Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Non-targeted metabolomics and transcriptomics reveal mechanisms of metabolic differences among roots, stems, and leaves of Cudrania tricuspidata caryophyllata

Zikang Ding
Huzhou University
Meng Zhao
Huzhou University
Xinyue Li
Huzhou University
Xiaomei Wang
Huzhou University
Zhongshan Zhang
01959@zjhu.edu.cn

Huzhou University

Research Article

Keywords: Cudrania tricuspidate, Metabolomics, transcriptomics, flavonoid

Posted Date: August 31st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3299607/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Cudrania tricuspidata is a traditional Chinese medicinal herb with a long history of application, and has a variety of active ingredients and pharmacological effects. In order to expand the source of new drugs, we analysed the metabolomic and transcriptomic profiles of different tissues of Cudrania. The genetic basis for the accumulation of active ingredients was analysed, providing a theoretical basis for high-quality cultivation of Cudrania.

Results

We detected a total of 1254 metabolites from the three tissues of Cudrania roots, stems, and leaves, and all metabolites were annotated and classified into eight categories by the KEGG database: steroids, lipids, antibiotics, vitamins and cofactors, nucleic acids, peptides, carbohydrates, and organic acids. Flavonoid-rich roots and stems of Cudrania were significantly different from the transcripts of leaves. GO and KEGG enrichment analyses revealed that the differential genes were mainly enriched in Photosynthesis - antenna proteins, Zeatin biosynthesis, Flavone and flavonol biosynthesis, Monoterpenoid biosynthesis pathway. The expression of flavonoid and flavonol biosynthesis-related genes was significantly up-regulated in roots and stems. From the perspective of the differences in metabolites among roots, stems and leaves of Cudrania, it can provide a basis for revealing the material basis of the differences in medicinal properties and efficacy of different parts.

Conclusion

These results indicate that Cudrania is a promising medicinal plant and provide important information for the identification of differential metabolites and major metabolic pathways among different parts of Cudrania, as well as a theoretical reference for the effective use of Cudrania resources.

Introduction

Cudrania is the dried rhizome of the *Cudrania tricuspidata*, sweet in flavour, warm in nature and non-toxic [1]. It is a traditional Chinese herbal medicine in China, with the effects of clearing heat and cooling the blood, and soothing the tendons and activating the collaterals [2]. It is mainly used for the treatment of women's "blood knot collapse, malaria and other diseases". In the clinical application in recent years, Cudrania is often used as an effective adjunctive Chinese medicine for anti-tumour radiotherapy. It is mainly shown to improve the clinical symptoms and signs of patients, improve the body's immunity, and no obvious adverse reactions have been seen [3]. *Gleanings from Materia Medica* records that Cudrania is mainly used to replenish deficiency. Cudrania has a wide range of uses, stems, roots and leaves can be used as medicine, and the fruits can be used as both medicine and food [4]. In the 1977 edition of the

Chinese Pharmacopoeia, Cudrania was only included under the heading of *Chuan po shi*. In the 1977 edition of the Chinese Pharmacopoeia, Cudrania was only included under the name *Chuan po shi*, and has not been included in the *Chinese Pharmacopoeia* since then, resulting in the lack of a uniform definition of the medicinal parts of Cudrania.

Cudrania contains a variety of components, mainly including flavonoids, xanthones, organic acids, polysaccharides and phenylpropanoid 5 types of components, of which cudweed flavonoids and xanthones are the representative components of cudweed, with a wealth of bio-activities, which can be antioxidant, anti-inflammatory, anti-tumor, neuroprotective and hepatoprotective [5]. Organic acids are widely found in the leaves, roots and fruits of plants. The fruits of the cudrania tree contain a wide range of organic acids in much higher amounts than those of the cudrania tree [6]. In addition to the above, there are alkaloidal components such as hydrastine, amino acids such as proline, glutamic acid, arginine, alanine and aspartic acid. Essential trace elements such as iron, zinc, manganese, strontium and copper, and possibly essential or harmless trace elements such as aluminum, silicon and tin [7–9].

At present, most of the domestic studies on the pharmacological effects of Cudrania cudrania are focused on anti-tumor [10–12]. Compositional studies are limited to flavonoid components and xanthone components, and the research on Cudrania is not deep enough [13]. Metabolomics technology, as an emerging research method, breaks through the limitations of traditional metabolite chemical composition detection and analysis, and allows high-throughput quantitative analysis of all metabolites in the organism, which not only systematically analyzes the differences in metabolite composition among different samples, but also detects substances difficult to be detected by conventional analytical methods, and systematically and completely elucidates the factors that affect plant growth and crop fruit quality, and is gradually being applied in the fields of agriculture, medicine, and food. Currently, there are fewer research reports on the metabolomics of Cudrania. Using metabolomics technology to systematically study the differences in metabolite compositions between different tissue parts of Cudrania is of great significance for carrying out the analysis and evaluation of the nutritional and medicinal qualities of Cudrania, enlarging the medicinal parts, searching for new sources of medicines, and guiding the development of Cudrania [14].

In this study, we used non-targeted metabolomics and transcriptomics technologies based on liquid chromatography-mass spectrometry (LC-MS) to analyze the composition of metabolites in roots, stems, and leaves of Cudrania, to identify the differences of metabolites in different parts of the plant and analyze the metabolic pathways of the metabolites, and to explore the effects of the different metabolites on the medicinal quality of Cudrania with the purpose of providing a basis for analysis and evaluation of the nutritional quality of Tsuga cerasus. This study is intended to provide a basis for analyzing and evaluating the nutrition and quality of Cudrania, and to provide a reference for the development of the whole plant of Cudrania [15].

Materials and Methods

Experimental Materials

The Cudrania material used in this study was obtained from Changxing County, China. The planting method used single rows and double planting, with 1.5 m row spacing and 1.0 m plant spacing. One branch and one root and two leaves of the three materials were harvested, respectively, from at least 30 Cudrania plants of each cultivar, and were immediately fixed with liquid nitrogen. Each sample included six independent biological replicates at the time of picking. Six biological replicates were used for each sample for metabolomics detection, and three biological replicates were randomly selected from six biological replicates for transcriptomic and biochemical component detection. All the samples were stored in a − 80 °Crefrigerator [16].

Metabolite Detection

The extraction of metabolites was carried out as follows: the freeze-dried crushed samples of 100 ± 0.1 mg were weighed and placed in a 2 mL centrifuge tube, with the addition of 400 µL methanol solution, vortexed for 1 min, extracted in an ultrasonic bath for 30 min, and then centrifuged for 10 min at 12,000 rpm and 4 °C. The supernatant was transferred to a new 2 mL centrifuge tube and then adding 150 µL of 2-chloro-l-phenylalanine (4 ppm) solution prepared with 80% methanol water (stored at 4 °C). The supernatant was filtered using a 0.22 µm PTFE membrane (Jinteng, Tianjin, China), and the filtrate was added to the flask for liquid chromatography–mass spectrometry (LC–MS) detection. Six biological replicates were performed for each sample [16].

Quality control sample processing was carried out, with quality control samples (QC) prepared from the same amount of mixed crushed samples of the four materials. They were extracted and detected using the same method as that used for the analytical samples, and the QC samples were repeated six times. In the process of instrument detection, one QC sample was inserted into every ten detection and analysis samples to monitor the repeatability of the whole analysis process.

The MS conditions were set as follows: Ion Source Gas1 (Gas1) and Ion Source Gas2 Gas2) were set as 60 psi, curtain gas (CUR) pressure was set as 30 psi, source temperature was 600 ◦C, IonSpray Voltage Floating (ISVF) was ± 5500 V. In MS only acquisition, the mass range of the in strument was from m/z 60–1000 Da, and the accumulation time for TOF MS scan was set at 0.20 s/spectra. In auto MS/MS acquisition, the mass range was set in m/z 25–1000 Da with an accumulation time for product ion scan of 0.05 s/spectra. The product ion scan was acquired using information dependent acquisition (IDA) with high sensitivity mode selected. And the parameters were set as follows: the collision energy (CE) was fixed at 35 V with ± 15 eV; declustering potential (DP) were 60 V (+) and − 60 V (−); exclude isotopes within 4 Da, and 10 candidate ions monitored per cycle [17].

Transcriptome Sequencing

Total RNA was extracted using the TRIzol reagent (Invitrogen, Waltham, MA, USA) according to the manufacturer's instructions. The quality-assured total RNA was subjected to cDNA library construction and transcript analysis. A total of 12 libraries were constructed and sequenced using the Illumina

platform based on the Paired-End 150 (PE150) strategy. Fragments per kilobase of transcripts per million fragments mapped (FPKM) were performed to calculate gene expression levels. Then, the DEGs were screened using the criteria of log2 (fold change) ≥ 1.5 and a p-value of < 0.01. Gene ontology (GO) enrichment analysis of the DEGs was carried out using the GOseq R software package, and KOBAS software was used to test the statistical enrichment of the DEGs in the Kyoto Encyclopedia of Genes and Genomes (KEGG) path [18–19].

Data Statistics and Analysis

ProteoWizard (v3.0.8789) was used to process the original data of the metabolomics. The XCMS package of R (v3.3.2) was used for peak identification, peak filtration, and peak alignment, and the data matrix including the mass-to-charge ratio (m/z), retention time, and peak area (intensity) was obtained. One-way analysis of variance (ANOVA) with the least significant difference (LSD) was performed using SPSS19.0. Heat mapping, PCA analysis, OPL-DS analysis, correlation analysis and pathway analysis were performed using the Micromax system.

Results

Metabolomics analysis

Qualitative and quantitative analysis of metabolites

We detected a total of 1254 metabolites from the three tissues of Cudrania roots (CR), stems (CB), and leaves (CL), and all metabolites were annotated and classified into eight categories by the KEGG database: steroids, lipids, antibiotics, vitamins and cofactors, nucleic acids, peptides, carbohydrates, and organic acids. The main metabolites are DL-Arginine L-Aspartic acid 6-Methylnicotinamide Trigonelline D-Saccharic acid D-(-)-Salicin D-(+)-Mannose Asparagine 2-Oxoglutaric acid Choline a-Lactose Parishin C Parishin E Sucrose Plantagoside D-(-)-Quinic acid Citric acid Fumaric acid DL-Malic acid DL-Stachydrine. Heat map analysis of the major metabolites of the three tissues showed that we found that the abundance of metabolites varied greatly. This also suggests that there are different medicinal values among the roots, stems and leaves of the Cudrania plant. It will provide a basis for the development and utilization of Cudrania plants (Fig. 1).

Principal Component Analysis (PCA)

The pre-processed LC-MS data were subjected to unsupervised PCA analysis to obtain the spatial distribution of roots, stems and leaves of the pre-processed LC-MS data were subjected to unsupervised PCA analysis to obtain the spatial distribution of roots, stems and leaves of Cudrania. As shown in the Fig. 2A, replicate samples from the same site were closer to each other, while there was a clear separation between samples from the Cudrania root, stem, and leaf groups. It was shown that PCA analysis had a good differentiation between roots, stems and leaves of Cudrania, indicating that there were significant differences in metabolome composition and content. PC1 explained 41.3% of the total variables and PC2

explained 25.3% of the total variables, with a cumulative contribution of 66.6%, which explains the overall variables well and is consistent with the results of the heat map analysis.

