Mechanistic analysis of the hardening process of the thorns on stems of Bougainvillea spectabilis Willdenow

Lina Sun
Guangxi Zhuang Autonomous Region Forestry Research Institute
Xinhua Wang
Guangxi Zhuang Autonomous Region Forestry Research Institute
Jinhua Li
Guangxi Zhuang Autonomous Region Forestry Research Institute
Jianying Gong
Guangxi Zhuang Autonomous Region Forestry Research Institute
Shuting Yang
Guangxi Zhuang Autonomous Region Forestry Research Institute
Kaitai Yang
Guangxi Zhuang Autonomous Region Forestry Research Institute
Er Chen
Guangxi Zhuang Autonomous Region Forestry Research Institute
Bing Li
Guangxi Zhuang Autonomous Region Forestry Research Institute
Zhixiang Lu
Shanghai Jiaotong University
Qi Chen
Nanning GoldTech Biotechnology Ltd
Lin Mao (49888178@qq.com)
Guangxi Zhuang Autonomous Region Forestry Research Institute

Research Article

Keywords:

Posted Date: August 31st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3268556/v1

Additional Declarations: No competing interests reported.

Abstract

Background

Bougainvillea spectabilis Willdenow is thorny woody vine or shrub. The rigidity of the thorns on the stems should be considered as a horticultural character. In order to find the genes and pathways related to the hardening process of the thorns on the stems of *B. spectabilis*, the eukaryotic unreferenced transcriptome sequencing analysis is applied to explore the 3 stages of the thorns hardening process.

Results

This study investigates the transcriptomic changes in *B. spectabilis* plants during the process of thorns hardening. 3 developmental stages from thorns formation (stage 1) to thorns hardening (stage 2 to stage 3) were examined. Total RNA was extracted from thorns and stems, and transcriptome libraries were constructed and sequenced using unreferenced Illumina sequencing. Gene function annotation was performed using various databases, resulting in 8937 co-annotated genes. The density distribution of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) depicted the overall gene expression patterns. Gene expression correlation analysis confirmed the reliability of the experiment, showing strong similarity among biological replicates. Differential expression analysis revealed that during thorns hardening, 1045 genes significantly up-regulated in thorns and 918 in stems at stage 2 compared to thorns formation (stage 1). At stage 3, as thorns became harder, 98 genes exhibited notable expression increase within thorns, and 46 genes up-regulated in stems, compared to stage 2. These findings highlight stage 2 as the period of highest gene expression activity during the thorns hardening process in *B. spectabilis*. Phenylpropanoid biosynthesis is a key step in the hardening process of the thorns of *B. spectabilis*. This transcriptome analysis offers insights into the molecular mechanisms underlying thorns development in this plant species.

Conclusion

The formation and hardening of thorns on the stem of *B. spectabilis* is a process in which lignin gradually accumulates in the thorns, and several genes are involved in the process. The phenylalanine ammonia-lyase, trans-cinnamate 4-monooxygenase, reductase4-coumarate-CoA ligase, cinnamoyl-CoA reductase, cinnamyl-alcohol dehydrogenase and peroxidase are the key genes for lignin synthesis and accumulation. The process involves the pathways-phenylpropanoid biosynthesis.

Background

Bougainvillea spectabilis Willdenow is an evergreen climbing shrub of the Centrospermae (Order), Nyctaginaceae (Family) [1]. It is native to South America, such as Brazil, Peru, Argentina's Chubut Province[2] and Bolivia, Ecuador, Paraguay[3]. It is widely cultivated in tropical and subtropical regions in

China. The introduction of *B. spectabilis* in China boasts a history spanning 100 years. Presently, in the tropical and subtropical regions of China, such as Guangdong, Hainan, Guangxi, Fujian, and Yunnan Province, *B. spectabilis* has achieved extensive cultivation[4]. *B. spectabilis* demonstrates remarkable adaptability and ease of cultivation, and colorful flowers. As such, it finds extensive utility within the realm of landscaping and ornamental horticulture. In the Guangxi region, flowers of *B. spectabilis* are observable year-round, exhibiting an extended and robust flowering period. Guangxi Zhuang Autonomous Region Forestry Research Institute introduce many germplasm resources of *B. spectabilis* from all over the country[5].

B. spectabilis grows as a woody vine or shrub with thorny stems. In the early stage of formation, the thorns on the stems of *B. spectabilis* are green in color and soft in texture. After that, the color gradually deepens and eventually becomes dark brown, and the texture becomes very hard. Ornamental evaluation of *B. spectabilis* involves multiple target traits[5].We think the rigidity of the thorns on the stems should be considered as a target trait to evaluate the horticultural character. We focus on the hardening process of the thorns on the stems of *B. spectabilis* with the aim of identifying the genes that regulate this process. This research will provide a scientific basis for the cultivation of soft-thorns or thorns-free *B. spectabilis*.

In this research, we searched for genes and pathways related to the hardening process of the thorns on the stems of *B. spectabilis* through eukaryotic unreferenced transcriptome sequencing analysis, because complete genome information of *B. spectabilis* is lacking in existing online databases.

Results

Overview of Transcriptome Analysis

The plants of *B. spectabilis* at 3 stages from thorns formation (stage 1) to thorns hardening (stage 2 to stage 3) (Fig. 1) are used for research, and the total RNA of thorns and stems was extracted. Three biological replicates of samples in each stage were used to construct transcriptome libraries and sequenced using the Illumina platform. The raw reads we obtained from 18 libraries ranged from 41266290 to 49987188. The percentage of bases (Q30(%)) with a base calling accuracy of over 99.9% for each treatment was over 90% (Table. 1).

The transcript sequence assembled by Trinity was used as the reference sequence for subsequent analysis. After hierarchical clustering by Corset, the longest Cluster sequence was obtained for subsequent analysis. The lengths of transcripts and clustered sequences were counted separately, and the results are shown in Fig. 2.

Gene function is annotated based on the NCBI, Pfam, KOG/COG, Swiss-Prot, KEGG Ortholog database and Gene Ontology. A total of 8937 genes are co-annotated by these databases (Fig. 3).

The FPKM density distribution reflects the gene expression pattern of each sample as a whole. The graph shows a non-standard normal distribution, and the area area is 1, which means that the sum of the probabilities is 1. The peak of the density distribution curve represents the largest number of genes at the expression level (Fig. 4).

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. Before performing differential expression analysis, the correlation of gene expression levels between samples should be checked. The Pearson correlation coefficient is used to represent the correlation of gene expression levels between samples, and the closer the correlation coefficient is to 1, the higher the similarity of expression patterns between samples is. A correlation coefficient between 0.8-1 is a very strong correlation. If the correlation coefficient between samples of biological repeats is lower than 0.8, it means that the correlation between samples is low. The correlation between pairwise comparisons of the three biological replicates at the same site in the same period is above 0.8, indicating that their respective gene expression levels are similar.

Differentially expressed gene analysis

Comparing to the thorns formation stage (stage 1), at the stage of thorns turning hard (stage 2), total 1045 genes significantly up-regulate expression in thorns (Fig. 5A), 918 genes significantly up-regulate expression in stems (Fig. 5B). As the thorns become hard and turn brown (stage 3), a total of 98 genes exhibit a noteworthy increase in expression within the thorns (Fig. 5A), while 46 genes show a significant up-regulation in expression within the stems, comparing with the stage 2 (Fig. 5B). The results indicate that stage 2 is the most active period of gene expression during the thorns hardening process of *B. spectabilis*.

