins with the conversion of phenylalanine to trans cinnamic acid through deamination, a process facilitated by PAL. Following this, C4H transforms *trans*-cinnamic acid into p-coumaric acid. These initial stages of the pathway are vital for generating a range of phenolic compounds, including those responsible for fruit pigment formation (Boerjan et al., 2003). In plants, carbon metabolism involves

distributing carbon compounds between source and sink tissues. Source tissues, typically leaves, generate sugars via photosynthesis, whereas sink tissues, like fruits, utilize these sugars for diverse metabolic functions (Koch, 1996). The translocation of sugars in-fluences the availability of precursors for pigment biosynthesis in developing fruits. ABC transporters participate in the transportation of anthocyanins, which serve as principal pigments accountable for the red, purple, and blue pigments observed in fruits. These transporters mediate the movement of anthocyanins from the cytoplasm to the vacuole, where they accumulate and contribute to the vibrant coloration of fruits (Goodman et al., 2004). ABC transporters also play a role in regulating flavonoid accumulation, including various pigments in fruits. They influence the intracellular concentration of flavonoids by controlling their transport across cellular membranes. This regulation impacts the overall pigment synthesis levels in fruits (Gomez et al., 2009).

Anthocyanins, carotenoids, and betalains represent three essential pigments found in plants. Anthocyanins are synthesized through the phenylpropanoid pathway, and in the cytoplasm, they give rise to anthocyanin derivatives, which are then deposited within the cytoplasm (Saslowsky et al., 2005; Zhao et al., 2017). The end product of the flavonoid metabolic pathway is anthocyanins. Several hormones, such as IAA and JA, regulate the biosynthesis of Anthocyanins. The synthesis of carotenoids, anthocyanins, and chlorophylls in gardenia is a complex process that involves several enzymes. ABA is known to influence fruit pigment synthesis, particularly during the ripening process. It is involved in controlling the biosynthesis of anthocyanins in fruits (Huang et al., 2012; Liu et al., 2015). These elements can control the expression of genes involved in anthocyanin production. Auxins, such as IAA, play roles in multiple aspects of plant development, including the growth of fruits. They can influence fruit pigment synthesis indirectly by regulating processes like cell expansion and differentiation. Differences in

Fig. 7. The diagram of Plant hormone Signal Transduction derived from KEGG. Without ABA, type 2C protein phosphatases (PP2Cs) bind to and inhibit SNF1-related kinases (SnRK2s), preventing ABA signaling. When ABA is present, it binds to the ABA receptor PYR/PYL/PCAR, enabling the receptor to interact with PP2Cs. This interaction releases SnRK2s, which then undergo phosphorylation either autonomously or via other kinases. The phosphorylated SnRK2s activate ABA-responsive element binding factors (ABFs), which play a role in controlling the expression of specific downstream genes of pigment synthesis as well as fruit ripening.

the types and quantity of pigments generated in fruits are frequently caused by variations in the genes linked to these mechanisms. Furthermore, both internal hormonal factors and external environmental conditions have varying degrees of influence on the synthesis and breakdown of these pigments. As a result, gardenia fruits display a variety of color variants in both their immature and mature forms.

Functional diversity was clarified by the COG categorization, which focused mainly on functions in pigment formation. Critical pathways affecting the pigments synthesis were revealed by pathway analysis using KEGG and GO. Cluster and pathway analyses offered a comprehensive understanding of metabolic patterns and pathway participation, which highlighted the significance of pathways like the manufacture of phenylpropanoids, plant hormone signal transduction, and secondary metabolite biosynthesis in plants. This combined transcriptome and metabolomic approach helps to enhance our comprehension of the molecular mechanisms governing pigment synthesis in gardenia fruits and provides essential information for future research.