PLS-DA analysis

By PLS-DA model analysis, the root, stem, and leaf groups of Cudrania can be completely separated, and the three tissue samples are distributed in different confidence intervals. The cross-validation parameters of the model $R^2X = 0.856$, $R^2Y = 1.000$, $Q^2 = 0.992$, and all three parameters are greater than 0.5, saying that the model has a good degree of interpretability and predictability (Fig. 2B). In order to verify whether the constructed model is overfitting, we can use the Permutation Test, OPLSDA's Permutation Test selects prediction accuracy as the test statistic. The test statistic is observed to be on the right side of the random distribution (observations are significantly larger than random values), Q2 = 0.981, with a p-value of 0.01, which is less than 0.05, suggesting that our PLS-DA discriminant model is significant and not overfitted. Three groups of roots, stems and leaves could be significantly differentiated, and there were also significant differences in metabolites between groups (Fig. 2C).

Differential metabolite analysis

Based on multivariate analysis of the VIP values of the first two principal components of the PLS-DA model, combined with univariate analysis of the multiplicity of differences (fold change, FC) and q-value values, the intersection of VIP $\geq$ 1.0, fold change $\geq$ 1.2 or $\leq$ 0.8333; and q-value < 0.05 was taken to obtain the shared ion, i.e., the difference ions, based on the VIP and FC values, the volcano diagram of differential metabolites was derived. The screening revealed a total of 649 differential metabolites compared to roots and stems, 742 differential metabolites compared to roots and leaves, and 611 differential metabolites compared to stems and leaves. It can be seen that the metabolite compositions of the various medicinal parts of Cudrania are very rich, and it is difficult to carry out a comprehensive study on the metabolites of Cudrania plants by traditional methods. Upon further analysis of the metabolites among the three parts in roots, stems and leaves, a total of 504 major differential metabolites were identified.

We made box plots for the top 25 ranked metabolomes, and the differential metabolites that accumulated more in roots were 2-oxo-2H-chromene-3-carboxylic acid, Gomisin N, 2-Methylbutyric acid, D-Mannitol 1-phosphate, 6,7-Dimethoxy-4-Methylcoumarin, Ciliatine, Linderane, Tricin O-malonylhexoside, Fargesin, Bergenin, Methyl cinnamate, Sodium Houttuyfonate, Ketoisophorone, 1-(1,8-dihydroxy-3,6-dimethyl-2-naphthyl)ethan-1-one, Flemiphilippinin A, 7-Demethylsuberosin, Coumarin-3-carboxylic acid, mainly flavonoids, polyphenols and coumarins. The differential metabolites that accumulated more in the stems were 2-Methoxy-4-vinylphenol. The differential metabolites that accumulate more in leaves were (1H-indol-3-yl)-3-[4-(trifluoromethyl)phenyl]acrylonitrile, Smyrindioloside, 1,11-Undecanedicarboxylic acid, 3-Butylidenephthalide, Miyabenol C, 1-(2-carboxy-2-methylpropyl)-4,6-dimethylbenzoic acid(Fig. 3).

Repeat relevance assessment

Correlation analysis is an effective method to identify metabolic similarities between different tissues. In our study, Pearson analysis was performed to find out the linear relationship between 100 differentially expressed metabolites and the results were displayed in a correlation matrix. By comparing the metabolite correlation analysis spectra for each group, we can observe if the metabolite correlation changes between groups. In this study, 577 metabolites showed significant correlations ($p < 0.001$). We found that among the top 5 metabolites with the highest correlations, organic acids and their derivatives, and peptides occupied a larger proportion relative to other biological classes, which may indicate that they can significantly influence the metabolism of other substances (Fig. 4). By analyzing the interactions between differential metabolites, we revealed the main relationships between anabolic and catabolic metabolism of metabolites as well as functional correlations between different compounds.

Metabolic pathway analysis

The purpose of enrichment analysis is to search for biological pathways that play a key role in a particular biological process, so as to reveal and understand the basic molecular mechanism of the biological process. For metabolomic data, before enrichment analysis, we need to select some metabolites to focus on, these metabolites usually need to be metabolites that are significantly different between subgroups (T-test, $p < 0.05$) and the metabolites need to be in the KEGG Compound Reaction Network (KEGG Compound Reaction Network, which is embedded in the KEGG Pathway). By enrichment analysis, the top 50 metabolic pathways with significant enrichment of differential metabolites were found, mainly including Degradation of flavonoids, Flavonoid biosynthesis, Nucleotide metabolism, and Phenylalanine metabolism, Biosynihesis of amino acids, Alanine, aspartate and glutamate metabolism are pathways that may play important roles in biology. Even if the metabolites of interest are significantly enriched in a pathway, we still don't know if these metabolites play a key role in this metabolic pathway, and topological analysis is able to calculate the magnitude of the role of the metabolite of interest in the metabolic pathway (measured by Impact). The metabolic pathways in the blue region are the prominent metabolic pathways in the ORA enrichment analysis, and the vertical coordinates demonstrate the Impact of these metabolic pathways in the topology analysis. We found that Degradation of flavonoids, Flavonoid biosynthesis, and Nucleotide metabolism were the pathways with more significant differences between the three tissues (Fig. 5).

Transcriptomics Analysis

QC analysis and functional database annotation

In order to further explore the molecular mechanisms underlying the differential effects of metabolism among different tissues of Cudrania, three materials, CB, CL, and CR, were analyzed by transcriptome sequencing. A total of 12 cDNA libraries were constructed and sequenced, and 77.69 Gb of clean data were obtained. The clean data for each sample amounted to 6.10 Gb, and the percentage of Q30 nucleobases was 93.08% or higher, indicating the high quality of the transcriptome data. Comparing the clean reads of each sample with the reference genome sequence of "Shuchazao", the comparative

efficiency was 82.91–84.84%. PCA was then performed based on the FPKM values of each gene expressed in the samples. The results of the principal component analysis of the three samples reproduced well, with significant differences among roots, stems, and leaves of Cudrania, and gene expression among different tissues was distinguished from each other, indicating that gene expression was significantly different among different tissues.

Of the 245205 genes obtained, only 44206 genes, or 18.03%, were simultaneously annotated to the seven major databases. In addition, a total of 175885 unigenes were annotated in the seven major databases, accounting for 71.73% of the total number of genes (Fig. 6A). Among them, the largest number of genes was annotated to the NR database, which was as high as 165859 genes, accounting for 67.64% of the total number of genes. The number of genes annotated to the NT database was 123,360 (50.31%). The number of genes annotated to the KEGG database was 78,952 (32.2%). The number of genes annotated to the SwissProt database was 123568 (50.39%). The number of genes annotated to the Uniprot database was 123,401 (50.33%). The number of genes annotated to the GO database was 81129 (33.09%). The proportion of genes annotated to the KOG database was 134824 (54.98%). Five databases, namely, NR, NT, KOG, SwissProt, and Uniprot, were selected for Venn diagrams (Fig. 6B).

Correlation check of gene expression levels and differential gene analysis

The correlation of gene expression levels between samples is an important indicator to test the reliability of the experiment and whether the sample selection is reasonable; the closer the correlation coefficient is to 1, the higher the similarity of expression patterns between samples. If there are biological replicates in the sample, the correlation coefficient between the biological replicates is usually high. The Encode program recommends that the squared Pearson correlation coefficient (R^2) be greater than 0.92 (under ideal sampling and experimental conditions). For specific experimental manipulations, we require that the R^2 between biological replicate samples be at least greater than 0.8. Intra- and inter-group correlation coefficients were calculated based on the TPM values of all genes in the sample and plotted as heat maps (Fig. 7A). It can be found that the correlation coefficients of each of the three groups, CB, CL and CR, are close to 1, indicating that the selected samples are reasonable and reliable, which is an important guarantee for the results of the differential gene analysis. A two-by-two comparative analysis of transcriptome data from different tissues of Cudrania revealed that the highest number of differentially expressed Unigenes (DEGs) were found in different tissues stem, leaf, and compared with root (Fig. 7B) Among them, stems were compared with leaves, of which 6741 were up-regulated and 11545 were down-regulated. Stems were compared with roots, of which 11,544 were up-regulated and 18,152 were down-regulated. Roots were compared with leaves, of which 12067 were up-regulated and 12748 were down-regulated. (Fig. 7C/D/E) Transcriptome data of different tissue samples A total of 245205 DEGs were obtained, with an error rate (FDR) of < 0.01. Overall gene expression heatmap analysis and hierarchical clustering analysis of DEGs from different tissues revealed that the stem transcript profiles of Cudrania were clustered into a single unit, while the roots and leaves could be clustered into a single unit, which

indicated that Cudrania stem transcript profiles were tissue-specific. The gene expression trends of different tissues could be roughly clustered into eight categories (Fig. 7F)

GO enrichment analysis of differential genes

A total of 175,885 Unigenes (about 71.73%) could be compared to the homologous sequences in the database by Unigenes data comparison. In order to further explore the potential molecular mechanisms of the accumulation of active constituents in roots of Cudrania, GO enrichment analysis of DEGs from different tissues was performed. 81,129 DEGs were enriched in 56 GO terms, categorized into 3 sections: molecular functions, biological processes and cellular components. Where leaves were not found to function as cellular components compared to roots, these terms, mainly annotated by DEGs, are closely related to metabolic regulation in Cudrania, suggesting that differences in metabolites between different tissues of Cudrania are transcriptionally regulated(Fig. 8).

KEGG enrichment analysis of differential genes

To further clarify the differential metabolic pathways, metabolic pathway enrichment analysis was performed on the differential genes based on KEGG enrichment analysis. The top 20 pathways with the highest enrichment degree are shown in Fig. 10. The transcriptome data were compared separately among different tissues of roots, stems and leaves, and a total of 78,952 differential genes were screened. KEGG enrichment analysis revealed that DEGs produced in stems compared to leaves were enriched in Photosynthesis - antenna proteins, Zeatin biosynthesis, Flavone and flavonol biosynthesis and Stilbenoid, diarylheptanoid and gingerol biosynthesis pathways. DEGs produced by stems compared to roots were enriched in Monoterpenoid biosynthesis, Photosynthesis - antenna proteins, Glucosinolate biosynthesis and Biosynthesis of secondary metabolites -unclassified pathways. DEGs produced by leaves and roots were enriched in Monoterpenoid biosynthesis, Valine, leucine and isoleucine biosynthesis, Photosynthesis-antenna proteins, Zeatin biosynthesis, Flavone and flavonol biosynthesis, Linoleic acid metabolism and Diterpenoid biosynthesis pathways.

Differential Gene Protein Interaction Network Analysis

We mainly apply the interactions in the STRING protein interactions database, and for the species included in the database, we directly extract the interactions of the target gene set (e.g., differential gene list) from the database to construct the network; for the species not included in the database, we firstly apply blastx (with the value set to 1e-10) to compare the target gene set sequences onto the protein sequences of the reference species included in the STRING database, and construct the interactions network using the protein interactions of the reference species. The size of a node in an interworking network graph is proportional to the degree of the node, i.e., the more edges are connected to the node, the larger its degree is, and the larger the node is, and these nodes are likely to be in a more central position in the network. The colour of the node is related to the clustering coefficient of the node, the

gradient of the colour from green to red corresponds to the value of the clustering coefficient from low to high (Fig. 11A).