GO enrichment analysis

Active gene expression means vigorous biosynthesis and metabolism in the organism. According to GO enrichment analysis, at the stage of thorns turning hard, the most gene-enriched biological process in thorns (C2) is metabolism, comparing with the thorns formation early stage (C1). At the same time, the molecular function exhibiting the highest degree of gene enrichment is catalytic activity (Fig. 6A). The degree of gene enrichment in stems sample is equivalent to that observed in thorns sample (Fig. 6B). Through the comparison of the gene enrichment analysis of the stage 3 and the stage 2, it is found that the genes in thorns (C3 VS C2) are mostly enriched in the biological process of metabolism and the molecular function of oxidoreductase activity, as the plants at the stage 2 (Fig. 6C). The level of gene enrichment in the stems (J3 VS J2) sample is comparable to what was noted in the thorns sample (Fig. 6D).

KEGG enrichment and RT-qPCR analysis and determination of lignin

Through KEGG enrichment analysis, we found that phenylpropanoid biosynthesis is a key step in the hardening process of the thorns of *B. spectabilis*. From stage 1 to stage 3, whether in thorns or stems, the most significantly enriched pathway of up-regulated genes is the phenylpropanoid biosynthesis (Fig. 7). From stage 1 to stage 2, many genes in phenylpropanoid biosynthesis up-regulated remarkably (Fig. 8). The above results are also verified by RT-qPCR. It indicates that PAL (EC:4.3.1.24, K10775, phenylalanine ammonia-lyase) (Fig. 9A), CYP73A (EC:1.14.14.91, K00487, trans-cinnamate 4-monooxygenase) (Fig. 9B), 4CL (EC:6.2.1.12, K01904, 4-coumarate-CoA ligase) (Fig. 9C), CCR (EC:1.2.1.44, K09753, cinnamoyl-CoA reductase) (Fig. 9D), CAD (EC:1.1.1.195, K00083, cinnamyl-alcohol dehydrogenase) (Fig. 9E) and peroxidase (EC:1.11.1.7, K00430) (Fig. 9F) are the key genes for phenylpropanoid biosynthesis pathway. The expression levels of these genes exhibited a notable up-regulation during the 2nd stage comparing with the 1st stage. The phenylpropanoid biosynthesis is the key step of lignin biosynthesis. The determination of the lignin content in thorns and stems at 3 stages also agrees with the results of transcriptome analysis and RT-qPCR (Fig. 10). To sum up, Phenylpropanoid biosynthesis is a key step in the hardening process of the thorns of *B. spectabilis*.

Discussion

From 1st stage to 3rd stage, a continuous accumulation of lignin in thorns of *B. spectabilis* is observed. It is confirmed that the hardening process of the thorns on the stems of *B. spectabilis* is the process of gradual accumulation of lignin.

Lignin constitutes a polymer resulting from the intricate polymerization of distinct lignin monomers. The process of lignin biosynthesis is orchestrated through the integration of the phenylalanine metabolic pathway and the dedicated lignin-specific pathway. The synthesis of lignin from phenylalanine can be divided into four steps. First, phenylalanine forms p-coumaric acid under the catalysis of phenylalanine ammonia-lyase (PAL) and cinnamic acid 4-hydroxylase (C4H). Next, p-coumaric acid forms caffeic acid under the catalysis of p-coumarate 3-hydroxylase (C3H), caffeic acid forms ferulic acid under the catalysis of caffeic acid O-methyltransferase (COMT), and ferulic acid forms ferulic acid under the catalysis of feruLate-5-hydroxylase (F5H) and caffeic acid O-methyltransferase (COMT) catalyzed to form 5-hydroxy-ferulic acid and snapic acid. Subsequently, the phenolic acids previously elucidated undergo a catalytic transformation facilitated by a sequence of enzymes. Commencing with 4-coumarate CoA ligase (4CL), the process advances through cinnamoyl-CoA reductase (CCR) and culminates in cinnamyl alcohol dehydrogenase/sinapyl alcohol dehydrogenase (CAD/SAD) mediated reactions. These orchestrated enzymatic steps culminate in the synthesis of distinctive compounds, namely p-coumaryl alcohol, caffeoyl alcohol, coniferyl alcohol, 5-hydroxy-coniferyl alcohol, and sinapyl alcohol. Ultimately, the trimeric units of lignin—namely, p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol—underwent a process of polymerization catalyzed by peroxidase (POD/PER/PRX) or laccase (LAC). This orchestrated transformation yielded three distinct high-molecular-weight lignin polymers: p-hydroxyphenyl lignin, guaiacyl lignin, and syringyl lignin[6]. From the formation to turning hardened of thorns on the stems of *B. spectabilis*, the genes encode PAL, CCR, 4CL, CAD and peroxidase are up-regulated during this process. At the same time, lignin gradually accumulates in thorns and stems. It is confirmed that the hardening

process of the thorns on stems of *B. spectabilis* depends on the lignin accumulation. The lignin accumulation starts from phenylpropanoid biosynthesis and it relies on the phenylpropanoid biosynthesis pathway.

When it comes to the regulation of lignin synthesis, current research mainly focuses on two types of transcription factors, NAC and MYB. In *Arabidopsis thaliana*, a three-level regulatory network composed of NAC and MYB jointly regulates the synthesis of lignin, cellulose and hemicellulose during secondary cell wall thickening in *A. thaliana*[7]. In addition to NAC and MYB transcription factors, lignin synthesis is also regulated by WRKY, bHLH, LBD, LIM and microRNA, etc[8]. Based on the transcriptome analysis of the 3 stages, we can further screen the regulatory factors that regulate the accumulation of lignin in thorns, in order to provide more scientific basis for the cultivation of soft thorns or thorn-free *B. spectabilis*.

Conclusions

Phenylpropanoid biosynthesis is a key step in the hardening process of the thorns of *B. spectabilis*. The formation and hardening of thorns on the stem of *B. spectabilis* is a process in which lignin gradually accumulates in the thorns, and several genes are involved in the process. The phenylalanine ammonia-lyase, trans-cinnamate 4-monooxygenase, reductase4-coumarate-CoA ligase, cinnamoyl-CoA reductase, cinnamyl-alcohol dehydrogenase and peroxidase are the key genes for lignin synthesis and accumulation. The process involves the phenylpropanoid biosynthesis pathways.

Materials and Methods

Plant material

B. spectabilis grows in Nanning, Guangxi Province in China.

Transcriptome analysis

In different stages of thorns formation on the stems, the thorns and stems of *B. spectabilis* are collected for transcriptome analysis. We choose 3 stages for analysis. The early stage of thorns formation, the stage in which thorns has been harden, and the stage that is between them.

Library preparation for Transcriptone sequencing

A total amount of 1.5 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First strand cDNA was synthesized using random hexamer primer and MMuLV Reverse Transcriptase (RNase H). Second strand cDNA synthesis was subsequently performed

using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3’ ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150 ~ 200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Clustering and sequencing

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer’s instructions. After cluster generation, the library preparations were sequenced on an Illumina Hiseq platform and paired-end reads were generated.

Data Analysis

Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and reads whose quality is low from raw data. At the same time, Q20, Q30 and GC content the clean data were calculated. All the downstream analyses were based on the clean data with high quality.