5. Conclusions

Gardenia fruits have gained significant attention due to their nutritional and medicinal properties. Gardenia fruits contain pharmacodynamically rich chemicals, including flavonoids, diterpenoids, triterpenoids, monoterpene glycosides, and organic acids. In this study, both transcriptomics and metabolomics analysis identified candidate genes and metabolites attributed to various biological pathways. The phenylpropanoid biosynthesis pathway is central to forming fruit pigments, particularly anthocyanins. The enzymatic reactions in this pathway, from phenyl-alanine to the synthesis of flavonoids, contribute to the diverse and vibrant colours observed in fruits. Other pathways, such as photosynthesis, plant secondary metabolite biosynthesis and plant hormone signaling, which are also involved in pigment formation, were also identified. Understanding the regulation of these pathways is crucial for manipulating fruit pigment formation in agriculture and enhancing nutritional and aesthetic qualities of fruits. The biosynthesis of plant secondary metabolites responsible for fruit pigment synthesis which are often starts with the phenylpropanoid pathway by producing several phenolic compounds. This study successfully identified several vital genes and metabolites responsible for pigment formation in gardenia fruits at premature and mature stages. These findings greatly contribute to the deep understanding of the molecular mechanisms underlying pigment formation in gardenia fruits and provide valuable insights for improving their overall quality.

Funding

Author contributions

Conceptualization: **K.L.**, **L.Y**. and **S.D.**; Methodology: **K.L.** and **L.Y.**; Validation: **K.L. L.Y.** and **L.Z.**; Formal analysis: **K.L.** and **L.Y.**; Investigation: **L.G.**, **L.Z.** and **X.F.** Data curation: **L.G** and **X.F.**; Writing—original draft preparation: **K.L., X.F** and **S.D.**; Writing—review and editing: **L.G. L.Z.** and **S.D.** All authors have read and agreed to the published version of the manuscript.

CRediT authorship contribution statement

Kangqin Li: Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization. **Lixin Yu:** Validation, Methodology, Formal analysis. **Liqin Gao:** Writing – review & editing, Investigation,

Data curation. **lingzhi Zhu:** Writing – review & editing, Validation, Investigation. **Xiaotao Feng:** Writing – original draft, Investigation, Data curation. **Shaoyong Deng:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

Ampomah-Dwamena, C., Thrimawithana, A. H., Dejnoprat, S., Lewis, D., Espley, R. V., & Allan, A. C. (2019). A kiwifruit (Actinidia deliciosa) R2R3-MYB transcription factor modulates chlorophyll and carotenoid accumulation. *New Phytologist, 221*(1), 309–325. https://doi.org/10.1111/nph.15362

Barry, C. S., McQuinn, R. P., & Thompson, A. J. (2008). Ethylene insensitivity conferred by the Green-ripe and Never-ripe 2 ripening mutants of tomato. *Plant Physiology, 148* (2), 949–957.

Bechtold, T., Mussak, R., Mahmud-Ali, A., Ganglberger, E., & Geissler, S. (2006). Extraction of natural dyes for textile dyeing from coloured plant wastes released from the food and beverage industry. *Journal of the Science of Food and Agriculture, 86* (2), 233–242. https://doi.org/10.1002/jsfa.2360

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology, 54*(1), 519–546. https://doi.org/10.1146/annurev. arplant.54.031902.134938

Boo, H. O., Hwang, S. J., Bae, C. S., Park, S. H., Heo, B. G., & Gorinstein, S. (2012). Extraction and characterization of some natural plant pigments. *Industrial Crops and Products, 40*, 129–135. https://doi.org/10.1016/j.indcrop.2012.02.042

Choi, H. J., Park, Y. S., Kim, M. G., Kim, T. K., Yoon, N. S., & Lim, Y. J. (2001). Isolation and characterization of the major colorant in Gardenia fruit. *Dyes and Pigments, 49* (1), 15–20. https://doi.org/10.1016/S0143-7208(01)00007-9

Delgado-Vargas, F., Jiménez, A. R., & Paredes-López, O. (2000). Natural pigments: Carotenoids, anthocyanins, and betalains—characteristics, biosynthesis, processing, and stability. *Critical Reviews in Food Science and Nutrition, 40*(3), 173–289. https:// doi.org/10.1080/10408690091189257