SSR Analysis

Simple Sequence Repeat (SSR) refers to a class of short tandem repeat sequences in which the repeat unit lengths and base compositions are basically the same, and only a few genes show minor differences. Eukaryotic genomes contain a large number of tandem repeats with very high repeat frequency and low sequence complexity, and the mutation status of these tandem repeats varies greatly among different species, different loci of the same species, or different alleles of the same locus. We used MISA software version 1.0 with default parameters; the corresponding minimum number of repeats for each unit size is: 1–10, 2–6, 3–5, 4–5, 5–5, 6 – 5 (e.g., 1–10, for mononucleotide repeats, the number of repeats is at least 10 to be detected; 2–6, for dinucleotide repeats, the minimum number of repeats is 6) (Fig. 11B).

Joint metabolomics and transcriptomics analysis

Flavonoid biosynthesis - Reference pathway

Flavonoids are a major class of plant secondary metabolites that serves a multitude of functions including pigments and antioxidant activity [20]. Flavonoids are synthesized from phenylpropanoid derivatives by condensation with malonyl-CoA. For example, condensation of p-coumaroyl-CoA (C6-C3) with three malonyl-CoA (C3) molecules results in naringenin chalcone with a diphenylpropane (C6-C3-C6) unit, which is converted to naringenin with the flavone (2-phenylchromen-4-one) backbone by conjugate ring closure (Fig. 12). These and further modifications yield a variety of structural forms including chalcones, flavanones, dihyroflavonols, and flavans, anthocyanins, flavones and flavonols, and isoflavonoids [21–22].

In order to gain a comprehensive understanding of the biosynthesis of flavonoids, we first investigated the biosynthesis of flavonoids. This includes the and downstream metabolites of flavonoid biosynthesis. As shown in the Fig. 13, compared with other synthesis pathways we noticed that the metabolites of Naringin, Naringenin, Pelargorudin, Dihydroquercetin, (+)-Catechin, and Dihydromyricetin were relatively higher in roots, suggesting that flavonoid biosynthesis in roots upstream may be more active [23]. The metabolites with higher content in stems were Phloretin, Phloritin, Pinocembin, Hesperetin, Neohesperidin, Luteolin, Quercetin, etc., while the metabolites with higher content in leaves were only Apigenin and Kaempferol.

The metabolism of flavonoids involves two secondary metabolite pathways [24], the phenylpropane metabolism pathway (Ko00940) and the flavonoid metabolism pathway (Ko00941, Ko00942, Ko00943, Ko00944). Based on the KEGG differential gene enrichment results, we found significant differences in both flavonoid and flavonol biosynthesis (Ko00944) pathways in the comparison of different tissues and organs. Cytochrome P450 monooxygenase (CYP) is the largest enzyme family in plants, which is widely

involved in many monooxygenation reactions such as hydroxylation, epoxidation, and demethylation, and plays a crucial role in plant growth, development, and defense [25–26]. In this biosynthetic pathway, we found that CYP75B1, an enzyme gene of the CYP75B subfamily, was significantly up-regulated, indicating that roots and stems contain more flavonoids compared to leaves. One of the studies showed that the CYP75B enzyme efficiently catalyzed the production of naringenin, which is consistent with the results of metabolomics differential metabolite analysis [27]. The expression of flavonoid-3-O-glucoside glucosyltransferase (K22794, FG3) was significantly down-regulated in roots and stems, suggesting that the weak expression of this gene promotes naringenin and quercetin expression [28].

Nucleotide Metabolism in Relation to Photosynthesis

A significant enrichment of the photosynthesis-haptoglobin synthesis pathway was found in a two-by-two comparison of KEGG data from root, stem, and leaf transcriptomes. This belongs to the metabolic pathway of photosynthesis. LHC family genes are mainly involved in the capture and transfer of light energy, with different expression patterns in organs such as roots, stems, and leaves, with the highest expression in leaves. Other studies suggest that LHC family genes may play a key role in drought stress tolerance [29–30]. LHC genes were significantly up-regulated in leaves and stems compared to roots, and significantly down-regulated in roots compared to leaves and stems, with Lhca5 being an uncommon gene (Fig. 14).

Significant differences in metabolic pathways of nucleotides were found in KEGG enrichment analysis of metabolomics across combinations. Nucleotides are important biomolecules in organisms and consist of a five-carbon sugar, a phosphate group and a nitrogen base. It has important structural properties and biological functions in living organisms [31]. ATP is the most important energy molecule in the cell, releasing energy through water interpretation of phosphate bonds, providing support for plant growth and photosynthetic processes [32]. Nucleotides accumulate metabolites significantly in leaves, which can lead to an increase in chlorophyll a, b and carotenoids in leaves, and also promotes an increase in chlorophyll a/b values, which improves the photosynthetic efficiency of leaves. Consistent with the results of enrichment analysis of metabolomics.

Discussion

The root of Cudrania is traditionally used as a medicinal part of the plant, scientifically known as *Chuan po shi* [33].The chemical constituents of the root include five types of constituents: flavonoids, xanthones, organic acids, polysaccharides, and phenylpropanoids [33–34]. Among them, flavonoids and xanthones are the representative constituents of Cudrania. At present, 82 kinds of flavonoids, 97 kinds of xanthones, 20 kinds of organic acids, 6 kinds of polysaccharides and 21 kinds of phenylpropanoids have been isolated and identified in Cudrania, in addition, Cudrania also contains alkaloids, amino acids, steroids and glycosides [35]. The stems of Cudrania mainly contain triterpenoids and flavonoids. Systematic cluster analysis revealed that the transcriptional profiles of the roots and leaves of Cudrania could be clustered into a single unit, while the stems were clustered into a single unit, suggesting that the roots and

leaves may have similar active ingredients and certain medicinal values [36]. The stems may be clustered into a single unit because of their unique terpene components [36–37]. The analysis of the tissue-specific distribution mechanism of active ingredient synthesis in Cudrania can clarify the synthesis pathways and occurrence sites of various active ingredients, and lay a theoretical foundation for the research of high-quality cultivation technology system of Cardinia based on the synthesis law of active ingredients.

Zeatin is a member of the cytokinin family, a class of phytohormones involved in various processes of growth and development in plants [38]. Most abundant cytokinin are adenine-type, where the N6 position of adenine is substituted with an isoprenoid, such as in zeatin, or an aromatic side chain, such as in kinetin. Zeatin can be synthesized in two different pathways: the tRNA pathway and the AMP pathway. In the tRNA pathway zeatin is a recycled product of isopentenylated tRNAs. In the AMP pathway zeatin is synthesized from an isopentenyl donor, dimethylallyl diphosphate (DMAPP), and AMP, ADP, or ATP by isopentenyl transferases [39]. After synthesis cytokinin can be glucosylated. This biosynthetic pathway also differed significantly in roots, stems, and leaves of Cudrania, and the related genes (C00008, ADP), (C00020,AMP), (C00015,UDP), and (C00147, Adenine) were highly expressed only in leaves, and this gene regulates the antioxidant enzyme activity and stress tolerance of the plant in order to affect photosynthesis[40].

The biosynthesis of monoterpenoids in the differential genes were mainly (K15086,TPS14), (K21925,TPS3), (K07385,TPS-CIN), (K15095), etc. The main flavonoids and terpenoids in the stem of Cudrania are mainly flavonoids and terpenoids. This may also be the main active substances that distinguish the stems of Cudrania from the roots and leaves [41–42].

The formation of quality of Chinese herbal medicines has an adversity effect. Under environmental stress conditions, plants can induce a series of signaling molecules and regulate the expression of downstream genes, thus regulating the accumulation of active ingredients, secondary metabolites, in medicinal plants. In this study, we found the expression of some genes related to plant-pathogen interactions, MAPK cascade signaling and hormone signaling, indicating that the expression of these stress-responsive genes can promote the accumulation of secondary metabolites to a certain extent [43].

Conclusion

We detected a total of 1254 metabolites from three tissues of roots, stems and leaves of Tsuga caryophyllum, and all metabolites were annotated and classified into eight categories by the KEGG database: steroids, lipids, antibiotics, vitamins and cofactors, nucleic acids, peptides, carbohydrates, and organic acids. The major metabolites were DL-Arginine, L-Aspartic acid, 6-Methylnicotinamide, Trigonelline, D-Saccharic acid, D-(-)-Salicin, D-(+)-Mannose, Asparagine, 2- Oxoglutaric acid, Choline, a-Lactose, Parishin C, Parishin E, Sucrose, Plantagoside, D-(-)-Quinic acid, Citric acid, Fumaric acid, DL-Malic acid, DL -Stachydrine. The results of the present study showed that the accumulation of active components and transcriptional profiles of Tsugaru roots differed significantly from those of the stem and leaf tissues, and the differential genes were mainly enriched in the pathways of flavonoid

biosynthesis, flavonoids and flavonols biosynthesis, nucleotide metabolism, monoterpene biosynthesis, zeatin biosynthesis, and photosynthesis-antenna proteins. Among them, flavonoid biosynthesis-related genes were significantly up-regulated in roots and stems, and photosynthesis-haptoglobin-related genes were highly expressed in leaves. Among the flavonoid metabolic pathways we noted relatively high levels of naringenin, naringenin, geranylgeranylgeranyl, dihydroquercetin, (+)-catechin and dihydromyricetin metabolites in roots. The metabolites with higher content in stems were Phloretin, Phloritin, Pinocembin, Hesperetin, Neohesperidin, Luteolin, Quercetin, etc., while the metabolites with higher content in leaves were only Apigenin and Kaempferol. These results indicate that Cudrania is a promising medicinal plant and provide important information for the identification of differential metabolites and major metabolic pathways among different parts of Cudrania, as well as a theoretical reference for the effective use of Cudrania resources. In this study, the candidate genes obtained from the screening were not verified for gene function to further clarify their roles in the accumulation of active ingredients, and the next step will be to analyses the function of the candidate flavonoid synthesis genes, terpene synthesis genes, stress-responsive genes and other key enzyme genes through the genetic transformation system.

Declarations

Authors' contributions

Z.K.D and M.Z wrote the main manuscript text and prepared all the figures. All authors reviewed the manuscript. Z.K.D and X.Y.L conceived and designed the research. Z.K.D and X.M.W analyzed data and Z.S.Z reviewed. All authors read and approved the manuscript.

Funding

This work was supported the Zhejiang Province Basic public welfare research project (LGN21D060001), Huzhou Science and Technology Plan Project (2023YZ57) .

Data Availability

All data generated during the current study are included in this published article. The SNPs data used for GWAS analysis submitted to National Agricultural Biotechnology Information Center (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA838136) with the following accession number PRJNA838136.