The left end (read1 files) of all fastq were pooled into one big left.fq file, and right end (read2 files) into one big right.fq file. Transcriptome assembly was accomplished based on the left.fq and right.fq using Trinity[9] with min kmer cov set to 2 by default and all other parameters set default. The assembled transcripts were hierarchically clustered to unigenes using shared reads and expression by Corset[10].

Gene function was annotated based on the NCBI non-redundant protein and nucleotide sequences. Gene function was also annotated by the Pfam (Protein family), KOG/COG (Clusters of Orthologous Groups of proteins), Swiss-Prot (A manually annotated and reviewed protein sequence database), KEGG Ortholog database and Gene Ontology.

SNP calling

Picard-tools v1.41 and samtools v0.1.18 were used to sort, remove duplicated reads and merge the bam alignment results of each sample. GATK3 software was used to perform SNP calling. Raw vcf files were filtered with GATK standard filter method and other parameters (cluster:3;WindowSize:35; QD < 2.0 or FS > 60.0 or MQ < 40.0 or SOR > 4.0 or MQRankSum < -12.5 or ReadPosRankSum < -8.0 or DP < 10).

SSR detection and primer design

SSR of the transcriptome were identified using MISA (http://pgrc.ipk-gatersleben.de/misa/misa.html), and primer for each SSR was designed using Primer3 (http://primer3.sourceforge.net/releases.php).

Quantification of gene expression levels Gene expression levels were estimated by RSEM[11] for each sample: 1. Clean data were mapped back onto the assembled transcriptome 2. Readcount for each gene was obtained from the mapping results.

Differential expression analysis

Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq (Differential expression sequence) R package (1.18.0). DESeq provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value < 0.05 found by DESeq were assigned as differentially expressed.

Prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two conditions was performed using the DEGSeq R package (1.20.0). The P values were adjusted using the Benjamini & Hochberg method. Corrected P-value of 0.005 and log2 (Fold change) of 1 were set as the threshold for significantly differential expression.

GO and KEGG enrichment analysis of differentially expressed genes

Gene Ontology (GO) enrichment analysis of differentially expressed genes was implemented by the GOseq R package, in which gene length bias was corrected. GO terms with corrected Pvalue less than 0.05 were considered significantly enriched by differential expressed genes. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-through put experimental technologies (http://www.genome.jp/kegg/). We used KOBAS software to test the statistical enrichment of differential expression genes in KEGG pathways.

PPI analysis of differentially expressed genes

PPI analysis of differentially expressed genes was based on the STRING database, which known and predicted Protein-Protein Interactions. For the species existing in the database, we construct the networks by extract the target gene list from the database; Otherwise, Blastx (v2.2.28) was used to align the target gene sequences to the selected reference protein sequences, and then the networks are built according to the known interaction of selected reference species.

RT-qPCR verification

To collect the thorns and stems of *B. spectabilis* in 3 stages for RT-qPCR (Reverse transcription fluorogenic quantitative Polymerase Chain Reaction) verification. The total RNA of each sample are extracted by HiPure HP Plant RNA Mini Kit (Magen Biotechnology Co., Ltd, R4165-02). The cDNA are

reverse transcribed from mRNA by HiScript II Q RT SuperMix for qPCR (+ gDNA wiper) (Vazyme Biotech Co., Ltd, R223-01). The cDNA and ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech Co., Ltd, Q711-02) are used for qPCR. Real time quantitative PCR analysis is conducted utilizing qTOWERE3 (AnalytikJena, German). The primers used in RT-qPCR are in the Table 2.

Table 1
List of data output quality

Sample	Raw Reads	Clean Reads	Error(%)	Q20(%)	Q30(%)	GC Content(%)
C1_1	46997486	46652954	0.03	97.38	93.08	43.9
C1_2	49987188	49659918	0.03	97.39	93.05	44.64
C1_3	47513884	47221972	0.03	97.38	93.01	44.68
C2_1	47716478	47405948	0.03	97.42	93.11	44.56
C2_2	45417610	45193820	0.03	97.34	92.87	44.71
C2_3	48635900	48360212	0.03	97.35	92.94	44.71
C3_1	45717148	45421208	0.03	97.34	92.95	44.4
C3_2	44137788	43892366	0.03	97.32	92.87	44.4
C3_3	47237234	46963976	0.03	97.37	92.94	44.29
J1_1	43726332	43502176	0.03	97.36	92.95	44.53
J1_2	45084210	44801296	0.03	97.28	92.84	44.65
J1_3	47441554	47145802	0.03	97.49	93.23	44.58
J2_1	46457002	46209204	0.03	97.38	92.98	44.39
J2_2	46184660	45934398	0.03	97.35	92.93	44.46
J2_3	41266290	41059518	0.03	97.46	93.19	44.68
J3_1	42274290	42065976	0.03	97.56	93.34	44.31
J3_2	43887862	43646514	0.03	97.56	93.38	44.61
J3_3	46624754	46386220	0.03	97.42	93.13	44.39

Table 2
The primers used in RT-qPCR

Gene name	Primers name	Sequence(5'to3')
actin	actin F	TAGACCCTCCTATCCAAACA
	actin R	TTTTCCAGCCTTCACTTATC
PAL	c59052_g5_i1 F	CTTGAGCCACCGTGAGAGTT
	c59052_g5_i1 R	ACCACCACCAGCACCAGAA
CYP73A	c49534_g1_i1 F	CCAGATAACAGAGCCAGACACA
	c49534_g1_i1 R	CCACCAGGCATTCACCAGAA
4CL	c77978_g1_i0 2F	GTGGCACAACAAGTGGATGG
	c77978_g1_i0 2R	GCGTAGAGCACACAGCAGTA
CCR	c48436_g0_i0 F	ACAACCAAGCCACCACTACC
	c48436_g0_i0 R	ACCATCCATGAAGCAATGAACC
CAD	c35663_g1_i0 F	TTGGTGTGATCGTTGGATGTTG
	c35663_g1_i0 R	CGCTGCTTGTTCTGATGCC
peroxiddase	c53159_g0_i0 F	AAGCAACCAGGCAGAGAAGG
	c53159_g0_i0 R	ATCAGCACAAGAGACGATTCCT

Determination of lignin content

The samples were dried at 80°C until constant weight, pulverized, passed through a 40-mesh sieve, and weighed for a certain amount (denoted as W). Acetylation of phenolic hydroxyl groups of lignin in samples. Following the acetylation of phenolic hydroxyl groups within lignin, a discernible absorption peak at 280 nm becomes evident. Notably, the absorbance measurement at 280 nm exhibits a direct positive correlation with the lignin content present. In this study, lignin content was characterized by the absorbance value of 280 nm. The formula for converting 280nm absorbance to Lignin content is as follows: Lignin (mg/m) = (ΔA-0.0068) ÷ 0.0347×V×10 − 3 ÷ W×T = 0.0294×(ΔA-0.0068) ÷ 0.002×50. (V: Total volume of the reaction; W: Sample quality (Dry weight); T: Dilution factor; A: Measurement of absorbance at 280nm; ΔA = Measurement of absorbance at 280nm of the sample- Absorbance value at 280nm for blank control).

Declarations

Ethics approval and consent to participate

All experiments on plants were performed according to the Guangxi Zhuang Autonomous Region Forestry Research Institute, national, and international guidelines, and legislation.

Consent for publication

Not applicable.

Availability of data and materials

All data generated and analyzed during this study are available from the corresponding author on reasonable request.

Competing interests

The authors declare no competing interests.