El-Shishtawy, R. M., Shokry, G. M., Ahmed, N. S., & Kamel, M. M. (2009). Dyeing of modified acrylic fibers with curcumin and madder natural dyes. *Fibers and Polymers, 10*, 617–624. https://doi.org/10.1007/s12221-010-0617-4

Gomez, C., Conejero, G., Torregrosa, L., Cheynier, V., Terrier, N., & Ageorges, A. (2009). In vivo grapevine anthocyanin transport involves vesicle-mediated trafficking and the contribution of anthoMATE transporters and GST. *The Plant Journal, 59*(6), 959–972.

Goodman, C. D., Casati, P., & Walbot, V. (2004). A multidrug resistance–associated protein involved in anthocyanin transport in Zea mays. *The Plant Cell, 16*(7), 1812–1826. https://doi.org/10.1105/tpc.022574

Havaux, M., & Niyogi, K. K. (1999). The violaxanthin cycle protects plants from photooxidative damage by more than one mechanism. *Proceedings of the National Academy of Sciences, 96*(15), 8762–8767. https://doi.org/10.1073/pnas.96.15.8762

Huang, Y., Dai, H., & Zheng, X. (2012). Plant hormone receptors: new perceptions. *Genetic and Molecular Research, 11*(4), 3125–3135.

Ichi, T., Higashimura, Y., Katayama, T., Koda, T., Shimizu, T., & Tada, M. (1995). Food chemical properties of crocetin derivatives in gardenia (*Gardenia jasminoides* Ellis) yellow color. DOI:10.3136/NSKKK.42.784.

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends in Plant Science, 18*(9), 477–483. https://doi.org/10.1016/j. tplants.2013.06.003

Kieffer, D. A., Piccolo, B. D., Vaziri, N. D., Liu, S., Lau, W. L., Khazaeli, M., & Adams, S. H. (2016). Resistant starch alters gut microbiome and metabolomic profiles concurrent with amelioration of chronic kidney disease in rats. *American Journal of Physiology-Renal Physiology, 310*(9), F857–F871. https://doi.org/10.1152/ajprenal.00513.2015

Koch, K. E. (1996). Carbohydrate-modulated gene expression in plants. *Annual Review of Plant Biology, 47*(1), 509–540. https://doi.org/10.1146/annurev.arplant.47.1.509

Kuai, B., Chen, J., & Hörtensteiner, S. (2018). The biochemistry and molecular biology of chlorophyll breakdown. *Journal of Experimental Botany, 69*(4), 751–767. https://doi. org/10.1093/jxb/erx322

Li, Y., Gao, Z. R., Zhang, C., Li, N., & Liu, C. (2015). Research progress on the molecular regulation mechanism of anthocyanin biosynthesis pathway. *Chinese Journal of Ecology, 34*(10), 2937. https://doi.org/10.3390/cimb45090458

Li, S., Wang, Y., Shen, Y., Zheng, S., & Liu, H. (2023). Transcriptome characterization of pigment-related genes in jujube (Ziziphus Jujuba Mill.) peel at different growth stages. *Biochemical Genetics, 61*(6), 2425–2442. https://doi.org/10.1007/s10528-023-10382-0

Liu, H. W. (2010). The current situation of Jiangxi Gardenia is worrying, and the market outlook is worthy of attention. *Zhongguo Xiandai Zhongyao, 12*(6), 49–51.