Ethics approval and consent to participate

Experimental research and field studies on plants, including the collection of plant material, in this study comply with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Keun H-K, Sanghae N. Antioxidant activities of volatile aroma components from *Cudrania tricuspidata* (Carr.) bureau extracts. J. Korean Soc. Food Sci. Nutr. 2012;41(11):1493-1501.
2. Jeon Y, Oh J, Cho M-S. Formulation optimization of sucrose-free hard candy fortified with *Cudrania tricuspidata* extract. Foods. 2021;10(10):2464-2464.
3. Song JP, Zhang HY, Qu LN, Yang F, Liu SF, Wang HH, et al. Protective effects of cudrania tricuspidata polysaccharides and lycium barbarium polysacharides on the immune function of tail-suspended mice. Space Med. Med. Eng. 2007;06:402-405.
4. Zheng ZP, Liang JY Hu LH. Water-soluble constituents of *Cudrania tricuspidata* (Carr.) Bur.. J. Integr. Plant Biol. 2006;48(8):996-1000.
5. Shi L, Dong Q, Ding K. Structure elucidation and immunomodulatory activity in vitro of a xylan from roots of *Cudrania tricuspidata.* Food Chem. 2014;152:291-296.
6. Jung GT, Ju IO, Choi SR, You DH, No JJ. Food nutritional characteristics of fruit of *Cudrania tricuspidata* in its various mat. Korean J Food Prese. 2013;20(3):330-335.
7. Yin YZ, Chen RD, Zhang DW, Qiao LR, Li JH, Wang RS, et al. Regio-selective prenylation of flavonoids by plant cell suspension cultures of *Cudrania tricuspidata* and Morus alba. J. Mol. Catal. B: Enzym. 2013;89:28-34.
8. Kim SJ, Lee KT, Youe WJ, Lee SS, Kim YS. Structure Analysis of Water-soluble Polysaccharides Extracted from The Unripe Fruit of *Cudrania tricuspidate*. J. Korean Wood Sci. Technol. 2014;42(6):740~746.
9. Zhang J, Sun H, Jiang KX, Song X, Wang X, Yang Y, et al. Cudraxanthone L inhibits gastric cancer by regulating the MAPK signalling and promoting FAS-mediated pathway. Biomed. Pharmacother. 2021;141:111876.
10. Jeong JY, Jo YH, Lee KY, Do S-G, Hwang BY, Lee MK. Optimization of pancreatic lipase inhibition by *Cudrania tricuspidata* fruits using response surface methodology. Bioorg. Med. Chem. Lett. 2014;24:2329–2333.
11. Youn KS, Kim JW. Antioxidant and angiotensin converting enzyme I inhibitory activities of extracts from mulberry (*Cudrania tricuspidata*) fruit subjected to different drying methods. J. Korean Soc. Food Sci. Nutr. 2012;41(10):1388-1394.
12. Seo MJ, Kang BW, Park JU, Kim MJ, Lee HH, Kim NH, et al. Effect of fermented *Cudrania tricuspidata* fruit extracts on the generation of the cytokines in mouse spleen cells. J. Life Sci. 2013;23(5):682-688.

13. Dong GX, Chen KS, Shi L, Song JP, Xu YT. Immunological enhancement in vitro of polysaccharides from roots of *Cudrania tricuspidata*. Current Immunology, 2005;25:247-255.
14. Zou YS, Hou AJ, Zhu GF, Chen YF, Sun HD, Zhao QS.. Cytotoxic isoprenylated xanthones from *Cudrania tricuspidata*. Bioorg. Med. Chem. 2004;12:1947-1953.
15. Shi L, Chen KS, Dong Q, Fang JN, Ding K. Separation, purification and structure elucidation of a polysaccharide from root of *Cudrania tricuspidata*. Front. Chem. in China, 2008;3:209-212.
16. Yue C. Peng H, Li W, Tong Z, Wang Z, Yang, P. Untargeted metabolomics and transcriptomics reveal the mechanism of metabolite differences in spring tender shoots of tea plants of different ages. Foods 2022, 11, 2303.
17. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods. 2015;12(4):357-360.
18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
19. Mao XZ, Cai T, Olyarchuk JG, Wei LP, Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics. 2005;21(19):3787-3793.
20. Markus G, Thomas H, Mari LB, Carlo R, Barbara F, Robert K, et al. Redirection of flavonoid biosynthesis through the downregulation of an anthocyanidin glucosyltransferase in ripening strawberry fruit. Plant Physiol. 2008;146(4):1528-1539.
21. Park S, Kim DH, Lee JY, Ha SH, Lim SH. Comparative analysis of two flavonol synthases from different-colored onions provides insight into flavonoid biosynthesis. J. Agric. Food Chem. 2017;65(26):5287-5298.
22. Wu Y, Guo J, Zhou Q, Xin Y, Wang G, Xu L. De novo transcriptome analysis revealed genes involved in flavonoid biosynthesis, transport and regulation in Ginkgo biloba. Ind. Crops Prod. 2018;124:226-235.
23. Moriguchi T, Kita M, Ogawa K, Tomono Y, Endo T, Omura M, Flavonol synthase gene expression during citrus fruit development. Physiol. Plantarum. 2002;114(2):251-258.
24. Thomas V. Phenylpropanoid biosynthesis. Mol. Plant. 2010;3:2–20.
25. Loke KK, Rahnamaie-Tajadod R, Yeoh CC, Goh HH, Noor NM. Transcriptome analysis of Polygonum minus reveals candidate genes involved in important secondary metabolic pathways of phenylpropanoids and flavonoids. Peerj. 2017;5:2938.
26. Buer CS, Imin N, Djordjevic MA. Flavonoids: new roles for old molecules. J. Integr. Plant Biol. 2010;52:98-111.
27. Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, et al. Flavonoids act as negative regulators of auxin transport in vivo in Arabidopsis. Plant Physiol. 2001;126:524-535.
28. Saito K, Matsuda F. Metabolomics for functional genomics, systems biology, and biotechnology. Annu. Rev. Plant Biol. 2010;61:463-489.

29. Zhang Y, Zhu Q, Zhang M, Guo Z, Yang J, Mo J, et al. Individual Cryptomeria fortunei Hooibrenk clones show varying degrees of chilling stress resistance. Forests. 2020;11:189.
30. Yamamoto Y, Aminaka R, Yoshioka M, Khatoon M, Komayama K, Takenaka D, et al. Quality control of photosystem II: impact of light and heat stresses. Photosynth. Res. 2008;98:589-608.
31. Liu WJ, Yuan S, Zhang NH, Lei T, Duan HG, Liang HG, et al. Effect ofwater stress on photosystem 2 in two wheat cultivars. Biol. Plantarum 2006;50:597-602.
32. Yoshioka M, Uchida S, Mori H, Komayama K, Ohira S, Morita N, et al. Quality control of photosystem II. Cleavage of reaction center D1 protein in spinach thylakoids by FtsH protease under moderate heat stress. J. Biol. Chem. 2006;281:21660-21669.
33. Chen L, Zhou Q, Li B, Liu SJ, Dong JX. A new flavonoid from Cudrania cochinchinensis. Nat. Prod. Res. 2015;29(13):1217-1221.
34. Jo YH, Kim SB, Liu Q, Lee JW, Hwang BY, Lee MK. Benzylated and prenylated flavonoids from the root barks of *Cudrania tricuspidata* with pancreatic lipase inhibitory activity. Bioorg. Med. Chem. Lett. 2015;25(17):3455-3457.
35. Song J, Xu K, Liu C, Wang T, Luan X, Zhu L, et al. Bioactive specialised metabolites from the endophytic fungus Xylaria sp. of Cudrania tricuspidate. Phytochemistry. 2022;196:113079.
36. Liu Z, Wei W, Zhou M, Gan C, Liu S. A New Prenylated Xanthone from Root Barks of Cudrania cochinchinensis. Chin. Herb. Med. 2013;5(2):101-103.
37. Zhang P, Feng Z, Wang Y. Flavonoids, including an unusual flavonoid-Mg2+ salt, from roots of Cudrania cochinchinensis. Phytochemistry. 2005;66(23):2759-2765.
38. Alejandro C-C, Fátima C, Sonia O, Manuel J, Dolores G, Francisco P. Postharvest abscisic acid treatment modulates the primary metabolism and the biosynthesis of t-zeatin and riboflavin in zucchini fruit exposed to chilling stress. Postharvest Biol. Technol. 2023;204:11257.
39. Eriko S, Takehiko O, Kentaro T, Mikiko K, Nobutaka K, Hitoshi S, et al. Uniconazole, a cytochrome P450 inhibitor, inhibits trans-zeatin biosynthesis in Arabidopsis. Phytochemistry. 2013;87:30-38.
40. Li Z, Wu Y, Hu J, Yang G, Wang Z, Sun J. Dissection of the response mechanism of alfalfa under phosphite stress based on metabolomic and transcriptomic data. Plant Physiol. Biochem. 2022;192:35-49.
41. Jiang H, Wang X. Biosynthesis of monoterpenoid and sesquiterpenoid as natural flavors and fragrances. Biotechnol. Adv. 2023;65:108151.
42. Hallahan DL. Monoterpenoid biosynthesis in glandular trichomes of labiate plants. Adv. Bot. Res. 2000;31:77-120.
43. Shambhavi S, Susmita S, Tuhin D, Ashish P, Manoj P. Recent perspective of non-coding RNAs at the nexus of plant-pathogen interaction. Plant Physiol. Biochem. 2023;201:107852.

Figures

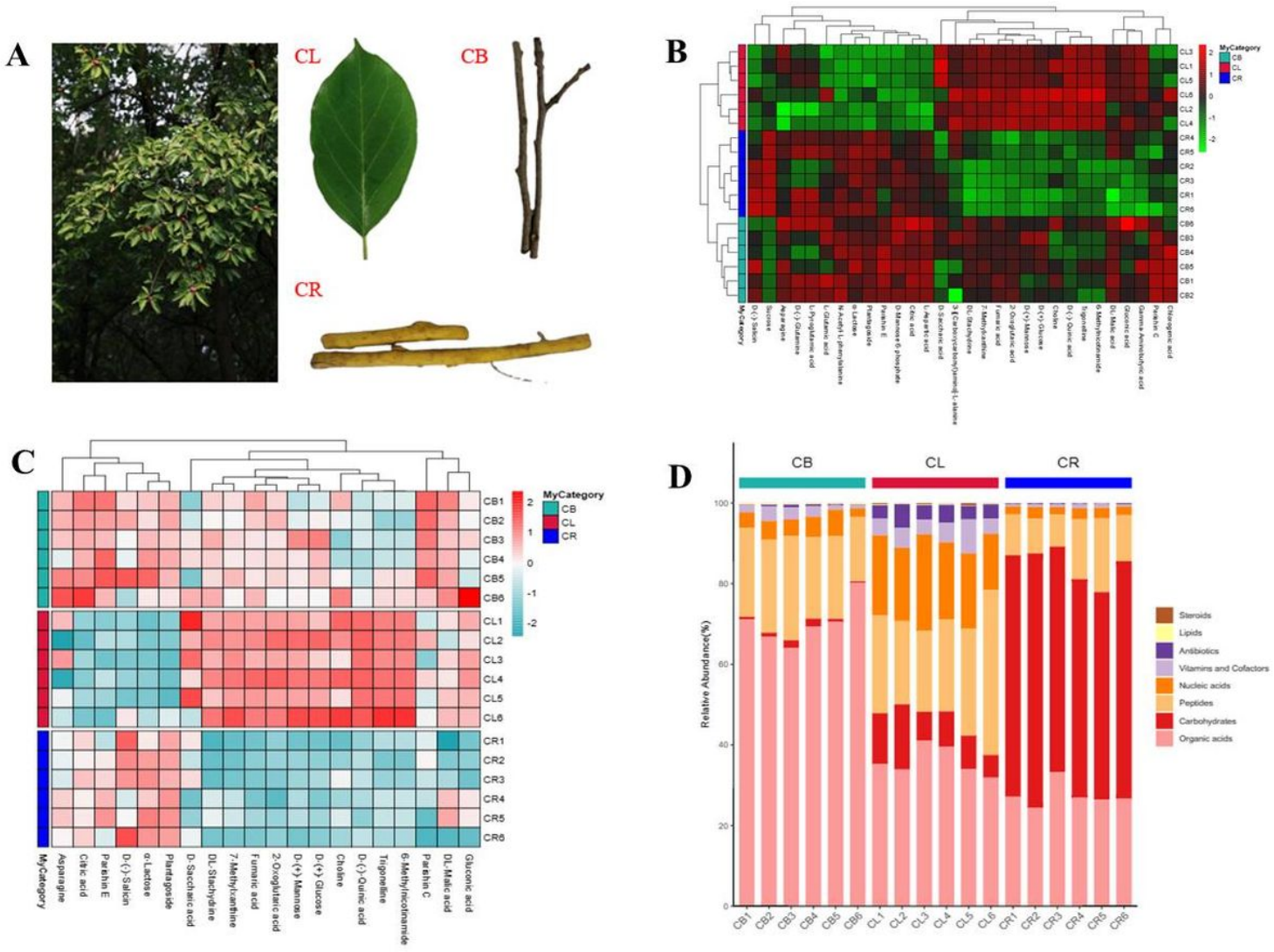


Figure 1

Different tissues of roots, stems and leaves and their associated metabolites in Cudrania plants (A) Different tissues of roots, stems and leaves of Cudrania (CL for leaves, CB for stems and CR for roots) (B) Percentage stacking bar chart of metabolites that play a biological role (C) Metabolite heat map clustering (D) Heat map of the top 20 metabolites