Funding

1. Natural Science Foundation of Guangxi Science and Technology Department (Project No.: 2020GXNSFAA297064)

2. Guangxi Forestry Science and Technology Promotion Demonstration Project (Contract No.: Guilin Scientific Research [2022] No. 11)

3. Guangxi Forestry Science and Technology Promotion Demonstration Project (Contract No.: 2023GXLK29).

Authors' contributions

Sun Lina and Lin Mao is responsible for the planning and design of the experiment. Sun Lina is responsible for the implementation of the experiment and the writing of the paper. Lu Zhixiang and Chen Qi Provide suggestions for the writing of the paper. Wang Xinhua, Li Jinhua, Gong Jianying, Yang Shuting, Yang Kaitai, Chen Er, Li Bing participated and assisted in some experiments.

Acknowledgements

Not applicable

Data Availability Statement

The data that support the findings of this study are available from the corresponding author, [author initials], upon reasonable request.

References

1. Flora of China., vol. 5. Beijing: Science Press; 2003.
2. Lim TK. Bougainvillea spectabilis. Edible Medicinal and Non Medicinal Plants: Volume 8, Flowers. Dordrecht: Springer Netherlands; 2014: 489–96.
3. Bautista MAC, Zheng Y, Boufford DE, Hu Z, Deng Y, Chen T. Phylogeny and Taxonomic Synopsis of the Genus Bougainvillea (Nyctaginaceae). Plants. 2022;11(13):1700.
4. Yanjing H. ISSR analysis of germplasm resources of *Bongainvillea Brasiliensis* Raeusch. Fuzhou: Fujian Agriculture and Forestry University; 2010.
5. Sun Lina LM, Li J, Qin T, Dingjian L. Liao Meilan: Comprehensive Evaluation on Ornamental Value of 30 Species of *Bougainvillea glabra* in Guangxi. Agricultural Res application. 2016;6:1–6.
6. Yoon J, Choi H, An G. Roles of lignin biosynthesis and regulatory genes in plant development. J Integr Plant Biol. 2015;57(11):902–12.
7. Nakano Y, Yamaguchi M, Endo H, Rejab NA, Ohtani M. NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. Front Plant Sci. 2015;6:288.
8. Behr M, Guerriero G, Grima-Pettenati J, Baucher M. A Molecular Blueprint of Lignin Repression. Trends Plant Sci. 2019;24(11):1052–64.
9. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644–52.
10. Davidson NM, Oshlack A. Corset: enabling differential gene expression analysis for de novo assembled transcriptomes. Genome Biol. 2014;15(7):410.
11. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.

Figures

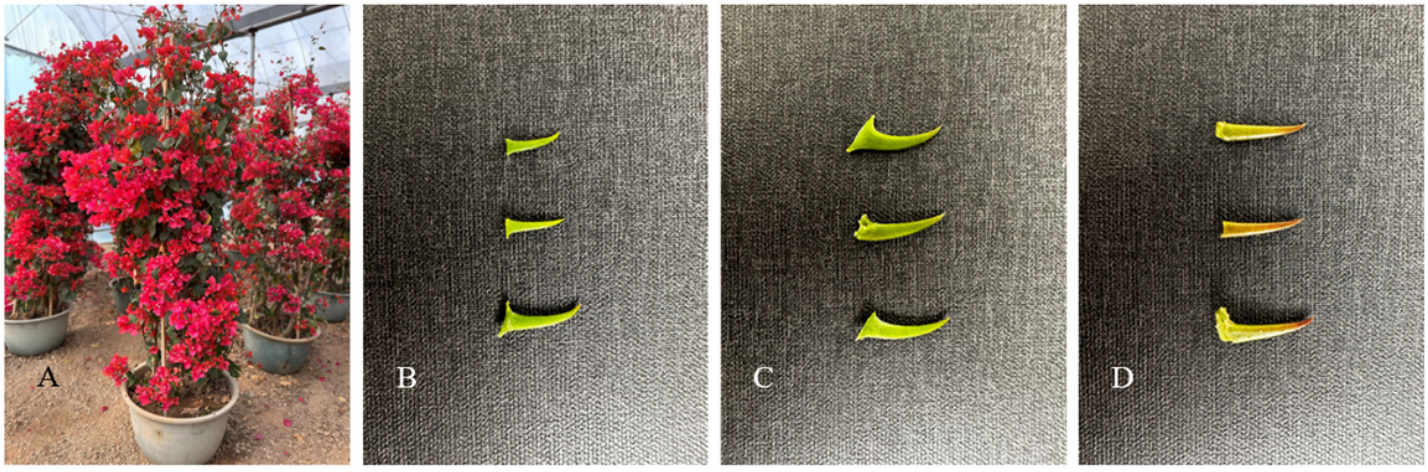

Figure 1

The plants of *B. spectabilis* at 3 stages from thorns formation to thorns hardening

A. The plants of *B. spectabilis*; B. The thorns at formation stage; C. The thorns at turning hard stage; D. The thorns at hardened stage.