Liu, J., Zhu, P., Zhao, C., Sui, S., Dong, Z., & Zhang, L. (2014). Study on the dyeing of wool fabrics with Phytolacca berry natural dyes. *Fibers and Polymers, 15*, 1601–1608. https://doi.org/10.1007/s12221-014-1601-1

Liu, M., Pirrello, J., Chervin, C., Roustan, J. P., & Bouzayen, M. (2015). Ethylene control of fruit ripening: Revisiting the complex network of transcriptional regulation. *Plant Physiology, 169*(4), 2380–2390. https://doi.org/10.1104/pp.15.01361

Lu, S., Zhang, Y., Zhu, K., Yang, W., Ye, J., Chai, L., … Deng, X. (2018). The citrus transcription factor CsMADS6 modulates carotenoid metabolism by directly regulating carotenogenic genes. *Plant Physiology, 176*, 2657–42676. https://doi.org/10.1104/pp.17.01830

Mao, X., Cai, T., Olyarchuk, J. G., & Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics, 21*(19), 3787–3793. https://doi.org/10.1093/bioinformatics/bti430

Nagata, Y., Watanabe, T., Nagasaka, K., Yamada, M., Murai, M., Takeuchi, S., & Taniuchi, N. (2016). Total dosage of gardenia fruit used by patients with mesenteric phlebosclerosis. *BMC complementary and alternative medicine, 16*, 1–11. https://doi.org/10.1186/s12906-016-1182-1

Navarro-Reig, M., Jaumot, J., García-Reiriz, A., & Tauler, R. (2015). Evaluation of changes induced in rice metabolome by Cd and Cu exposure using LC-MS with XCMS and MCR-ALS data analysis strategies. *Analytical and Bioanalytical Chemistry, 407*, 8835–8847. https://doi.org/10.1007/s00216-015-9042-2

Ogata H, Goto S, Sato K, et al. KEGG: kyoto Encyclopedia of Genes and Genomes [J]. (1999). Nucleic Acids Research. 27(1), 29–34.

Oren, E., Tzuri, G., Vexler, L., Dafna, A., Meir, A., Faigenboim, A., & Gur, A. (2019). The multi-allelic APRR2 gene is associated with fruit pigment accumulation in melon and watermelon. *Journal of Experimental Botany, 70*(15), 3781–3794. https://doi.org/10.1093/jxb/erz182

Pan, Y., LI, L.-Y., Wang, Y., & Song, X.-H. (2019). Investigation report on distribution and comprehensive utilization of Gardenia jasminoides cultivar resources in China. *Natural Product Research and Development, 31*(10), 1823–1830.

Park, K. H., Kim, T. Y., Park, J. Y., Jin, E. M., Yim, S. H., Choi, D. Y., & Lee, J. W. (2013). Adsorption characteristics of gardenia yellow as natural photosensitizer for dye-sensitized solar cells. *Dyes and Pigments, 96*(2), 595–601. https://doi.org/10.1016/j.dyepig.2012.10.005

Peng, K., Yang, L., Zhao, S., Chen, L., Zhao, F., & Qiu, F. (2013). Chemical constituents from the fruit of *Gardenia jasminoides* and their inhibitory effects on nitric oxide production. *Bioorganic & medicinal chemistry letters, 23*(4), 1127–1131. https://doi.org/10.1016/j.bmcl.2012.11.099

Pyle, R. M. (2003). Nature matrix: Reconnecting people and nature. *Oryx, 37*(2), 206–214. https://doi.org/10.1017/S0030605303000383

Sangare, R., Madhi, I., Kim, J. H., & Kim, Y. (2023). Crocin attenuates NLRP3 inflammasome activation by inhibiting mitochondrial reactive oxygen species and ameliorates monosodium urate-induced mouse peritonitis. *Current Issues in Molecular Biology, 45*(3), 2090–2104. https://doi.org/10.3390/cimb45030134

Saslowsky, D. E., Warek, U., & Winkel, B. S. (2005). Nuclear localization of flavonoid enzymes in Arabidopsis. *Journal of Biological Chemistry, 280*(25), 23735–23740. https://doi.org/10.1074/jbc.M413506200

Shahid, M., & Mohammad, F. (2013). Perspectives for natural product-based agents derived from industrial plants in textile applications–a review. *Journal of Cleaner Production, 57*, 2–18. https://doi.org/10.1016/j.jclepro.2013.06.004

Shen, J., Gao, P., & Ma, H. (2014). The effect of tris (2-carboxyethyl) phosphine on the dyeing of wool fabrics with natural dyes. *Dyes and Pigments, 108*, 70–75. https://doi.org/10.1016/j.dyepig.2014.04.027

Siwawej, S., & Jarayapun, A. (1998). The food color extraction from gardenia fruit (*Gardenia jasminoides* Ellis forma var. gardiflora Makino): the appropriate extraction conditions. 520-528.