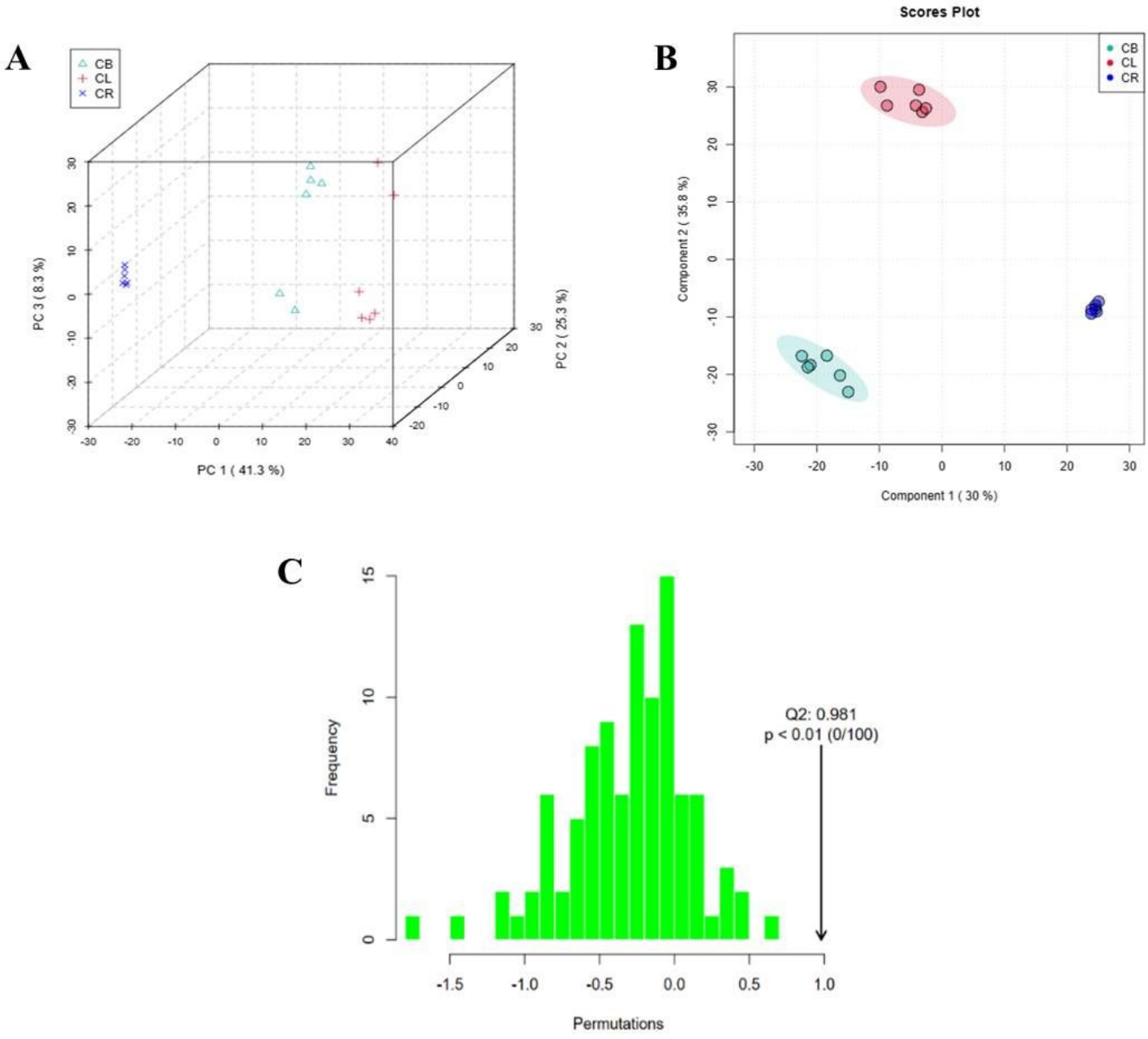


Figure 2

PCA analysis and OPLS-DA test (A) PCA (B) PLS-DA point cloud diagram (C) Distribution of the test statistic (Q2) for the OPLS-DA permutation test and the p-value



Figure 3

Metabolite Differences Box Plot

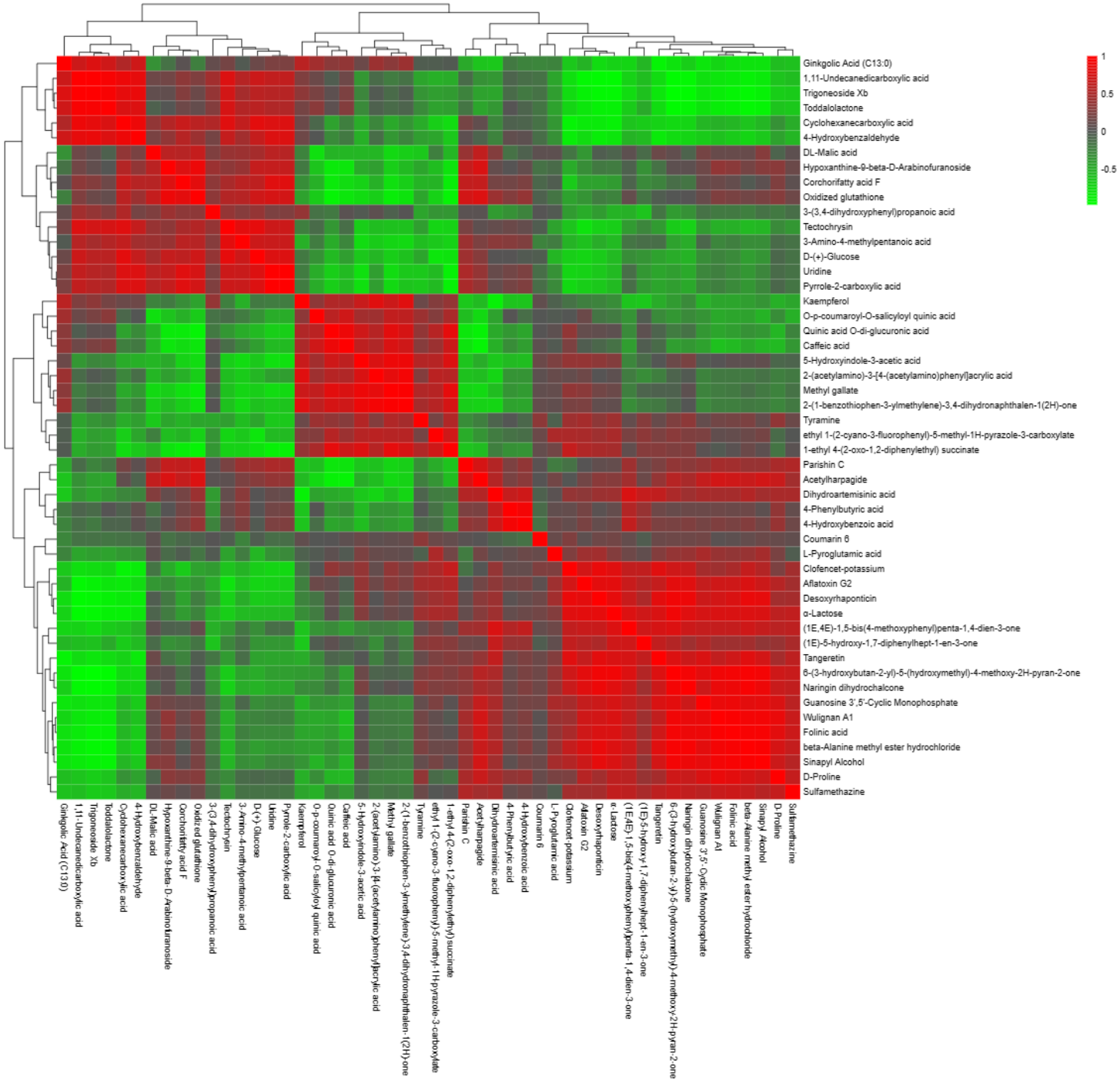


Figure 4

Pearson correlation heat map (Matrix of correlation coefficients for the top 100 metabolites in total, with correlation coefficients indicated by colour, red positive and green negative.)