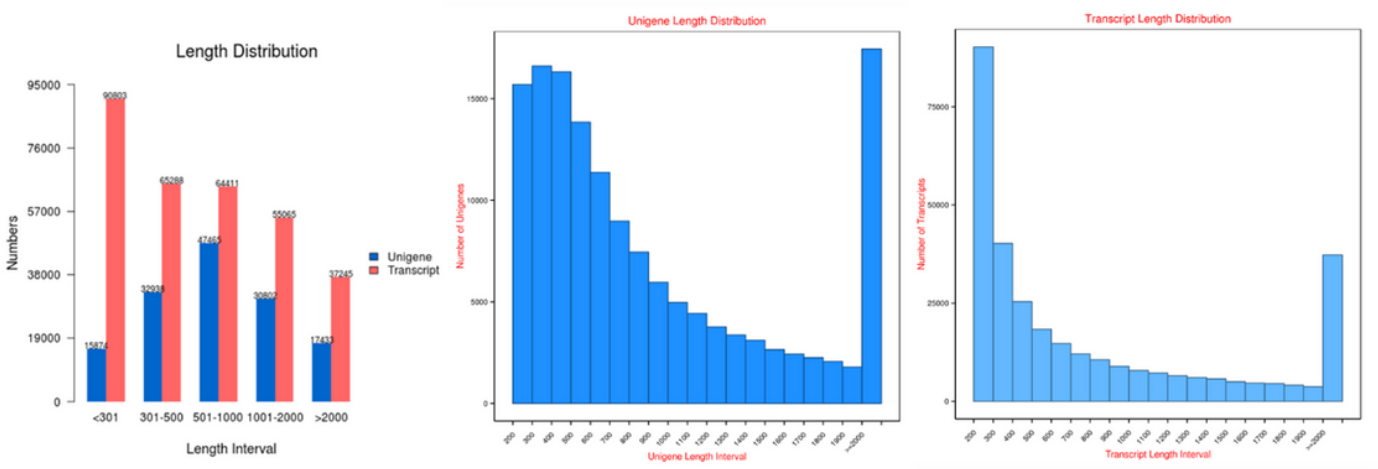

Figure 2

Spliced transcripts and gene sequence length distribution map

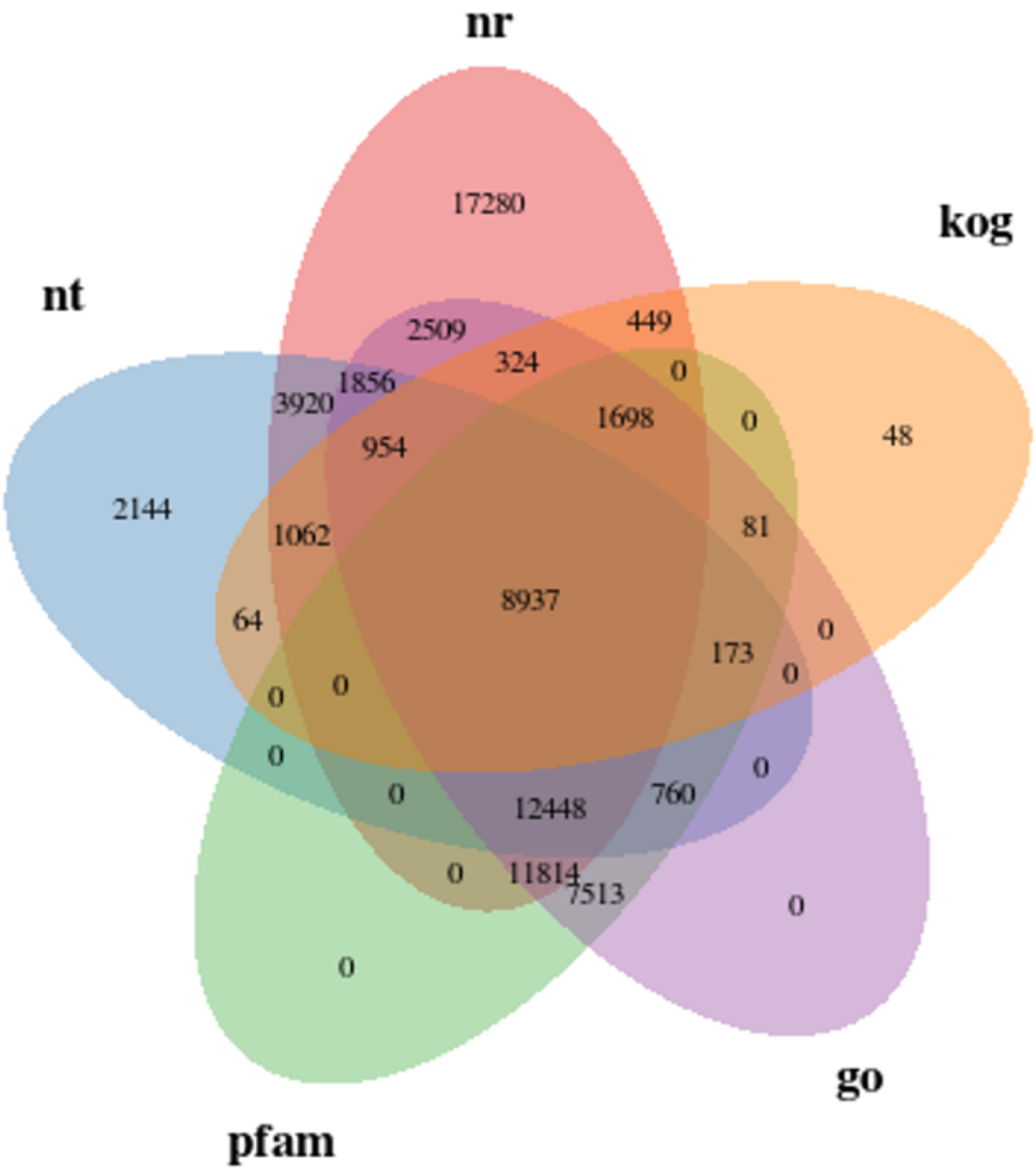

Figure 3

Number of genes annotated by each database

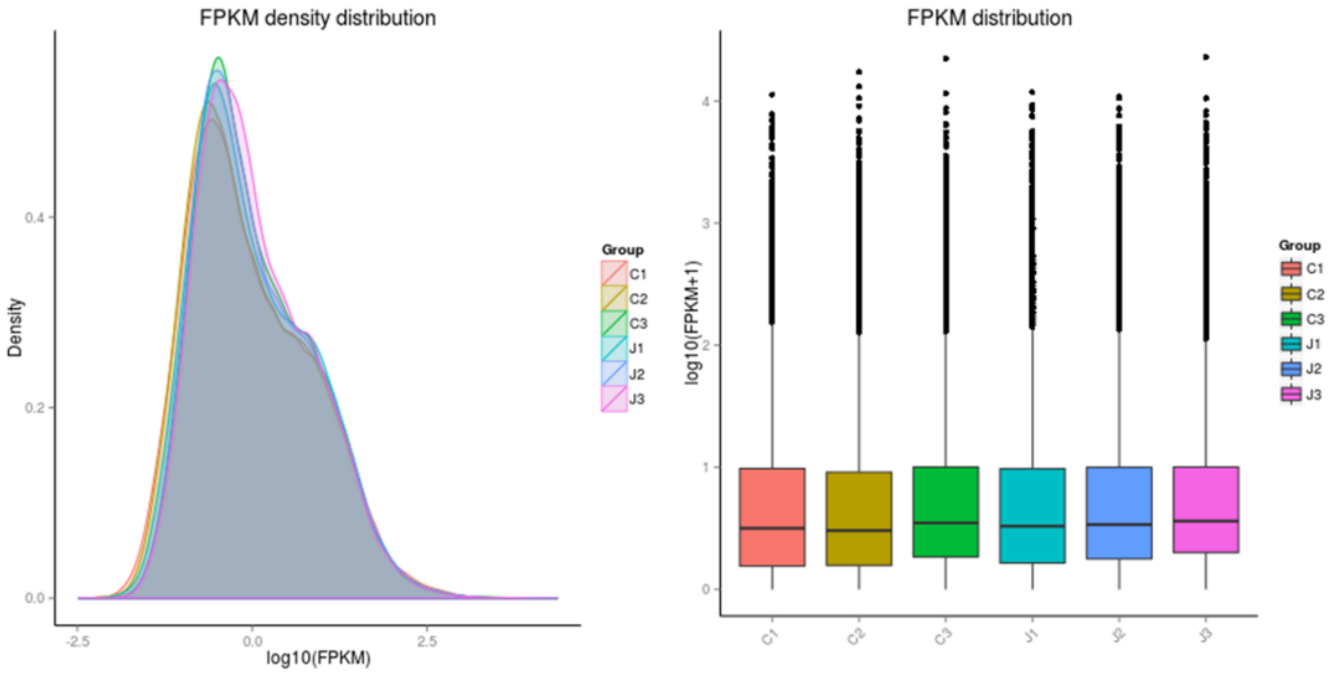

Figure 4

Comparison chart of gene expression levels in different parts in different periods

Note: C1: The thorns at formation stage; C2: The thorns at turning hard stage; C3: The thorns at hardened stage; J1: The stems at thorns formation stage; J2: The stems at thorns turning hard stage; J3: The stems at thorns hardened stage.