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical chemistry, 78*(3), 779–787. https://doi.org/10.1021/ac051437y

Sundaramoorthy, J., Park, G. T., Lee, J. D., Kim, J. H., Seo, H. S., & Song, J. T. (2015). Genetic and molecular regulation of flower pigmentation in soybean. *Journal of the Korean Society for Applied Biological Chemistry, 58*, 555–562. https://doi.org/10.1007/s13765-015-0077-z

Sun, T., Tadmor, Y., & Li, L. (2020). Pathways for carotenoid biosynthesis, degradation, and storage. *Plant and Food Carotenoids: Methods and Protocols, 3–23*. https://doi.org/10.1007/978-1-4939-9952-1_1

Sun, T., Rao, S., Zhou, X., & Li, L. (2022). Plant carotenoids: Recent advances and future perspectives. *Molecular Horticulture, 2*(1), 3. https://doi.org/10.1186/s43897-022-00023-2

Tanaka, Y., Sasaki, N., & Ohmiya, A. (2008). Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *The Plant Journal, 54*(4), 733–749. https://doi.org/10.1111/j.1365-313X.2008.03447.x

Tangpao, T., Chung, H. H., & Sommano, S. R. (2018). Aromatic profiles of essential oils from five commonly used Thai basils. *Foods, 7*(11), 175. https://doi.org/10.3390/foods7110175

Tian, S. L., Li, L., Shah, S. N. M., & Gong, Z. H. (2015). The relationship between red fruit colour formation and key genes of capsanthin biosynthesis pathway in Capsicum annuum. *Biologia plantarum, 59*(3), 507–513. https://doi.org/10.1007/s10535-015-0529-7

Todaka, D., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2015). Recent advances in the dissection of drought-stress regulatory networks and strategies for development of drought-tolerant transgenic rice plants. *Frontiers in Plant Science, 6*, Article 121920. https://doi.org/10.3389/fpls.2015.00084

Tsanakas, G. F., Manioudaki, M. E., Economou, A. S., & Kalaitzis, P. (2014). De novo transcriptome analysis of petal senescence in *Gardenia jasminoides* Ellis. *BMC Genomics, 15*, 1–16. https://doi.org/10.1186/1471-2164-15-554

Tuan, P. A., Kumar, R., Rehal, P. K., Toora, P. K., & Ayele, B. T. (2018). Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals. *Frontiers in Plant Science, 9*, Article 362906. https://doi.org/10.3389/fpls.2018.00668

Wang, T. (2015). Research progress in effective constituents and pharmacological effects of Gardenia Jasminoides. *China Pharmacist, 18*(10), 1782–1784.

Wang, T., Li, C., Wu, Z., Jia, Y., Sun, S., & Wang, X. (2017). Abscisic acid regulates auxin homeostasis in rice root tips to promote root hair elongation. *Frontiers in Plant Science, 8*, Article 265957. https://doi.org/10.3389/fpls.2017.01121

Wang, C. K., Han, P. L., Zhao, Y. W., Ji, X. L., Yu, J. Q., You, C. X., & Hao, Y. J. (2020). Auxin regulates anthocyanin biosynthesis through the auxin repressor protein *MdIAA26*. *Biochemical and Biophysical Research Communications, 533*(4), 717–722. https://doi.org/10.1016/j.bbrc.2020.09.065

Wang, W., Shao, F., Deng, X., Liu, Y., Chen, S., Li, Y., & Zhang, X. (2021). Genome surveying reveals the complete chloroplast genome and nuclear genomic features of the crocin-producing plant *Gardenia jasminoides* Ellis. *Genetic Resources and Crop Evolution, 68*, 1165–1180. https://doi.org/10.1007/s10722-020-01056-6

Wenping, X., Shiming, L., Siyu, W., & Tang, H. C. (2017). Chemistry and bioactivity of Gardenia jasminoides. *Journal of Food and Drug Analysis., 25*(1), 43–61.