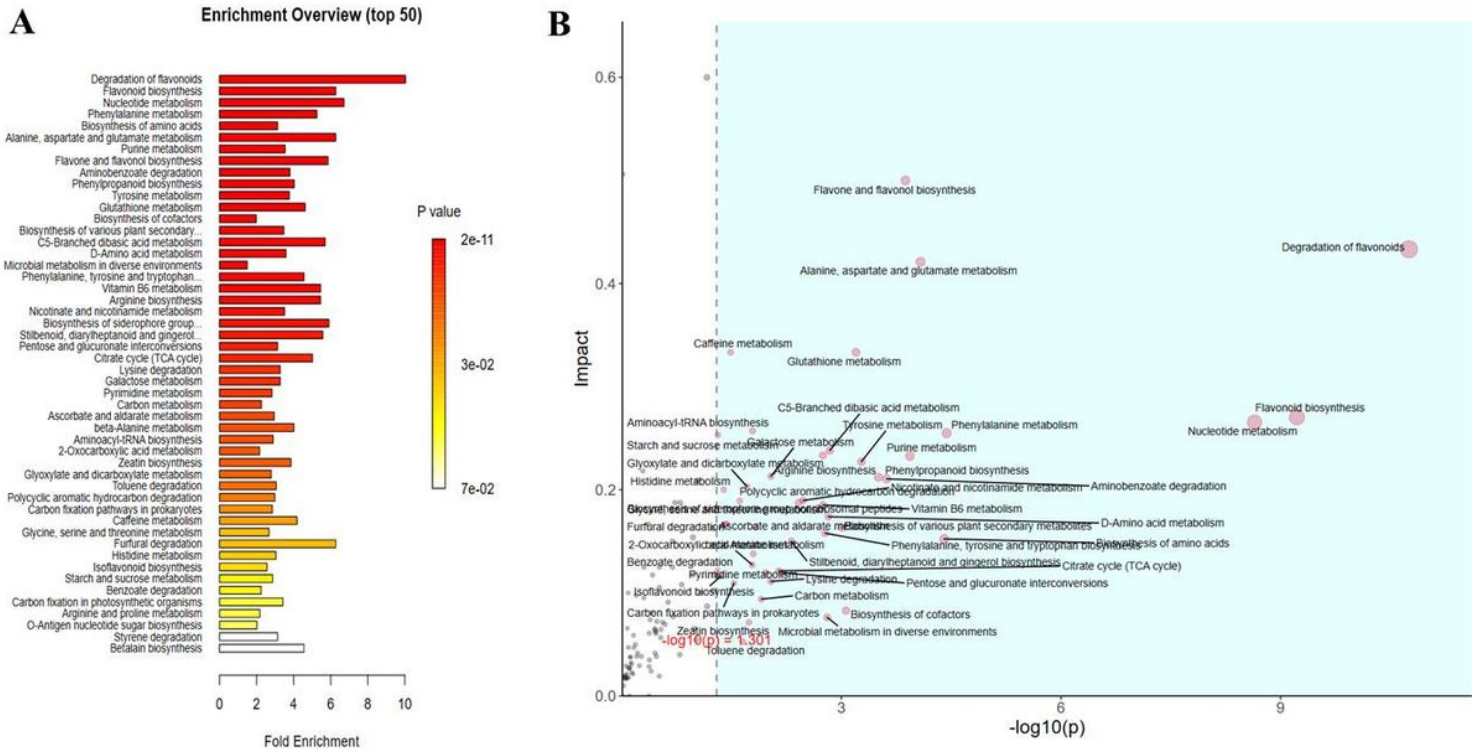


Figure 5

Enrichment analysis and topology analysis (A) ORA enrichment analysis (The horizontal coordinate is the fold enrichment, which is the number of observed metabolites/theoretical metabolites in the metabolic pathway. p-value magnitude is indicated by the colour; the darker the colour, the smaller the p-value.) (B) ORA enrichment analysis and topology analysis(Horizontal coordinates are ORA analysis p-values, blue areas are significant (p<0.05); vertical coordinates are topological analysis Impact.)

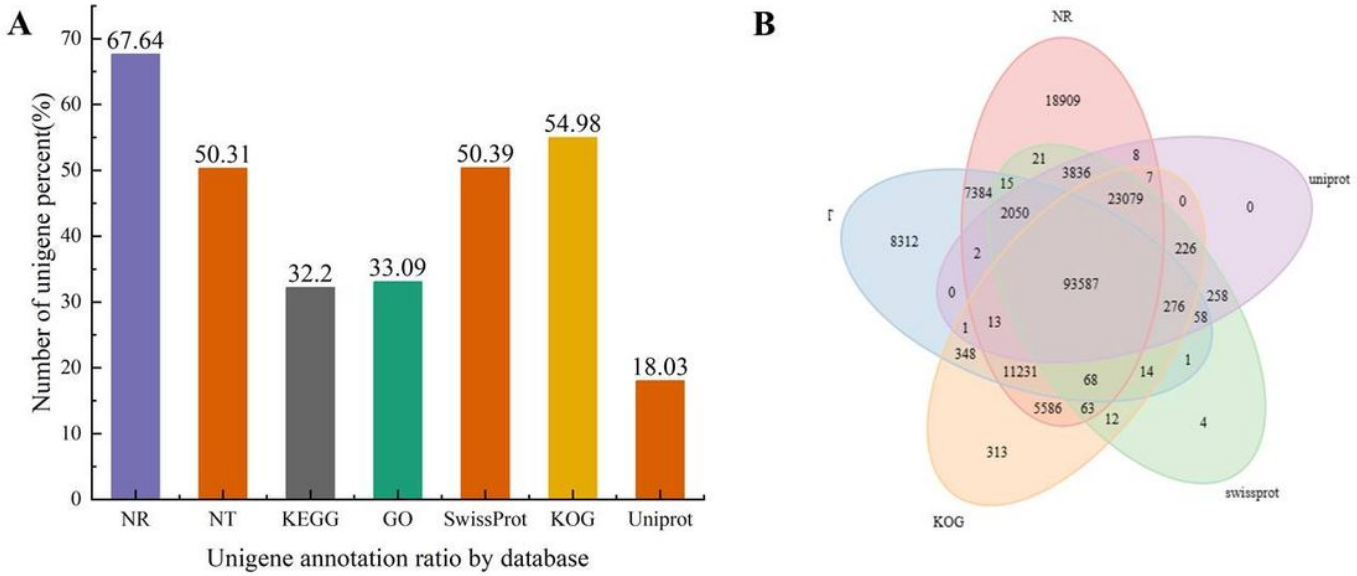


Figure 6

Proportion of Unigene annotations and results for functional databases (A) Unigene annotation ratio by database (B) Venn diagram of annotation results of 5 Unigene databases

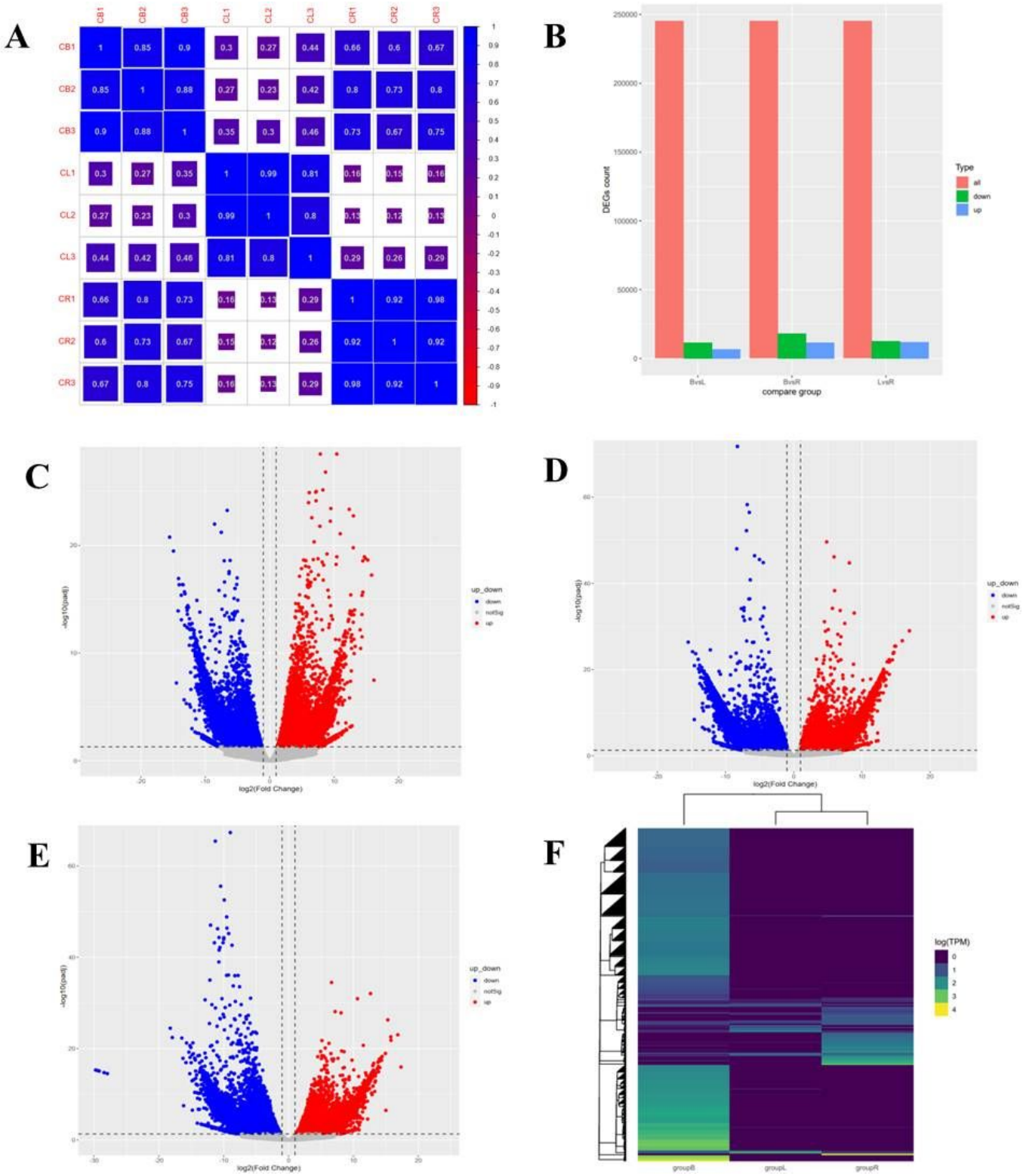


Figure 7

Correlation coefficients between samples and differences in genes between combinations (A) Heat map of correlation coefficients between samples (B) Statistical histogram of the number of differential genes in the differential comparison combinations (C) Stem(CB) and Leaf(CL) Differential Gene Volcano Maps (D) Stem(CB) and Root(CR) differential gene volcano maps (E) Leaf(CL)and Root(CR) differential gene volcano maps (F) Heatmap for clustering differentially expressed genes