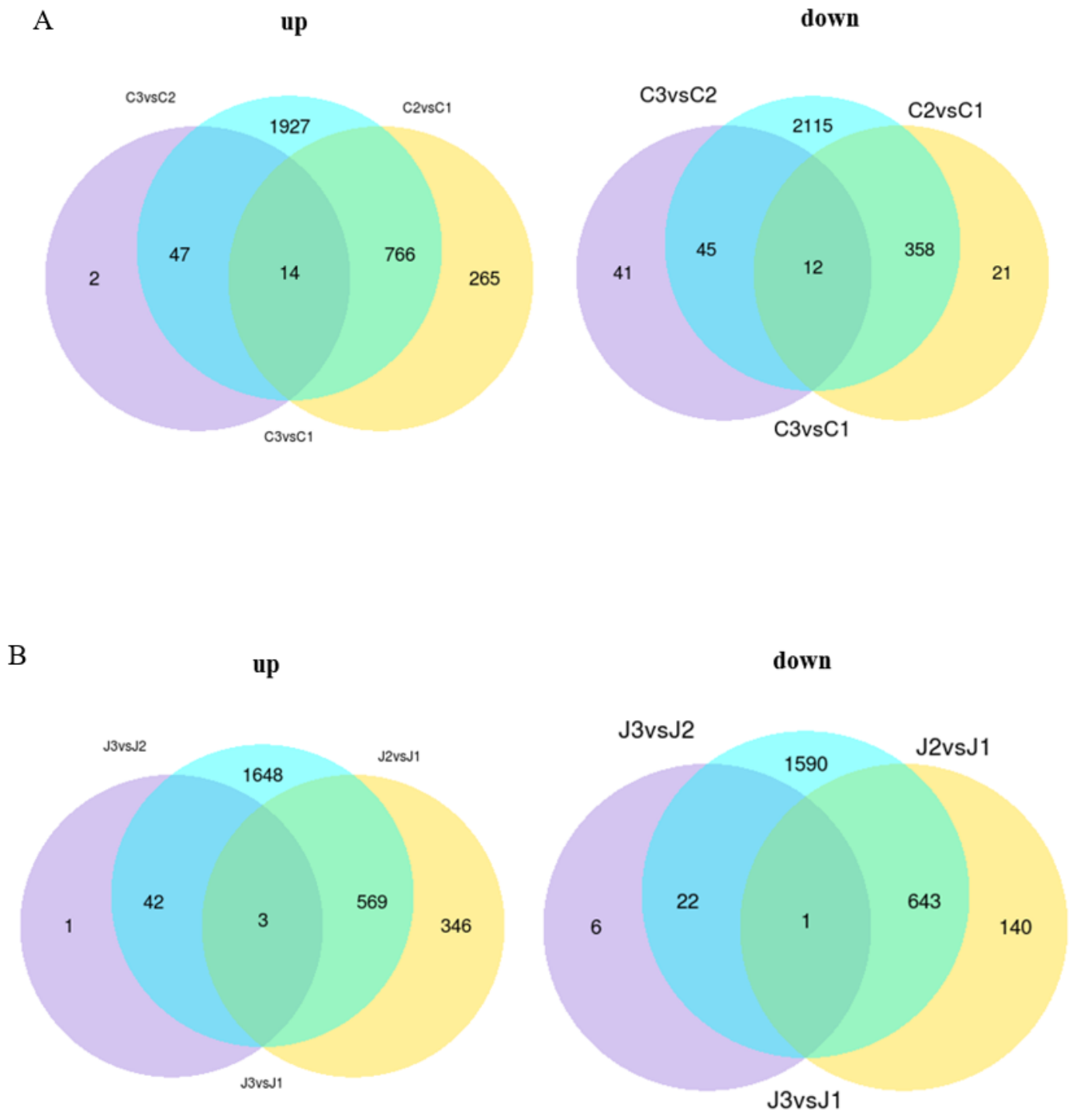

Figure 5

The differences in gene expression at 3 stages

A. The differences in gene expression of thorns; B. The differences in gene expression of stems.

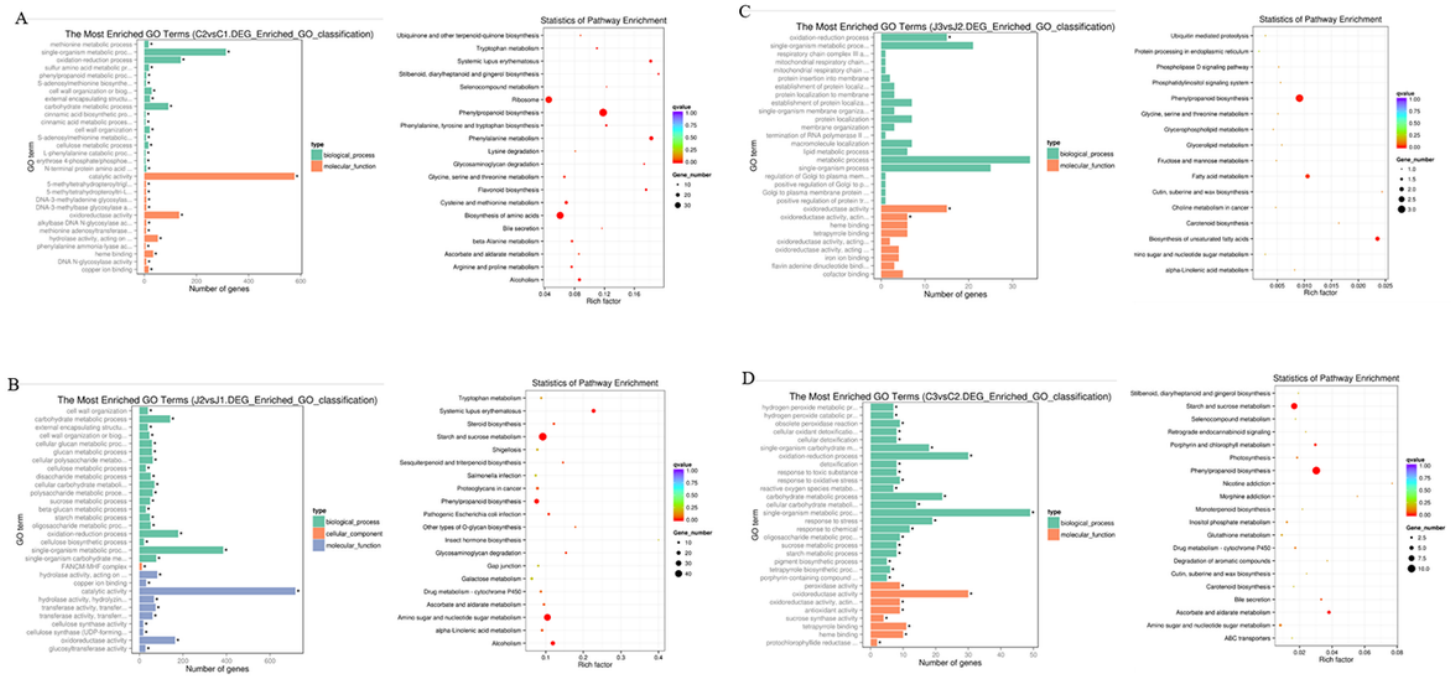

Figure 6

GO Gene enrichment analysis

A. GO Gene enrichment analysis in C2 VS C1; B. GO Gene enrichment analysis in J2 VS J1; C. GO Gene enrichment analysis in C3 VS C2; D. GO Gene enrichment analysis in J3 VS J2.