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., & Querengesser, L. (2007). HMDB: The human metabolome database. *Nucleic Acids Research, 35*(suppl_1), D521–D526. https://doi.org/10.1093/nar/gkl923

Xia, J., & Wishart, D. S. (2011). Web-based inference of biological patterns, functions and pathways from metabolomic data using Metabo Analyst. *Nature Protocols, 6*(6), 743–760. https://doi.org/10.1038/nprot.2011.319

Xie, Z., Nolan, T. M., Jiang, H., & Yin, Y. (2019). AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in Arabidopsis. *Frontiers in Plant Science, 10*, Article 437723. https://doi.org/10.3389/fpls.2019.00228

Yang, C. X., Zhang, T., Xu, M., Zhu, P. L., & Deng, S. Y. (2016). Insights into biosynthetic genes involved in the secondary metabolism of *Gardenia jasminoides* Ellis using transcriptome sequencing. *Biochemical Systematics and Ecology, 67*, 7–16. https://doi.org/10.1016/j.bse.2016.05.011

Zelena, E., Dunn, W. B., Broadhurst, D., Francis-McIntyre, S., Carroll, K. M., Begley, P., & Kell, D. B. (2009). Development of a robust and repeatable UPLC– MS method for the long-term metabolomic study of human serum. *Analytical Chemistry, 81*(4), 1357–1364. https://doi.org/10.1021/ac8019366

Zhang, M., Chen, W. D., Li, Y. Y., Zhang, C., Chai, Z. H., Li, Y. F., & Wang, X. R. (2021). Complete chloroplast genome of a wild-type *Gardenia jasminoides* ellis (rubiaceae) adapted to island climate. *Mitochondrial DNA Part B, 6*(2), 313–314. https://doi.org/10.1080/23802359.2020.1861997

Zhang, L., Ai, Y., Chen, Y., Li, C., Li, P., Chen, J., & Liu, Q. (2023). Elucidation of geniposide and crocin accumulation and their biosysnthsis related key enzymes during *Gardenia jasminoides* fruit growth. *Plants, 12*(11), 2209. https://doi.org/10.3390/plants12112209

Zhao, D., Wang, R., Meng, J., Li, Z., Wu, Y., & Tao, J. (2017). Ameliorative effects of melatonin on dark-induced leaf senescence in gardenia (Gardenia jasminoides Ellis): Leaf morphology, anatomy, physiology and transcriptome. *Sci. Rep., 7*, 10423. https://doi.org/10.1038/s41598-017-10799-9

Zhou, Y., Zhang, J., Tang, R.-C., & Zhang, J. (2014). Simultaneous dyeing and functionalization of silk with three natural yellow dyes. *Ind Crops Prod., 64*, 224–232. https://doi.org/10.1016/j.indcrop.2014.09.041

Zhou, D., Shen, Y., Zhou, P., Fatima, M., Lin, J., Yue, J., & Ming, R. (2019). Papaya CpbHLH1/2 regulate carotenoid biosynthesis-related genes during papaya fruit ripening. *Horticulture Research, 6*. https://doi.org/10.1038/s41438-019-0162-2

Zhou, Z., Ying, Z., Wu, Z., Fu, S., & Liu, Z. (2021). Anthocyanin genes involved in the flower coloration mechanisms of Cymbidium kanran. *Frontiers in Plant Science, 12*, Article 737815. https://doi.org/10.3389/fpls.2021.737815