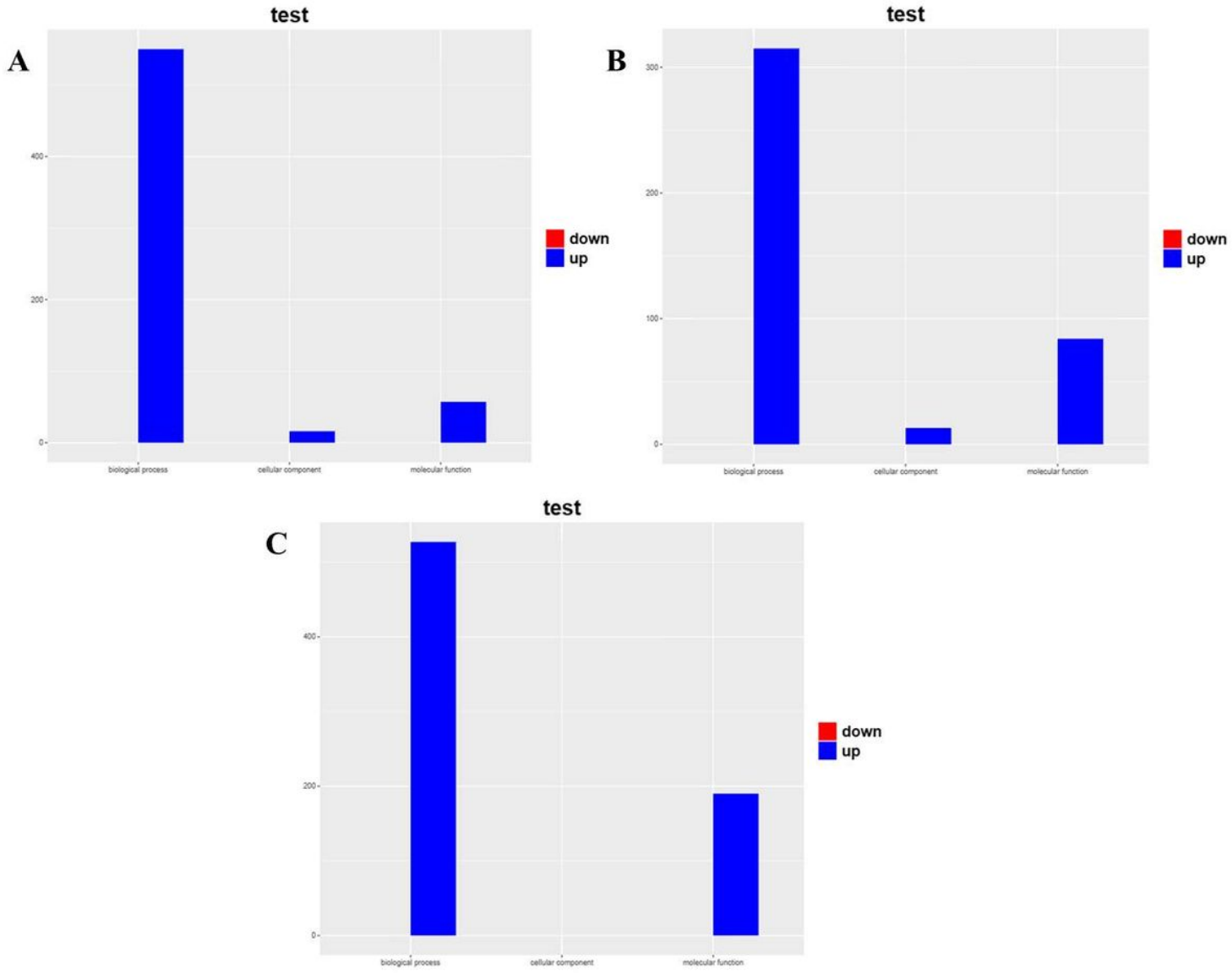


Figure 8

Enrichment histogram of differential genes (A) Stem and leaf differential gene GO enrichment histograms (B) Stem and Root differential gene GO enrichment histograms (C) Leaf and Root differential gene GO enrichment histograms

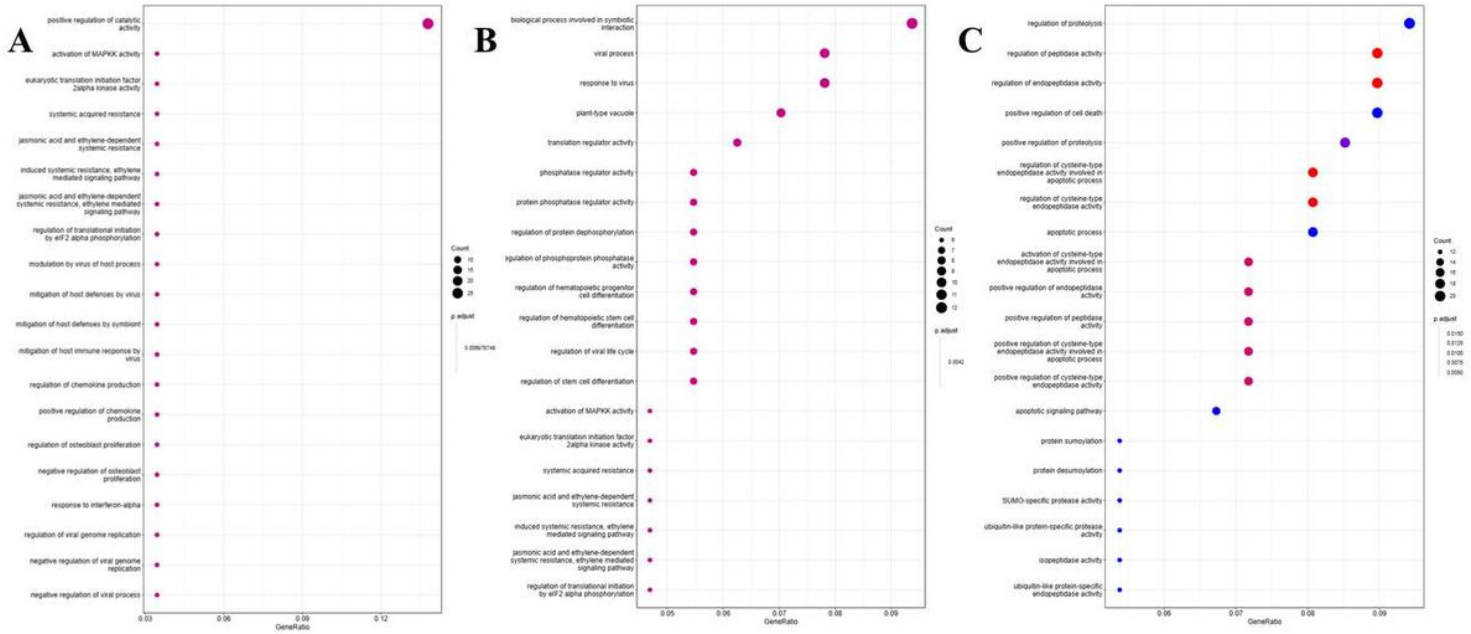


Figure 9

GO enrichment analysis of differential genes (A) GO enrichment analysis of differentially expressed genes in stems and leaves (B) GO enrichment analysis of differentially expressed genes in stems and roots (C) GO enrichment analysis of differentially expressed genes in leaves and roots

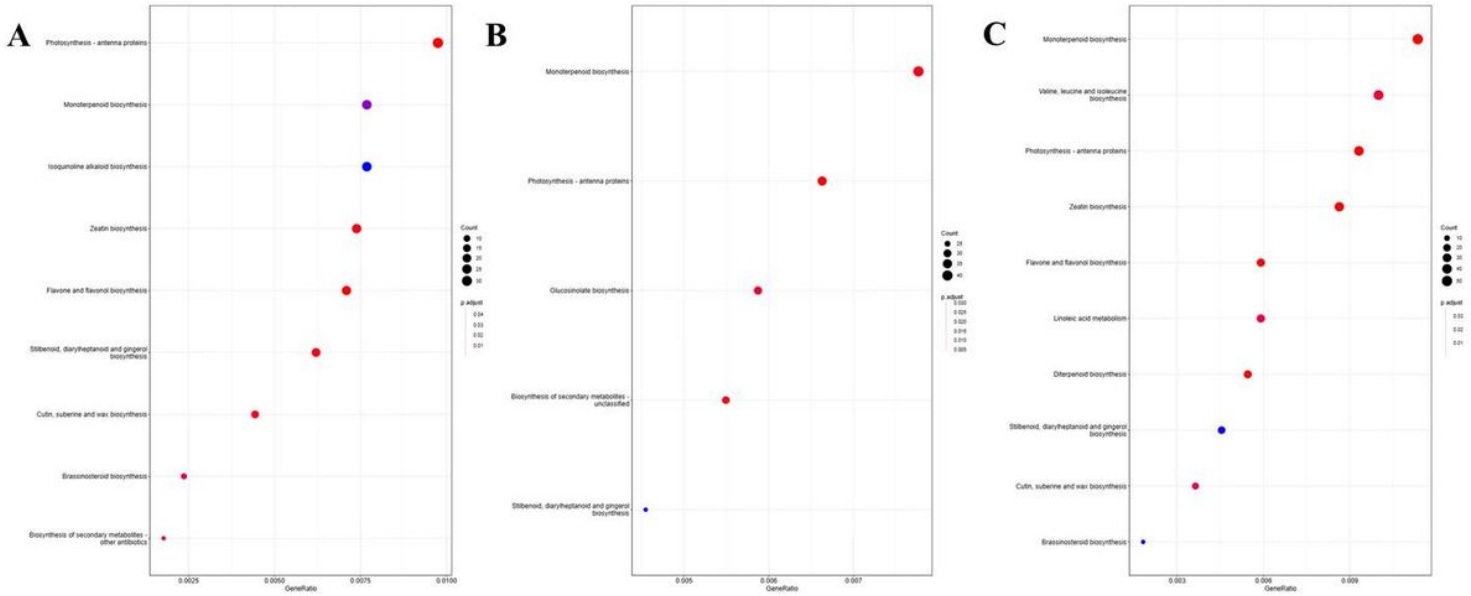


Figure 10

KEGG enrichment analysis of differentially expressed genes (A) KEGG enrichment analysis of differentially expressed genes in stems and leaves (B) KEGG enrichment analysis of differentially expressed genes in stems and roots (C) KEGG enrichment analysis of differentially expressed genes in leaves and roots

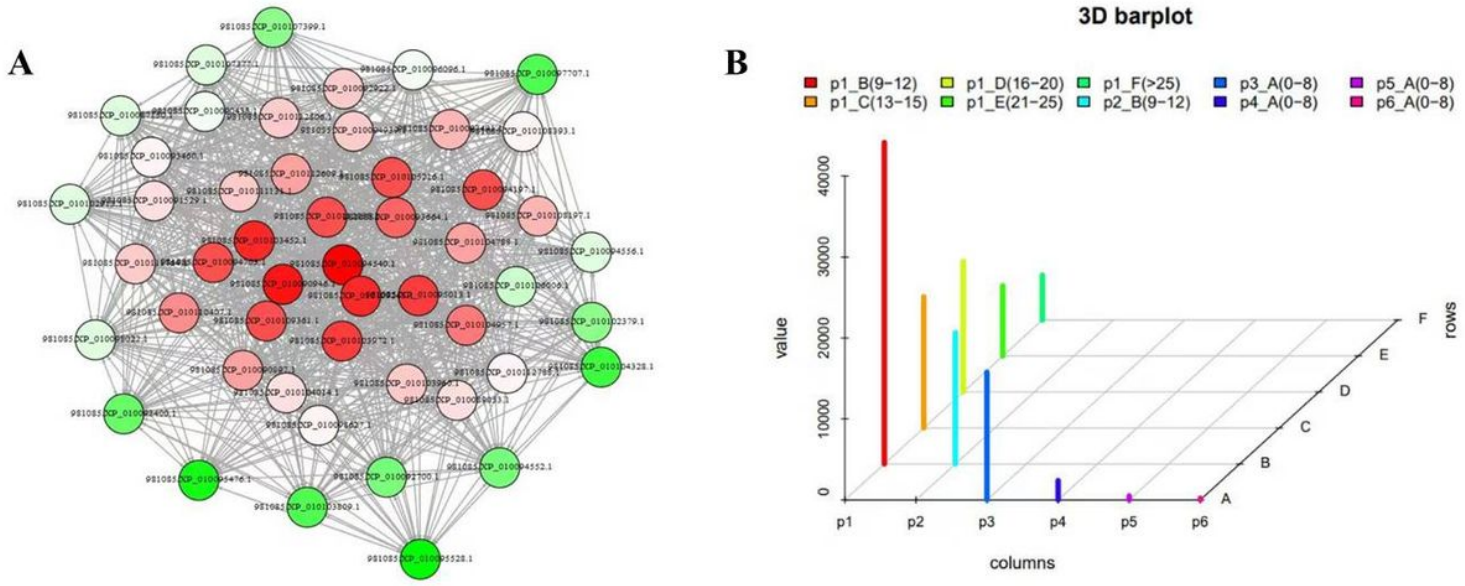


Figure 11

Differential Gene Protein Network Interaction Map (A) and Distribution of SSR species (B)