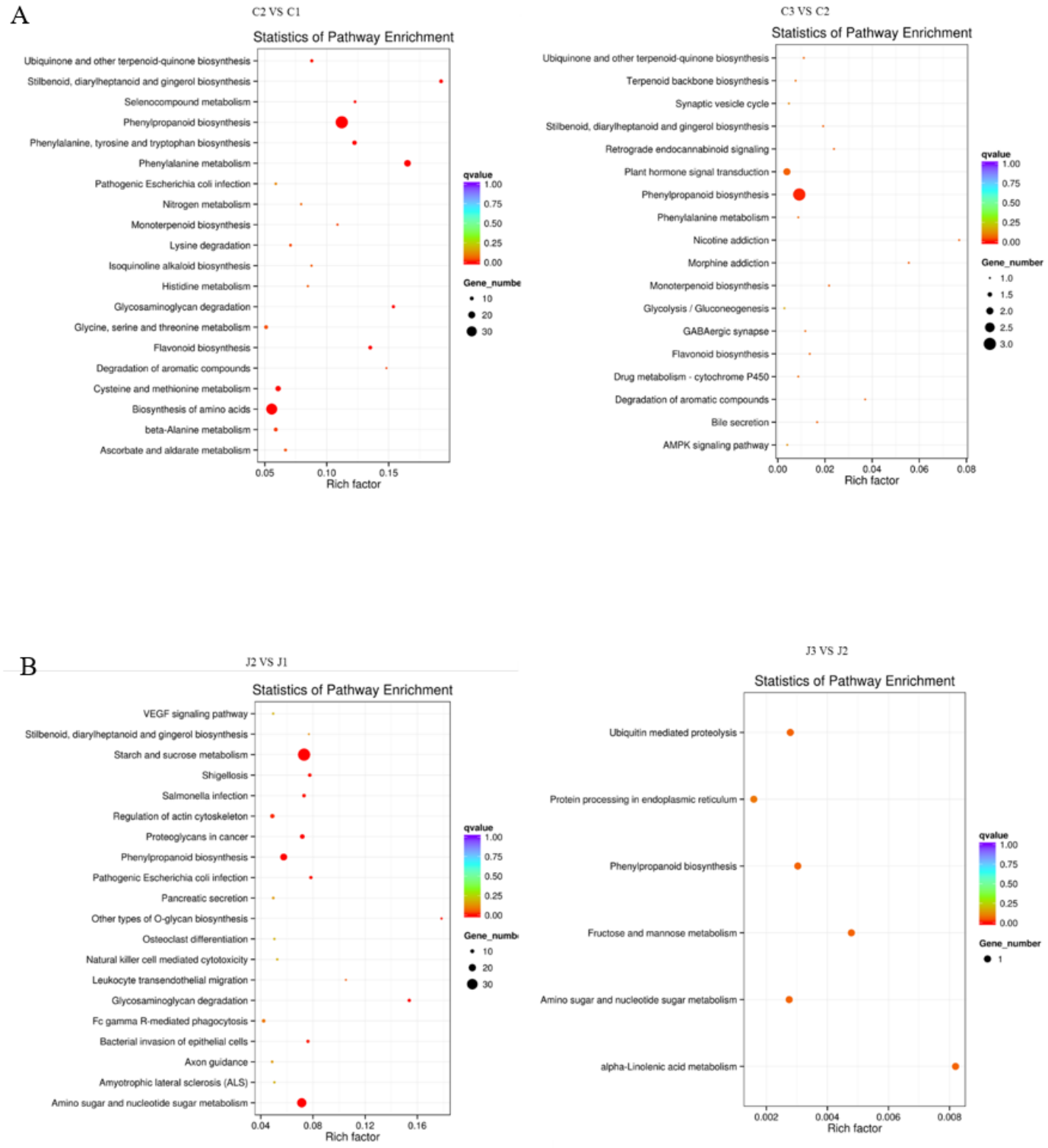

Figure 7

KEGG Gene enrichment analysis in stems

A. KEGG Gene enrichment analysis in thorns; B. KEGG Gene enrichment analysis in stems.

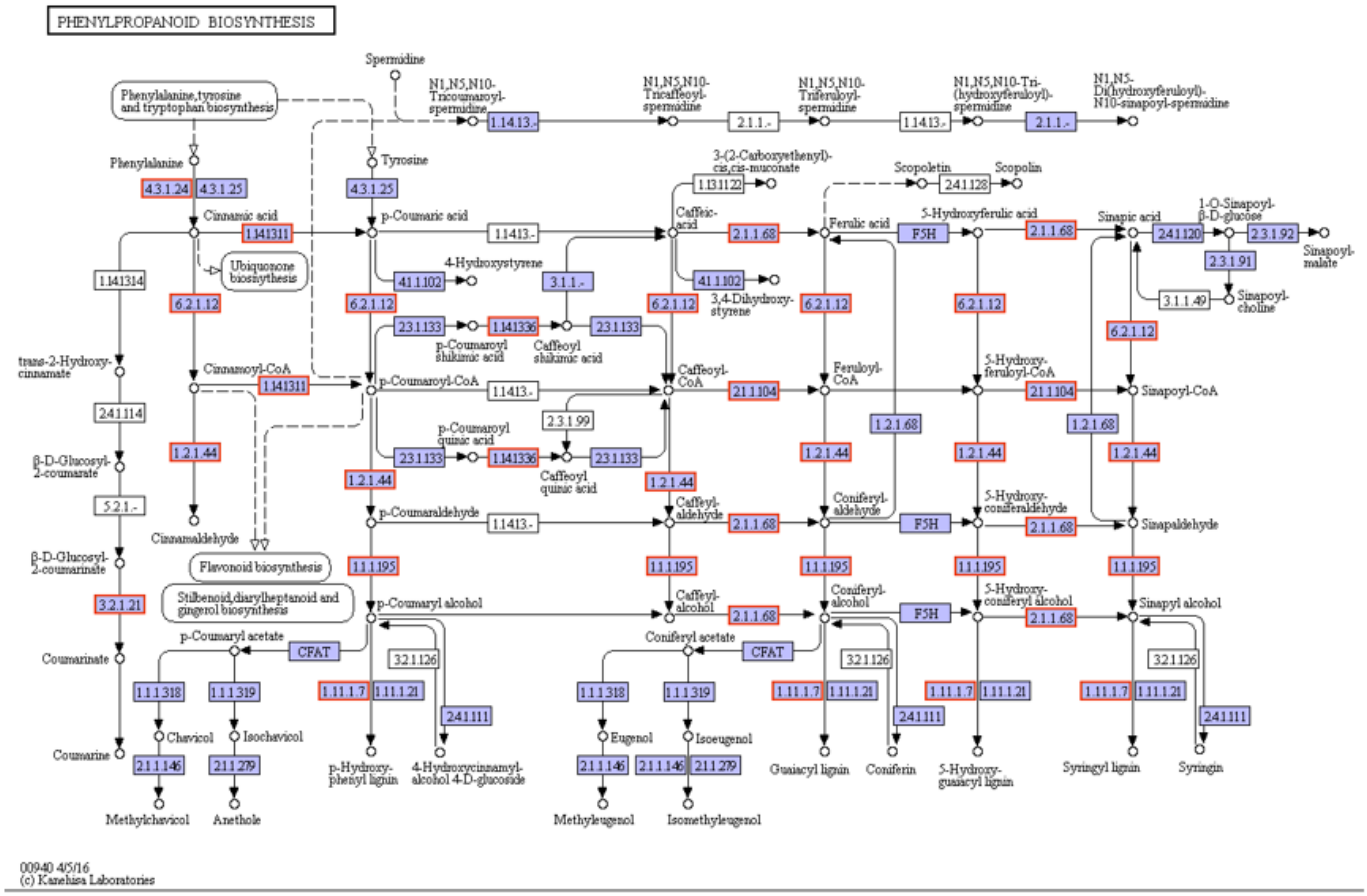

Figure 8

The up-regulated genes in phenylpropanoid biosynthesis (C2 VS C1)

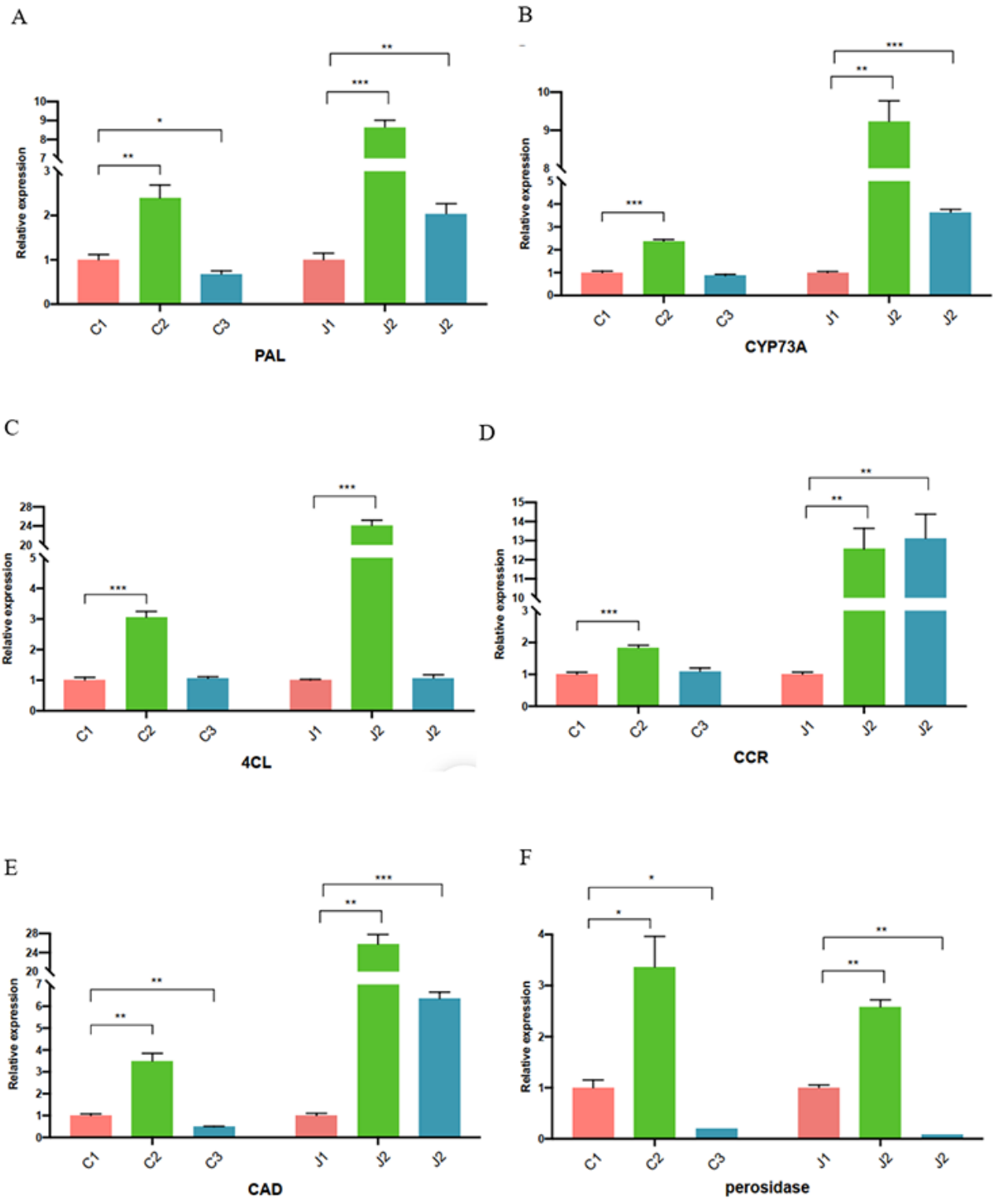

Figure 9

RT-qPCR verification

A. The expression levels of PAL gene in thorns and stems at 3 stages; B. The expression levels of CYP37A gene in thorns and stems at 3 stages; C. The expression levels of 4CL gene in thorns and stems at 3 stages; D. The expression levels of CCR gene in thorns and stems at 3 stages; E. The expression levels of

CAD gene in thorns and stems at 3 stages; F. The expression levels of peroxidase gene in thorns and stems at 3 stages.

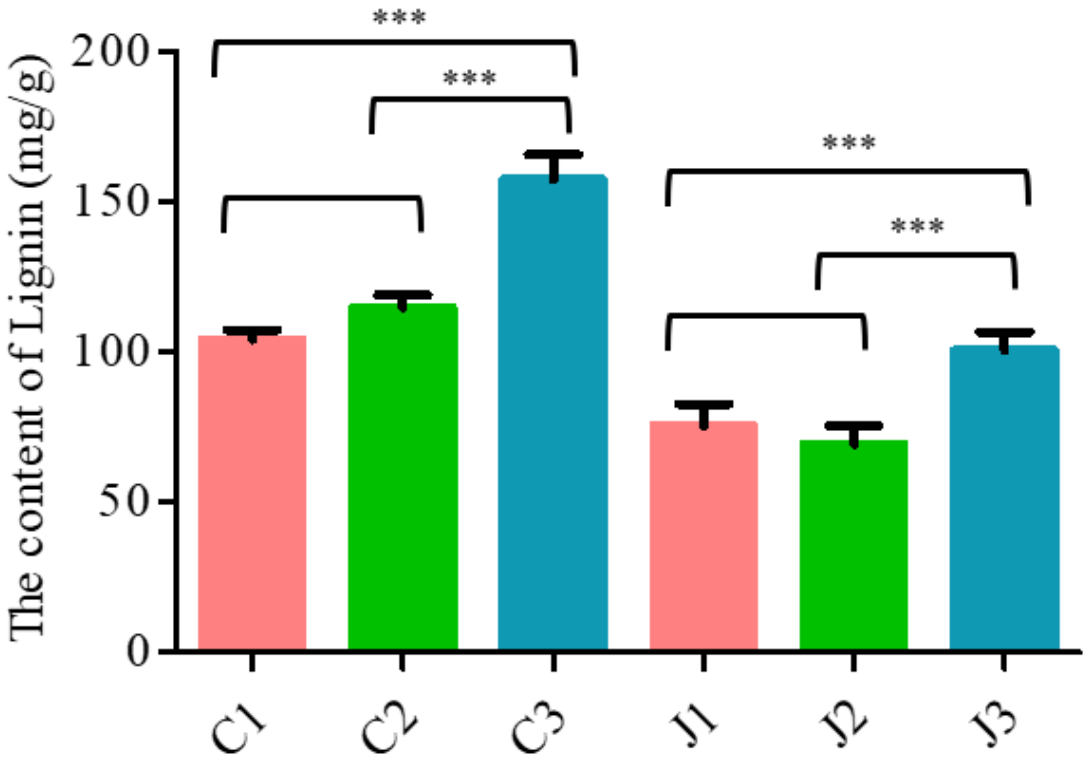

Figure 10

The content of lignin in each part of each stage

Supplementary Files

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

- AdditionalTables.docx