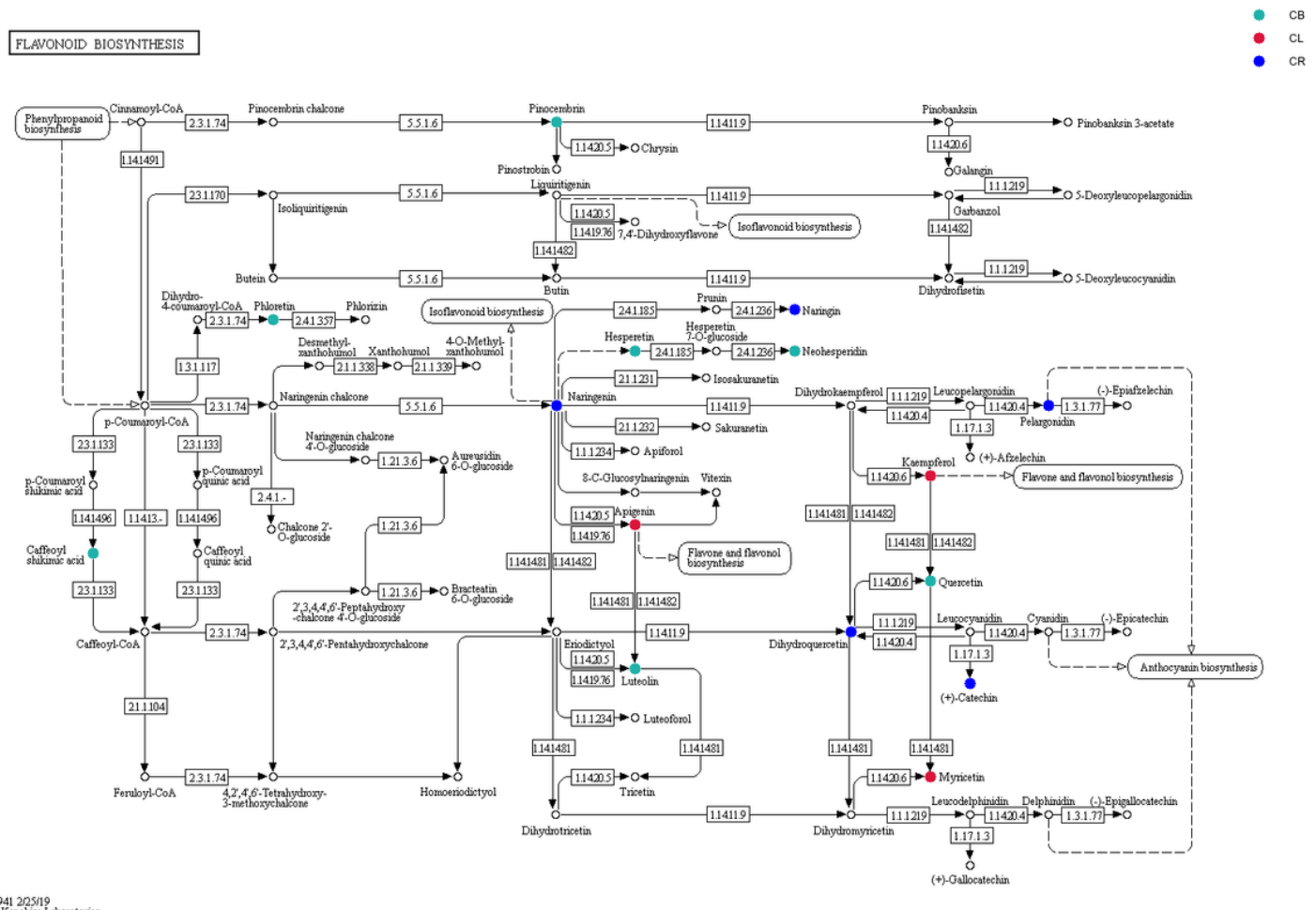


Figure 12

Flavonoid biosynthesis - Reference pathway

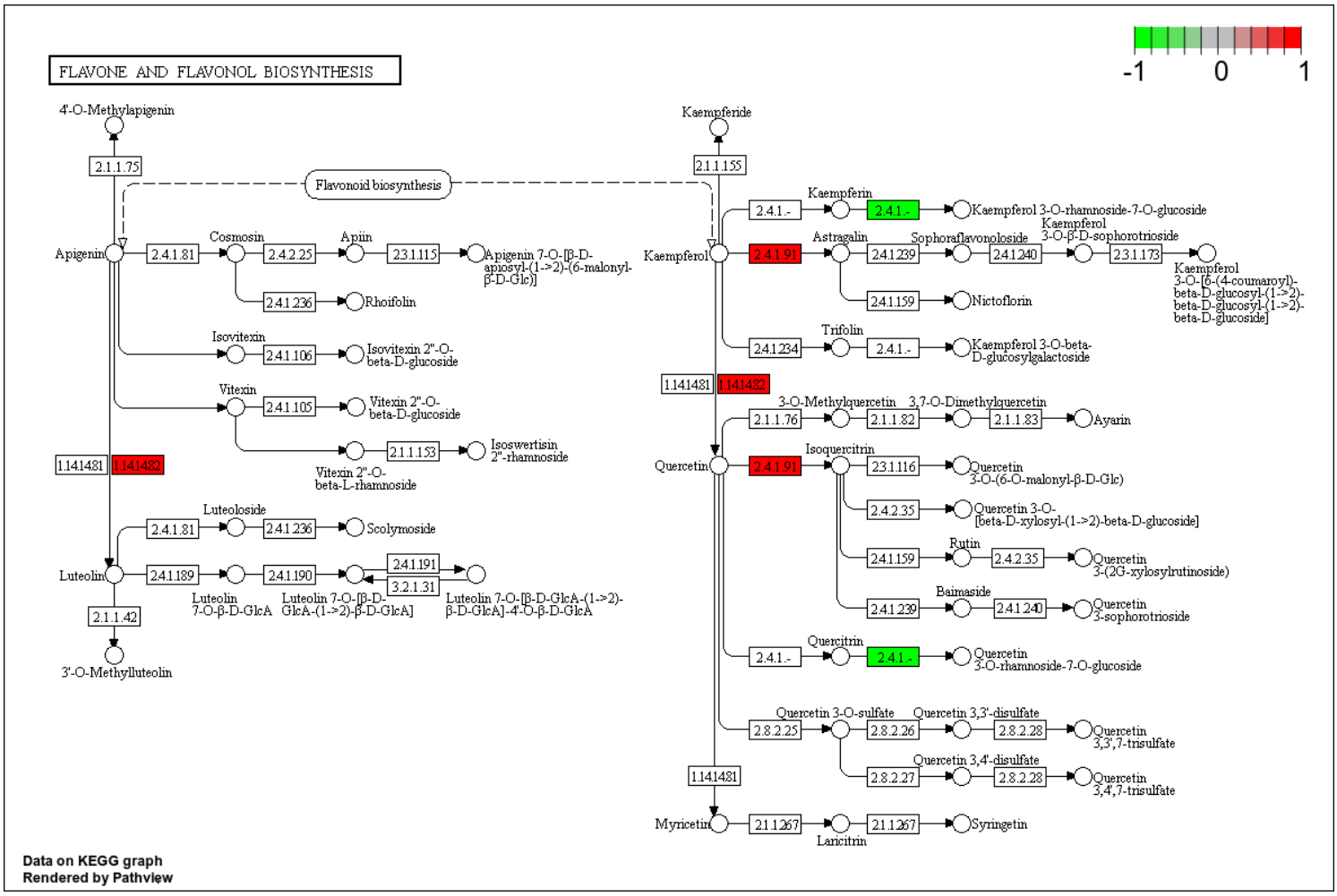


Figure 13

Flavone and flavonol biosynthesis- Reference pathway

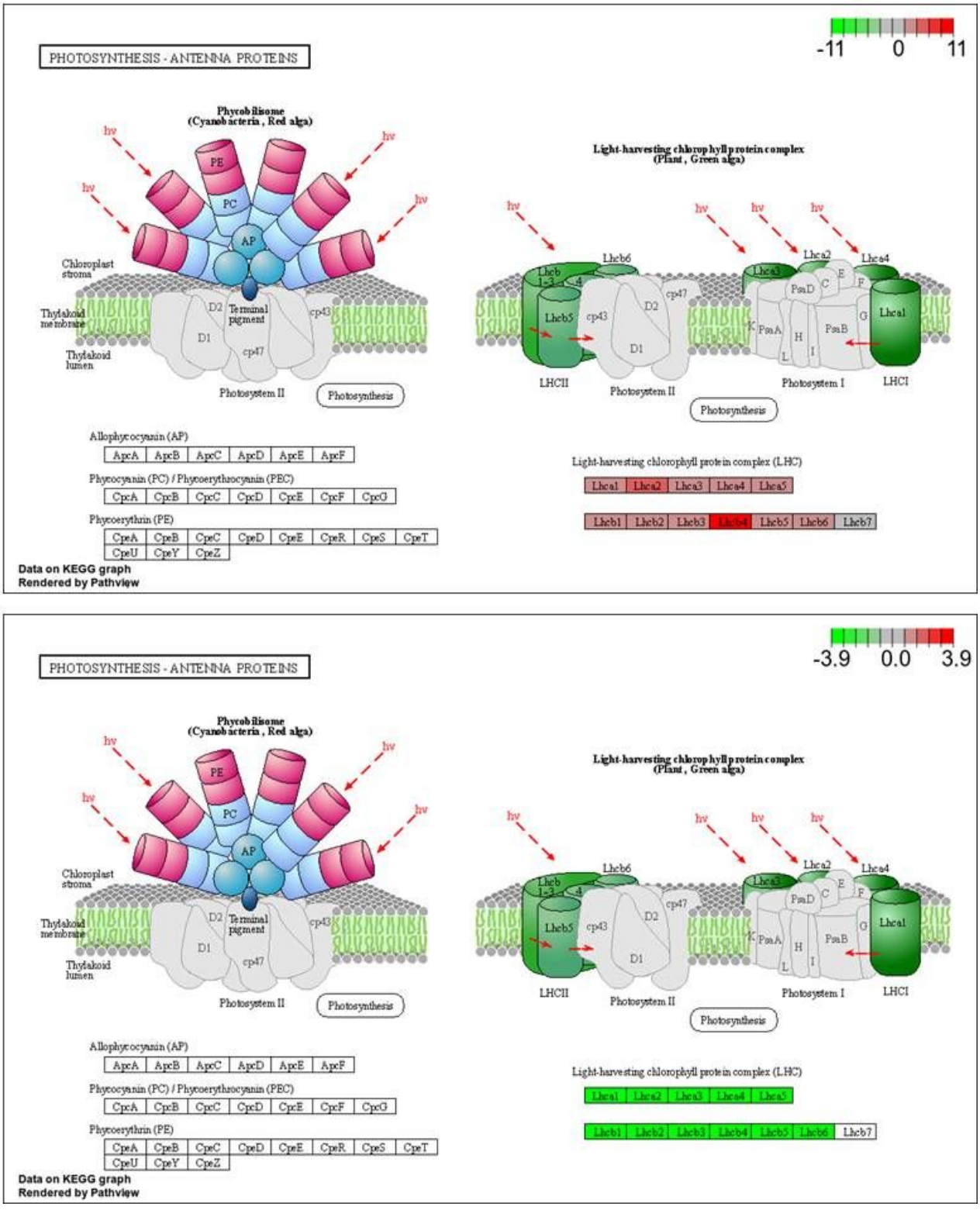


Figure 14

Photosynthesis-antenna protein biosynthesis pathway